



MANUAL METODOLÓGICO

VII EME Encuesta de Microemprendimiento 2022



SUBDIRECCIÓN TÉCNICA

DEPARTAMENTO DE ESTADÍSTICAS DEL TRABAJO

SUBDEPARTAMENTO DE ESTADÍSTICAS ESTRUCTURALES DEL TRABAJO

DEPARTAMENTO DE METODOLOGÍA E INNOVACIÓN ESTADÍSTICA

Manual Metodológico VII Encuesta de Microemprendimiento

Junio 2023

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

Correo electrónico: ine@ine.cl

Santiago de Chile

Índice

1. INTRODUCCIÓN.....	6
2. ANTECEDENTES DEL ESTUDIO.....	8
2.1. Objetivos del producto estadístico.....	9
2.1.1. Objetivo general.....	9
2.1.2. Objetivos específicos de la encuesta.....	9
3. MARCO DE REFERENCIA.....	11
3.1. Marco conceptual.....	11
3.1.1. Marco conceptual emprendedores.....	11
3.1.2. Economía informal.....	12
3.1.2.1. El Sistema de Cuentas Nacionales (SCN 2008) y el sector informal.....	13
3.1.2.2. Sector informal y Economía no observada.....	15
3.2. Referencias nacionales e internacionales.....	16
4. DISEÑO ESTADÍSTICO.....	19
4.1. Universo.....	19
4.2. Población objetivo.....	19
4.3. Marco estadístico.....	19
4.3.1. Marco muestral de la EME.....	19
4.3.1.1. Cobertura geográfica.....	19
4.3.1.2. Estratificación geográfica.....	20
4.3.1.3. Conformación del marco muestral.....	20
4.3.1.4. Estratificación del marco muestral.....	23
4.4. Cobertura Geográfica.....	24
4.5. Unidades estadísticas.....	24
4.6. Definición de variables.....	24
4.7. Nomenclaturas utilizadas.....	25
4.8. Diseño estadístico por muestreo.....	26
4.8.1. Diseño muestral.....	26
4.8.2. Cálculo y distribución del tamaño muestral.....	27
4.8.3. Selección de unidades muestrales.....	32
4.8.3.1. Selección de viviendas.....	32
4.8.3.2. Selección de informantes.....	33
4.8.3.3. Cambios entre muestra diseñada y muestra seleccionada.....	33
4.9. Medidas de precisión.....	34
4.9.1. Medidas de precisión proceso de recolección.....	34
4.9.2. Medidas de precisión proceso de análisis.....	40

4.10.	Estimación de parámetros.....	41
4.10.1.	Formula estimadores	42
5.	METODOLOGÍA DE RECOLECCIÓN DE DATOS.....	43
5.1.	Técnica y Métodos de recolección de datos	44
5.1.1.	Técnicas de recolección de datos.....	44
5.1.2.	Medio de recolección de información utilizado.	44
5.1.3.	Período de referencia.....	44
5.1.4.	Fechas de recolección de datos y procesos complementarios.....	45
5.1.5.	Equipo de trabajo.....	47
5.1.6.	Termino proceso de recolección.	47
5.2.	Diseño del instrumento de captura de información.....	47
5.3.	Método de capacitación para la recolección de datos.....	50
5.3.1.	Capacitaciones encuestadores, encargados de grupos y supervisores.....	50
5.4.	Métodos de supervisión.....	50
5.4.1.	Supervisión directa	51
5.4.2.	Supervisión indirecta.....	51
6.	METODOLOGÍA DE PROCESAMIENTO DE DATOS.....	52
6.1.	Integrar bases de datos.....	52
6.2.	Codificación automática.....	53
6.2.1.	Codificación CAENES y CIUO.....	53
6.2.1.1.	Proceso de codificación automática	53
6.2.1.2.	Selección de casos a auditar	53
6.2.2.	Codificación y reclasificación especifique.....	54
6.3.	Revisar y validar	54
6.3.1.	Nominación y tratamiento de valores atípicos.	54
6.3.1.1.	Detección	54
6.3.1.2.	Tratamiento valores outliers.....	55
6.3.2.	Grilla de suficiencia	55
6.3.3.	Validaciones de primer y segundo nivel.....	56
6.4.	Edición e imputación	56
6.4.1.	Imputaciones provenientes de validaciones de primer nivel	56
6.4.2.	Imputaciones provenientes de validaciones de segundo nivel	57
6.4.3.	Imputaciones provenientes de prueba piloto.....	57
6.5.	Derivación de nuevas variables	57
6.5.1.	Depuración Gastos, Ingresos y Ganancias (GIG).....	57

6.5.1.1.	Análisis de respuestas en GIG	58
6.5.1.2.	Objetivo del proceso de depuración	60
6.5.1.3.	Depuración y estimación de ganancias (casos 1 y 2).....	61
6.5.1.4.	Imputación de ganancia anual (caso 4)	62
6.5.1.5.	Construcción de variable ganancia final	62
6.5.2.	Depuración variables trabajo no remunerado (TNR)	63
6.6.	Cálculo de ponderadores	64
6.6.1.	Cálculo ponderador base	64
6.6.2.	Ajuste por elegibilidad desconocida	66
6.6.3.	Ajuste por no elegibilidad	66
6.6.4.	Ponderador ajustado por falta de respuesta.....	67
6.6.5.	Suavizamiento del ponderador ajustado por no respuesta	69
6.6.5.1.	Calibración	69
7.	METODOLOGÍA DE ANÁLISIS.....	73
7.1.	Análisis de los resultados.....	73
7.1.1.	Análisis descriptivo	73
7.1.2.	Comparación con fuentes externas	73
7.2.	Preparar resultados y aplicar protocolos de calidad	74
7.2.1.	Preparación de data previo al proceso de anonimización	74
7.2.2.	Implementación del proceso de enmascaramiento.....	76
8.	ALCANCE DE LA PRESENTACIÓN DE RESULTADOS	77
8.1.	Productos estadísticos a publicar	77
8.1.1.	Total de Personas Microempendedoras.....	77
8.1.2.	Proporción de personas microempendedores por cuenta propia.....	77
8.1.3.	Proporción de personas microempendedores informales	77
8.2.	Difusión de los resultados	77
8.2.1.	Bases de datos	78
8.2.2.	Publicaciones y anuarios.....	78
8.2.3.	Metodología.....	78
8.2.4.	Manuales.....	78
8.2.5.	Formularios.....	78
8.2.6.	Metadatos	79
8.3.	Accesibilidad a los datos y marco legal	79
8.3.1.	Marco legal	79
9.	CONCEPTOS CLAVES.....	80

10.	REFERENCIAS BIBLIOGRÁFICAS	87
11.	Anexo 1: Prueba piloto.....	90
Tabla 1:	Sectores institucionales del SCN 2008	14
Tabla 2:	Unidades y actividades de producción	16
Tabla 3:	Experiencias Internacionales en la aplicación de encuestas a micronegocios....	17
Tabla 4:	Cantidad de UPM especiales y excluidas, por región	22
Tabla 5:	Cantidad de estratos de muestreo por región y área	24
Tabla 6:	Criterios utilizados en el cálculo del tamaño muestral.....	30
Tabla 7:	Errores esperados asociados al parámetro de interés según tamaño muestral objetivo.....	31
Tabla 8:	Total de viviendas seleccionadas según región y mes de levantamiento.....	31
Tabla 9:	Muestra seleccionada por región y submuestra	34
Tabla 10:	Clasificación de códigos de disposición final según resultado final en la Hoja de Ruta.....	35
Tabla 11:	Encuestadores y supervisores por región.....	47
Tabla 12:	Cambios cuestionario Prueba Piloto VII EME	49
Tabla 13:	Razones de supervisión indirecta obligatoria.....	51
Tabla 14:	Ausencia de respuesta en preguntas de gastos, ingresos y ganancias	58
Tabla 15:	Ausencia de respuesta en preguntas de gastos, ingresos y ganancias	59
Tabla 16:	Prevalencia por situación de consistencia gastos, ingresos y ganancias	60
Tabla 17:	Listado de Actividades de Trabajo No Remunerado consultados VII EME.....	63
Tabla 18:	Distribución de individuos según la clasificación de elegibilidad.	67
Tabla 19:	Total de unidades elegibles, unidades que responden y tasas de respuesta....	68
Tabla 20:	Total de microemprendedores estimados a partir de la ENE – periodo MAM 2022	70
INDICE DE FIGURAS		
Figura 1:	Relación entre Economía no observada y el Sector Informal	15
Figura 2:	Flujo para la evaluación de la calidad de las estimaciones.....	41
Figura 3:	Dispersión del Ponderador ajustado por falta de respuesta.....	69
Figura 4:	Gráfico de caja para los diferentes ponderadores.....	72
INDICE DE DIAGRAMAS		
Diagrama 1:	Línea de tiempo de trabajo de recolección de datos.....	45
Diagrama 2:	Esquema de participación de los informantes en la ENE y EME.	46
Diagrama 3:	Flujo operativo general del proyecto VII EME.....	46
Diagrama 4:	Etapas de la actualización del cuestionario	48

1. INTRODUCCIÓN

El Instituto Nacional de Estadísticas (INE) en el marco del convenio suscrito con el Ministerio de Economía, Fomento y Turismo, desde el año 2013 tiene a su cargo 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 todas las actividades económicas.

A diferencia de otras encuestas a empresas o negocios, la EME es una encuesta a hogares de segunda fase¹, cuya población objetivo son las personas que trabajan como 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. Esta encuesta se denomina como “Encuesta mixta” ya que permite obtener información de dos unidades de análisis: el microemprendedor y su negocio.

El proceso estadístico de la última versión de la encuesta se expone al público a través del presente. 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 del estudio desde donde se obtienen los criterios que le dan sustento. Además, se incluye experiencia de otros países en esta medición.
- El capítulo IV presenta el diseño muestral de la encuesta, en el 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 del proceso de recolección 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 de campo.
- El capítulo VI de metodología de procesamiento de datos, ahonda en la 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

¹ 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.

ponderadores.

- En el capítulo VII, metodologías de análisis, se describen los métodos estadísticos utilizados para analizar e interpretar los datos. Junto al proceso de anonimización de las bases de datos.
- Finalmente, en el último capítulo, se exponen los productos y cómo acceder a ellos.

2. ANTECEDENTES DEL ESTUDIO.

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 INE, el diseño, recolección, procesamiento y análisis de la EME, uno de sus principales instrumentos de aproximación a la realidad de los microemprendedores en Chile. Esta encuesta se realiza de forma bienal desde el año 2009.

La 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 visibilizados por las encuestas tradicionales a empresas.

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

La EME se aplica a un subuniverso de la Encuesta Nacional de Empleo (ENE) correspondiente a los microemprendedores. Para las dos primeras versiones de la EME, la diferencia en la aplicación entre la ENE y la EME era superior a un año, lo que incidía en una mayor tasa de no respuesta y elevados cambios en la situación ocupacional de los informantes. Para poder solucionar este problema, a partir del año 2013 el INE se hace cargo del proceso estadístico de la encuesta, reduciendo el tiempo entre encuestas a dos meses.

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 contribuyó con la disminución de la tasa de rechazo y el desgaste de los informantes, mejorando las tasas de respuesta y logro de la encuesta.

Entre las modificaciones al cuestionario, destaca 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 de este tipo de unidades 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 estimaciones² 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 microempresaria. La estructura de la información recopilada se repite para cada uno de los empleos generados por el informante, por ello desde la V versión del estudio, 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 microempresario.

En la VI EME, se incorporó 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 aproximarse al tiempo destinado a las actividades de trabajo de cuidado y trabajo doméstico no remunerado.

En la versión actual se incorporaron preguntas referidas a la pandemia del COVID-19 y al uso de tecnologías de la información. Lo que permite capturar y estudiar el impacto de la pandemia del COVID-19 en el funcionamiento del microempresariado, sus ganancias, sus trabajadores y la persistencia en el tiempo, y además poder estimar la propagación del uso de tecnologías de información y la prevalencia de estas.

2.1. Objetivos del producto estadístico

2.1.1. Objetivo general

Caracterizar los microempresarios 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.

2.1.2. Objetivos específicos de la encuesta

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

1) Observar las formas en las que operan y funcionan los micronegocios.

Identificar la escala de funcionamiento de los negocios y estudiar la organización y planificación del trabajo de los microempresarios, como también los factores utilizados en el proceso productivo (insumos, maquinarias, trabajadores, entre otros) y la capacidad de poder realizar una separación los factores de capital y de trabajo.

² 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.

2) 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 en detallar la situación de formalidad de los negocios y sus dueños.

3) 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.

4) Determinar los recursos productivos de los micronegocios.

Otra virtud de la encuesta es que permite determinar el patrimonio de los microemprendedores, debido a que captura información de herramientas, equipos y vehículos que los micronegocios destinan para fines de su actividad económica. De tal manera que se puede saber sus valores y posibles inversiones futuras en nuevos insumos.

5) Estimar los ingresos y gastos de los micronegocios.

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.

6) Identificar la generación de empleo de las microempresas y sus características.

Algunos de los aspectos que permite estudiar la EME son 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.

7) Estudiar el impacto del COVID-19 en los micronegocios.

En esta versión de la EME, se consideran una serie de preguntas que permiten indagar sobre impactos del COVID-19 en el funcionamiento del microemprendimiento, sus ganancias, sus trabajadores y la persistencia en el tiempo.

3. MARCO DE REFERENCIA

El siguiente apartado contiene el marco conceptual que sustenta el estudio sobre 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, la parte de ellos que pertenece al sector informal. De esta manera en este subapartado se expone sobre los conceptos de emprendedores y su lugar en la población ocupada en particular en el caso de Chile, y en segundo lugar algunos de los conceptos que conjugan 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,2018). 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 asalariados 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 una remuneración. 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 quienes 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, se decidió restringir el análisis a un subconjunto de los trabajadores independientes, que son los que trabajan en una unidad económica de hasta 10 trabajadores denominada como microempresa. Esto debido a que la mayor parte de ellos (97,6% utilizando datos de la IV EME), corresponde a microemprendedores, además de la baja comparabilidad de las estimaciones de los emprendedores de mayor impacto con otras fuentes de información. Por ejemplo, según datos de la IV EME los emprendedores dueños de grandes empresas generaban alrededor de 500.000 empleos mientras que según registros administrativos (Formulario 1887) este grupo generó casi 5 millones de empleos (alrededor de 10 veces más).

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” (OIT, 2002).

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 de actividad ante las autoridades tributarias nacionales. Estas se desarrollan frecuentemente 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 o beneficio. Este amplio abanico de características conjuga al sector informal y constituyen las razones por las cuales muchas veces, las 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 carece de 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 áreas beneficia el cumplimiento de sus objetivos.

3.1.2.1. El Sistema de Cuentas Nacionales (SCN 2008) y el sector informal.

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 la siguiente tabla:

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 (Empresas del sector informal).		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 cuasisociedades³. 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 primordialmente para la generación de ocupación e ingresos y no se encuentran constituidas en sociedad, excluidas las cuasisociedades.

El Sistema de Cuentas Nacionales (2008), aborda a través de uno de sus capítulos, aspectos relacionados con la economía informal. En él, 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.

³ 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).

3.1.2.2. 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 sólo 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.

En muchos países es posible encontrar los distintos grupos ejemplificados en la figura 1. Si bien gran parte de las empresas del sector informal pueden tener algún tipo de registro, a partir del pago de algún impuesto o la obtención de un permiso, es posible que no cumplan todas las exigencias administrativas y legales, en este caso, estas empresas quedarían clasificadas en el 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 el Manual de Informalidad “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, 2013a). 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

A nivel regional existen varias 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 ocupación. En cambio, la encuesta mixta independiente implica el levantamiento de una encuesta adicional que indague 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.

Las recomendaciones internacionales promueven la realización de este último tipo para no sobrecargar y complicar el levantamiento de la encuesta base. Sin embargo, muchos países

prefieren optar por una encuesta mixta modular para reducir costos. En el caso de la EME se optó por realizar una encuesta mixta independiente para no sobrecargar el trabajo de la ENE.

A partir de las experiencias revisadas se puede apreciar que no existen una definición única de la población objetivo, ni tampoco del diseño y del marco muestral. Algunos 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 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 los Hogares	Independiente	Unidad económica perteneciente al sector de los hogares	2012	2022
3	Colombia	Encuesta de Micronegocios	Independiente	Unidad económica de hasta 9 personas ocupadas	2019	2022
4	México	Encuesta Nacional de Micronegocios	Independiente	Unidad económica no agrícolas de hasta 10 o 15 asalariados	1993	2012
5	Perú	Encuesta Nacional de Hogares	Modular	Unidad económica del sector informal	2012	2023

Fuente: Elaboración propia, año 2023

Una de las experiencias más longevas a nivel internacional corresponde a la Encuesta de Micronegocios de México (ENAMIN) realizada desde 1992 hasta 2012 con una periodicidad bianual. La ENAMIN es una encuesta mixta independiente aplicada a una subpoblación de la Encuesta Nacional de Ocupación y Empleo (ENOE), la cual abarca todas las actividades no agrícolas, sin prejuizar si ellas son formales o informales⁴. Esta encuesta se ha considerado en varias oportunidades en la EME en la fase de diseño del cuestionario y también de la población objetivo. Al igual que la ENAMIN, la EME realiza una encuesta mixta independiente. Además, en el año 2013 al momento de rediseñar el cuestionario de la EME, la experiencia de la ENAMIN sirvió como base para reestructura el cuestionario, diseñar los módulos y contenidos de ellos. Por ejemplo, Los actuales módulos de “Situación ocupacional”, “Antecedentes laborales”, “Situación de formalidad” y “Generación de empleo” fueron creados en gran parte a partir de esta encuesta. Por otra parte, en la V EME (2017) se efectuó un gran cambio en la población objetivo desde

⁴ La ENAMIN 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).

emprendedores a microemprendedores, esto permitió incorporar una planilla que indaga en una serie de características de los trabajadores, estos cambios también se basaron en la experiencia de la ENAMIN.

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 (ENAH), desde la que se extrae el marco muestral de la ENAMEH⁵. La utilidad de la ENAMEH para la EME se relaciona con la reestructuración de las preguntas de gastos ingresos y ganancias realizadas en la VI versión de la encuesta. El orden de las preguntas proviene de la experiencia de esta encuesta, como también la separación de las preguntas de ingresos agrícolas y no agrícolas y la incorporación de un mayor detalle de los gastos del negocio.

Finalmente, una de las experiencias más recientes y completas 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 sólo a trabajadores independientes del sector informal, logran estimar la contribución de este sector al PIB. Si bien no se han realizado grandes cambios en base a esta experiencia, la estimación del PIB proveniente del sector informal es una utilidad importante que se le puede dar a la encuesta por lo que la experiencia peruana es de gran utilidad para este fin.

⁵ 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).

4. DISEÑO ESTADÍSTICO

Este capítulo explica todos los elementos que componen el diseño estadístico entre los cuales destacan el universo, la población objetivo, el marco estadístico, la cobertura geográfica, las unidades estadísticas, la definición de variables, las nomenclaturas utilizadas, el diseño muestral, las medidas de precisión y los parámetros y la estimación de parámetros. A continuación, se muestra el detalle de cada uno de estos componentes.

4.1. Universo

Corresponde a todas las personas que trabajan de forma independiente y que son dueños de una microempresa que residen en territorio nacional. Es decir, trabajadores por cuenta propia y empleadores dueños de una unidad económica de hasta 10 trabajadores y/o trabajadoras (incluyendo a la persona informante dueña de la unidad económica).

4.2. 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 persona informante dueño de la unidad económica), denominadas en su conjunto como personas microempresedoras, que residen en viviendas particulares ocupadas dentro del territorio nacional.

4.3. Marco estadístico

Desde el segundo semestre del año 2019, el INE pone a disposición el marco muestral de Viviendas construido con base en el Censo de Población y Vivienda de 2017 (MMV 2017). Además, cuenta con protocolos para llevar a cabo procesos de actualización continua y permanente, los cuales ya se han puesto en práctica y, gracias a ello, desde el segundo trimestre del año 2021 se encuentra disponible el MMV 2020, versión actualizada del MMV 2017, a partir de enumeraciones en terreno para unidades geográficas priorizadas y de verificaciones en gabinete para aquellas unidades que no han presentado cambios importantes en el último periodo, haciendo uso de todos los registros administrativos con los que cuenta el INE para tal fin.

A continuación, se describen las características del marco muestral utilizado para la selección muestral de la VII EME. Es importante considerar que, al ser la EME una encuesta en segunda fase de la ENE, esta considera las mismas características del Marco muestral de la ENE, con el objetivo de sintetizar este documento se omitirá la información de la estratificación de la ENE, la cual se encuentra detallada en el Informe Metodológico del Diseño Muestral VII EME.

4.3.1. Marco muestral de la EME

4.3.1.1. Cobertura geográfica

La cobertura es una propiedad estadística que caracteriza al marco muestral. La falta de cobertura denota el error de no incluir ciertos elementos (o unidades completas) de la población objeto de estudio, según el marco muestral que se ha definido. Este error puede ser intencionado

o no. Un ejemplo de error no intencionado es la exclusión de ciertas unidades en el proceso de conteo e identificación de las viviendas, previo a la selección.

La VII EME posee un diseño muestral bifásico, por lo tanto, comparte las propiedades de cobertura de dos marcos muestrales: el primero, utilizado para la selección de las viviendas de la ENE (período MAM 2022); y el segundo; utilizado para la selección de los “microemprendedores”.

La ENE utiliza como insumo la totalidad del marco muestral maestro. Sin embargo, debido a que se realizan exclusiones por motivos operativos, el marco de selección deja fuera 35.287 viviendas, correspondientes a 0,57% del total nacional. El proceso de exclusión se explica con mayor detalle en el apartado 4.3.1.3.

Finalmente, en la elaboración del marco muestral de la VII EME, se excluyen intencionadamente todas las viviendas que no poseen un “microemprendedor”, es decir, que no poseen unidades elegibles.

4.3.1.2. Estratificación geográfica

La estratificación del marco muestral de la ENE, correspondiente a la primera fase, tiene por objetivo crear una división geográfica que cumpla con los requisitos mínimos de UPM para disponer de suficientes unidades al realizar las rotaciones de unidades propias de la encuesta y que refleje, en alguna medida, la aglomeración espacial del mercado laboral.

Se establece como requisito para los estratos un mínimo de 36 UPM en cada uno de ellos, de modo que existan suficientes UPM para completar una rotación completa. Adicionalmente, la estratificación se realiza por separado para el área urbana y para el área rural, de modo que el segmento rural de una comuna podrá compartir estrato con otras, mientras el segmento urbano constituye un estrato por sí mismo.

Finalmente, se obtuvieron 163 estratos geográficos distribuidos en todo el territorio nacional, salvo las zonas de exclusión. Cabe mencionar que existen 11 estratos que no cuentan con el mínimo de UPM requerido, ya que no tenían otras comunas con las cuales agruparse.

En la segunda fase, la VII EME está estratificada de forma natural de acuerdo con las 16 regiones que posee el país.

4.3.1.3. Conformación del marco muestral

Con la inclusión de los estratos geográficos, el marco muestral MMV 2020 ya cuenta con toda la información necesaria para ser utilizado en la ENE, según el diseño muestral definido. Sin embargo, es necesario definir las UPM que se van a excluir del estudio. Para ello, en primer lugar, se identifica a las UPM susceptibles de exclusión, es decir, las áreas especiales del MMV 2017, como también, aquellas que pertenecen a zonas que se encontraban excluidas en el estudio previo. Para decidir cuáles de las UPM identificadas serán finalmente excluidas, el equipo técnico

el INE trabajó, en coordinación con los equipos regionales de recolección, de modo que la decisión incluya también las capacidades instaladas de cobertura⁶.

El resultado fue la exclusión de 355 UPM, que corresponde a 1% de las UPM a nivel nacional, donde diez de las 345 comunas que cubre el marco muestral maestro se ven excluidas en su totalidad, observándose una concentración relevante en la región de Los Lagos, donde la provincia de Palena se ve totalmente excluida. Asimismo, las viviendas del marco excluidas del estudio alcanzan 0,6% a nivel nacional. La **Tabla 4: Cantidad de UPM especiales y excluidas, por región** muestra la distribución regional de UPM y viviendas marco, especiales y excluidas.

⁶ A cada equipo regional le fue remitido el correspondiente listado de UPM susceptibles de exclusión, para que, en función de sus conocimientos del terreno determinaran si era necesario excluir la UPM o no del estudio.

Tabla 4: Cantidad de UPM especiales y excluidas, por región

Región	UPM	UPM especiales	% UPM Especiales	UPM excluidas	% UPM excluidas	Viviendas Marco	Viviendas marco especiales	% Viv. Marco especiales	Viviendas marco excluidas	% Viv. Marco excluidas
Nacional	35.149	973	2,8%	355	1,0%	6.145.493	87.831	1,4%	35.287	0,6%
Arica y Parinacota	405	9	2,2%	8	2,0%	72.602	828	1,1%	736	1,0%
Tarapacá	570	22	3,9%	13	2,3%	107.473	2.046	1,8%	1.195	1,1%
Antofagasta	965	10	1,0%	4	0,4%	186.917	878	0,4%	359	0,2%
Atacama	607	11	1,8%	6	1,0%	105.554	924	0,8%	511	0,5%
Coquimbo	1.738	45	2,6%	0	0,0%	278.060	4.030	1,4%	0	0,0%
Valparaíso	3.687	38	1,0%	28	0,8%	673.541	3.413	0,5%	3.921	0,6%
Metropolitana	12.054	17	0,1%	0	0,0%	2.352.958	1.543	0,1%	0	0,0%
O'Higgins	2.150	71	3,3%	6	0,3%	331.870	6.331	1,9%	524	0,2%
Maule	2.545	171	6,7%	5	0,2%	386.081	15.778	4,1%	457	0,1%
Ñuble	1.225	17	1,4%	8	0,7%	179.827	1.526	0,8%	726	0,4%
Biobío	3.126	75	2,4%	24	0,8%	553.538	6.709	1,2%	2.064	0,4%
La Araucanía	2.398	162	6,8%	49	2,0%	355.092	14.679	4,1%	4.494	1,3%
Los Ríos	977	77	7,9%	23	2,4%	143.005	6.937	4,9%	2.023	1,4%
Los Lagos	2.096	187	8,9%	167	8,0%	316.294	16.847	5,3%	16.895	5,3%
Aysén	272	44	16,2%	11	4,0%	41.682	3.996	9,6%	1.177	2,8%
Magallanes	334	17	5,1%	3	0,9%	60.999	1.366	2,2%	205	0,3%

Fuente: Instituto Nacional de Estadísticas.

Las viviendas seleccionadas y encuestadas en la ENE para el periodo MAM 2022, son utilizadas como marco de muestreo para la VII EME, pues permite identificar las viviendas donde reside al menos un microemprendedor.

En la segunda fase, se clasifican las viviendas en dos grupos, de acuerdo con si éstas contienen o no, en el período de referencia, al menos un microemprendedor. Las viviendas que no poseen microemprendedores se descartan.

Luego, se revisa detalladamente el listado de microemprendedores para descartar todas las unidades que no cumplen los criterios técnicos para ser clasificadas efectivamente como un microemprendimiento, mitigando así los problemas posteriores de levantamiento debido a la existencia de casos fuera de muestra; formando así el marco de muestreo para la VII EME, sólo con aquellas viviendas con presencia de microemprendedores.

4.3.1.4. Estratificación del marco muestral

El diseño muestral de la ENE considera información tanto de los estratos geográficos, descritos en el apartado anterior, como de la estratificación socioeconómica presente en el MMV 2020. Los estratos de muestreo se conforman a partir del cruce entre la estratificación geográfica y socioeconómica, siempre que los estratos geográficos cumplan ciertas condiciones. Se establece que, para desagregar un estrato geográfico según estratificación socioeconómica (NSE) debe:

- Contener más de 150 UPM.
- Poseer suficiente heterogeneidad en su interior en relación con la condición socioeconómica, esto es:
 - El estrato geográfico no debe agrupar más de 80% de las UPM en un solo NSE o;
 - debe presentar una proporción de UPM en torno a 40% en un NSE y, los dos restantes en conjunto, alrededor de 60%.
- Asegurar que todos los estratos resultantes cumplan con un mínimo de 36 UPM.

Según lo anterior, se obtiene un total de 232 estratos de muestreo, de los cuales 122 provienen directamente desde la estratificación geográfica, y 110 incorporan la información de la estratificación socioeconómica. La distribución regional de los estratos se muestra en la **Tabla 5: Cantidad de estratos de muestreo por región y área**.

Tabla 5: Cantidad de estratos de muestreo por región y área

Región	Área		Total
	Urbano	Rural	
Total	153	79	232
Arica y Parinacota	3	1	4
Tarapacá	4	2	6
Antofagasta	5	2	7
Atacama	5	3	8
Coquimbo	8	5	13
Valparaíso	22	9	31
Metropolitana	33	10	43
O'Higgins	8	7	15
Maule	13	10	23
Ñuble	6	4	10
Biobío	15	5	20
Araucanía	8	4	12
Los Ríos	7	5	12
Los Lagos	11	8	19
Aysén	2	2	4
Magallanes	3	2	5

Fuente: Instituto Nacional de Estadísticas.

4.4. Cobertura Geográfica

La Encuesta de Microemprendimiento se realiza a lo largo del territorio nacional, de acuerdo con las 16 regiones que posee el país, las áreas urbana y rural, y las exclusiones del marco estadístico.

4.5. Unidades estadísticas

La unidad de análisis de la VII EME son los microemprendedores, entendidos como trabajadores por cuenta propia y empleadores con hasta 10 trabajadores incluyéndose. 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.

4.6. Definición de variables

En esta sección se definen las principales variables que delimitan el estudio y que son utilizadas en el proceso estadístico de la encuesta. Las variables son las siguientes:

Persona microempresaria: 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). Para identificar este grupo en la ENE se consideraron todos informantes clasificados como independientes (CISE=1 o 2) que además trabajan en una microempresa (b15_1=1 o 2).

Microempresa: unidad económica de hasta 10 personas trabajadoras incluido el dueño o dueña de la unidad. Se entiende también como microemprendimiento, micronegocio o actividad.

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. Para ser clasificado en esta categoría debe tener al menos un trabajador asalariado (f2_f=1) trabajando de forma permanente (f2_i<>86).

Trabajador por cuenta propia: trabajador independiente sin relaciones de subordinación o dependencia, 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.

4.7. Nomenclaturas utilizadas

En el proceso productivo de la encuesta fueron considerados 5 clasificadores los cuales se presentan enseguida.

Clasificación Internacional de la Situación en el Empleo (CISE): Esta clasificación fue elaborada en la decimoquinta Conferencia Internacional de Estadísticos del Trabajo (XV CIET) en el año 1993. Esta clasifica los distintos empleos utilizando dos criterios básicos: quién asume el riesgo económico de la actividad o empleo; y el tipo de autoridad (o dependencia) que tienen los trabajadores.

Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES): clasificador elaborado por el Instituto Nacional de Estadísticas (INE), el cual corresponde a una adaptación del Clasificador Chileno de Actividades Económicas CIIU.CL 2012. Este clasificador incluye criterios para facilitar el trabajo de codificación de glosas provenientes de encuestas de hogares.

Clasificador Internacional Uniforme de Ocupaciones (CIUO): Es una herramienta para organizar los empleos en una serie de grupos definidos claramente en función de las tareas que comporta cada empleo. Está destinada tanto a los usuarios del campo de la estadística como a los usuarios orientados al cliente. La versión utilizada es la actualización desarrollada en el 2008 en la decimoctava Conferencia Internacional de Estadísticos del Trabajo (XVIII CIET).

Clasificador Internacional Normalizada de educación (CINE): forma parte de la familia internacional de Clasificaciones Económicas y Sociales de las Naciones Unidas. Dentro de estas, la CINE representa una clasificación de referencia que permite ordenar los

programas educativos y sus respectivas certificaciones por niveles de educación y campos de estudio. Su elaboración es el resultado de un acuerdo internacional adoptado formalmente por la Conferencia General de los Estados Miembros de la Unesco. La versión considerada en la EME es la misma que la implementada en la Encuesta Nacional de Empleo la cual corresponde al año 1997.

Clasificación de Actividades de uso del tiempo para América Latina y el Caribe (Cautal): Esta clasificación permite diferenciar las actividades desarrolladas por las personas en actividades productivas o actividades personales (no productivas). Dentro de las actividades permite distinguir cuáles de ellas corresponden a trabajo remunerado y no remunerado. Esta distinción considera la frontera de producción del Sistema de Cuentas Nacionales (SCN) y es consistente con las recomendaciones de la resolución de la XIX Conferencia Internacional de Estadísticos del Trabajo (CIET) de la Organización Internacional del Trabajo (OIT).

4.8. Diseño estadístico por muestreo

El diseño muestral se compone de todos los elementos necesarios para construir una muestra que sea representativa del fenómeno de estudio. Para ello se decidió realizar 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 es que éstas 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.

4.8.1. Diseño muestral

El objetivo del diseño muestral fue determinar y distribuir un número de viviendas y microemprendedores que permitiera obtener estimaciones del parámetro de interés. En el caso de la VII 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 establecidos⁷ a nivel nacional y regional.

El diseño muestral de la VII EME contó de dos fases, siendo la primera fase la constitución del marco estadístico de la EME y la segunda fase la selección de la muestra a participar en la EME.

⁷ Los niveles de precisión son: error absoluto de 1,4% y error relativo de 1,6%, a nivel nacional. Para las regiones, el error absoluto máximo es 8,1% y, el error relativo no supera el 10,5% correspondiente a la Región de Magallanes.

4.8.2. Cálculo y distribución del tamaño muestral

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 cálculo de la variable de interés para la determinación del tamaño muestral se realizó con la información de la ENE levantada en los trimestres MAM 2019 y MAM 2021. La información de ambos trimestres se fusiona; conformando un escenario híbrido, el cual surge de la necesidad de incorporar en el estudio un escenario en condiciones normales y otro con efecto “Pandemia”, desde donde es posible obtener una estimación de la cuasi-varianza poblacional, para obtener un tamaño muestral inicial bajo muestreo aleatorio simple monoetápico.

Luego, el tamaño inicial calculado recibe una serie de ajustes asociados al efecto de diseño, ajuste por finitud de la población y el aumento del tamaño muestral objetivo para cubrir la posible pérdida de unidades por concepto de no logro.

Dado que este escenario considera dos trimestres en distintos años, el tamaño objetivo que se asigna a nivel regional (corregido por finitud) es aproximadamente el promedio entre los dos tamaños definidos en cada trimestre. Luego, el tamaño muestral objetivo a nivel nacional es de 5.959 viviendas, que permite realizar una estimación del parámetro de interés con un error absoluto esperado de 1,4% y; un error relativo de 1,6%. Con este tamaño y utilizando el promedio entre las tasas de no respuesta entre la VI EME (considera un escenario normal de levantamiento) y la ENE MAM 2021 (considera escenario con efecto pandemia) se obtiene un tamaño con sobremuestreo de 7.850 viviendas.

Se utilizó como parámetro de interés la proporción de microemprendedores por cuenta propia sobre el total de microemprendedores y sus estadísticos asociados, obtenidos desde la ENE en los trimestres mencionados anteriormente.

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 a encuestar por unidad primaria de muestreo en la región r en la ENE trimestre MAM 2019 y MAM 2021.
- n_{r0} : Número de UPM levantadas en trimestre MAM 2019 y MAM 2021 (ENE) en la región r .
- $n_{r0} \cdot \bar{m}_{r0} = m_{r0}$: Número de viviendas logradas en trimestre MAM 2019 Y MAM 2021 (ENE) en la región r .
- λ : Porcentaje de selección asignado para cada región (70,0% o 80,0%).

$SE(\hat{p}_{r0})$: 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.

$t_{1-\alpha/2, v-1}$: Cuantil de una distribución t de Student con nivel de confianza $1-\alpha$ (para efectos de la simulación se utiliza un valor fijo $t=2$).

Etapa 1:

Utilizando la fusión de las bases de datos del levantamiento de la ENE en los trimestres MAM 2019 y MAM 2021, se obtienen las estimaciones a nivel regional de: parámetro de interés regional (\hat{p}_{r0}), error estándar asociado ($SE(\hat{p}_{r0})$) y total de unidades logradas (m_{r0}). La cuasi-varianza poblacional $S(p_{r0})^2$ se calcula a partir del error estándar obtenido, utilizando la siguiente ecuación:

$$S(p_{r0})^2 = \frac{n \cdot \bar{m}_{r0}}{Def f(p_{r0})} \cdot SE(p_{r0})^2 \quad (1)$$

Etapa 2:

A partir del error estándar, se obtienen los errores absolutos por región (ea_{0r}) y luego, los errores relativos (e_{0r}). Posteriormente, se establece el nivel de confianza de 95% para obtener el cuantil de una distribución t de Student que, para efectos de la simulación, se utiliza un valor de $t=2$. Con estos y la cuasi-varianza del parámetro de interés derivada en (1), se obtienen los tamaños muestrales iniciales a nivel regional considerando un muestreo aleatorio simple monoetápico. Este cálculo se observa en la ecuación (2):

$$m_{0,r} = \frac{2^2 \cdot S_r^2}{ea_{r1}^2} \quad (2)$$

Donde,

$m_{0,r}$: Número de viviendas a nivel regional, bajo muestreo aleatorio simple.

Etapa 3

Debido a que el diseño muestral es complejo, es necesario considerar la incorporación del efecto diseño, dado que el muestreo complejo implica una pérdida de eficiencia en las estimaciones con respecto al muestreo aleatorio simple. El efecto del diseño utilizado corresponde al efecto considerado para el cálculo de las simulaciones de tamaño muestral del nuevo diseño de la ENE (ENE 2020) donde se propuso un efecto fijo ($Deff$) de 1,7 para cada región. Por esta razón, el tamaño de muestra inicial, bajo muestreo aleatorio simple, se incrementa a partir del efecto de diseño según la siguiente ecuación (3):

$$m_{1,r} = m_{0,r} \cdot Deff \quad (3)$$

Donde,

$m_{1,r}$: Número de viviendas a nivel regional ajustado por efecto de diseño.

$Deff$: Efecto de diseño fijo para cada región.

Este tamaño corresponde a 70,0% de las viviendas disponibles en el marco de selección, donde reside al menos un microemprendedor, con excepción de algunas regiones, donde se asigna 80,0%. Este aumento de la muestra se realiza con el objetivo de disminuir el error muestral.

Etapa 4

Los tamaños muestrales regionales obtenidos, deben corregirse por un ponderador que da cuenta de la condición de población finita (fracción de muestreo, FRm_r). Esta fracción se obtiene a partir de la ecuación (4):

$$FRm_r = \frac{m_{1r}}{MM_r} \quad (4)$$

Etapa 5

Obtenida la fracción de muestreo a nivel regional, se determina el tamaño muestral “corregido por finitud”, según la ecuación (5):

$$m_{2,r} = m_{1,r} \cdot \left(\frac{1}{1 + \frac{m_{1,r}}{M_r}} \right) \quad (5)$$

Etapa 6

Una vez obtenidos los tamaños objetivos regionales, se aplica un factor que corresponde a la tasa de no respuesta (tnr). Dicha tasa se obtiene a partir del promedio entre la tasa de no respuesta de la VI EME y el trimestre ENE MAM 2021, y es considerada con el propósito de contrarrestar la pérdida de muestra por no respuesta por parte de los encuestados.

La fórmula para calcular el tamaño de muestra con sobremuestreo viene dada por la siguiente ecuación (6):

$$m_{3,r} = \frac{m_{2,r}}{1 - tnr} \quad (6)$$

$m_{3,r}$: Número de viviendas con sobremuestreo a nivel regional.

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 6, se presentan definiciones y consideraciones utilizadas.

Tabla 6: Criterios utilizados en el cálculo del tamaño muestral

Parámetro	Descripción
Variable de diseño	I : Variable Bernoulli que considera los siguientes valores $I = \begin{cases} 0, & \text{si el microemprendedor no es cuenta propia} \\ 1, & \text{si el microemprendedor es cuenta propia} \end{cases}$
Parámetro asociado	Proporción de microemprendedores por cuenta propia sobre el total de microemprendedores en la región r (p_{r0})
Estimador asociado	Estimador de razón: $r = \frac{\text{Estimación del número de microemprendedores por cuenta propia en la región}}{\text{Estimación del número de microemprendedores en la región}}$
Niveles de estimación	Nacional. Regional.
Errores de muestreo	Nacional: El error absoluto no debe superar 1,4% y; el error relativo 1,6%. Regional: Los errores absolutos no superan 8,1% y; los errores relativos 10,5%.

La siguiente tabla presenta los errores esperados, asociados a la estimación del parámetro de interés, para un tamaño muestral objetivo de 7.850 viviendas⁸ a nivel nacional.

⁸ 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 5.959 viviendas.

Tabla 7: Errores esperados asociados al parámetro de interés según tamaño muestral objetivo

Nivel	Estimación Microemprendedores MAM 2019-2021	Efecto diseño propuesto	Error absoluto propuesto VII EME	Error relativo propuesto VII EME	Tamaño objetivo ajustado VII EME	Tamaño con sobremuestra VII EME
Nacional	86,9%	1,7	1,4%	1,6%	5.959	7.850
Arica y Parinacota	88,8%	1,7	4,9%	5,5%	239	298
Tarapacá	86,5%	1,7	7,8%	9,1%	197	229
Antofagasta	87,7%	1,7	5,4%	6,2%	202	251
Atacama	86,9%	1,7	5,2%	6,0%	214	268
Coquimbo	85,6%	1,7	4,9%	5,8%	370	493
Valparaíso	85,2%	1,7	3,6%	4,2%	681	938
Metropolitana	87,2%	1,7	2,9%	3,3%	1.030	1.472
O'Higgins	86,2%	1,7	5,0%	5,8%	316	415
Maule	86,7%	1,7	4,0%	4,6%	418	535
Ñuble	81,8%	1,7	6,3%	7,7%	240	292
Biobío	89,3%	1,7	2,9%	3,3%	602	775
La Araucanía	89,0%	1,7	4,8%	5,4%	360	483
Los Ríos	86,4%	1,7	5,3%	6,2%	262	315
Los Lagos	86,3%	1,7	5,0%	5,8%	461	619
Aysén	86,1%	1,7	4,4%	5,1%	222	286
Magallanes	77,3%	1,7	8,1%	10,5%	145	181

Fuente: Instituto Nacional de Estadísticas.

Una vez obtenido el tamaño muestral requerido para cada región, se debe distribuir a nivel de submuestra. Dado que la muestra ENE está subdividida en tres meses o períodos de levantamiento, con el objetivo de disminuir el tiempo transcurrido entre el levantamiento de información de la ENE y la EME y, con ello tener una menor atrición⁹, la muestra de la VII EME se distribuye independientemente en tres meses de levantamiento: mayo, junio y julio, de acuerdo con el mes de levantamiento de la ENE, marzo, abril y mayo, respectivamente.

Según lo anterior, el tamaño regional se distribuye en tres partes iguales, o lo más similares posibles. La tabla 8 expone la distribución de la muestra según mes de levantamiento y región.

Tabla 8: Total de viviendas seleccionadas según región y mes de levantamiento

⁹ 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).

Mes de levantamiento VII EME				
Nivel	Mayo	Junio	Julio	Total viviendas seleccionadas
Nacional	2.683	2.575	2.592	7.850
Arica y Parinacota	92	107	99	298
Tarapacá	77	78	74	229
Antofagasta	84	84	83	251
Atacama	90	88	90	268
Coquimbo	172	139	182	493
Valparaíso	340	319	279	938
Metropolitana	519	464	489	1.472
O'Higgins	143	138	134	415
Maule	179	178	178	535
Ñuble	98	97	97	292
Biobío	259	243	273	775
La Araucanía	165	159	159	483
Los Ríos	105	105	105	315
Los Lagos	209	215	195	619
Aysén	90	101	95	286
Magallanes	61	60	60	181

Fuente: Instituto Nacional de Estadísticas.

4.8.3. Selección de unidades muestrales

La selección de unidades muestrales se realiza en 2 etapas. En la primera etapa se seleccionan las viviendas que contienen al menos un microemprendedor y luego, en la segunda etapa se seleccionan los microemprendedores al interior de los hogares que contienen las viviendas seleccionadas.

4.8.3.1. Selección de viviendas

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

La selección de viviendas es implementada mediante el paquete "Sampling" del software R, fijando una semilla aleatoria, a fin de que pueda ser replicado en cualquier momento.

Para calcular la probabilidad de inclusión de la j -ésima vivienda dentro de la r -ésima región, se definen:

$m_{3,r}$: Número de viviendas a seleccionar en la región r .

M_r : Número de viviendas que contiene el marco de selección en la región r .

Y para la selección de viviendas ($m_{3,r}$) el software ejecuta los siguientes pasos:

Paso 1

En primera instancia, se ordenan en forma ascendente todas las viviendas según región, estrato, comuna, id_upm, y número de orden de vivienda dentro de la upm.

Paso 2

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

$$k = M_r / m_{3,r}.$$

Paso 3

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

Paso 4

Posteriormente se suma sucesivamente el período "k" al arranque "A" para obtener distintos valores, los que dan origen a la selección de unidades de la siguiente forma: "A", "A + k", "A + 2k", "A + 3k", ..., "A + ($m_{3,r} - 1$)k".

La primera vivienda seleccionada es "A" y es un número entero, la segunda es el redondeo de "A + k", la tercera es el redondeo de "A + 2k" y así sucesivamente, hasta la $m_{3,r}$ selección, dada por el redondeo de "A + ($m_{3,r} - 1$)k".

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

$$P_r(j|r) = \frac{m_{3,r}}{M_r} \quad (7)$$

4.8.3.2. 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.

4.8.3.3. Cambios entre muestra diseñada y muestra seleccionada.

Como se mencionó anteriormente, la muestra de la VII EME (2022) fue construida el año 2021 por medio del promedio de dos periodos 2019 y 2021 para el trimestre de referencia

de MAM. Este promedio consideraba una recuperación en la muestra recolectada el 2021 debido a las mejores tasas de respuestas observadas en la ENE a finales del 2021. Sin embargo, el aumento de la tasa de respuesta en la ENE fue menor al proyectado. En consecuencia, no fue posible seleccionar la muestra diseñada, la cual ascendía a 8.565 viviendas a nivel nacional. Se seleccionaron finalmente un total de 7.850 unidades, generando un déficit de 715 unidades, lo que equivale a un 8,3% a nivel nacional. Las regiones más afectadas fueron la Metropolitana, Los Lagos y Tarapacá, donde el déficit fue 347 (-19,1 %), 144 (-18,9 %) y 66 viviendas (-22,4 %), respectivamente (ver Tabla 9).

Tabla 9: Muestra seleccionada por región y submuestra

Nivel	Submuestra 1		Submuestra 2		Submuestra 3		Total		Diferencia	
	Dis.	Selec.	Dis.	Selec.	Dis.	Selec.	Dis.	Selec.	n	%
Nacional	2.863	2.683	2.853	2.575	2.849	2.592	8.565	7.850	-715	-8,3
Arica y Parinacota	100	92	99	107	99	99	298	298	0	0,0
Tarapacá	99	77	98	78	98	74	295	229	-66	-22,4
Antofagasta	84	84	84	84	83	83	251	251	0	0,0
Atacama	90	90	89	88	89	90	268	268	0	0,0
Coquimbo	172	172	172	139	171	182	515	493	-22	-4,3
Valparaíso	340	340	340	319	339	279	1.019	938	-81	-7,9
Metropolitana	607	519	606	464	606	489	1.819	1.472	-347	-19,1
O'Higgins	147	143	147	138	146	134	440	415	-25	-5,7
Maule	179	179	178	178	178	178	535	535	0	0,0
Ñuble	98	98	97	97	97	97	292	292	0	0,0
Biobío	259	259	258	243	258	273	775	775	0	0,0
La Araucanía	165	165	164	159	164	159	493	483	-10	-2,0
Los Ríos	105	105	105	105	105	105	315	315	0	0,0
Los Lagos	255	209	254	215	254	195	763	619	-144	-18,9
Aysén	102	90	102	101	102	95	306	286	-20	-6,5
Magallanes	61	61	60	60	60	60	181	181	0	0,0

Fuente: Instituto Nacional de Estadísticas.

Nota: Dis. refiere a "Diseñada" y Selec. Refiera a "Seleccionada".

El color rojo representa que la muestra seleccionada fue menor a la muestra diseñada

El color verde representa que la muestra seleccionada fue mayor a la muestra diseñada

4.9. Medidas de precisión

En este apartado se muestran algunas medidas de precisión diseñadas para los procesos de recolección y procesamiento de datos.

4.9.1. Medidas de precisión proceso de recolección

La clasificación de los resultados finales se hizo a partir de los códigos de disposición final registrados en la hoja de ruta. A través de ellos fue posible distinguir informantes elegibles, no elegibles y de elegibilidad desconocida. La Tabla 10 presenta la clasificación de elegibilidad y categorización de resultado según los códigos registrados en la hoja de ruta.

Tabla 10: Clasificación de códigos de disposición final según resultado final en la Hoja de Ruta

Clasificación elegibilidad	Categorización de resultado	Códigos de disposición final (CDF)
Elegible	Entrevistados (E)	11, 12
	Rechazo (R)	21, 22
	No contacto (NC)	31, 32, 33, 34, 35, 71, 72, 73, 74, 75, 76, 77, 39
	Otros motivos de no entrevista (presencial y telefónica) (O)	41, 42, 43, 44, 48, 49, 81, 82, 83
Elegibilidad desconocida	Elegibilidad desconocida (ED)	52, 53, 54, 59
No Elegible	No Elegible (NE)	61, 62, 63, 64, 65, 66, 67, 68, 69

Fuente: INE, VII Encuesta de Microemprendimiento, 2022

Los resultados de la gestión realizada en terreno puede ser medida a partir de diferentes indicadores que permiten a su vez explicar e interpretar los resultados del trabajo de campo.

Estos indicadores son construidos a través de los datos capturados mediante la Hoja de Ruta (HR) de la encuesta. La HR corresponde a un instrumento inserto al principio del cuestionario de la EME y tiene por objetivo llevar una bitácora sobre los intentos de contacto con el informante, ya sean visitas presenciales o llamadas telefónicas. Este registro esta estandarizado según los Códigos de Disposición Final (CDF), los cuales consisten en una clasificación de situaciones típicas que pueden ocurrir al momento de contactar a un informante. Los CDF son particularmente importantes ya que son el dato concreto que permite calcular los indicadores operativos y también son un insumo esencial en la construcción de los factores de expansión.

En concreto, los CDF permiten construir los siguientes indicadores:

- Tasa de No Elegibilidad
- Tasa de Contacto
- Tasa de Cooperación
- Tasa de Rechazo
- Tasa de Respuesta
- Tasa de Logro

A continuación, se explica el detalle del flujo de gestión operativa, detallando cada concepto con el cual se construyen los indicadores.

No Elegibilidad

Al hacer un intento de contacto con el informante, pueden ocurrir tres situaciones:

- Elegible
- Elegibilidad desconocida
- No elegible

Los casos no elegibles hacen referencia a situaciones del informante o de la vivienda que, de haberse conocido al momento de la conformación del marco muestral, habrían sido excluidos por no cumplir con alguna de las condiciones definidas para ser parte del estudio. Estas condiciones son:

1. Que la vivienda sea particular
2. Que resida al menos un microemprendedor en la vivienda

Los casos no elegibles proporcionan una aproximación a posibles debilidades en cuanto a la exactitud del marco de selección empleado, lo cual no implica necesariamente errores en las fuentes de información utilizadas para su conformación, dado que parte de los casos considerados como no elegibles se originan por cambios esperables en la situación de las personas, como cambios de domicilio, fallecimientos, entre otros. Aunque también se incluyen aquellos casos clasificados incorrectamente dentro de la categoría de trabajadores independientes¹⁰.

En general, todo marco muestral tendrá una proporción de unidades que finalmente serán consideradas como no elegibles, para lo cual el INE realiza un trabajo de depuración con el objeto de minimizar estos problemas para la VII EME.

La VII EME considera como no elegibles a los informantes o viviendas en las siguientes situaciones:

- (61) Establecimiento, oficina de gobierno u otra organización
- (62) Institución (por ejemplo: hospital, cárcel, asilo de ancianos, etc.)
- (63) Dormitorio colectivo (por ejemplo: militar, de trabajo, internado, hostel, residencial, etc.)
- (64) Vivienda en demolición, incendiada, destruida o erradicada
- (65) Vivienda desocupada (por ejemplo: en venta, arriendo, etc.)
- (66) Vivienda de veraneo o de uso temporal
- (67) Edificación no calificable (por ejemplo: en construcción, en reparaciones. etc.)
- (68) No existe informante elegible (Informante fuera de marco, cambio de domicilio)

¹⁰ En detalle, corresponde a casos que fueron calificados como microemprendedores en la Encuesta Nacional de Empleo (ENE), pero que al re contactarse en la EME, los informantes reportan nunca haber sido microemprendedores. Es decir, hubo una clasificación errónea.

(69) Otra razón para ser considerada no elegible

Finalmente, existirá una pequeña proporción de casos que se denominade elegibilidad desconocida”, al no poder ser identificada la vivienda en terreno, estar ubicada en un área peligrosa u otra situación, incluida la posibilidad de que la vivienda no haya sido gestionada.

Contacto

En el caso que se determine que un informante es elegible, la gestión realizada para el logro de la encuesta podría dar como resultado dos situaciones:

- Contacto
- No contacto

El no contacto implica que no fue posible obtener algún tipo de interacción con las personas residentes en la vivienda. Resulta un indicador relevante para evaluar las dificultades en cuanto a gestiones asociadas a lograr contactar a un informante, en el entendido de que lograr interactuar con este es un elemento imprescindible y necesario para obtener el logro de una encuesta, permitiendo cuantificar la magnitud del problema de moradores ausentes existente en cada región y, con ello, comparar los resultados de otras encuestas para evaluar la eficacia de la gestión realizada.

La VII EME considera como no contacto a los informantes o viviendas en las siguientes situaciones:

- (31) Vivienda ocupada, sin moradores presentes
- (32) Se impide el acceso a la vivienda
- (33) Informante no ubicable, se contacta a otra persona
- (34) Informante ausente en periodo de levantamiento de la encuesta

Cooperación

En el caso que se haya logrado contactar a un informante, pueden ocurrir dos situaciones:

- Cooperación
- No Cooperación

La cooperación dice relación con los casos en los cuales fue posible lograr interactuar con el informante seleccionado y este “coopero” con el estudio, dando respuesta a las preguntas de la encuesta de manera total o suficiente para considerarla como una encuesta lograda, excluyendo a su vez los casos de rechazos y otras situaciones que generaron como resultado que el informante no respondiera la encuesta o la respondiera de manera insuficiente.

La VII EME considera como cooperación las siguientes situaciones:

- (11) Entrevista completa
- (12) Entrevista parcial

Rechazo

En el caso de que se un informante sea elegible y haya sido contactado, pueden ocurrir dos situaciones:

- Rechazo
- No Rechazo

El rechazo es una de las causales que afectan los niveles de respuesta y que debe ser considerada en el diseño de cada estudio con el objeto de reducir aspectos como la sobrecarga de preguntas o encuestas a las mismas personas, evitando así el agotamiento de los informantes.

Se considera rechazo cuando se logra contacto con los informantes, pero estos declinan a cooperar con el estudio y dar respuesta a las preguntas de la encuesta, incluidos los casos en los que la entrevista es interrumpida durante su desarrollo por alguna razón. No equivale a todos los casos de no cooperación pues las personas podrían no cooperar por otras razones, como por problemas de idioma, de salud u otros.

En la VII EME se considera como rechazo las siguientes situaciones:

(21) Informante del hogar rechazo la entrevista

(22) Informante directo rechazo la entrevista

Respuesta

En el caso de que se un informante sea elegible y haya sido contactado, pueden ocurrir dos situaciones:

- Respuesta
- No respuesta

Se considera como respuesta a los casos en los cuales, después de toda la gestión realizada, fue lograda la aplicación de una encuesta al informante de manera completa o suficiente para considerarla como una encuesta lograda. Su valor para el caso de la VII EME equivale al total considerado también como Cooperación. No obstante, se diferencian en que la cooperación, como se verá más adelante, se interpreta en base a los contactos realizados, en tanto la respuesta se interpreta en base al total de unidades potencialmente elegibles. Finalmente, es en base a los resultados obtenidos a nivel de informantes que se agrega la información para estimar niveles de logro asociados a las viviendas y comparar con la muestra objetivo considerada en el diseño muestral.

La VII EME considera como respuesta las siguientes situaciones:

(11) Entrevista completa

(12) Entrevista parcial

Tasa de no elegibilidad a nivel de personas (TNE).

Este indicador se refiere a la proporción de unidades muestrales no elegibles respecto al total de la muestra de personas, lo que se calcula en base a la siguiente ecuación:

$$TNE = \frac{NE}{\text{Muestra Total}} \times 100$$

Tasa de contacto a nivel de personas (CON1)

Este indicador se refiere a la proporción de casos de la muestra en que el encuestador logra hablar, o tener la atención del entrevistado, pudiendo obtener o no la entrevista. Por lo tanto, se entiende como el acceso al informante. La siguiente ecuación permite realizar dicho cálculo:

$$CON1 = \frac{E + R + O}{E + R + NC + O + ED} \times 100$$

Tasa de Cooperación a nivel de personas (COOP2)

Este indicador se refiere a la proporción de unidades muestrales entrevistadas, el cual se calcula en base a todas las unidades elegibles que fueron contactadas, como lo muestra la siguiente ecuación:

$$COOP2 = \frac{E}{E + R + O} \times 100$$

Tasa de rechazo a nivel de personas (REF1)

Este indicador se refiere a la proporción de unidades que rechazan participar en el estudio, incluyendo aquellos casos que interrumpen la entrevista de forma definitiva. Lo anterior se calcula sobre el total de casos potencialmente elegibles, como lo señala la siguiente ecuación:

$$REF1 = \frac{R}{E + R + NC + O + ED} \times 100$$

Tasa de respuesta no ponderada a nivel de personas (RR1np).

Este indicador se refiere a la proporción de encuestas logradas sobre el total de casos potencialmente elegibles. A continuación, se indica la ecuación que permite realizar dicho cálculo.

$$RR1np = \frac{E}{E + R + NC + O + ED} \times 100$$

Tasa de logro sobre muestra objetivo

Este indicador se refiere al cumplimiento de la muestra objetivo, en él se presenta la tasa de logro obtenido en base al objetivo muestral diseñado para cada región del país. La forma de calcular este indicador es la siguiente:

$$Tasa\ de\ logro = \frac{E}{Muestra\ Objetivo} \times 100$$

4.9.2. Medidas de precisión proceso de análisis

Para evaluar la precisión de las estimaciones realizadas en las EME se adoptó el estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares del INE¹¹, el cual entrega indicaciones que permiten definir si las estimaciones son fiables, poco fiables o no fiables, y establece un criterio de publicación de tabulados.

La figura 2 muestra el flujo con el detalle de los filtros para ser clasificado en cada uno de los grupos de fiabilidad. Para poder publicar una estimación se debe cumplir con dos condiciones excluyentes para cualquier tipo de estimación: tener un tamaño muestral¹² de al menos 60 casos muestrales (n) y un número de grados de libertad¹³ mayor o igual a 9. En caso contrario, el flujo termina y las estimaciones deben ser clasificadas como “Estimación no fiable”.

Si cumplen las dos condiciones anteriores, se presentan dos posibilidades para avanzar en el flujo se dependiendo de si el tipo de estimador es de proporción o razón entre 0 y 1 o no. En el primer caso, el error estándar debe cumplir la siguiente condición, dependiendo del nivel de la estimación asociada de p :

$$Máximo\ ee\ tolerable = \begin{cases} \sqrt[3]{p^2}/9 & ; p < 0,5 \\ \sqrt[3]{(1-p)^2}/9 & ; p \geq 0,5 \end{cases} \quad (1)$$

En caso de que el error estándar (ee) se encuentre en dicho intervalo, el estimador debe considerarse como *estimación fiable*. Pero si el estimador no cumple con dicho criterio, debe consignarse como *estimación poco fiable* dada su baja confiabilidad.

El resto de las estimaciones, se evalúan mediante el coeficiente de variación, que debe ser menor igual a 15% para ser clasificado como *estimación fiable*. Cuando el coeficiente de variación sea mayor a 15% y menor o igual a 30%, la estimación debe venir acompañada con una nota que indique que la *estimación poco fiable*; finalmente, si el coeficiente de

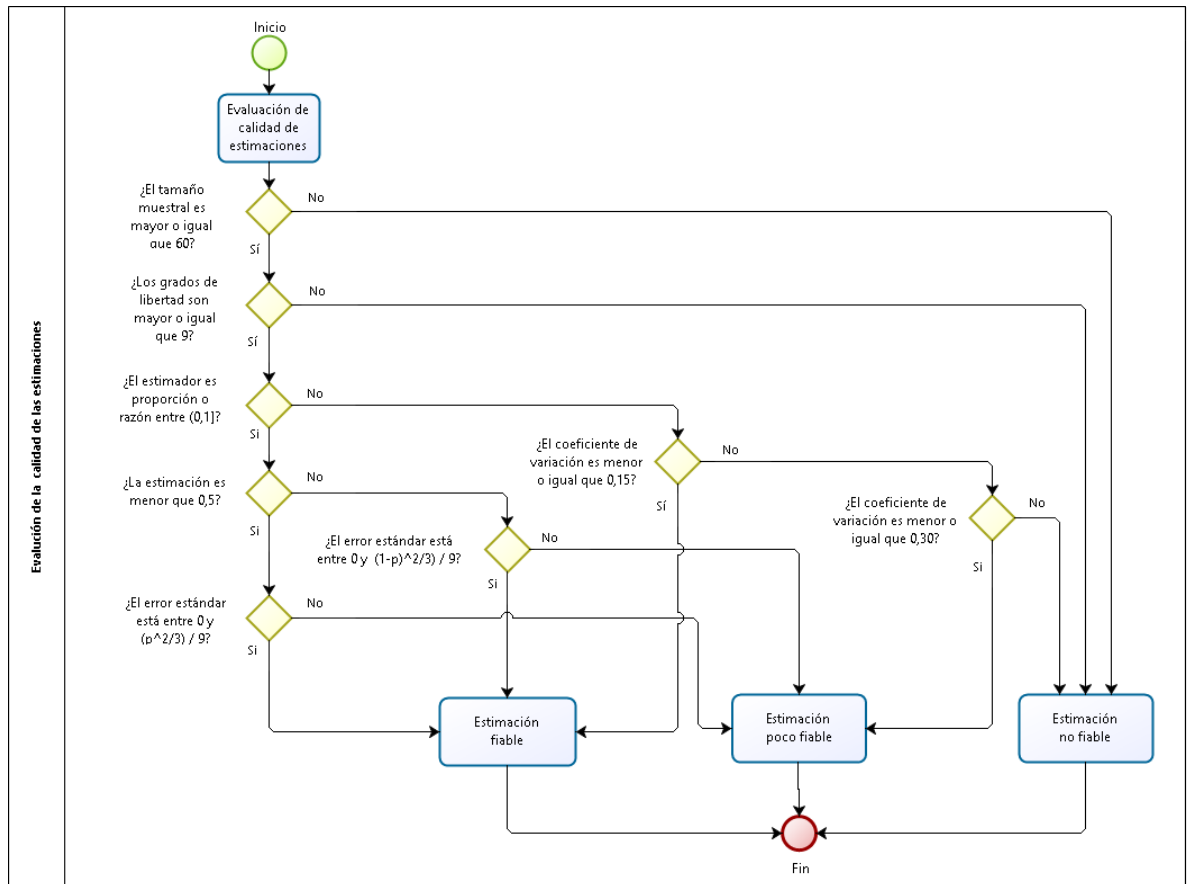
¹¹ Disponible en: <https://www.ine.cl/docs/default-source/institucionalidad/buenas-prácticas/clasificaciones-y-estandares/estándar-evaluación-de-calidad-de-estimaciones-publicación-27022020.pdf>

¹² Sin pérdida de generalidad, al señalar el término “tamaño muestral” se hará referencia al total de unidades de análisis consideradas para obtener las estimaciones. Con esto se busca comprender las diferentes unidades de análisis que se pueden encontrar en las encuestas de hogares por muestreo, a saber: viviendas, hogares y/o personas

¹³ Los grados de libertad (gl) se definen como el número de observaciones independientes que son libres de variar (número de individuos en los datos) menos el número de parámetros estimados (número de relaciones impuestas a los datos). Los gl constituyen uno de los elementos a considerar para poder emplear las distribuciones de frecuencia en las labores de inferencia. Más aún, dan cuenta de la precisión con la que se están estimando los parámetros a través de la muestra, ocurriendo que, mayores gl conducen a estimaciones más precisas.

variación supera 30% la estimación debe ser publicada con una nota que indique que la *estimación no fiable*.

Figura 2: Flujo para la evaluación de la calidad de las estimaciones



Fuente: Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares (INE, 2020)

Publicación de tabulados

El estándar determina que un tabulado es publicable si al menos el 50% de sus estimaciones son fiables (Figura 2).

Publicación de gráficos

Si bien el estándar no se pronuncia sobre los criterios para la publicación de gráficos se recomienda aplicar el mismo criterio que el indicado para los tabulados, pero sin considerar las categorías de “no sabe/no responde” u “otro especifique”.

4.10. Estimación de parámetros

En este apartado se describen las fórmulas para calcular los principales indicadores de la encuesta. La nomenclatura utilizada para la descripción de los indicadores es la siguiente:

Factor_EME: Ponderador para la persona Microemprendedora

P_c : es una variable binaria que representa con un 1 a todos los individuos que cumplen una determinada característica c y con un 0 a quienes no la cumplen.

P : Variable Bernoulli que considera los siguientes valores

$$P = \begin{cases} 1; & \text{Si el microemprendedor cumple la característica } c \\ 0; & \text{En otro caso} \end{cases}$$

Y_c : es una variable binaria que asigna el valor de la variable Y para los encuestados que cumplen la característica C y un 0 a quienes no la cumplen.

Y : Variable Bernoulli que considera los siguientes valores

$$Y = \begin{cases} Y_c * 1 = Y_c; & \text{Si el microemprendedor cumple la característica } c \\ Y_c * 0 = 0; & \text{En otro caso} \end{cases}$$

4.10.1. Formula estimadores

Antes de describir los parámetros, es importante indicar que las estimaciones realizadas a través de la encuesta de microemprendimiento no consideran a los encuestados que cambiaron de situación ocupacional, esto debido a que ya no mantienen su condición de trabajador independiente. Sin embargo, a estos encuestados se les realiza algunas preguntas sobre su situación actual capturadas a través del módulo A. Las respuestas de estas preguntas se encuentran disponibles en la base de datos.

A continuación, se describen las fórmulas de los principales indicadores utilizados:

- Total: Corresponde al conjunto de todos los elementos sobre los cuales se observa una o más características. Por ejemplo, total de personas microemprendedoras. El cual se obtiene como el total de los individuos de la muestra que no cambiaron de situación ocupacional ($a_1=1$ o 2) ponderado por los factores de expansión. Donde M corresponde a los microemprendedores que cambiaron de categoría ocupacional y N al total de microemprendedores.

$$t = \hat{T}_y = \sum_{i=1}^M F_i \quad i \in M \text{ y } M \in N$$

La EME genera estimaciones para los siguientes indicadores de totales:

- Número de microemprendedores
- Número de microemprendedores informales
- Número de microemprendedores sin acceso a la seguridad social
- Número de microemprendedores por categoría ocupacional
- Número de microemprendedores por categoría ocupacional desagregada
- Media: Corresponde a la suma de los atributos de una característica o variable continúa ponderada por el factor de expansión de la unidad estudiada (individuos, empresas, etc.), la cual luego es dividida por el número total de elementos estudiados.

$$\bar{y} = \hat{y} = \frac{\sum_{i=1}^M y_i * F_i}{\sum_{i=1}^M F_i} \quad i \in M \text{ y } M \in N$$

Algunos ejemplos de estimadores de media (y_i) en la EME son:

- ganancia promedio de los microemprendedores
- Tiempo promedio de trabajo global
- Tiempo promedio de trabajo no remunerado
- Tiempo promedio de trabajo doméstico no remunerado
- Tiempo promedio de trabajo de cuidado no remunerado

- Proporción o tasas: Fracción de unidades que estén clasificados dentro de alguna clase definida sobre el total de un grupo mayor

$$p = \hat{p} = \frac{\sum_{i=1}^C F_i}{\sum_{i=1}^N F_i} \quad i \in C \text{ y } C \in M$$

Donde C corresponde al total de Microemprendedores que cumple la característica y N corresponde a los microemprendedores que cambiaron de categoría ocupacional.

- Tasa de microemprendedores por cuenta propia
- Tasa de microemprendedores informales
- Tasa de microemprendedores sin acceso a la seguridad social

- Razón: Definido como el cociente de dos totales o medias, por ejemplo, promedio de tiempo destinado al trabajo no remunerado de mujeres sobre el promedio del tiempo destinado al trabajo no remunerado de los hombres.

$$r = \hat{r} = \frac{\hat{Y}}{\hat{X}} = \frac{Y}{X} \quad i \in M \text{ y } M \in N$$

Donde \hat{Y} corresponde al promedio de tiempo destinado al TNR por mujeres y \hat{X} al de hombres.

5. METODOLOGÍA DE RECOLECCIÓN DE DATOS

El presente capítulo detalla las metodologías implementadas en el proceso de recolección de datos. El primer subapartado retrata las 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. El tercer subapartado detalla el método de capacitación, de encuestadores, supervisores y analistas; y finalmente en el cuarto subapartado se describen 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. Técnicas de recolección de datos.

Para poder estudiar a la población objetivo del estudio se optó por realizar una encuesta a hogares. Esta técnica de recolección de datos permite obtener información de todos los microemprendimientos inclusive de aquellos que se desarrollan al margen de la formalidad, los cuales no están incluidos de los directorios de las encuestas a establecimientos.

5.1.2. Medio de recolección de información utilizado.

A diferencia de las versiones anteriores, la recolección de la VII EME se realizó utilizando el método de captura CAPI¹⁴ (computer asisted personal interview), a través de dispositivos móviles de captura (DMC) y del uso de la plataforma Survey Solutions¹⁵, que incluyó la hoja de ruta, el cuestionario y las observaciones de terreno.

Respecto del modo de aplicación, se realizaron entrevistas presenciales y telefónicas dependiendo de las características y disponibilidad de la persona informante. En el caso de Microemprendedores de alto impacto (Empleadores con gran cantidad de trabajadores o actividades económicas específicas) se priorizó la realización de la encuesta en modalidad presencial. Es importante indicar que previo al levantamiento oficial se realizó una prueba piloto que estudió el impacto del cambio de medio (desde papel a DMC) y el modo (Presencial v/s telefónico), sin encontrar impactos significativos. Para más detalles ver Anexo 1.

En cuanto al tipo de informante, la EME utiliza informante directo, es decir, cada persona seleccionada responde por sí misma.

5.1.3. Período de referencia

El período de referencia de donde se obtiene la muestra es el trimestre marzo – mayo de 2022 de la Encuesta Nacional de Empleo. El trabajo de campo de la EME se realizó durante el trimestre mayo, junio, julio de 2022.

Adicionalmente, el cuestionario de la VII 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, mayo 2021 a abril 2022, desde marzo 2020, pandemia COVID-19 y todo el tiempo que ha trabajado como microempendedor.

¹⁴ Se ocupó el método de captura PAPI (paper and pencil interview), utilizando un formulario en papel para situaciones excepcionales que impedían aplicar el DMC.

¹⁵ El sistema Survey Solutions corresponde a un software diseñado por el Banco Mundial para el desarrollo de encuestas.

5.1.4. Fechas de recolección de datos y procesos complementarios

El proceso de recolección de datos inició con la capacitación de los encuestadores, encargados de grupo y supervisores el 25 de abril del 2022. Inmediatamente terminada la capacitación comenzó la recolección de la primera de las tres submuestras que conformaron la muestra de la VII EME.

A pocos días de que se iniciara el proceso de recolección de información de la segunda submuestra, 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 junio y julio de 2022 y a lo largo del país.

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



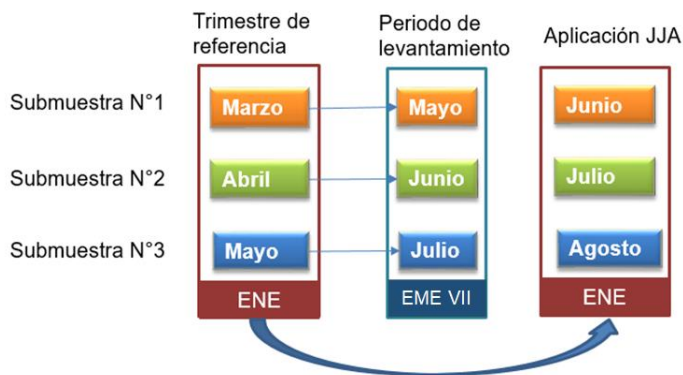
Fuente: INE, VII Encuesta de Microemprendimiento, 2022

La muestra de la VII EME estuvo dividida en tres submuestras, cada una de ellas se encuestó a un grupo específico de informantes. Esta subdivisión del trabajo de recolección fue realizada para no interferir la recolección de la ENE, para ello el trabajo de campo inició dos meses después de la recolección de cada submuestra de la ENE¹⁶.

El siguiente diagrama resume en qué momento se encuestó a cada integrante de la muestra dependiendo del mes en fue entrevistado por la ENE.

¹⁶ Una vez seleccionada una vivienda a participar en la ENE esta es entrevistada 6 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 meses. Dado esto, la mayor parte de la EME se aplica sobre muestra activa de la ENE, de manera que 5/ 6 de la muestra serán vueltos a entrevistar, por lo que la recolección de datos debe cuidar que éste no interrumpa el ciclo de visitas realizadas por la ENE.

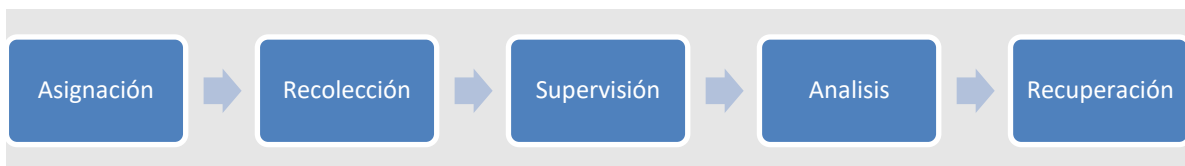
Diagrama 2: Esquema de participación de los informantes en la ENE y EME.



Fuente: INE, VII Encuesta de Microemprendimiento, 2022

A rasgos generales, el flujo consistía en una asignación de carga de trabajo al personal de recolección realizada por los encargados de grupo (asignación), quienes luego procedían a recolectar la información (recolección). Una vez recolectada la información, se supervisaban¹⁷ las encuestas logradas (supervisión) y, en caso de que ocurriera alguna situación que pusiera en cuestión la calidad de la información recolectada, las encuestas volvían a los encargados de grupo, quienes luego se las devolvían a los recolectores. Si las encuestas supervisadas no presentaban ningún problema, pasaban a ser analizadas por el equipo técnico. Luego, el equipo técnico realizaba una nueva validación de consistencia de las encuestas para identificar encuestas que no cumplieran con los criterios de calidad suficientes (consistencia y completitud), si la encuesta no cumplía estos criterios, estas eran devueltas al equipo de recolección, sino se clasificaban como entrevista finalizada. En los casos que las encuestas fueran devueltas por supervisión o por el análisis del equipo técnico, los recolectores volvían a contactar a los informantes en un periodo de tiempo dedicado a la recuperación de información.

Diagrama 3: Flujo operativo general del proyecto VII EME



Fuente: INE, VII Encuesta de Microemprendimiento, 2022

¹⁷ Además de la supervisión indirecta, que se realizaba posterior a la recolección de la encuesta, la metodología de supervisión consideraba una supervisión directa, la cual consistía en un acompañamiento presencial al personal de recolección realizado por el equipo de supervisión. Ambos tipos de supervisiones se realizaban por medio de una pauta predeterminada, las cuales diferían entre ellas ya que tenían diferentes objetivos, mientras la supervisión directa se enfocaba en la evaluación del desempeño de la aplicación de la encuesta, la supervisión indirecta, la s

5.1.5. Equipo de trabajo

A lo largo del país fueron 68 encuestadores los que se encargaron de recolectar la información de la VII EME, a ellos se sumaron 8 supervisores, 18 encargados de grupo, 2 encargados de área y 1 encargado nacional que monitorearon el avance del trabajo y la calidad del mismo. A continuación, en la tabla 11, se expone la distribución geográfica de los encuestadores y supervisores que tomaron parte de esta versión de la Encuesta de Microemprendimiento.

Tabla 11: Encuestadores y supervisores por región

Región	Recolector de datos	Encargados de grupo	Supervisores	Encargados de área	Encargado Nacional
Arica y Parinacota	3	1			
Tarapacá	3	1			
Antofagasta	2	1			
Atacama	2	1			
Coquimbo	4	1			
Valparaíso	8	2	1		
Metropolitana	14	3	4	2	1
O'Higgins	3	1			
Maule	4	1	1		
Biobío	6	1	1		
Ñuble	3	1			
La Araucanía	4	1			
Los Ríos	3	1			
Los Lagos	5	1	1		
Aysén	2	1			
Magallanes	2	1			
Nacional	68	18	8	2	1

Fuente: INE, VII Encuesta de Microemprendimiento, 2022

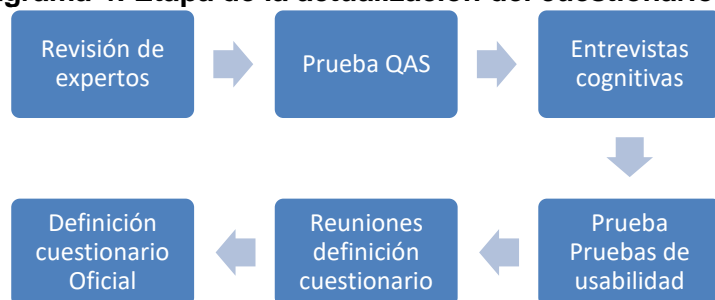
5.1.6. Terminó proceso de recolección.

La última etapa correspondiente al proceso de recolección corresponde validación de la coherencia y la consistencia de las respuestas de los informantes, ante la presencia de alguna inconsistencia se devuelve la encuesta a terreno. Esta etapa concluye luego de que la encuesta pasa la grilla de suficiencia aplicada por parte del equipo técnico.

5.2. Diseño del instrumento de captura de información

En este apartado abordaremos el diseño del cuestionario el cual varias etapas que permiten construir el nuevo cuestionario de recolección. a pilotear y en el siguiente las otras dos etapas de la actualización del cuestionario.

Diagrama 4: Etapa de la actualización del cuestionario



Fuente: INE, VII Encuesta de Microemprendimiento, 2022

La primera parte corresponde a la actualización del cuestionario piloto, que comienza a inicios del año 2021 y se puede dividir en la aplicación de cuatro técnicas de evaluación de cuestionarios: revisiones de expertos, prueba QAS entrevistas cognitivas y pruebas de usabilidad.

La revisión de expertos tenía por objetivo determinar nuevas áreas de información que pudieran entregar información relevante sobre nuevos fenómenos que afecten a las personas microempendedoras. Esto se tradujo en el incremento del número de preguntas del cuestionario de la VII EME. Estas preguntas se relacionan con cuatro temas: (1) Impacto de la pandemia covid-19 en microemprendimientos, (2) Tecnologías de información (grado de digitalización de los negocios), (3) Análisis de consistencia en preguntas de gastos, ingresos y ganancias del negocio, y (4) Ajuste en la medición de trabajo no remunerado.

Las pruebas QAS, consistieron en la aplicación de pruebas estandarizadas que permiten recopilar de forma sistemática la opinión de expertos respecto a potenciales problemas en el diseño de preguntas. Específicamente, se llevaron a cabo dos pruebas QAS durante agosto de 2021 a dos analistas de la institución. Se pusieron a prueba 31 de 91 preguntas del cuestionario (32% del total de preguntas).

Las entrevistas cognitivas, se dieron teniendo lugar una serie de entrevista cualitativas para comprobar que las preguntas del cuestionario cumplieran con los estándares cognitivos requeridos o, dicho de otra forma, que los informantes pudieran comprender y responder las preguntas según los objetivos y la información esperada de cada una. En particular, se realizaron 12 entrevistas cognitivas durante septiembre de 2021. Estas permitieron poner a prueba la incorporación de nuevas preguntas, así como también evaluar la prevalencia de ciertas dificultades en el proceso de respuesta que fueron identificadas en aplicaciones pasadas del cuestionario.

Las Pruebas de usabilidad, consistieron en dos aplicaciones del cuestionario en pequeña escala, con el objetivo de poner a prueba la incorporación de la plataforma Survey Solution (SuSo) como medio de recolección de datos. Se pusieron a prueba aspectos como el flujo de respuesta, validaciones por pregunta y uso de la hoja de ruta. En concreto, las pruebas

se aplicaron entre octubre y diciembre del año 2021 a 20 informantes (por prueba) recolectados a través de un muestreo tipo bola de nieve.

El cuestionario que surgió de este proceso conservó gran parte la estructura y contenido. El número de módulos y el contenido de ellos no cambió significativamente (ver tabla 17). Debido a la identificación de nuevas necesidades, el número total de preguntas aumentó de 97 a 123, incluyendo 30 preguntas nuevas y eliminando 4 de la versión anterior.

Luego de la definición del cuestionario preliminar, se realizó una prueba piloto en la Región Metropolitana entre los días 10 de enero y 18 de febrero de 2022, la que consideró un período de cinco semanas para la recolección de información y una para la recuperación. La muestra total viviendas fue de 896 y la muestra total de casos fue de 1.007 personas microempendedoras. Uno de los principales hallazgos del cuestionario piloteado fue el aumento en la duración del cuestionario en promedio de 30 a 35 minutos.

Para la definición del cuestionario oficial, se realizaron algunas reuniones con el Ministerio de Economía, Fomento y Turismo, en las cuales se realizaron las últimas modificaciones con el objetivo reducir la extensión del cuestionario, lo que implicó eliminar algunas preguntas.

A continuación, se muestra la tabla 12 con el resumen del proceso, la cual compara el número de preguntas por modulo para la VI EME, VII EME Piloto y VII Oficial:

Tabla 12: Cambios cuestionario Prueba Piloto VII EME

AREAS DE ESTUDIO	VI EME	VII EME (Piloto)	VII EME (L. Oficial ¹⁸)	Diferencia VII EME Oficial – VII EME Piloto	Diferencia VII EME Oficial – VI EME Oficial
A. Situación Laboral	12	13	13	0	1
B. Historia Laboral	9	11	9	-2	0
C. Características del Negocio o Actividad por Cuenta propia	9	12	9	-3	0
D. Gastos, Ingresos y Ganancias	12	16	13	-3	1
E. Situación de Formalidad	10	11	11	0	1
F. Generación de empleo	4	5	5	0	1
G. Relación con el sistema Financiero	13	15	10	-5	-3
H. Herramientas y Equipos	4	4	4	0	0
I. Tecnología de la información	4	10	9	-1	5

¹⁸ La diagramación completa del cuestionario se encuentra en el sitio web: <https://www.ine.gob.cl/estadisticas/sociales/mercado-laboral/microemprendimiento>, en la sección de formularios, documentos.

J. Capacitación	4	4	4	0	0
K. Entorno del negocio	13	17	15	-2	2
L. Trabajo No Remunerado	3	5	5	0	2
Total	97	123	107	-16	10

Fuente: INE, VII Encuesta de Microemprendimiento, 2022

5.3. Método de capacitación para la recolección de datos

La VII versión de la EME tuvo un proceso de capacitación, orientado hacia los encuestadores, encargados de grupo y supervisores, realizado tanto en regiones como en la Región Metropolitana. A continuación, se comentan algunos de los aspectos centrales del proceso.

5.3.1. Capacitaciones encuestadores, encargados de grupos y supervisores

Las capacitaciones tuvieron como objetivo entregar al personal de trabajo de campo todas las herramientas y brindar 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 cinco días, donde distintos equipos de la institución colaboraron para efectuar las jornadas de capacitación de manera simultánea en distintas macrozonas del país.

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 68 encuestadores y 8 supervisores, 2 encargados de área y un encargado a nivel nacional.

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

- Manual de trabajo de campo
- Cuestionario
- Dispositivo Móvil de Captura (DMC)
- Credencial
- Kit de materiales (mochila, tabla, lápices, jockey institucional, entre otros)
- Kit COVID-19 (mascarillas, alcohol gel, entre otros)

5.4. Métodos de supervisión

La supervisión tiene por objetivo controlar la veracidad y calidad de los datos obtenidos en la recolección mediante los procedimientos instalados en el diseño del trabajo de campo para el proyecto, permitiendo desplegar medidas correctivas durante el proceso de recolección.

Las supervisiones se llevaron a cabo en paralelo a la recolección de los datos y tenían dos modalidades: directas e indirectas.

5.4.1. 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 a la reinstrucción del encuestador.

Las supervisiones directas, se realizaron de manera periódica a lo largo de todo el proyecto, poniendo especial énfasis en aquellos recolectores que requirieran una mayor retroalimentación. Se dio mayor énfasis a recolectores que se encontraran en las siguientes situaciones:

- Alta tasa de rechazo (superior a la media del equipo).
- Problemas de ubicación de viviendas seleccionadas.
- Inconsistencias reiteradas en el llenado de formularios.
- Evidentes dudas que pudiesen presentar los recolectores.

Las supervisiones directas se materializaron en un formulario de supervisión, donde se registró la interacción recolector-informante en términos de: apertura de la vivienda, aplicación del cuestionario y cierre de visita. Ante cualquier dificultad observada, se retroalimentó al recolector.

5.4.2. 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.

Las supervisiones indirectas, también se realizaron de manera periódica a lo largo de todo el proyecto y se efectuaban en función de seis razones principales: por cambio de situación ocupacional, por elección aleatoria, por selección del encargado de grupo, por entrevista parcial o por no elegibilidad (ver **Tabla 13**). Estas supervisiones indirectas obligatorias permitían minimizar los errores en el reporte o registro de la información y, por ende, asegurar la calidad de los datos recopilados.

Tabla 13: Razones de supervisión indirecta obligatoria

Tipo de supervisión	Tópicos	Responsables
Supervisión indirecta obligatoria	Informante cambia de situación ocupacional (a1=3)	Informantes que dejan de ser microemprendedores fueron supervisados para verificar esta situación

Supervisión aleatoria por sistema	Entrevistas escogidas aleatoriamente para supervisión
Selección por encargado de grupo	Entrevistas que el encargado de grupo enviaba a supervisión, a raíz de alguna inconsistencia distintas a las nombradas o por alguna situación sospechosa
Entrevista parcial	Toda entrevista interrumpida fue supervisada
No elegibilidad	Códigos de disposición final: 61, 62, 63, 64, 65, 66, 67, 68 y 69

Fuente: INE, VII Encuesta de Microemprendimiento, 2022

6. METODOLOGÍA DE PROCESAMIENTO DE DATOS

En el caso de la EME el procesamiento de la base de datos se organiza en base al quinto proceso del Modelo Genérico del Proceso Estadístico (GSBPM), el cual se divide en cinco subprocesos: integración de bases de datos, codificación automática, revisar y validar, edición e imputación y derivación de nuevas variables. Cada uno de los subprocesos fue adaptado a las necesidades de la EME. El objetivo del procesamiento es contar con las bases de usuario integradas, codificadas, validadas e imputadas. En concreto, el procesamiento de los datos da como producto dos bases de datos:

- Base de Datos Full: es la base de datos del cuestionario completo, es decir, contiene todas las variables de todos los módulos de la VII EME. Por ejemplo, gastos, inscripción en el SII, trabajo no remunerado, entre otros. Cada fila representa un informante de la EME.
- Base de Datos Empleo: base de datos de la pregunta F2 del Módulo F “Generación de Empleo”. Esta base contiene variables relacionadas a los trabajadores de los microemprendedores, como sus características sociodemográficas, las horas trabajadas o sus salarios. Cada fila representa a un trabajador de un informante determinado de la EME.

6.1. Integrar bases de datos

El subproceso de integración de bases de datos implicó la combinación de dos bases internas, con el fin de tener una base única a la cual se le pudieran aplicar los siguientes subprocesos. Las bases que se combinaron son:

- Base de Datos Cero: correspondiente a la base de datos que proviene directamente de la recolección, sin ningún tipo de procesamiento. Es una base de datos de uso interno, extraíble de la plataforma Survey Solutions (SuSo). La base de datos utilizada para la integración corresponde a la base descargada posterior al término

del trabajo de recolección.

- Base de Datos de Directorio: corresponde a una base de datos con información sociodemográfica de los informantes. Es una base de datos de uso interno, utilizada con la más extrema confidencialidad estadística e integrada a la base de datos con tal de incorporar variables relevantes para posteriores análisis con los datos de la EME.

La integración de ambas bases de datos da como producto la Base de Datos Integrada.

6.2. Codificación automática

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”.

6.2.1. Codificación CAENES y CIUO

6.2.1.1. 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 ¹⁹.

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 (CAENES²⁰) y el Oficio (Grupo Ocupacional) a nivel de subgrupo principal según el Clasificador Chileno de Ocupaciones (CIUO 08.CL²¹ en adelante CIUO).

6.2.1.2. 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 (A1=2) y a los casos que mantuvieron su actividad económica (A1=1) pero la actividad económica declarada, difería de la declarada en la ENE. El detalle de los resultados y la metodología de este procedimiento se encuentra disponible en el reporte de calidad.

¹⁹[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)

²⁰[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)

²¹http://webanterior.ine.cl/docs/default-source/publicaciones/2018/ciuo-08.cl-clasificador-chileno-de-ocupaciones.pdf?sfvrsn=50bd58d2_4.

6.2.2. **Codificación y reclasificación especifique**

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 26 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 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.

6.3. **Revisar y validar**

Este subproceso tuvo por objetivo la examinación de los datos a modo de identificar potenciales problemas, errores y discrepancias como: valores atípicos, falta de respuesta de los ítems y codificación errónea, en un procedimiento iterativo a partir de reglas predefinidas. Es fundamental considerar que este subproceso se encarga de la identificación de problemas en los datos, mientras que es en el subproceso siguiente en el que se realizan las modificaciones a tales discrepancias.

La realización de la actividad requirió de la Base de Datos Clasificada, producto del subproceso anterior, así como el documento de tratamiento de valores atípicos, la grilla de suficiencia y la documentación técnica de revisión y validación de datos. También se necesitó el manual metodológico y la ficha metodológica de la operación estadística.

La actividad buscó revisar y validar los datos a nivel de microdatos, a modo de asegurar el cumplimiento de los criterios de calidad estipulados. En específico, se efectuaron tres tareas para la revisión a nivel de microdatos: nominación y tratamiento de valores atípicos (outliers), la identificación de cumplimiento de la grilla de suficiencia, y la ejecución de validaciones en primer y segundo nivel.

6.3.1. **Nominación y tratamiento de valores atípicos.**

6.3.1.1. **Detección**

Se entiende como valor atípico (u outlier) a toda aquella observación que, encontrándose dentro de los rangos establecidos para una variable, resultan atípicas, ya sea por su distancia respecto del resto de las observaciones dentro de la misma variable, como por su relación con valores reportados para otros atributos de la misma observación. Es decir, corresponde a una observación no ajustada al modelo de la mayoría de los datos, pudiendo encontrarse en la cola de la distribución estadística, o lejos del centro de los datos.

Para la identificación de valores atípicos se aplicó el modelo estadístico Bacon (blocked adaptive computationally efficient outlier nominators) en las variables seleccionadas. Para ello se consideraron 4 variables independientes entre sí y correlacionadas con las variables de estudio sexo (2 categorías), Número de trabajadores asalariados (10 categorías),

registro en el SII (6 categorías), Actividad Económica (8 categorías). En el anexo 2 se explica con mayor detalle la metodología diseñada.

6.3.1.2. **Tratamiento valores outliers**

Una vez identificado como outlier, la validación del dato depende del análisis sucesivo de cinco etapas (metodología similar a la seguida en la VI EME). El detalle de las etapas es el siguiente:

1. Para cada caso identificado como outlier se revisó todas las observaciones provenientes desde recolección. En caso de contar con una observación que justifique el valor identificado como atípico. De ser así, se mantenían los valores registrados en la base de datos. En caso contrario, se continuaba con el segundo paso.
2. En caso de no justificar el valor atípico, 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. De no ocurrir lo constatado en el paso anterior, se analizó la pregunta c1 (actividad principal del negocio). Se asumió que el dato es válido siempre que la actividad principal justifique el outlier debido a la naturaleza del negocio (por ejemplo, servicios de transporte y gasto elevado en gasolina).
4. De no cumplirse lo señalado en el paso anterior, se construye el ratio: variables escrutada/ganancia, analizando si dicho ratio se encuentra por debajo el percentil 5 o por sobre el percentil 95. De no ser así, el dato es validado. Con este ejercicio se buscó controlar por el tamaño del negocio.
5. 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.

6.3.2. **Grilla de suficiencia**

El propósito de la grilla de suficiencia es establecer criterios de logro necesarios y mínimos para asegurar la calidad del levantamiento de los datos y otorgar el carácter de lograda a cada encuesta. El análisis de cumplimiento de la grilla de suficiencia fue realizado de forma periódica por el equipo técnico, entregando retroalimentación constante al equipo de recolección a través de reportes de rechazos a encuestas.

Una encuesta lograda fue considerada como completa si contaba con información consistente (establecida en los procedimientos de análisis) y válida (respuestas distintas a NS y NR) en algunas variables clave que permitían construir los Indicadores principales de la encuesta.

6.3.3. Validaciones de primer y segundo nivel

Como ejercicio de control de calidad de la información levantada, el equipo técnico diseñó e implementó una serie de reglas de validación de los datos. El propósito radica en que las unidades encuestadas tuvieran completitud en la información pesquisada y consistencia entre ésta. En este contexto se desarrollan una serie de validaciones de primer y segundo nivel, cuyo detalle se presenta a continuación.

Validadores de primer nivel: permiten identificar problemas de flujo, es decir, cuando un informante no responde una pregunta que debía ser respondida, o presenta una respuesta sin tener que hacerlo. Ambos tipos de problemas fueron abordados para todas las variables del cuestionario, tanto en la Base de Datos Full como en la Base de Datos Empleo.

Validadores de segundo nivel: permiten identificar problemas de coherencia global de la encuesta. En concreto, se aplicaron 75 validadores que tenían como objetivo identificar, para luego corregir, estos problemas de consistencia.

Al igual que con la grilla de suficiencia, se realizó un reporte periódico de las inconsistencias identificadas en los validadores de segundo nivel. En específico, cada reporte de rechazos alertaba al equipo de recolección sobre la necesidad de recuperar información en terreno para las preguntas involucradas en cada validador. Para finalizar, es importante mencionar que el procedimiento de recuperación se aplicó para 33 de los 75 validadores mencionados anteriormente. El resto de las inconsistencias se solucionaron mediante un proceso de edición.

Una vez terminado este subproceso, se obtuvo la Base de Datos Revisada y Validada.

6.4. Edición e imputación

Este subproceso tiene por objetivo realizar las modificaciones provenientes de problemas, errores y/o discrepancias identificadas en el subproceso anterior: revisar y validar. La realización de la actividad requirió de la Base de Datos Revisada y Validada en el subproceso anterior. También se necesitó el manual metodológico y la ficha metodológica de la operación estadística.

En específico, hay cuatro tareas a cumplir para la revisión a nivel de microdatos: (1) imputaciones provenientes de validaciones de primer nivel, (2) imputaciones provenientes de validaciones de segundo nivel, (3) imputación de casos que participaron en la prueba piloto, y (4) Tratamiento de inconsistencia en el número de socios.

6.4.1. Imputaciones provenientes de validaciones de primer nivel

Como fue mencionado, los validadores de primer nivel identificaron problemas de flujo: (1) cuando un informante no respondió una pregunta que debía ser respondida, o (2) presentó una respuesta sin tener que hacerlo. Para el primer caso, se imputó 96 "Sin dato" en aquellas preguntas que debieron ser respondidas, pero no lo fueron. Finalmente, en el caso de preguntas que fueron respondidas cuando no debían serlo, se imputó un missing value (NA).

6.4.2. Imputaciones provenientes de validaciones de segundo nivel

Como se señaló en el apartado correspondiente al subproceso Revisar y Validar, los validadores de segundo nivel permiten identificar problema de coherencia global de la encuesta. Para los 75 validadores (presentados en ANEXOS 4) hubo dos procedimientos de tratamiento de acuerdo a la naturaleza de la inconsistencia. El primero de ellos consistió en devolver la encuesta a terreno para la corrección del problema identificado. Como se mencionó anteriormente, 33 de los validadores contemplaron este procedimiento de solución. Por otro lado, para los 42 validadores restantes se realizó una imputación de una o más variables que permitiera solucionar la inconsistencia.

6.4.3. Imputaciones provenientes de prueba piloto

Es importante mencionar que aquellos informantes que participaron en la prueba piloto de la VII EME, y además responden mantener su actividad económica (a1=3) en el levantamiento oficial, sólo responden el módulo D de gastos, ingresos y ganancias. Por este motivo, es necesario imputar la información correspondiente al resto de módulos de acuerdo a lo respondido por el informante en la prueba piloto.

6.5. Derivación de nuevas variables

Este subproceso tuvo por objetivo depurar dos tipos de variables: las variables de gastos, ingresos y ganancias y las variables de trabajo no remunerado. El procedimiento está orientado a mitigar los problemas de falta de respuesta e inconsistencia de la información reportada, y como producto, generar variables procesadas lista para uso del usuario. Se usó como insumo la Base de Datos Editada e Imputada.

6.5.1. Depuración Gastos, Ingresos y Ganancias (GIG)

Una de las áreas más importantes de indagación de la EME es la sección de gastos, ingresos y ganancias de los negocios, en donde este último es uno de los principales indicadores utilizados para describir la situación económica de los micronegocios.

La carencia de respuestas en las encuestas de hogares es una situación habitual y no contar con estrategias para su tratamiento puede limitar la fase de inferencia estadística, impidiendo corregir errores no muestrales en las variables e indicadores sociales y económicos.

Debido a la variabilidad de los ingresos y la distinta periodicidad de los gastos efectuados por los microemprendedores, desde la VI EME se decidió indagar por un periodo de referencia más amplio, pasando a preguntar respecto de los últimos 12 meses, permitiendo calcular un monto de ganancia anual. Adicionalmente, se incluyeron variables de estimación directa de la ganancia, a través de las cuales se les solicitó a los informantes que indicarán el monto exacto obtenido el mes anterior y en los últimos doce meses. En este contexto, es factible construir la variable de ganancia anual a partir de la estimación resultante de los ingresos menos los gastos o bien a través de la información auto-declarada. Esta estrategia permite, por un lado mantener tasas de no respuesta relativamente bajas, pero por otra parte generar ciertas inconsistencias en las estimaciones realizadas, ya que los informantes entregan respuestas contradictorias.

A modo de exponer la metodología de edición e imputación empleada, esta sección se divide en cuatro apartados: análisis de respuesta en GIG, objetivos del proceso de depuración, depuración y estimación e imputación ganancia anual.

6.5.1.1. Análisis de respuestas en GIG

A partir de las preguntas de la séptima EME es posible obtener una ganancia autoreportada por un informante (pregunta d6b), o estimarla a través de otras variables (preguntas d1, d2, y d5 (o d4)).

En este contexto, se observaron dos problemas que a continuación, se detallan:

Problema 1: Ausencia de Respuesta

Existen cuatro problemas de ausencia de respuesta, los cuales se presentan por medio de la siguiente tabla, las casillas blancas corresponden a casos con respuesta válida y las negras a omisión de información.

Tabla 14: Ausencia de respuesta en preguntas de gastos, ingresos y ganancias

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

Fuente: INE, VII Encuesta de Microemprendimiento, 2022

El caso 1 describe a un informante con respuestas en todas las preguntas en estudio, donde además de tener una auto-declaración de ganancias, fue posible calcularla. El caso 2 muestra informantes con respuestas en las preguntas de ingresos y gastos mas no en las de ganancias auto-declaradas, 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. La tabla 15 resume los estados de respuesta según los casos anteriormente descritos:

Tabla 15: Ausencia de respuesta en preguntas de gastos, ingresos y ganancias

Situación	VII EME	
	N	Porcentaje respecto del total
Caso 1: Responde ingresos, gastos y ganancias	6112	92,9%
Caso 2: Responde ingresos y gastos, pero no las ganancias	120	1,8%
Caso 3: No responde ingresos o gastos, pero si las ganancias	284	4,3%
Caso 4: No responde ingresos o gastos, ni tampoco ganancias	63	1%
Total	6579	100%

Fuente: INE, VII Encuesta de Microemprendimiento, 2022

Además del problema de ausencia de respuesta, en los casos con respuesta en las preguntas de gastos, ingresos y ganancias, se evidenciaron inconsistencias en la información declarada.

Problema 2: Inconsistencias entre ganancias estimadas y auto-declaradas

Al analizar los casos que sí presentan respuesta a todas las preguntas en estudio (92,7 %), se puede evidenciar una diferencia entre la ganancia estimada por medio de los ingresos y gastos declarados respecto de la ganancia auto-declarada por los informantes. Esta situación es esperable, dado que los dueños de micronegocios pocas veces llevan una contabilidad completa o conocen en detalle los gastos o ingresos de su negocio.

Para esta revisión, se siguió la metodología de análisis empleada en la VI EME, la cual clasifica las ganancias –auto-declaradas (G.D) y estimadas (G.E) - en tramos construidos a partir del salario mínimo legal, evaluando si ambos valores pertenecen o no al mismo tramo. A través de este ejercicio se pueden ubicar a los informantes en los siguientes grupos:

- **Información consistente (caso 1.1.1):** ganancias en el mismo tramo e ingreso mayor a ganancia auto-declarada.
- **Inconsistencia leve (caso 1.1.2):** ganancias en un tramo de diferencia e ingreso mayor a ganancia auto-declarada.
- **Inconsistencia severa (casos 1.2):**
 - **Caso 1.2.1:** Ingreso menor a la ganancia auto-declarada y menos de dos tramos de diferencia.
 - **Caso 1.2.2:** Ganancia auto-declarada mayor en dos o más tramos a ganancia estimada.
 - **Caso 1.2.3:** Ganancia estimada mayor en dos o más tramos a ganancia autodeclarada.

La tabla 16 resume los estados de inconsistencia según los casos anteriormente descritos:

Tabla 16: Prevalencia por situación de consistencia gastos, ingresos y ganancias

Situación	VII EME	
	N	Porcentaje respecto del total
1.1.1 Información consistente en ganancias	3054	50%
1.1.2 Inconsistencia leve en ganancias (un tramo de diferencia)	1112	18,2%
1.2.1 Inconsistencia severa: Ingreso menor a ganancia y tramo de diferencia menor a 2	977	16%
1.2.2 Inconsistencia severa (G.E. <G.D)	870	14,2%
1.2.3 Inconsistencia severa (G.E. >G.D)	99	1,6%
Total	6112	100%

Fuente: INE, VII Encuesta de Microemprendimiento, 2022

6.5.1.2. **Objetivo del proceso de depuración**

El objetivo principal del proceso de depuración e imputación de ingresos, gastos y ganancias es construir una variable de ganancia anual que pueda reflejar de la mejor manera la situación económica del negocio. Para ello es necesario contar con estrategias de mitigación para tratar los errores no muestrales. En detalle, este proceso tiene los siguientes fines:

1. Resguardar la precisión de estimaciones y evitar los errores de medición: desde la IV EME uno de los principales indicadores para describir la situación económica de los micronegocios es la ganancia. Para construirlo es necesario realizar procedimientos de depuración, estimación e imputación de gastos, ingresos y ganancias. Carecer de una fase de procesamiento con estrategias ad hoc para la depuración e imputación de estas variables, podría afectar los niveles de desagregación (región y macro zona) de las estimaciones finales.

2. Reducir los errores no muestrales de las preguntas de gastos, ingresos y ganancias: para tratar las inconsistencias en las preguntas de gastos, ingresos y ganancias, propias de las encuestas de ingresos, se realizan mecanismos de edición y depuración de datos enfocados en corregir los problemas de consistencia interna de las estimaciones. Los que provienen de distintas fuentes de errores, entre las que destacan, por ejemplo, los problemas de comprensión del concepto de ingreso, que muchas veces es confundido con el de ganancia.

3. Mejorar la consistencia interna de la información: el procedimiento de depuración mejora la consistencia entre la ganancia estimada y auto-declarada, lo que permite también resguardar la comparabilidad con otras fuentes de captura.

4. Mantener comparabilidad con versiones anteriores: si bien la VI EME introdujo cambios relevantes en el cuestionario modificando la metodología de indagación de gastos, ingresos y ganancias producto de la variación de fraseos de preguntas, alternativas de respuesta y la ampliación del periodo de referencia, parte de las inconsistencias observadas en esta versión también se evidenciaron en la EME anterior. Por ello, es importante

considerar los procedimientos de depuración e imputación aplicados en las versiones pasadas como punto de partida para la creación de una metodología actualizada de depuración e imputación de datos.

6.5.1.3. **Depuración y estimación de ganancias (casos 1 y 2)**

Antes de construir la ganancia final se deben realizar dos procedimientos. El primero de éstos consiste en la depuración de las respuestas cuando se presentan inconsistencias severas entre lo estimado y lo auto-declarado (Caso 1). Luego, para quienes cuentan con respuestas válidas de los ingresos y los gastos, y carecen de la declaración de ganancia, se realiza la estimación de ganancia (Caso 2), contrastando esta última con información auxiliar.

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

El proceso de depuración propone la corrección en tres casos:

- **Caso 1.2.1:** ganancias estimadas y auto-declaradas consistentes (o con una inconsistencia leve) pero con ingreso menor a la ganancia auto-declarada. 977 casos detectados.
- **Caso 1.2.2:** casos con ganancias estimadas menores en más de un tramo a la autodeclarada. 870 casos detectados.
- **Caso 1.2.3:** casos con ganancias estimadas mayores a las auto-declaradas. 99 casos detectados.

Depuración casos 1.2.1 y 1.2.2: A partir de lo observado en terreno en las visitas técnicas a regiones y el análisis de los datos, se evidencia que muchos informantes confundían el concepto de ingresos con el de ganancias, declarando la información de ganancias en ambos. Por lo tanto, la depuración a realizar en estos casos consistió en asignar a la variable de ganancia estimada depurada el ingreso, y no calcular esta información a través de la resta de los ingresos y los gastos. Este procedimiento sólo se realizó en aquellos casos en que la nueva ganancia estimada depurada no fuese mayor a la ganancia declarada por el informante.

Depuración casos 1.2.3: A la luz de la experiencia recogida en la VI EME, se propone que en los casos pertenecientes al problema 1.2.3 la variable de ganancia depurada sea igual a la ganancia auto-declarada.

Caso 2: Estimación de ganancias en casos que no las auto-declaran

En esta sección, se realiza una propuesta de estimación de ganancias cuando se tiene información completa sobre los ingresos y gastos del negocio, y el informante no auto-declara un monto de ganancias. Esto representa 1,8% de los informantes (120 casos). Para calcular la ganancia estimada depurada, al igual que en el caso 1, se evalúa si el informante confunde los conceptos de ingresos y ganancias. Para determinar si esto ocurre, se analiza si la ganancia estimada es menor a cero o no. Además, se construye una variable de ganancia donada instrumental (la cual sólo es utilizada para comparar la información entregada por el informante) que se compara con la ganancia estimada y el ingreso declarado por el informante.

Caso 3: Sólo ganancia declarada

En el caso 3 (4,3%) se decidió considerar a la ganancia declarada como la ganancia final del negocio.

6.5.1.4. **Imputación de ganancia anual (caso 4)**

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 (INE, 2018b), EPF (INE, 2018d), entre otras. En estos casos, se imputó la variable de ganancia final por medio de un informante con las mismas características de aquel que no tenía información. La información donada se obtuvo a partir de la mediana de la ganancia final resultado de la solución de los casos 1, 2 y 3.

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 ocho variables: sexo, edad, educación, horas trabajadas, situación en la ocupación (CISE), tamaño del negocio (según número de trabajadores), tipo de inscripción ante el SII y registro contable.

A partir de estas variables se construyeron 11 niveles o perfiles de informantes, los cuales resultan de combinaciones de estas 8 variables. El primer nivel se construye a través de combinaciones desagregadas de estas 8 variables: las cuales presentan 2 categorías de sexo, 11 distintos grupos de edad, 6 niveles de educación, 15 jornadas de trabajo, 2 categorías ocupacionales, 5 de tamaño de la unidad económica, 7 tipos de Inscripción ante SII, 6 tipos de contabilidad.

El proceso de identificación de donantes comienza buscando en el primer nivel definido (Nivel 1) donantes que den su información de ganancias. Si no se encuentran donantes que tengan el mismo perfil del receptor se avanza 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 se avanza 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 termina una vez encontrado donantes para todos los receptores de información (Caso 4).

Para esta versión el número mínimo de donantes es 3, siguiendo la experiencia de la VI EME.

6.5.1.5. **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 auto-declarada dependiendo del caso. En líneas generales en los casos en que se cuenta con respuesta de gastos ingresos y ganancias, la variable de ganancia final tomará el mayor valor entre ambas ganancias. Cabe decir que en los casos 1.1.1 y 1.1.2 la ganancia estimada depurada es la ganancia estimada, por lo que se contrastaría la ganancia estimada y la ganancia auto-declarada.

En los casos que no se cuenta con información de ganancia auto-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 2.1) y el ingreso en los casos que si confundan estos conceptos (Casos 2.2 y 2.3).

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.

6.5.2. Depuración variables trabajo no remunerado (TNR)

Desde la VI versión de la EME, el módulo L del cuestionario recoge información respecto al trabajo no remunerado. Este módulo recoge información respecto sobre una serie de actividades relacionadas a trabajo doméstico y de cuidados. A cada informante se le preguntó por la participación en cada una de estas actividades, en el día de referencia asignado. En caso de respuesta positiva, el encuestado responde cuántas horas y minutos dedicó a dicha actividad en el día de referencia.

El proceso de creación de las variables se inició transformando los valores a minutos diarios.

El detalle de estas actividades se encuentra en la Tabla 17:

Tabla 17: Listado de Actividades de Trabajo No Remunerado consultados VII EME

N°	Actividad	Tipo de trabajo
TD1	¿Cocinó, puso o recogió la mesa, lavó la loza o limpió la cocina?	Trabajo Doméstico No remunerado
TD2	¿Limpió u ordenó el interior de la vivienda, jardín, balcón o sacó la basura?	
TD3	¿Cuidó mascotas o plantas en su hogar? Excluya trabajo en huertas.	
TD4	¿Lavó, planchó, guardó o cosió ropa? Excluya llevar a la lavandería o reparación de ropa.	
TD5	¿Compró alimentos, ropa o artículos para el hogar?	
TC1	¿Bañó, mudó, vistió o acostó a algún integrante de su hogar?	Trabajo de Cuidado No remunerado
TC2	¿Ayudó o dio de comer a algún integrante de su hogar?	
TC3	¿Llevó al colegio, ayudó en las tareas escolares, contó cuentos o leyó a algún integrante del hogar?	
TC4	¿Dio medicamentos, realizó algún tratamiento por alguna enfermedad o lesión, o acompañó a un centro de salud a algún integrante del hogar?	

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 influenciar 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 95 de la distribución en cada una de las nueve variables base que forman los indicadores mencionados en el párrafo anterior, asignándole el valor del percentil 95 a todos los casos que superan este límite.

6.6. Cálculo 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 seis ponderadores o ajustes: el cálculo ponderador base, ajuste por elegibilidad desconocida, ajuste por no elegibilidad, ajuste por no respuesta, suavizamiento y la calibración.

6.6.1. 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, condicional a que la vivienda de residencia fue seleccionada y lograda en la ENE en el trimestre MAM 2022. 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 2022.
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.

1. Probabilidad de selección y entrevista de las viviendas en la ENE- Trimestre MAM 2022

El diseño muestral de la Encuesta Nacional de Empleo corresponde a un muestreo probabilístico, estratificado y bietápico, donde las unidades primarias son las UPM; mientras que, las unidades de segunda etapa son las viviendas particulares.

Las unidades primarias de muestreo son seleccionadas en forma proporcional al tamaño, mientras que al interior de cada UPM las unidades secundarias se seleccionaron de forma sistemática y con igual probabilidad de selección. El factor de expansión que se extrae desde la ENE para la muestra de la VII EME es el factor de expansión ajustado por no respuesta para aquellas viviendas que contestan la ENE en el periodo MAM 2022 llamado w_{hij}^{NR22} .

²² Más información en https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/contingencia-covid-19/separata-t%C3%A9cnica-4-contingencia-covid-19.pdf?sfvrsn=37f5a108_6 .

2. Probabilidad de selección de los microemprendedores

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

Así, la probabilidad de seleccionar una vivienda que posee al menos un microemprendedor viene dada por la ecuación (8):

$$p_{Rj}^v = \frac{m_R^{micro}}{M_R^{micro}} \quad (8)$$

Donde:

R : Subíndice que representa la región de pertenencia. $R = 1, \dots, 16$.

j : Subíndice que representa la vivienda j .

p_{Rj}^v : Probabilidad de seleccionar la vivienda j perteneciente a la región R , según el listado de viviendas del marco EME, que poseen al menos un microemprendedor.

M_R^{micro} : Total de viviendas seleccionadas del marco EME, provenientes del periodo MAM de la ENE, con al menos un microemprendedor en la región R .

m_R^{micro} : total de viviendas seleccionadas desde el marco de la EME con al menos un microemprendedor en la región R .

Una vez seleccionada la vivienda se seleccionan los microemprendedores. La probabilidad de seleccionar al microemprendedor 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 ecuación (9):

$$p_{Rjklm}^{micro|v} = \frac{S_{Rjlm}^{micro}}{T_{Rjlm}^{micro}} \quad (9)$$

Donde:

T_{Rjlm}^{micro} : Total de microemprendedores identificados en la EME, en la vivienda j , hogar l , rama de actividad m , perteneciente a la región R .

S_{jlm}^{micro} : Total de microemprendedores seleccionados, en la vivienda j , hogar l , rama de actividad m , perteneciente a la región R .

Luego, la probabilidad condicional de seleccionar el microemprendedor k , en la vivienda j , de la región R , viene dada por la ecuación (10):

$$p_{Rjk}^{micro} = p_{Rj}^v \cdot p_{Rjklm}^{micro|v} \quad (10)$$

Luego, el ponderador base es obtenido a partir del factor ajustado por no respuesta en el trimestre MAM de la ENE y la probabilidad de seleccionar un microemprendedor desde la EME. El ponderador base está definido por la ecuación (11):

$$F_{Rjk}^{base} = w_{hij}^{NR} \cdot \left(\frac{1}{p_{Rjk}^{micro}} \right) \quad (11)$$

6.6.2. Ajuste por elegibilidad desconocida

Inicialmente, todas las unidades tienen un estatus conocido de elegibilidad debido a que el marco de selección es el conjunto de microemprendedores entrevistados en la ENE en el periodo marzo, abril y mayo 2022.

No obstante, durante el trabajo de campo, este estado no se puede comprobar en algunas unidades, clasificándolas como elegibilidad desconocida. Los pesos de las personas residentes en estas viviendas son eliminados y redistribuidos en las personas que residen en las viviendas que, si se conoce su elegibilidad, proporcionalmente al interior de cada estrato.

El ponderador que contiene los pesos a redistribuir $R_{h,known}$, se calcula como la razón entre la estimación del total de personas que residen en viviendas de la población y la estimación del total de personas que residen en viviendas con elegibilidad conocida (viviendas elegibles y no elegibles). Ambas estimaciones se realizan con el ponderador de selección de microemprendedor, es decir,

$$R_{h,known} = \frac{\sum_{j \in \Omega_R} \sum_{k \in \Theta_j} F_{Rjk}^{base}}{\sum_{j \in \Omega_R} \sum_{k \in \Theta_{j,known}} F_{Rjk}^{base}} \quad (12)$$

Donde,

- Θ_j Es el conjunto de micro emprendedores k seleccionados al interior de cada vivienda j en la región R .
- $\Theta_{j,known}$ Es el conjunto microemprendedores seleccionados en la vivienda j clasificada como elegibles o no elegibles.
- Ω_R Es el conjunto de estratos h de la EME (regiones).
- F_{Rjk}^{base} Es el ponderador de selección del k –ésimo microemprendedor de la vivienda j ubicado en la región R .

Esta razón es equivalente para todas las unidades muestrales pertenecientes a un mismo estrato.

Luego, el ponderador de selección de microemprendedor corregido por elegibilidad desconocida F'_{Rjk} está dado por:

$$F'_{Rjk} = R_{R,known} \cdot F_{Rjk}^{base} \quad (13)$$

6.6.3. Ajuste por no elegibilidad

Originalmente, la muestra de la Encuesta de Microemprendimiento contempla 8.577 personas residentes en viviendas particulares. De este total, resultaron 17 personas con elegibilidad desconocida, una vez finalizado el trabajo de campo. De las 8.560 personas

restantes 8.251 resultaron elegibles (logradas y no logradas) y; 309, no elegibles (ver tabla 10).

Dado el objetivo de generar inferencias hacia la población de microemprendedores que reside en viviendas particulares (elegibles) en esta etapa, se realiza un ajuste que consiste en descartar aquellas personas que no conforman la población objetivo de la encuesta; asignándoles un valor blanco o “missing” en el ponderador.

Tabla 18: Distribución de individuos según la clasificación de elegibilidad.

Clasificación elegibilidad	Cantidad personas	Porcentaje
Elegible	8.251	96,20%
No elegible	309	3,60%
Elegibilidad desconocida	17	0,20%
Total	8.577	100,00%

Fuente: Instituto Nacional de Estadísticas.

6.6.4. Ponderador ajustado por falta de respuesta

El ajuste por no respuesta consiste en aplicar un ponderador 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.

De un total de 7.850 viviendas seleccionadas, se seleccionaron 8.577 microemprendedores. De éstos, 8.251 fueron clasificados como elegibles (96,2%) de los cuales 6.942 respondieron la encuesta. Por lo tanto, la tasa de respuesta de la EME ajustada por elegibilidad es 84,1%.

El método utilizado para compensar la falta de respuesta es el de estratificación mediante “propensity score”. De acuerdo, a lo indicado por Valliant, Dever, & Kreuter (2013), este método consiste en modelar la probabilidad de respuesta en la VII EME como la realización de un proceso de variables latentes ($R_i^* = x_i^T \beta + u_i$), es decir, un conjunto de variables que inciden en la “motivación” (R^*) de participar o no en la encuesta. Así, mediante un conjunto de variables conocidas para quienes responden y quienes no responden se busca estimar la probabilidad de responder en la encuesta ($P(R_i^* > \theta)$) y posteriormente, ordenar de menor a mayor estas probabilidades conformando 6 grupos de igual frecuencia denominados sextiles.

El modelo que se utiliza para realizar el ajuste por no respuesta es el modelo logístico y se resume según 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 ajustes²³, se procede a estimar el ponderador ajustado por falta de respuesta, dado por la ecuación (14):

$$\hat{R}_c^{NR} = \frac{\sum_{k \in S_c} F'_{Rjk}{}^r}{\sum_{k \in S_{c,R}} F'_{Rjk}{}^r} \quad (14)$$

Donde:

- c : Subíndice de la celda de ajuste por falta de respuesta. $c = 1, \dots, 6$.
- S_c : Total de microemprendedores seleccionados y elegibles en la celda c .
- $S_{c,R}$: Total de microemprendedores seleccionados en la celda c y que responde la encuesta.
- $F'_{Rjk}{}^r$: Estimación del total de microemprendedores que responden, utilizando el ponderador ajustado por elegibilidad.
- $F'_{Rjk}{}^r$: Corresponde a la estimación de microemprendedores utilizando el ponderador ajustado por elegibilidad para la persona k , de la vivienda j , en la región R .

Luego, la expresión del ponderador ajustado por no respuesta, está dada por la ecuación (15):

$$F_{Rjk}^{NR} = F'_{Rjk}{}^r \cdot \hat{R}_c^{NR} \quad (15)$$

De acuerdo con la metodología antes expuesta, son 6 las celdas en las cuales se realizan los ajustes por falta de respuesta. En la siguiente tabla se presentan las tasas de respuesta para cada una de estas celdas, así como también el factor de ajuste por no respuesta (\hat{R}_c^{NR}).

Tabla 19: Total de unidades elegibles, unidades que responden y tasas de respuesta

Celda Ajuste	Total Elegibles	Total Responde	Tasa de Respuesta	\hat{R}_c^{NR}
Total	8.251	6.942	84,1%	1,04
1	1.376	768	55,8%	1,82
2	1.375	1.099	79,9%	1,27

²³ Las celdas de ajuste son 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 “sextiles” de respuesta, mediante un modelo logístico, en que se predice la probabilidad de responder, y en base a esta probabilidad ordenada en forma ascendente, se generan seis grupos de igual cantidad de individuos.

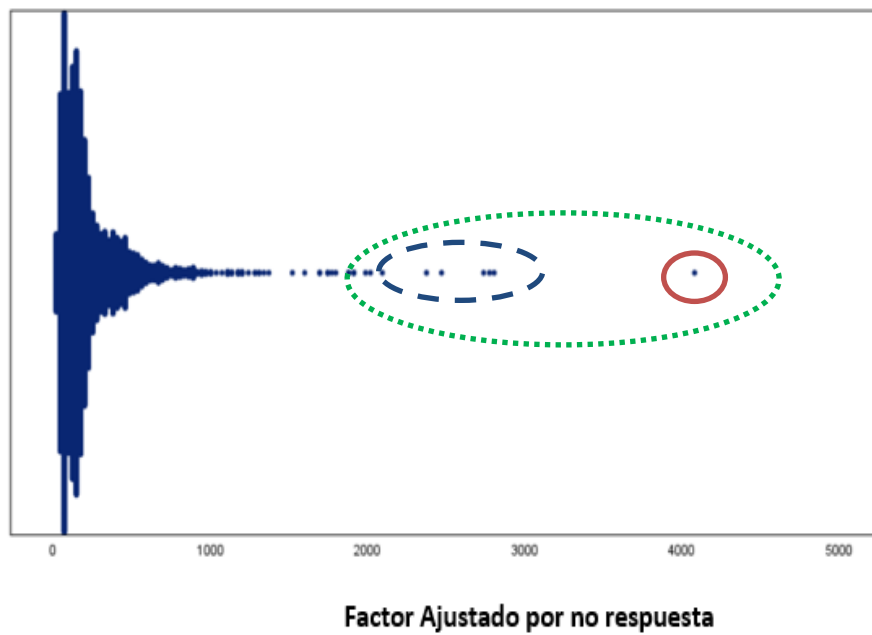
3	1.375	1.189	86,5%	1,14
4	1.375	1.267	92,1%	1,08
5	1.375	1.290	93,8%	1,07
6	1.375	1.329	96,7%	1,04

Fuente: Instituto Nacional de Estadísticas.

6.6.5. Suavizamiento del ponderador ajustado por no respuesta

Obtenido el factor de microemprendedores ajustado por no respuesta, se evalúa la pertinencia de realizar suavizamiento. En la figura 3, al observar los factores ordenados, permite identificar valores más altos que el resto de los ponderadores.

Figura 3: Dispersión del Ponderador ajustado por falta de respuesta



Fuente: Instituto Nacional de Estadísticas.

Según la figura 3 se decide realizar el suavizamiento del ponderador ajustado por falta de respuesta. Este suavizamiento se realiza por región y rama de actividad económica.

El método utilizado para el suavizamiento del factor corresponde a un método mixto (MM) conformado por el método de contracción a la media (CM) y el método R-K. El detalle de esta metodología se encuentra en el Informe Metodológico del Diseño Muestral VII EME.

6.6.5.1. Calibración

La estimación utilizada en la calibración del ponderador de la EME se obtiene a partir de los siguientes pasos:

Primero, se considera toda la información (referente a microemprendedores) levantada para la ENE en el período MAM 2022, es decir, todos los microemprendedores que fueron clasificados como tales, sin importar que, a futuro puedan cambiar de condición.

Se debe obtener la estimación del total de los microemprendedores a partir del levantamiento de la ENE al mes de abril y proyectarlas al mes de junio de 2022 que corresponde al mes central del levantamiento de la EME a nivel de región, área y sexo.

Se calcula un nuevo factor de expansión F_{hij}^{2cal} , obtenido según el punto 2, como se observa en la ecuación (16):

$$F_{hij}^{2cal} = F_{hijk}^{cal} \cdot \frac{p_{hs}^{junio}}{p_{hs}^{abril}} \quad (16)$$

Donde:

F_{hijk}^{cal} : Factor calibrado del periodo MAM 2022 de la ENE.

p_{hs}^m : Total de población por sexo y tramo de edad, del estrato ENE h, proyectada al 15 del mes m (abril) y a junio de 2022.

Cabe destacar, que, para efectos de generar estos nuevos stocks, F_{hij}^{2cal} , estos poderadores son redondeados a números enteros.

La Tabla 20, presenta el total de microemprendedores estimado a partir de la publicación de la ENE en el periodo MAM 2022; y según el total de personas estimadas con la información levantada en MAM 2022, pero con proyecciones actualizadas a la fecha de levantamiento de la EME. En esta tabla, se observa que, el total de “microemprendedores” estimados y publicados oficialmente son 2.086.024 personas. Sin embargo, al actualizar las proyecciones de población este total asciende a 2.089.383, lo que equivale a un incremento de 0,16% a nivel nacional.

Tabla 20: Total de microemprendedores estimados a partir de la ENE – periodo MAM 2022

Región	Sexo	Total Micro emprendedores	
		Factor Expansión Oficial ENE - MAM 2022	Factor Expansión Información ENE - ajustado a junio 2022
Total nacional		2.086.024	2.089.383
Total nacional	Hombre	1.236.463	1.238.447
	Mujer	849.561	850.936
Arica y Parinacota	Hombre	18.865	26.991
	Mujer	12.360	22.686
Tarapacá	Hombre	26.912	37.882
	Mujer	22.623	31.398
Antofagasta	Hombre	37.805	18.449
	Mujer	31.336	13.691
Atacama	Hombre	18.433	45.502
	Mujer	13.667	32.493
Coquimbo	Hombre	45.395	119.667

Región	Sexo	Total Micro emprendedores	
		Factor Expansión Oficial ENE - MAM 2022	Factor Expansión Información ENE - ajustado a junio 2022
Valparaíso	Mujer	32.406	86.856
	Hombre	119.451	50.573
Metropolitana	Mujer	86.705	33.787
	Hombre	530.427	71.750
O'Higgins	Mujer	390.762	45.583
	Hombre	50.481	87.180
Maule	Mujer	33.721	56.444
	Hombre	71.629	81.153
Ñuble	Mujer	45.484	45.273
	Hombre	34.398	66.474
Biobío	Mujer	16.621	30.420
	Hombre	87.101	9.678
La Araucanía	Mujer	56.385	8.002
	Hombre	81.076	12.386
Los Lagos	Mujer	45.226	7.229
	Hombre	66.381	531.342
Los Ríos	Mujer	30.369	391.339
	Hombre	26.069	26.085
Aysén	Mujer	16.688	16.706
	Hombre	9.671	18.907
Magallanes	Mujer	7.988	12.380
	Hombre	12.369	34.428
	Mujer	7.220	16.649

Fuente: Instituto Nacional de Estadísticas

Finalmente, el ponderador calibrado actualizado a las proyecciones de junio 2022 es un insumo para la calibración del factor. Esta información auxiliar se utiliza como un vector de totales poblacionales, llamados marginales de calibración, y para esta encuesta fueron consideradas las siguientes desagregaciones:

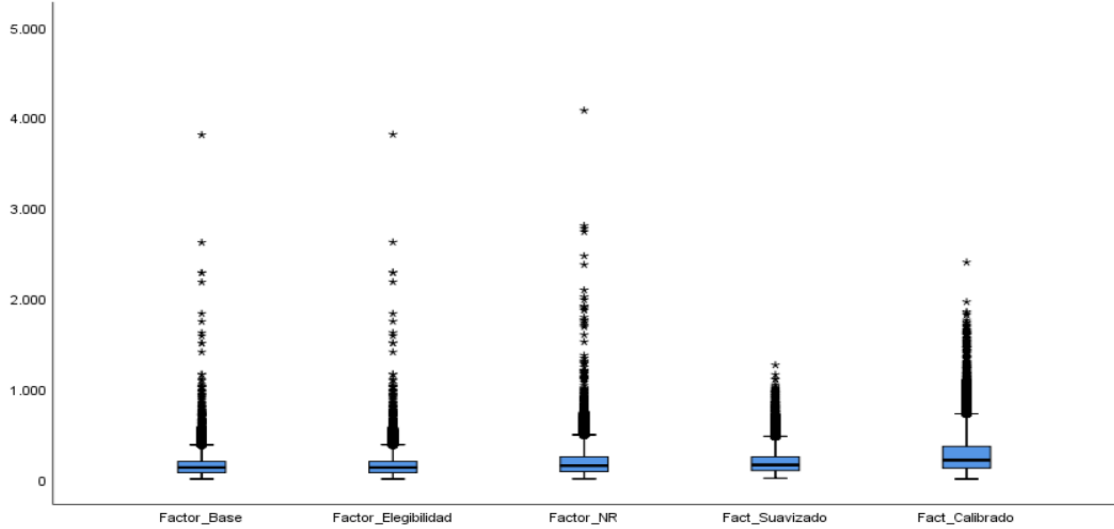
- A nivel nacional: Sexo y 11 tramos etarios: [15-24], [25-29], [30-34], [35-39], [40-44], [45-49], [50-54], [55-59], [60-64], [65-69] [70-100].
- A nivel regional: Población total según sexo.
- A nivel regional: Población total área urbana y sexo.

Una vez calibrado el factor se procede a revisar los coeficientes de variación de cada una de las marginales, cuyo valor máximo es 13,6%²⁴. A partir del proceso de calibración se obtiene una estimación de 2.089.383 personas a nivel nacional.

En la figura 4 se observan, de manera descriptiva, los ajustes del ponderador de selección de microemprendedor, hasta llegar al proceso de calibración.

²⁴ El umbral de calidad del coeficiente de variación se definió en 15% al interior de cada marginal.

Figura 4: Gráfico de caja para los diferentes ponderadores.



Fuente: Instituto Nacional de Estadísticas

7. METODOLOGÍA DE ANÁLISIS

En este apartado se describen los análisis realizados para estudiar el comportamiento de los resultados del producto estadístico. Así como la preparación de los resultados y la aplicación de los protocolos de calidad.

7.1. Análisis de los resultados

Para analizar los resultados se realizaron dos métodos, un primer análisis descriptivo y un segundo en el que se compararon algunos indicadores con otras fuentes de información.

7.1.1. Análisis descriptivo

Se realizaron análisis descriptivos con el fin de resumir los resultados arrojados por la VII 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.

7.1.2. 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.

- **VI 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, la formalidad del negocio, el financiamiento principal entre otras.
- **ENUT 2015.** Esta versión de la EME al igual que la anterior contiene un módulo de trabajo no remunerado, por lo que se recurrió a la Encuesta Nacional sobre 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 diarias dedicadas al trabajo no remunerado como tal, al trabajo doméstico no remunerado, al trabajo de cuidado no remunerado y al trabajo global, así como las principales diferencias entre la estimación de la ENUT y la de la EME.
- **ENE MAM 2022.** La EME obtiene su marco muestral a partir de la Encuesta Nacional

de Empleo. Esta situación llevó que ambas encuestas compartieran informantes, generándose la oportunidad de comparar algunos resultados relevantes entre ambos estudios, en especial lo que tenía que ver con la caracterización de los microemprendedores. Las variables analizadas fueron categoría ocupacional, rama actividad, lugar de trabajo, registro ante el Servicio de impuestos internos (SII), contabilidad, separación de gastos, uso de plataformas web y/o aplicaciones móviles por parte de los informantes, revisándose la estructura de éstas y la consistencia entre las respuestas entregadas en la EME y ENE considerando las diferencias metodológicas y temporales en la captura de las preguntas de ambas mediciones.

7.2. Preparar resultados y aplicar protocolos de calidad

En el caso de la EME la aplicación de los últimos protocolos de calidad (distintos a los descritos en el apartado de procesamiento), consideran la aplicación del proceso de anonimización de las bases de datos implementado por el INE. Este proceso 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 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 auditoria del proceso y liberación de datos (pasos 11 y 12).

7.2.1. Preparación de data previo al proceso de anonimización

A continuación, se detalla la preparación de los datos previo al proceso de anonimización (la cual abarca los pasos 1 al 5).

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 sólo 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 6.942 personas, de los cuales solo 6.579 mantuvieron su condición de trabajador independiente ($a_1=1,2$; o $a_1=3$ y $a_3=9$). De ellos, 5.872 trabajaban por cuenta propia y 707 tenían al menos a una persona trabajando de forma asalariada permanente. Además, contiene un total de 504 variables.

Producto de este proceso se excluyeron 48 variables, de las cuales 22 correspondían a identificadores del hogar y encuesta, 9 a variables de la hoja de ruta, 4 a variables relacionadas con la ENE y 13 a variables de registro de hora de inicio y término de la aplicación de módulos del cuestionario.

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 de publicación de datos implementada por la Institución. Al estar disponible a todo público se requiere la máxima confidencialidad de la información. Esta Base se encuentra 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 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

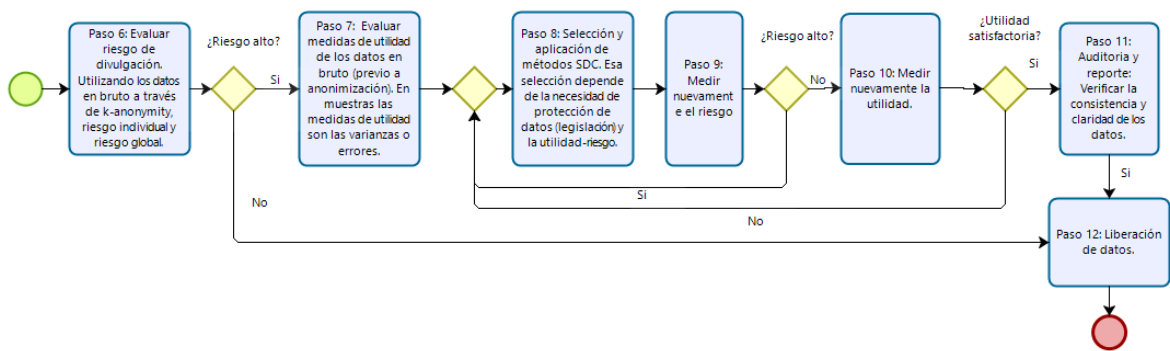
El objetivo de la EME es poder realizar una profunda caracterización de los microemprendedores y sus negocios. Una de las principales áreas de investigación de estas unidades son la informalidad, la generación de empleo, la rama de actividad económica, el acceso al sistema financiero y algunos análisis 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.

7.2.2. 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 auditoría y liberación de datos.

Ilustración 1: Proceso de enmascaramiento



Fuente: INE, VII Encuesta de Microemprendimiento, 2022

8. ALCANCE DE LA PRESENTACIÓN DE RESULTADOS

El siguiente apartado detalla los productos estadísticos que fueron 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 a publicar

La Encuesta de Microemprendimiento realiza una serie de preguntas que construyen más de 20 indicadores. A continuación, se presentan los principales indicadores producidos por la encuesta:

8.1.1. Total de Personas Microempendedoras

Este indicador estima el número de empleadores o trabajadores por cuenta propia dueños de una microempresa (unidad económica con hasta 10 trabajadores, incluyendo al informante que es dueño de la unidad económica). Este número no considera a los trabajadores que dejaron de ser independientes (Pregunta a1 categoría 3) entre la Encuesta Nacional de Empleo (Fase 1) y la aplicación de la Encuesta de Microemprendimiento (Fase 2).

8.1.2. Proporción de personas microempendedores por cuenta propia

Este indicador se construye por medio del ratio entre el total de microempendedores por cuenta propia sobre el total de microempendedores. Los microempendedores por cuenta propia son trabajadores que trabajan solo o junto a un socio, un familiar no remunerado o bien emplean a un trabajador remunerado de forma temporal. Para crear esta variable se utilizó la variable CISE.

8.1.3. Proporción de personas microempendedores informales

Este indicador se construye por medio del ratio entre el número de microempendedores informales sobre el total de microempendedores. Los microempendedores informales considera a todos los microempendedores que no se encuentran registrados ante el Servicio de Impuestos Internos (SII) y no llevan un sistema contable que les permita separar los gastos del negocio de los gastos del hogar. Para crear esta variable se consideraron la pregunta de registro ante el SII (E3) y la pregunta sobre el tipo de contabilidad (E1) y si separa gasto (E2).

8.2. Difusión de los resultados

La difusión de la Base de datos de la Encuesta junto con su documentación se realizó en junio del 2023 por medio de los sitios web del Ministerio de Economía²⁵ y el Instituto

²⁵ Link: <https://www.economia.gob.cl/category/estudios-encuestas/emprendimiento>

Nacional de Estadísticas²⁶. El detalle de los documentos y productos publicados de describen a continuación:

8.2.1. **Bases de datos**

Al igual que la versión anterior de la Encuesta se ponen 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.

Ambas bases de datos se encuentran disponibles en 3 formatos: CSV, Stata y Rds.

8.2.2. **Publicaciones y anuarios**

Al momento de la divulgación de la base de datos, se publicó la Síntesis de la VII Encuesta de Microemprendimiento, la cual resume 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.

8.2.3. **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 VII Encuesta de Microemprendimiento y el documento de Cambio de Factor de Expansión.

8.2.4. **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, además de otros.

8.2.5. **Formularios**

Se encuentra disponible en formato digital (pdf) el formulario de levantamiento el cual fue aplicado por medio del dispositivo móvil de captura (DMC) a los encuestados seleccionados a participar en el estudio.

²⁶ Link: <https://www.ine.gob.cl/estadisticas/sociales/mercado-laboral/microemprendimiento>

8.2.6. **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.

8.3. **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: proteger la identidad de las personas y/o empresas que participan de los estudios realizados por la Institución y 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".

9. CONCEPTOS CLAVES

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

Asalariado: es la persona que trabaja para un empleador, jefe o patrón y percibe una remuneración en forma de sueldo, salario, comisión, propinas, pagos a destajo o en especie.

Asalariado privado: esta categoría incluye a todos los asalariados que no trabajan en instituciones o empresas del Estado, independiente del sector económico donde se desempeñan.

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

Bases de datos: es una colección de información organizada de forma que un programa de ordenador pueda seleccionar rápidamente los fragmentos de datos que necesite.

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): 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.

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.

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, 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”

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.

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.

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.

Fuerza de trabajo potencial (Inactivos potencialmente activos): personas en edad de trabajar que, durante la semana de referencia, no estaban ocupadas ni desocupadas y que habían llevado a cabo «actividades de búsqueda» pero no estaban «actualmente disponibles para trabajar»; o que no llevaron a cabo «actividades de búsqueda», pero que si estaban «actualmente disponibles para trabajar».

Horas efectivas: cuantifican el tiempo dedicado por las personas al desempeño de actividades económicas que contribuyen a la producción de bienes y servicios durante un período de referencia corto (la semana de referencia de la encuesta).

Horas habituales: corresponden a las horas comúnmente trabajadas en un período de referencia determinado (promedio de los últimos tres meses).

Informante Directo: corresponde a la persona que responde una operación estadística entregando información de sí mismo, es decir, de sus propias características o actividades en relación con un estudio en particular.

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).

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.

Muestra teórica: total de unidades seleccionadas en la muestra, incluyendo las unidades asignadas como sobremuestra.

Muestra objetivo (Mo): total de unidades seleccionadas en la muestra para garantizar los niveles de precisión diseñados para cada estudio. La muestra objetivo es equivalente a la muestra teórica, por tanto, se excluyen las unidades asignadas como sobremuestra.

Muestra Lograda: unidades muestrales a encuestar en un período determinado y donde se obtiene una entrevista (completa o parcial).

Ocupados (trabajadores): 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: comprende a todos los trabajadores dependientes que carecen de acceso al sistema de seguridad social (pensión y salud) por concepto de su vínculo laboral. Para el caso de los trabajadores independientes, se considera que poseen una ocupación informal si la empresa, negocio o actividad que desarrollan pertenece al sector informal. Familiar no remunerado son todos clasificados como informales.

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.

Población en edad de trabajar (PET): población actualmente residente en el país, de 15 años y más.

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.

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.

Sector Informal: conjunto de unidades económicas que carecen de registro en el Servicio de Impuestos Internos (SII) o no llevan una contabilidad a través de la cual puedan separar los gastos del negocio de los del hogar.

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.

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.

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 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.

Unidad Económica: corresponde a un grupo de personas o bien personas autónomas que trabajan en una actividad dedicada 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.

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. REFERENCIAS BIBLIOGRÁFICAS

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

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

De Mel (2009). Measuring microenterprise profits: Must we ask how the sausage is made?

Fafchamps (2012). Using PDA consistency checks to increase the precision of profits and sales measurement in panels.

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 (2018a), Manual conceptual y metodológico, diseño muestral, Encuesta Nacional de Empleo (ENE). Santiago. Recuperado de: https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodolog%C3%ADa-y-dise%C3%B1o-muestral.pdf?sfvrsn=94e53c33_11

INE Chile. (2018b). 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

INE Chile. (2018c). 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

INE Chile. (2018d). 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)

INE Chile. (2019a). Fortalecimiento de la Encuesta Nacional de Empleo (ENE). Misión de Asistencia Técnica Conjunta al Instituto Nacional de Estadísticas de Chile. Organización Internacional del Trabajo (OIT) y Comisión Económica para América Latina y el Caribe (CEPAL) Recuperado de: https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/publicaciones/informe-ine-oit-fortalecimiento-de-la-encuesta-nacional-de-empleo.pdf?sfvrsn=739eb98f_3

INE Chile. (2019b). Manual de Implementación NDGM 2019. Instituto Nacional de Estadísticas de Chile.

INE Chile. (2020). Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares. Instituto Nacional de Estadísticas de Chile. Disponible en "<https://www.ine.cl/docs/default-source/institucionalidad/buenas-pr%C3%A1cticas/clasificaciones-y-estandares/est%C3%A1ndar-evaluaci%C3%B3n-de-calidad-de-estimaciones-publicaci%C3%B3n-27022020.pdf>"

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

INEGI México (2012). Encuesta Nacional de Micronegocios. Metodología. Disponible en https://www.inegi.org.mx/contenidos/programas/enamin/2012/doc/enamin12_met_dis_mu_estrat.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

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

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 (2013a), La medición de la informalidad: Manual estadístico sobre el sector informal y el empleo informal. Recuperado de: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_222986.pdf

OIT (2013b), Resolución I, Resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo de la 19ª Conferencia Internacional de Estadísticos del Trabajo. Ginebra. Recuperado de: 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 20ª Conferencia Internacional de Estadísticos del Trabajo. Ginebra. Recuperado de: 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 de 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. Anexo 1: Prueba piloto

Además de evaluar la comprensión y funcionamiento de las nuevas preguntas incluidas en el cuestionario, la prueba piloto tenía por objetivo estimar el impacto del cambio en el medio (desde encuesta en papel a dispositivo de captura móvil) y del modo de recolección (telefónico) de la encuesta, en los principales indicadores.

Para poder evaluar el impacto en los cambios metodológicos mencionados se construyeron 3 grupos de estudio:

- Grupo de control: consideró medio papel y modalidad presencial
- Grupo de tratamiento 1: consideró medio dispositivo móvil y modalidad presencial
- Grupo de tratamiento 2: consideró medio dispositivo móvil y modalidad telefónica.

Luego se realizaron algunos análisis para medir los efectos del cambio en el medio de recolección (PAPI vs CAPI) y el modo de recolección (CATI) en los principales indicadores de la encuesta tales como: tasa de microemprendedores por cuenta propia, tasa de informalidad, acceso a la seguridad social, actividad económica, ganancias del negocio y motivación emprendedora; dependiendo del modo y medio de recolección de la encuesta. Además, se evaluaron efectos en las tasas de respuesta a la unidad y el ítem entre los distintos grupos, la duración del cuestionario

A partir de los análisis realizados se desprenden los siguientes hallazgos que abarcan al cuestionario de recolección, el modo de recolección y la metodología de visitas:

- 1) No se evidencia un impacto en los principales indicadores del cambio en medio y modo. Las estimaciones no muestran efectos significativos en la modalidad o medio de levantamiento.
- 2) No se observa mayor diferencia en la respuesta al ítem y en los indicadores de inconsistencias por modalidad.
- 3) Tampoco se evidencia diferencias significativas en la duración de la encuesta.
- 4) No obstante, se evidencia una menor tasa de respuesta y una mayor tasa de rechazo en la aplicación telefónica, la cual si bien puede ser revertida en una visita presencial puede incidir en el no logro muestra objetivo.

Con el fin de hacer más eficiente el proceso de recolección se estableció una nueva metodología de visitas que considera un primer contacto vía teléfono, el cual tuvo dos objetivos: la aplicación inmediata de la encuesta o bien la programación de una cita.