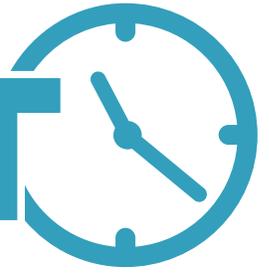




II ENUT

Encuesta Nacional sobre
Uso del Tiempo



Documento **METODOLÓGICO**



II ENCUESTA SOBRE USO DEL TIEMPO

Período de información: septiembre 2023 – diciembre 2023

Publicación: enero 2025

Instituto Nacional de Estadísticas – Chile

Subdepartamento de Recolección de Estadísticas Sociales

Subdepartamento de Supervisión y Verificación de Estadísticas Sociales

Subdepartamento Diseño de Marcos y Muestras

Subdepartamento de Estadísticas Estructurales del Trabajo

Subdirección Técnica

Leonardo González Allendes

Jefatura de Departamento

David Niculcar Castillo

Jefatura de Subdepartamento

Carla Arellano Mendoza

Jefatura de Proyecto y

Coordinador Equipo Técnico

Agustín Arce Garcés

Analista Socioeconómico Senior

Verónica Valdivia Medina

Analistas Socioeconómicos

Joaquín Galdames Hernández

Paula Herrera Domínguez

Juan León Gaete

Santiago Ortúzar Lyon

Martín Venegas Márquez

Diseño Muestral

Iván Touron Romero

Ximena Espinosa Pavez

Marta Cisternas Astudillo

Encargado Nacional de Recolección

Nicolás Montecinos Guzmán

Encargadas de Área de Recolección

Ninoska Chamorro Muñoz

Oriela Barraza Aravena

Encargada Nacional de Supervisión

Berta Cid Cárdenas

Analista de Implementación de Diseños

Bárbara Martínez Correa

Dirección: Morandé 801, Santiago de Chile

Teléfono: (56) 23246 3830

Sitio web: www.ine.gob.cl

Correo electrónico: ine@ine.gob.cl

Facebook: @ChileINE

Twitter: @INE_Chile

Instagram: @chile.ine

Contenido

1. Presentación.....	8
2. Introducción.....	11
3. Contexto metodológico	15
3.1. Antecedentes de la operación estadística.....	15
3.2. Antecedentes históricos ENUT en Chile	18
3.3. Objetivos de la II Encuesta Nacional sobre Uso del Tiempo.....	18
3.4. Diferencias metodológicas respecto a la I ENUT.....	19
3.4.1. Modo de recolección de la información.....	20
3.4.2. Módulo de trabajo de cuidados (TC).....	21
3.4.3. Lineamientos de calidad	23
3.4.4. Factores de expansión	26
3.4.5. Módulo de simultaneidad.....	28
4. Marco de referencia.....	30
4.1. Marco Conceptual.....	30
4.1.1. Concepto de trabajo	30
4.1.2. Sistema de Cuentas Nacionales.....	31
4.1.3. Frontera general de la producción en hogares.....	32
4.1.4. Concepto ampliado de “trabajo”	35
4.2. Dimensiones de análisis II ENUT.....	38
4.2.1. Trabajo en la ocupación y actividades conexas.....	38
4.2.2. Trabajo no remunerado (Fuera de la frontera de producción del SCN)	39

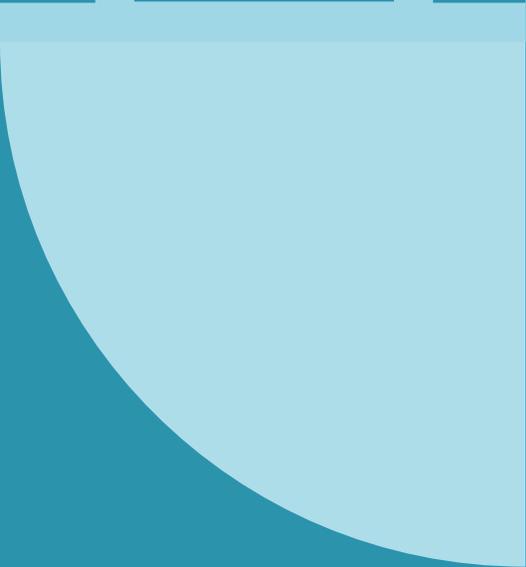
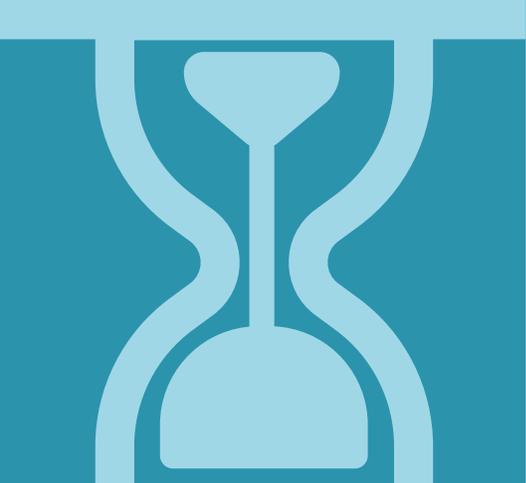
4.2.3. Actividades personales (Fuera de la frontera de la producción del SCN).....	40
4.3. Recomendaciones internacionales para medir uso del tiempo	41
4.4. Experiencias internacionales sobre Uso del Tiempo.....	42
5. Diseño estadístico.....	47
5.1. Población objetivo	47
5.2. Cobertura geográfica.....	47
5.3. Unidades estadísticas	49
5.3.1. Unidad de observación	49
5.3.2. Unidad de análisis.....	49
5.3.3. Unidad de muestreo.....	49
5.4. Definición de variables	50
5.5. Clasificaciones estadísticas	53
5.5.1. Clasificador Internacional Uniforme de Ocupaciones (CIUO-08.CL) 53	
5.5.2. Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES).....	57
5.5.3. Clasificación Internacional Normalizada de la Educación (CINE) 2011	59
5.5.4. Clasificación de Actividades de Uso del Tiempo para América Latina y el Caribe (CAUTAL).....	60
5.6. Marco estadístico	63
5.6.1. Características del MMV 2020	63
5.6.2. Marco muestral II ENUT	64
5.7. Diseño muestral	67

5.7.1.	Dominios de estudio	67
5.7.2.	Estrategia muestral.....	67
5.7.3.	Cálculo y distribución del tamaño muestral	68
5.7.4.	Etapas para el cálculo del tamaño muestral para la II ENUT 70	
5.7.5.	Resumen del proceso de simulación de tamaño muestral	75
5.7.6.	Selección de las unidades de primera etapa.....	81
5.7.7.	Selección de las unidades de segunda etapa (viviendas particulares ocupadas).....	82
5.8.	Factores de expansión	84
5.8.1.	Ponderador de selección de Unidades de Primera Etapa (UPM) 85	
5.8.2.	Ajuste por omisión de comunas	86
5.8.3.	Ponderador de selección de viviendas.....	86
5.8.4.	Ponderador corregido por elegibilidad.....	89
5.8.5.	Ajuste por no respuesta	93
5.8.6.	Suavizamiento del ponderador	99
5.8.7.	Calibración de los factores de expansión.....	102
5.9.	Estimación de la varianza	106
5.9.1.	Métodos convencionales para estimación de la varianza	106
5.9.2.	Varianza con muestras complejas	109
5.9.3.	Algoritmo de cálculo en muestras complejas.....	111
5.9.4.	Variables que identifican el diseño muestral complejo....	118
5.9.5.	Coefficiente de variación	120
6.	Metodología de recolección de datos.....	123

6.1.	Características generales de la recolección de datos	123
6.1.1.	Características organizacionales de equipos participantes de la recolección de datos	123
6.1.2.	Modo de recolección de la información	131
6.1.3.	Periodo de referencia y de recolección	131
6.1.4.	Ejecución operativa de la recolección de una submuestra operativa	133
6.1.5.	Indicadores de desempeño operativo	139
6.1.6.	Estrategias de sensibilización	144
6.2.	Diseño de los instrumentos de captura de información.....	145
6.2.1.	Tecnología de producción de la información.....	145
6.2.2.	Cuestionario del Hogar (CH).....	148
6.2.3.	Cuestionario sobre Uso del Tiempo (CUT)	150
6.3.	Prueba y evaluación de los instrumentos de captura de información.....	153
6.3.1.	Prueba Piloto II ENUT	153
6.3.2.	Hitos previos a la Prueba Piloto II ENUT	156
6.3.3.	Ajustes para versión oficial.....	157
6.4.	Método de capacitación.....	159
6.4.1.	Capacitación del equipo de recolección	159
6.5.	Supervisión de la recolección	161
6.5.1.	Tareas del equipo de supervisión.....	162
6.5.2.	Supervisión directa	163
6.5.3.	Supervisión indirecta	165
6.5.4.	Indicadores de supervisión	167

6.6.	Sistema informático	168
6.7.1.	Funcionalidades asociadas a la ejecución de recolección 168	
6.7.2.	Funcionalidades asociadas a la supervisión de recolección 170	
7.	Metodología de procesamiento de datos	173
7.1.	Estructura de trabajo	173
7.1.1.	Entregas por submuestra	173
7.1.2.	Nomenclatura de las bases de datos.....	174
7.2.	Subprocesos del procesamiento.....	175
7.2.1.	Integración.....	176
7.2.2.	Clasificar y codificar	178
7.2.3.	Revisar y validar	181
7.2.4.	Editar e imputar.....	185
7.2.5.	Derivar nuevas variables y unidades	194
7.2.6.	Calcular ponderadores	196
7.2.7.	Finalizar los archivos de datos	196
8.	Metodología de análisis de resultados	198
8.1.	Preparar borradores de resultados	198
8.2.	Validar los resultados	198
8.3.	Interpretar y explicar los resultados	200
8.4.	Aplicar control a la divulgación.....	201
8.5.	Finalizar resultados.....	204
9.	Difusión de resultados.....	206
9.1.	Productos estadísticos para publicar	206

9.2.	Accesibilidad a los datos y marco legal.....	207
10.	Glosario	212
11.	Referencias	218
12.	Anexos.....	230
12.1.	GSBPM en la producción estadística	230
12.2.	Construcción del marco estadístico	235
12.2.1.	Estratificación socioeconómica de las UPM	236
12.2.2.	Áreas especiales	238
12.2.3.	Actualización de los límites urbanos censales (LUC) ...	239
12.2.4.	Características de las áreas geográficas.....	241
12.2.5.	Cambios en las definiciones urbano y rural.....	243
12.2.6.	Comparación MMV 2020 Y MM 2015	248
12.2.7.	Conformación de las Unidades Primarias de Muestreo	250
12.3.	Anexo 1 – Afiche II ENUT.....	252
12.4.	Anexo 2 – Díptico II ENUT	253
12.5.	Anexo 3 – Carta de apertura hogar II ENUT	254



1

PRESENTACIÓN

1. Presentación

El Instituto Nacional de Estadísticas (INE) presenta a la ciudadanía la Metodología de la II Encuesta Nacional sobre Uso del Tiempo (II ENUT). El trabajo de recolección de datos en terreno para este estudio se llevó a cabo entre el 14 de septiembre y el 29 de diciembre de 2023. La ENUT es una encuesta dirigida a hogares que permite analizar cómo las personas distribuyen su tiempo en las actividades de la vida cotidiana, enfocándose principalmente en el trabajo no remunerado, el trabajo en la ocupación y otras actividades personales. La ENUT recopila información en los principales centros urbanos de cada región del país, mediante entrevistas directas a personas de 12 años en adelante.

La metodología de la ENUT, cuya primera versión fue realizada por el INE en 2015, sigue las recomendaciones para medir el uso del tiempo y el trabajo no remunerado de la Comisión Económica para América Latina y el Caribe (CEPAL), Naciones Unidas (ONU) y la Organización Internacional del Trabajo (OIT). Se emplean los clasificadores internacionales de actividades específicos para encuestas de este tipo: Clasificación de Actividades de Uso del Tiempo para América Latina y el Caribe (CAUTAL) y Clasificación Internacional de Actividades para Estadísticas sobre el Uso del Tiempo (ICATUS), cuyo marco conceptual se basa en el Sistema de Cuentas Nacionales (SCN).

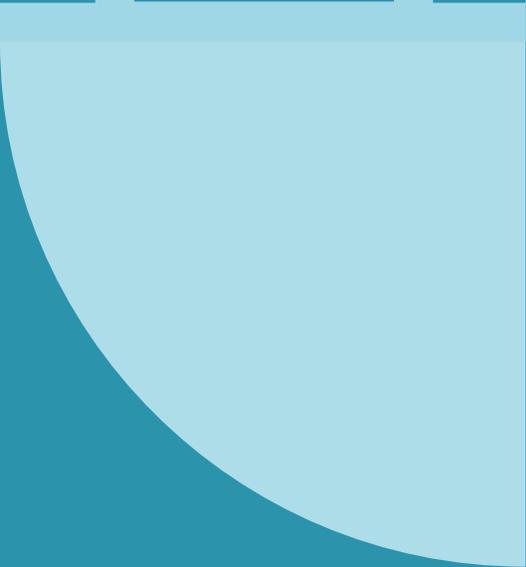
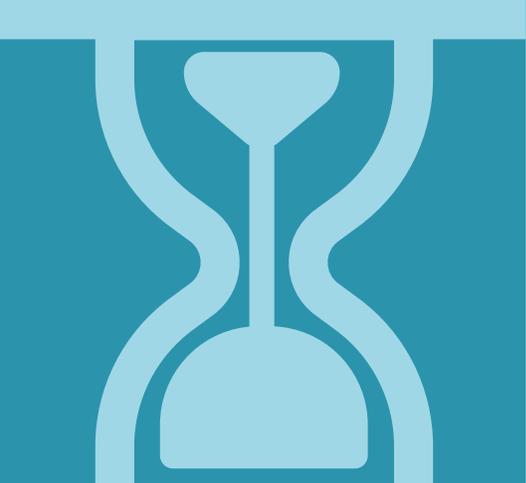
La ENUT es el único estudio a nivel nacional que aborda la dimensión del trabajo no remunerado, una actividad fundamental para el desarrollo económico y el bienestar de la población, que generalmente tiende a ser invisibilizada a través de la medición de los indicadores tradicionales sobre participación en el mercado laboral.

En cuanto a las desigualdades de género, la información proporcionada por la ENUT permite a las estadísticas oficiales de Chile cumplir con

varios compromisos internacionales, tales como los asumidos en las Conferencias de la Mujer, la Convención sobre la Eliminación de Todas las Formas de Discriminación contra la Mujer (CEDAW), la Plataforma de Acción de Beijing y Beijing+5, entre otros. Además, constituye una fuente valiosa para el diseño de programas y políticas públicas orientadas a la promoción de la autonomía económica de las mujeres en el país.

Para brindar a la ciudadanía toda la información metodológica sobre el proceso de producción estadística de la II ENUT, en este documento se describe detalladamente cada etapa del estudio, incluyendo las innovaciones y mejoras incorporadas en esta versión. Entre las que destacan la inclusión de preguntas sobre nacionalidad, pertenencia a pueblos originarios, género y orientación sexual, así como un fortalecimiento en la medición de los cuidados. Además, destaca en el ámbito de la recolección de la encuesta, la implementación del uso de dispositivos móviles para la captura (DMC) de datos en terreno.

Finalmente, el INE extiende un agradecimiento a todas las personas que participaron en este importante estudio, especialmente a las y los ciudadanos que colaboraron brindando la información solicitada. También agradecemos a las personas que trabajaron directamente en el proyecto, a los equipos del INE que brindaron soporte en distintas áreas del proceso productivo y a quienes forman parte de las Direcciones Regionales de esta institución.



2

INTRODUCCIÓN

2. Introducción

El INE, tiene como misión articular el Sistema Estadístico Nacional y producir las estadísticas oficiales del país. Acerca de sus objetivos, uno de los más relevantes, es el de constituirse como la entidad pública que provee de información confiable, oportuna y de calidad, a la toma de decisiones del país.

En este ámbito, la ENUT es un estudio socioeconómico, basado en una encuesta probabilística por muestreo aplicado a hogares, que busca recopilar datos sobre el tiempo que las personas dedican a diversas actividades dentro de un período determinado. Su foco se encuentra en las actividades relacionadas con el concepto amplio de trabajo, especialmente aquellas por las que no se reciben pagos en dinero o beneficios, como son el trabajo doméstico y el trabajo de cuidados no remunerados.

Con el establecimiento de los Objetivos de Desarrollo Sostenible (ODS) por parte de la Asamblea General de las Naciones Unidas en septiembre de 2015, se generó como país la necesidad de actualizar el ODS 5.4.1, compromiso que supone medir la proporción de tiempo destinada al trabajo doméstico y asistencial no remunerado, desglosada por sexo, edad y ubicación geográfica. Hacia esa meta, la ENUT significa una acción clave como fuente de información y diagnóstico de la situación del país en el cumplimiento de este objetivo.

Dentro de las características de este tipo de encuestas, se destaca la relevancia de contar con un alto nivel de participación de los hogares encuestados, por tanto, el tiempo para visitar a los domicilios y entrevistar a todos los integrantes de 12 años o más, fue dispuesto en un margen de dos semanas. Asimismo, se recolecta la información en dos fases, en la primera la información entregada por el(a) informante sobre

la composición de los integrantes del hogar y su situación sociodemográfica y laboral, impacta en cuáles y qué preguntas se administrarán al resto de los integrantes al momento de realizar el cuestionario individual de uso del tiempo, sumando otra complejidad a la aplicación. Por todos esos motivos, la construcción del instrumento, la metodología de recolección y el procesamiento de los datos demandó grandes esfuerzos en su diseño y puesta en marcha.

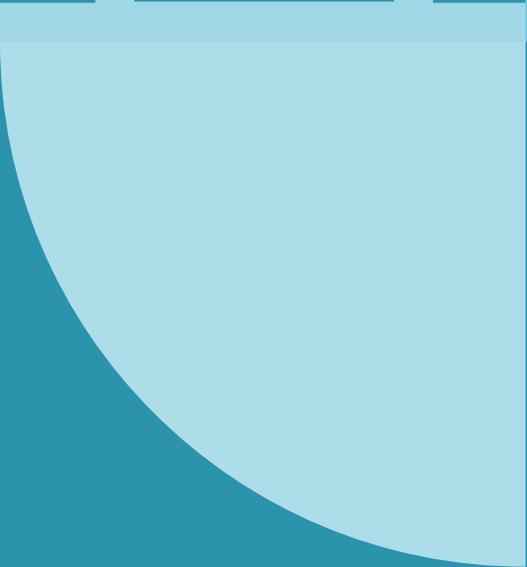
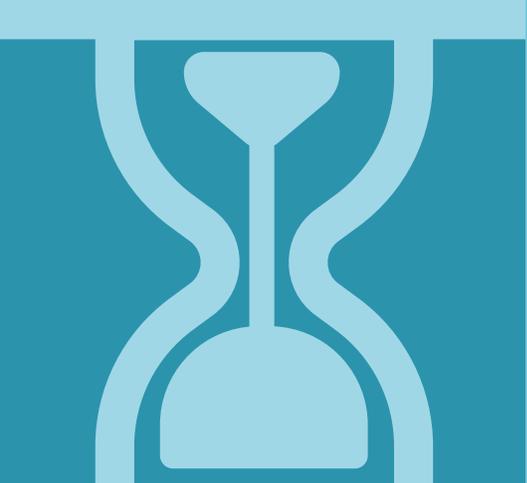
En esta segunda versión de la ENUT se gestionaron además diversas mejoras al proceso de producción estadística, relacionadas con ajustes a los cuestionarios, la introducción de dispositivos móviles en la recolección, nuevos ítems para ampliar la captura del fenómeno de uso del tiempo, y la automatización y optimización del proceso de codificación de datos. Todo desde un diseño de producción congruente al Modelo Genérico del Proceso de Producción Estadística (GSBPM, por sus siglas en inglés) e incorporando las buenas prácticas y estándares que han sido aplicadas en otros estudios al interior de la institución.

A continuación, el informe aborda las diferentes etapas del proceso de producción estadística, que permiten dar cumplimiento al objetivo de contar con datos representativos de la participación y tiempo destinados a las actividades de la vida diaria en el país.

La Metodología forma parte de una serie de documentos de referencia que el INE pone a disposición de la ciudadanía, para aclarar los procedimientos y decisiones metodológicas de la II ENUT, en la búsqueda del mejor y correcto uso de los datos. Así, el presente informe se distribuye de la siguiente forma:

- **Capítulo 3:** se presentan los antecedentes y objetivos de la operación estadística.

- **Capítulo 4:** se da a conocer el marco conceptual en el que se enmarca el estudio, sus dimensiones de análisis y referencias nacionales e internacionales.
- **Capítulo 5:** se describen las características del diseño muestral, desde los objetivos, a la estimación del tamaño muestral. Además de explicar la construcción de los factores de expansión.
- **Capítulo 6:** se describe el diseño y ejecución del proceso de recolección de datos de la II ENUT.
- **Capítulo 7:** en este apartado se abordan los ocho subprocesos que componen el procesamiento de datos del estudio. Esto incluye la clasificación de los datos, los procedimientos para detectar inconsistencias, los diversos mecanismos para corregirlos (ediciones, imputaciones) y los cálculos y ajustes de ponderadores.
- **Capítulo 8:** se describen los métodos mediante los cuales se analizaron los datos, incluyendo el control a la divulgación de datos.
- **Capítulo 9:** presenta los productos de difusión del estudio y la accesibilidad a los datos.



3

**CONTEXTO
METODOLÓGICO**

3. Contexto metodológico

3.1. Antecedentes de la operación estadística

La calidad de vida de las personas y su bienestar se ven afectados por el uso que hacen del tiempo, según qué actividades prioricen y la forma en que las organizan. El tiempo, recurso lineal y finito, cuyo uso está mediado por condicionantes socioeconómicas, de género, el ciclo de vida y la cultura, entre otros factores. En este contexto, las encuestas de Uso del Tiempo proponen como medición básica las actividades que realizan las personas y el tiempo que dedican a ellas, permitiendo conocer la manera en que una población organiza y distribuye este recurso.

A pesar de que existe un ordenamiento estándar del tiempo a través de un año calendario en 365 días, semanas de siete días y 24 horas por día, el tiempo puede ser entendido como un recurso cuya disposición presenta un alto contenido social y dentro de él, la componente género reviste un nivel de importancia alto, siendo en sí mismo un medio que facilita o restringe la posibilidad de desarrollar ciertas actividades, que impactan en el bienestar económico y subjetivo de las personas.

Los cambios en el mercado del trabajo han afectado las exigencias sociales respecto a la determinación del tiempo. El aumento de la participación femenina en el mercado laboral, sin que por ello se redistribuyan significativamente las tareas domésticas o de cuidados, conlleva en mayor proporción que las mujeres deban asumir una doble o triple jornada de trabajo, adicionando a las labores en el mercado laboral los quehaceres domésticos y de cuidado de otros integrantes del hogar.

El fenómeno descrito suele presentar además de un sesgo socioeconómico, un sesgo de género, en tanto las mismas oportunidades de inserción al mercado laboral se ven afectadas por la

posibilidad de acceso a redes de apoyo, familiar, público o de mercado, para la realización de las labores domésticas y de cuidado.

En este sentido, la brecha de participación laboral femenina en Latinoamérica, por ejemplo, es una de las más altas del mundo. El tercer trimestre de 2022 la tasa de empleo asciende al 51,8% en mujeres, en los hombres es del 74,5%, con una brecha de género de 23 puntos porcentuales en la región (OIT 2023b). Asimismo, la crisis económica y sanitaria del Covid 19 afectó en mayor proporción a las mujeres. Los indicadores de participación laboral femenina en el 2020 mostraban valores similares a los de 1990, tras una caída de 16 puntos durante la pandemia, en comparación a los 10 puntos de retroceso de la tasa masculina (Bustelo et al. 2021). La distancia salarial también figura como otro factor, en palabras de Urquidi y Chalup (2023), las mujeres obtienen *“un ingreso por hora menor que el de los hombres en 14 de los 18 países observados”* en la región. Todos estos factores, además de la debilidad de las redes de apoyo, perpetúan los roles de género.

Las encuestas de uso del tiempo son la mejor estrategia para medir los tiempos que la población dedica a las distintas actividades de interés, en un período de referencia determinado. Con la medición del tiempo que las personas destinan al trabajo en los hogares sin mediar remuneración, la II ENUT avanza en el desafío de generar información sobre las distintas formas de trabajo, independientemente de si su realización es con fines de mercado o para el consumo de los propios hogares. Adicionalmente, al reportar información sobre el tiempo destinado a actividades como el cuidado personal, la vida social o los estudios, se da cuenta de la contraparte del tiempo destinado al trabajo, área central en el bienestar en su dimensión de conciliación de la vida familiar y vida laboral.

A partir de lo anterior, los datos de la II ENUT permiten alinearse con las mediciones ya realizadas en la mayor parte de los países de la región, y

caracterizar las brechas existentes entre hombres y mujeres, así como la de otros grupos de la población en cuanto a tiempo destinado a distintas actividades de trabajo o personales. Estas particularidades hacen de la ENUT un insumo clave para la evaluación y diseño de políticas públicas.

En relación específicamente a las desigualdades de género, la información proporcionada posibilita que las estadísticas oficiales de Chile cumplan con distintos compromisos internacionales, como los adquiridos en las Conferencias de la Mujer, la Convención sobre la eliminación de todas las formas de discriminación contra la mujer (CEDAW), la Plataforma de Acción de Beijing, Beijing+5, entre otras.

La primera versión del estudio, realizada en 2015, sirvió como fuente de información para diversos programas en torno a la promoción de la autonomía económica de las mujeres en el país, así como antecedente para el proyecto de ley de sala cuna universal y la ley de responsabilidad parental y pago efectivo de deudas de pensiones de alimentos, entre otros.

Para la segunda versión, con el objetivo de robustecer las dimensiones de análisis abarcadas por el estudio, se convocó a una mesa técnica para la detección de necesidades; instancia en la que participaron organismos públicos (Ministerio de Economía, Ministerio de la Mujer y Equidad de Género, Ministerio del Trabajo y Previsión Social, Ministerio de Desarrollo Social y Familia y Banco Central de Chile) y organizaciones no gubernamentales (Organización Internacional del Trabajo y Comisión Económica para América Latina y el Caribe) vinculadas a la ENUT. Los resultados de esta Mesa consolidaron un instrumento de medición que pudiera responder a las necesidades de información actuales para el análisis, seguimiento y formulación de políticas públicas del país y que esté alineados con las buenas prácticas internacionales sobre la medición de estadísticas sobre uso del tiempo.

3.2. Antecedentes históricos ENUT en Chile

La II ENUT es la segunda experiencia oficial de medición de uso del tiempo que desarrolla el INE en Chile. La primera corresponde a la Encuesta Nacional sobre Uso del Tiempo 2015 (ENUT 2015).

La II ENUT mantiene tanto el alcance nacional urbano de la versión 2015, como gran parte de sus dimensiones de análisis. Entre ellas: trabajo en la ocupación y traslados asociados a esa actividad, Búsqueda de empleo, Trabajo de cuidados no remunerados, Trabajo doméstico no remunerado, Quehaceres domésticos y de cuidados para otros hogares, Trabajo voluntario, Cuidados personales, Educación, Ocio y vida social, y finalmente, Satisfacción con el uso del tiempo. Esta nueva versión innova incluyendo el Módulo de Simultaneidad para medir este fenómeno, la incorporación de preguntas asociadas al acceso a servicios de cuidado por parte del hogar (en el Módulo de Ayudas que recibe el hogar), la homologación de las preguntas que miden ocupación en base a la Encuesta Nacional de Empleo (ENE), la incorporación de la pregunta sobre nacionalidad, junto a preguntas que indagan el género de las personas de 5 años o más, y la orientación sexual de los integrantes del hogar de 15 años o más.

3.3. Objetivos de la II Encuesta Nacional sobre Uso del Tiempo

Los objetivos de la II ENUT son:

Objetivo general

Obtener información de cómo las personas de 12 años o más utilizan su tiempo, principalmente, en torno al trabajo no remunerado, trabajo en la ocupación y actividades personales; además de proveer información

relacionada a su calidad de vida y bienestar respecto al uso del tiempo, en las principales zonas urbanas del país.

Objetivos específicos

- i. Identificar las brechas de género en el tiempo dedicado a trabajo doméstico y de cuidados no remunerado.
- ii. Medir la carga global de trabajo de la población de 15 años o más, a través del tiempo que se dedica al trabajo en la ocupación y en el trabajo no remunerado.
- iii. Medir el uso del tiempo de las personas de 12 años o más en actividades personales.
- iv. Medir la satisfacción de la población de 12 años o más de las principales zonas urbanas de Chile con distintos aspectos de su uso del tiempo.
- v. Identificar el acceso y la necesidad de servicios e instituciones de cuidado para la población dependiente priorizada en el estudio.

3.4. Diferencias metodológicas respecto a la I ENUT

Es importantes entender que **la ENUT es una encuesta estructural**, que tiene como prioridad entre una versión y otra el mejoramiento de las mediciones en términos de calidad y precisión en la captura del fenómeno de interés. Dado lo anterior, entre la I ENUT (2015) y la II ENUT (2023) existen una serie de cambios y modificaciones metodológicas, abordándose a continuación aquellas con mayor impacto.

3.4.1. Modo de recolección de la información

La I ENUT fue levantada en PAPI¹ en su totalidad (ambos cuestionarios), los que posteriormente fueron digitados de manera centralizada en las oficinas del INE Santiago. La digitación de los formularios en papel se realizó a través del Sistema Operativo de Gestión de los Hogares (SGOH), el cual contó con una interfaz similar al formulario en papel, para así evitar errores o distorsiones. Este sistema a su vez consideró reglas de validación de primer nivel para evitar que se ingresaran rango de datos inválidos.

En la II ENUT, como se detalla en el capítulo 6, la captura de la información fue realizada a través de Dispositivos Móviles de Captura (DMC) con el programa *Survey Solutions*, lo cual significó una mejora en términos de eficiencia de tiempo y recursos humanos, al no tener que contar con un proceso intermedio adicional entre la recolección y la obtención de los datos brutos por parte del equipo técnico.

La captura en DMC significó también una reducción de los errores no muestrales, al poder implementar inmediatamente filtros, saltos de preguntas y establecer rangos de respuestas válidos a ciertas preguntas. De esta forma, la validación de primer nivel reducida incorporada en el proceso de digitación en la I ENUT, fue mucho más robusta e inmediata en los DMC de la II ENUT, entregando mensajes directos de posibles errores o inconsistencias a las personas encuestadoras.

¹ *Computer Assisted Personal Interviewing (CAPI)*, en este caso refiere al uso de dispositivo móvil de captura (tablet) para el registro de la información entrevistada. *Paper and Pencil Interviewing (PAPI)*, refiere a la utilización de un instrumento de recolección en papel y el registro de la información a través de lápiz

3.4.2. Módulo de trabajo de cuidados (TC)

i. Diferencias al preguntar sobre el Trabajo de Cuidados por subpoblaciones

Otro gran cambio en el Módulo de Trabajo de Cuidados No Remunerados entre versiones del estudio es la forma en la cual se presentan las preguntas de cuidado. En el cuestionario 2015 se puede observar que las preguntas se realizan de acuerdo a las siguientes subpoblaciones:

- Personas que requieren cuidados de salud permanentes.
- Cuidados de niños y niñas de 0 a 4 años.
- Cuidados a niños y niñas de 5 a 14 años.
- Cuidado a personas de 15 a 65 años.
- Cuidado a personas de 66 años o más.

Por otro lado, en la II ENUT el cuestionario no separa las preguntas por subpoblaciones, sino que, permite la identificación de la persona sujeta de cuidado según el número de línea que ocupa en el hogar, así como también identificar si esta persona está en situación de dependencia funcional o no, dependiendo de lo que se haya informado en el módulo sobre Personas en Situación de Dependencia Funcional (D) del CH.

La diferencia entre ambas versiones es importante a la hora de poder medir los tiempos y participación en las tareas de cuidado a integrantes del hogar, toda vez que, al considerar los tiempos del cuestionario 2015, se podrían duplicar tiempos si ayudó con tareas escolares a un niño de 0 a 4 años y a otro de 5 a 14 años, mientras que, en el cuestionario 2023, es posible registrar el tiempo total de cuidado en ambos niños del hogar, ya que en una misma actividad, se puede registrar el tiempo de cuidado agregado o por separado, permitiendo reconstruir las subpoblaciones que se requieran en el análisis, sin estar limitado a las definidas en la construcción del instrumento de recolección.

ii. Paso de Personas en Situación de Discapacidad (PESD) a Personas en Situación de Dependencia Funcional (PSDF)

Existe una diferencia fundamental a nivel conceptual para entender el nuevo módulo de trabajo de cuidado a integrantes del hogar. La diferencia radica en el cambio de enfoque al preguntar por el trabajo de cuidado para las PESD en la I ENUT, a preguntar por el trabajo de cuidado para las **PSDF** en la II ENUT.

En este contexto, resulta esencial comprender esta diferencia en un nivel macro. La Convención de las Naciones Unidas sobre los Derechos Humanos de las Personas con Discapacidad (Naciones Unidas, 2006b) plantea que la **discapacidad** es un **concepto en evolución** que surge de la interacción entre las personas con deficiencias y las barreras impuestas por el entorno y las actitudes sociales. Por su parte, el Servicio Nacional de la Discapacidad ², define que las personas en situación de dependencia funcional son aquellas que, debido a limitaciones en su capacidad física, mental o de movilidad, requieren el apoyo de terceros para realizar ciertas tareas. Asimismo, la Ley N° 20.422 define la dependencia funcional como una condición de carácter permanente en la que las personas, por deficiencias físicas, mentales o sensoriales que afectan su autonomía, necesitan la asistencia de otras personas o ayudas significativas.

Desde el enfoque de la dependencia funcional, los servicios de apoyo, asistencia personal y de cuidado que reciben quienes están en esta situación se han convertido en un aspecto crucial para analizar el trabajo de cuidados. La literatura distingue dos categorías principales de servicios de asistencia para personas en situación de dependencia: el cuidado informal y el cuidado formal. El cuidado informal es aquel

² (Servicio Nacional de la Discapacidad, 2017)

brindado por personas cercanas al beneficiario, como familiares o amigos, de manera voluntaria y sin remuneración ni intervención de organizaciones. Por otro lado, el cuidado formal consiste en servicios de atención profesional especializados, ofrecidos por instituciones públicas o entidades privadas, destinados a suplir las necesidades de cuidado que superan las capacidades individuales o familiares (Roger-García, 2009). Tradicionalmente, el cuidado informal ha desempeñado un rol predominante en la asistencia a personas dependientes, recayendo esta responsabilidad principalmente en el ámbito familiar, y, en particular, en las mujeres (Díaz, 2007).

Por este motivo, el Módulo de Trabajo de Cuidados (TC) en el CUT de la II ENUT se enfoca en la asistencia y el trabajo de cuidados no remunerados al interior del hogar que requieren las PSDF, mientras que el Módulo Ayudas que Recibe el Hogar (A) del CH, pregunta sobre los servicios de asistencia de instituciones privadas o públicas a PSDF.

Si bien PESD y PSDF son conceptos relacionados, no son lo mismo, y si se desea realizar comparaciones entre ambas versiones de las ENUT, se debe tener en cuenta esta incorporación.

3.4.3. Lineamientos de calidad

En relación con los nuevos lineamientos y estándares de calidad estadística, en la II ENUT se incorporaron preguntas referentes a nacionalidad, pertenencia a pueblos originarios, género y orientación sexual.

i. Nacionalidad

Sobre nacionalidad, se incorpora esta pregunta que no estaba presente en el cuestionario 2015, permitiendo identificar si la persona informante

se trata de una persona chilena, extranjera o chilena con una segunda nacionalidad.

ii. Pueblos Originarios

Sobre pueblos originarios, en el cuestionario 2015, se pregunta: *“En Chile la ley reconoce 9 pueblos indígenas, ¿pertenece o es descendiente de alguno de ellos...?”*, lo que, de acuerdo con la Ley indígena 19.253, podría llevar a responder que sólo los inscritos en el registro de indígenas de CONADI pertenece o es descendiente de los pueblos indígenas. Por otro lado, el cuestionario de la II ENUT pregunta específicamente por si la persona *“se considera perteneciente a un pueblo indígena u originario de Chile”*, más allá de la Ley.

iii. Estándar de preguntas sobre Sexo, Género y Orientación Sexual (SGOS)

El estándar sobre SGOS, se trata de una iniciativa institucional que, en línea con las recomendaciones de las Naciones Unidas (Naciones Unidas, 2019) busca dar orientaciones conceptuales, armonizar la producción estadística y aportar una medición válida y confiable de las variables de sexo, género y orientación sexual, dentro del ámbito de encuestas dirigidas a hogares y a censos de población.

A diferencia de la II ENUT, la I ENUT no contó con preguntas sobre género y orientación sexual, teniendo en común sólo la pregunta sobre sexo.

En esta nueva versión la pregunta sobre género va inmediatamente después de la pregunta sobre Sexo en el Cuestionario del Hogar (CH), y se pregunta específicamente: *“¿Con cuál género se identifica [NOMBRE]?”*, esta pregunta se realiza para las personas de 5 años o más (respondida por el informante idóneo) y las alternativas de respuestas siguen las recomendaciones del estándar de medición de SGOS.

Por otro lado, la pregunta sobre Orientación sexual se encuentra al final del Cuestionario de Uso del Tiempo (CUT) y se aplica sólo a personas de 15 años o más (informante directo). Donde la persona informante debe registrar su respuesta en el dispositivo móvil de captura (DMC) de forma directa.

Información más detallada sobre el estándar para la medición de estas características, puede ser consultada en la documentación que la institución publicó en el apartado de directrices metodológicas³.

iv. Cambio en el Diseño Muestral

Un cambio fundamental entre ambas versiones de la ENUT es la diferencia significativa de cobertura, estructura y precisión de las Unidades Primarias de Muestreo (UPM) utilizadas en la I ENUT y la II ENUT. Estas diferencias quedan expuestas en extenso en la sección 7.2.9 sobre Comparación del Marco Muestral de Viviendas (MMV) del año 2020 y el Marco Muestral del 2015 (MM) del 2015. Las principales diferencias, se pueden resumir en los siguientes puntos:

El MM 2015 consideraba dos componentes principales, el Marco Muestral de Manzanas para ciudades y grandes centros urbanos, y el Marco Muestral de Secciones para áreas más periféricas (Resto de área urbana, RAU, y rural) el cual se encontraba basado en el Censo del 2002 con actualizaciones menores. La actualización del MM 2015 fue limitada, lo que dejó en evidencia problemas de cobertura y precisión en áreas con cambios demográficos importantes o con unidades de viviendas nuevas.

El MMV del 2020 reemplaza el esquema de dos componentes del MM 2015 por un listado único que diferencias entre las áreas urbanas de las rurales, sin incluir la división RAU, con cobertura a nivel nacional y un

³ www.ine.gob.cl/calidad-estadistica/directrices-metodologicas.

enfoque en áreas homogéneas y actualizadas. El MMV 2020 se diseñó a partir del Censo del año 2017, incorporando nuevas áreas de exclusión (ADA y ACO), y ofrece una mayor precisión al incluir actualizaciones periódicas de los límites urbanos, una estratificación económica detallada, y áreas especiales. Esto permite una mayor cobertura en zonas previamente excluidas, reduciendo el sesgo en las estimaciones.

El enfoque del MMV del 2020 en base al Censo 2017, es de revisiones anuales, aportando un marco más preciso y flexible, lo cual ayuda a reducir el riesgo en estimaciones compensando ciertas desactualizaciones.

3.4.4. Factores de expansión

Salvaguardando las diferencias en el diseño muestral entre ambas versiones de la ENUT, el proceso de construcción de los factores de expansión es similar entre ellas. Ambas versiones parten de un factor base, el cual corresponde al factor teórico de las unidades de primera y segunda etapa, y luego se incorporan distintos ajustes (por omisión de comunas, elegibilidad, no respuesta y suavizamiento), hasta llegar a un factor calibrado a los stocks poblacionales con la información más actualizada al año de publicación. Sin perjuicio de estas similitudes, existen algunas diferencias producto de los esfuerzos institucionales por mejorar los procesos estadísticos e implementar innovaciones metodológicas. Las principales diferencias metodológicas se listan a continuación:

- i. **Factor teórico de primera etapa.** Considerando el cambio de marco muestral entre ambas versiones de la encuesta, las unidades de primera etapa son distintas: la I ENUT emplea manzanas y la II ENUT emplea Unidades primarias de muestreo construidas a partir de un número aproximado de viviendas

esperadas dentro de una zona geográfica (UPM)⁴. Esto tiene un impacto directo en la construcción de los factores, ya que cambia el significado del factor base. Como el factor teórico se define como el inverso de la probabilidad de selección, en la I ENUT el factor teórico se traduce en cuántas manzanas en la población son representadas por una manzana en la muestra. En cambio, en la II ENUT el factor teórico se traduce en cuántas UPM en la población son representadas por una UPM en la muestra.

- ii. **Cantidad de factores de expansión.** En la I ENUT, existen tres factores de expansión, uno para cada nivel de análisis: hogares, integrantes del hogar y personas que responden CUT. En la II ENUT, existen solo dos factores de expansión: hogar/integrantes del hogar y personas que responden CUT. Esto quiere decir que, para esta versión de la ENUT, un factor de expansión se puede emplear para análisis a nivel de hogar, así como también a nivel de integrantes del hogar. Este factor sintético se ha denominado factor de expansión CH (fe_ch) y es de especial utilidad para efectuar análisis con la información del cuestionario del hogar.
- iii. **Calibración a los stocks poblacionales.** El proceso de calibración a los *stocks* poblacionales cambia en fuente y en procedimiento. En cuanto a fuente, la I ENUT utilizaba las proyecciones poblacionales en base al CENSO 2002, en cambio, la II ENUT utiliza las proyecciones poblacionales en base al CENSO 2017. En cuanto a procedimiento, se calibran los factores de expansión de la II ENUT considerando los avances metodológicos que ha tenido la institución en esta temática. En concreto, se incorpora una

⁴ Esto por un esfuerzo de homologar el tamaño de las unidades primarias de selección o muestreo en término de unidades de segunda fase (viviendas en este caso) en su interior, de tal manera de mantener probabilidades de selección similares para estas últimas.

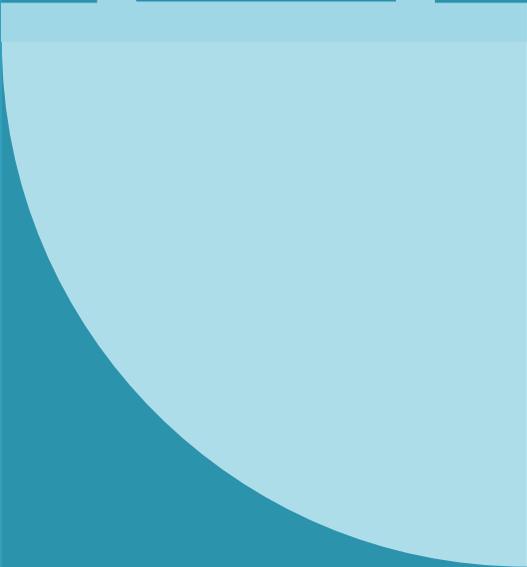
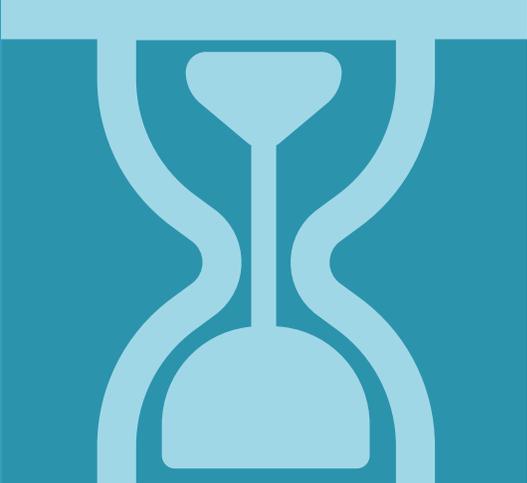
calibración por tramos de edad desagregados, con el objetivo de abordar la potencial subrepresentación de población de adultos jóvenes en encuestas relacionadas al mercado laboral (Alvarado & Pizarro, 2019). En concreto, la I ENUT calibraba por dos tramos de edad ([0-11] y [+12]), en tanto, la II ENUT considera la calibración por ocho tramos de edad para el factor de expansión CH ([0-11], [12-14], [15-24], [25-34], [35-44], [45-54], [55-64], [65+]) y seis tramos de edad para el factor de expansión CUT ([12-24], [25-34], [35-44], [45-54], [55-64], [65+]).

3.4.5. Módulo de simultaneidad

Una innovación con respecto a versiones previas de la ENUT en Chile y en otras regiones, es la incorporación de un módulo de Simultaneidad. Este módulo permite medir con mayor precisión el uso del tiempo, ampliando el alcance de los datos recopilados para incluir actividades que se realizan de manera simultánea.

Su función es permitir el registro y estudio de la realización de actividades simultáneas o múltiples en distintas subpoblaciones. El reconocimiento de acciones de esta clase no solo contribuye a una recopilación de datos más profundo del fenómeno de estudio, sino también visibiliza los efectos que puede tener este tipo de práctica en los niveles de satisfacción, y las características de los hogares en que ocurre más frecuentemente.

Por último, cabe mencionar que las acciones simultáneas consultadas en este módulo son de dos clases, actividades y de traslado, tomando como base las categorías permitidas en los listados de actividades del mismo Cuestionario sobre Uso del Tiempo (CUT).



4

**MARCO DE
REFERENCIA**

4. Marco de referencia

4.1. Marco Conceptual

En esta sección se presenta el marco conceptual que configura las dimensiones abordadas por la II ENUT, y cuya base son las directrices provenientes del Sistema de Cuentas Nacionales, elaboradas conjuntamente por los siguiente cinco organismos internacionales: Organización de las Naciones Unidas, Banco Mundial, Fondo Monetario Internacional, Comisión Europea y la Organización para la Cooperación y el Desarrollo Económico.

4.1.1. Concepto de trabajo

La forma en que se organiza el trabajo es uno de los elementos centrales de la fisonomía que adquiere la sociedad, en tanto representa un factor determinante del establecimiento de las jerarquías sociales, las diferencias de género y la conformación de identidades de los sujetos (Gómez, 2010). Por otro lado, los cambios en el mercado del trabajo movilizan transformaciones en la forma en que las personas organizan su tiempo para las distintas actividades, alterando con ello la organización de la vida doméstica (Aguirre, 2024).

En los últimos años las transformaciones del mundo del trabajo en la ocupación se han visto fuertemente impulsadas por la globalización y el desarrollo tecnológico, decantando en relaciones de trabajo más flexibles.

Por todo lo anterior, y dado que el trabajo en su sentido amplio desempeña un rol articulador de la vida social, el estudio del tiempo destinado a las formas que adopta, y a su influencia en otros tipos de actividades, resulta esencial para caracterizar las condiciones de vida de

la población sin dejar fuera aspectos fundamentales vinculados a la organización de la sociedad.

4.1.2. Sistema de Cuentas Nacionales

Uno de los elementos pendientes a nivel internacional en el desarrollo de las estadísticas sobre uso del tiempo, es la armonización metodológica y conceptual de los estudios bajo esta temática. En un intento por superar este déficit, distintos organismos internacionales como Naciones Unidas o la Comisión Económica para América Latina y el Caribe (CEPAL) han propuesto estrategias que permiten la comparabilidad de información entre países.

Esta aspiración ha implicado que se proponga al Sistema de Cuentas Nacionales (SCN) como un marco de referencia para la conceptualización, medición y clasificación de las actividades en las encuestas de uso del tiempo. En esencia, el SCN organiza varias actividades humanas a partir del eje del trabajo, distinguiendo entre actividades personales y productivas, y dentro de las productivas, entre el trabajo realizado para el mercado, por un lado, y el que tiene por objeto el consumo de los propios hogares por el otro.

Este alineamiento de los marcos conceptuales también se ve reflejado en la resolución estadística de la Organización Internacional del Trabajo (OIT) relacionada con la medición de “las estadísticas del trabajo, ocupación y subutilización del trabajo” aprobada en el marco de la 19° Conferencia de Estadísticos del Trabajo del año 2013 (OIT 2013) y que propone un marco integrador y alineado al SCN 2008 sobre la medición de las distintas formas de trabajo. Para una correcta comparación entre países, se destaca el uso como referencia de los clasificadores internacionales de actividades para las encuestas sobre uso del tiempo

(CAUTAL⁵ e ICATUS⁶) en el diseño y construcción de los instrumentos de recolección y análisis, los que permiten un mayor alcance en el análisis económico, social y de género a los niveles más agregados de estos clasificadores como secciones o divisiones, tales como Trabajo doméstico, Educación o Empleo.

4.1.3. Frontera general de la producción en hogares

En el marco conceptual del SCN, el significado de “producción” constituye un elemento central que actúa como criterio para sistematizar diversas actividades humanas. Según este, producir debe entenderse como “(...) un proceso físico, realizado bajo la responsabilidad, control y gestión de una unidad institucional, en el que se utilizan mano de obra y activos para transformar insumos de bienes y servicios en productos de otros bienes y servicios”, (SCN, 2008, 1.40).

El trabajo representa un insumo de cualquier proceso productivo, y en relación con la estructura del SCN, delimita los alcances de la frontera general de producción (Gómez, 2010). Esa frontera, relacionada a un concepto particular de trabajo analizado desde una perspectiva integral, incorpora tanto el que se efectúa con fines de mercado, como el de la actividad no remunerada para el consumo del propio hogar, debido a su impacto en el mercado laboral y en los hogares. En consecuencia, a partir del SCN 2008 “(...) incluye conceptos y definiciones que permiten un tratamiento del trabajo no remunerado acorde con el núcleo central de las cuentas nacionales”, (Gómez, 2010, p. 13).

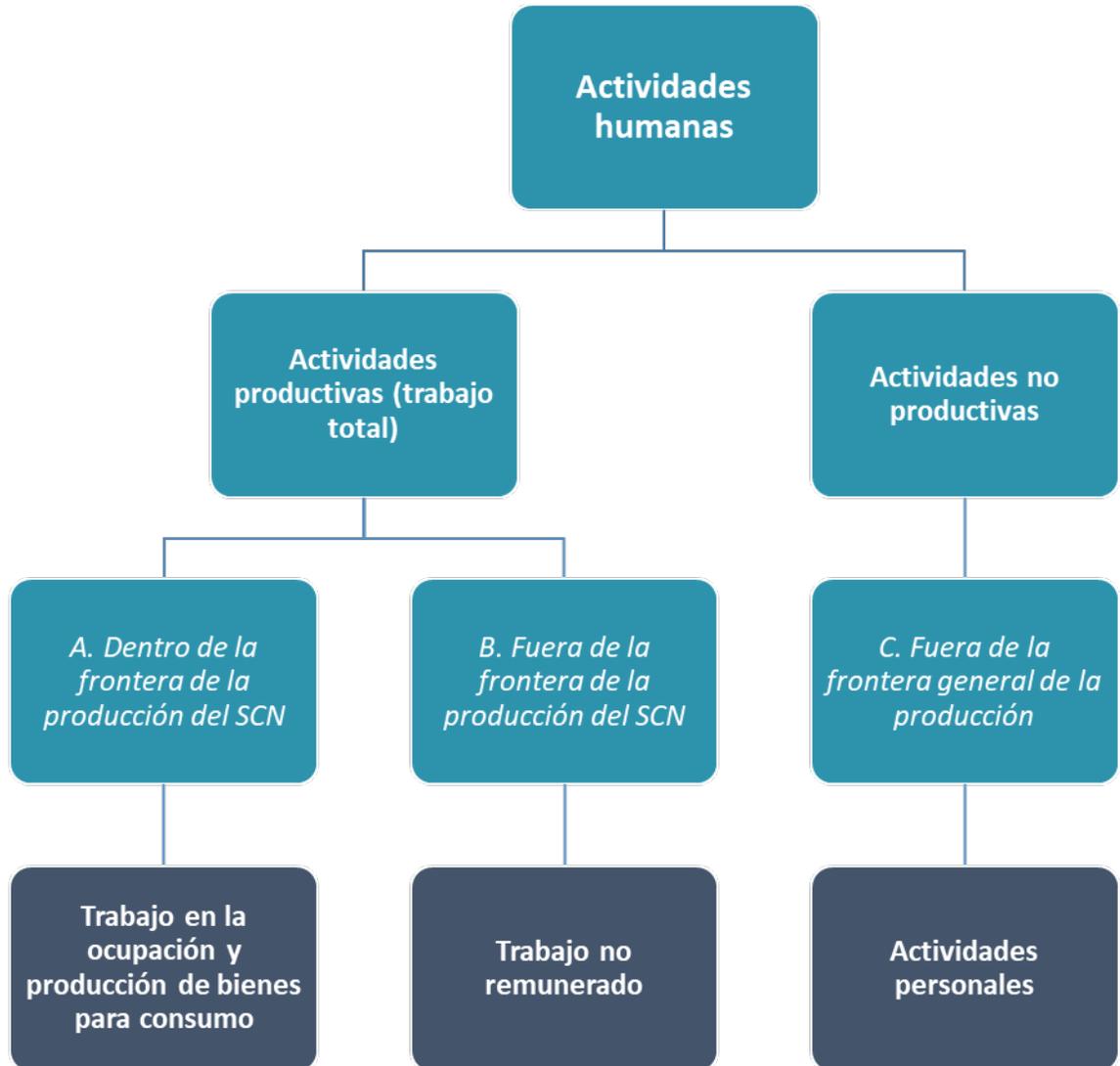
⁵ CAUTAL: Clasificación de Actividades de Uso del Tiempo para América Latina y el Caribe. www.cepal.org/es/publicaciones/39624-clasificacion-actividades-uso-tiempo-america-latina-caribe-cautal

⁶ ICATUS: Clasificación Internacional de Actividades para Estadísticas sobre el Uso del Tiempo de 2016. <https://unstats.un.org/unsd/demographic-social/time-use/icatus-2016>.

Este giro conceptual en cómo se comprende la producción, implica una valorización del trabajo más allá de su papel en el mercado, visibilizando el aporte productivo del realizado en hogares, y el papel que este asume en el bienestar de las personas. A partir de lo anterior, el marco conceptual de la clasificación CAUTAL utilizado en la ENUT, “(...) *incorpora criterios económicos fundamentados en el SCN y parte de distinguir dos grandes conjuntos: las actividades productivas, vinculadas a la producción de bienes y servicios, y las actividades no productivas o personales*” (CAUTAL, 2016, p. 13).

La Figura 1 corresponde a un diagrama sencillo presente en el propio documento de Clasificación de Actividades de Uso del Tiempo para América Latina y el Caribe (CAUTAL) que expone el criterio de clasificación en base al SCN, y sirve de referencia para entender la lógica de categorización de las actividades humanas:

Figura 1. Marco conceptual para la clasificación de Actividades de Uso del Tiempo para América Latina y el Caribe (CAUTAL)



Nota: el punto "B. Fuera de la frontera de la producción del SCN", se encuentra, por definición dentro de la frontera general de la producción.

Fuente: CAUTAL (2016)

El primer paso en el ordenamiento de las actividades es establecer si las actividades son productivas o no, es decir, si se encuentran vinculadas a la producción de bienes y servicios. Las actividades no productivas o personales por su parte abarcan las actividades de aprendizaje, convivencia social, entretenimiento y deportivas, y cuidados personales relacionados al uso del tiempo libre.

Entre las actividades productivas, se encuentra en primer lugar, las que clasifican como trabajo en la ocupación y producción de bienes para consumo, siendo aquellas acciones que producen bienes o servicios para el auto consumo, o el trabajo en la ocupación a cambio de un pago o beneficio. Por otra parte, se encuentran aquellas identificadas como trabajo no remunerado, que corresponden a toda acción que produzca servicios sin que exista de por medio un pago o beneficio, e incorporan principalmente a actividades domésticas y de cuidado de personas en el propio hogar, para otros hogares o a la comunidad.

4.1.4. Concepto ampliado de “trabajo”

En la 19° Conferencia Internacional de Estadísticos del Trabajo (19° CIET), la OIT define trabajo como “(...) *todas las actividades realizadas por personas de cualquier sexo y edad con el fin de producir bienes o prestar servicios para el consumo de terceros o el autoconsumo*” (OIT, 2013, p. 64), independiente de la legalidad, formalidad o la mediación de pago por la actividad realizada, en línea con las fronteras de la producción establecidas por el SCN 2008.

Dentro de las precisiones más relevantes al momento de hablar del concepto o de su vínculo con otros tipos de actividades, se encuentra la delimitación del espectro de actividades posibles de clasificar como trabajo. Tradicionalmente “(...) *se realiza un uso indistinto de las nociones de trabajo y empleo*” (Aguirre, 2024) restringiendo el alcance del concepto a las actividades que tienen reconocimiento social asociado al carácter

del mercado de trabajo. Sin embargo, cuando se amplía la noción de trabajo desde el empleo (trabajo en la ocupación, según la 19° CIET) a un concepto que supera los criterios de mercado, se puede establecer que las diversas actividades realizadas por la población son posibles de separar en productivas, caracterizadas por el cumplimiento del criterio de tercera persona, es decir “(...) *si su realización puede encomendarse a otro individuo con los mismos resultados deseados*”, y en no productivas o actividades personales, (Naciones Unidas, 2006a, p. 196).

Con esta nueva conceptualización, la 19° CIET identifica cinco formas de trabajo:

- a) El **trabajo en la ocupación**, realizado para terceros a cambio del pago en dinero o de beneficios.
- b) El **trabajo en formación no remunerado**, realizado para terceros con la finalidad de adquirir experiencia.
- c) El **trabajo voluntario**, no remunerado y realizado para terceros.
- d) El **trabajo de producción de bienes y servicios para el autoconsumo**, no remunerado y realizado para el uso propio, de la persona o el hogar.
- e) Otras actividades productivas no definidas.

A partir de esos criterios, es viable concebir el **trabajo como cualquier actividad que realiza una persona con la finalidad de obtener bienes y servicios**, convirtiéndolo en un insumo de la producción, con independencia de si es remunerado o no. Sentido en el que Gómez plantea que “(...) *desde el punto de vista económico, el trabajo es la fuente de valor y de la riqueza económica*”, (Gómez, 2010, p. 25).

Esta diferenciación es crucial, ya que el aporte de los distintos tipos de actividades al bienestar de las personas y sociedades se sustenta tanto en el trabajo realizado en contextos de mercado, como en las actividades

productivas desarrolladas fuera de este, en el marco de relaciones privadas y de organizaciones sociales (Aguirre, 2024).

Desde la mirada de las estadísticas oficiales, el tránsito desde concepciones estrechas hacia un concepto integrador del trabajo se puede identificar desde la 18° Conferencia Internacional de Estadísticas del Trabajo, en su primera resolución referente a la medición del tiempo trabajado, en la que se indica:

“El tiempo de trabajo también puede estar relacionado a actividades que están más allá de la frontera de producción del SCN, como los servicios producidos y consumidos dentro del mismo hogar y las actividades de los trabajadores voluntarios en hogares que producen servicios para el propio uso final del hogar. En este caso, el objetivo es producir estadísticas del tiempo de trabajo sobre trabajo doméstico no remunerado y trabajo voluntario necesarias para la producción de cuentas satélite y para una comprensión y un enfoque más amplios de las políticas económicas, sociales y del mercado laboral”, (OIT, 2008, p. 42).

En ese ámbito, la medición del tiempo destinado a la realización de los distintos tipos de trabajo, aparece como un elemento fundamental para dar cuenta de los objetivos señalados por la OIT, de evaluar, por un lado, el involucramiento de todas las formas de actividades productivas que contribuyan al desarrollo económico, sustento de los hogares y bienestar de los individuos y sociedad, y por el otro, del estudio del impacto económico y social que genera la participación de los distintos subgrupos poblacionales en las diversas formas de trabajo (OIT, 2013).

4.2. Dimensiones de análisis II ENUT

La II ENUT cuenta con dos instrumentos o cuestionarios, al igual que su versión 2015: el Cuestionario del Hogar (CH), cuya entrevista es formulada a una persona identificada como informante idóneo⁷, y el Cuestionario sobre Uso del Tiempo (CUT), que debe ser respondido por todas las personas de 12 años o más del hogar.

El objetivo del CUT es registrar cuánto tiempo se destina a las principales actividades de la vida diaria, además de medir componentes del bienestar subjetivo de las personas entrevistadas, respecto a la distribución de su tiempo dentro de otras dimensiones.

A continuación, son descritas las dimensiones de análisis abordadas en la medición de uso del tiempo, los tipos de actividades que las componen (afines a la estructura CAUTAL) y su organización en los instrumentos de medición.

4.2.1. Trabajo en la ocupación y actividades conexas

La primera dimensión de análisis y segmentación de la CAUTAL incluye las actividades de trabajo relacionadas con la producción de bienes y servicios considerada dentro de la frontera del SCN. En él se agrupa a todas las actividades que hacen posible, facilitan o fomentan la producción, incluyendo los traslados o viajes.

Para la medición del tiempo destinado a actividades de trabajo dentro de la frontera general de la producción, fueron incorporadas las siguientes secciones dentro del Módulo de Trabajo en la Ocupación (TO):

⁷ Persona integrante del hogar de 18 años o más con conocimiento de las características socioeconómicas de quienes que conforman su hogar.

- I. **Trabajo en la ocupación:** tiempo destinado a actividades relacionadas a la ocupación, incluyendo los traslados.
- II. **Búsqueda de trabajo.**

4.2.2. Trabajo no remunerado (Fuera de la frontera de producción del SCN)

Esta dimensión de análisis corresponde a la segunda sección de CAUTAL e involucra el trabajo no contabilizado en la frontera de producción del SCN, pero que forma **parte de la frontera general de la producción**. Incluye los servicios no remunerados para el propio hogar u otros hogares y aquellos voluntariados no remunerados.

La medición del tiempo destinado a las actividades de trabajo fuera de la frontera de la producción del SCN se organiza en las siguientes subdimensiones:

Cuidados no remunerados a miembros del hogar: el módulo de Trabajo de cuidados no remunerados (TC) del CUT se estructura de acuerdo con las edades o características de salud de los integrantes del hogar, distingue por naturaleza del trabajo de cuidado, y si la acción fue dirigida a sujetos de cuidado permanente o a otros miembros del hogar. Este módulo comprende las siguientes secciones:

- I. Cuidados esenciales a integrantes del hogar.
- II. Cuidados relativos a la enseñanza de integrantes del hogar.
- III. Otras actividades de cuidados a integrantes del hogar.

Trabajo no remunerado en quehaceres del propio hogar: el módulo de Trabajo doméstico no remunerado (TD) del CUT circunscribe una serie de acciones sin pagos económicos de por medio. Se subdivide en las secciones:

- I. Preparación y servicio de comida.

- II. Limpieza de la vivienda.
- III. Limpieza y cuidado de ropa y calzado.
- IV. Mantenimiento y reparaciones menores en el propio hogar.
- V. Administración y abastecimiento del hogar.
- VI. Cuidado de mascotas y plantas.

Trabajo no remunerado y de ayuda para otros hogares: el módulo de Trabajo Voluntario (TV), en su sección 1 de “Quehaceres domésticos y de cuidados para otros hogares”, incluye actividades de ayuda doméstica y de cuidado a personas de hogares diferentes al del respondiente. Por su parte, la sección de “Trabajo voluntario” consulta por el ejercicio de actividades de este tipo.

4.2.3. Actividades personales (Fuera de la frontera de la producción del SCN)

La tercera dimensión de análisis y sección de CAUTAL corresponde a las actividades no productivas. El tiempo que se invierte en las actividades de trabajo, marca el tiempo dedicado a otro tipo de actividades.

Sobre esta área en la II ENUT 2023 se consulta lo siguiente:

Cuidados personales y necesidades fisiológicas: módulo Cuidados Personales (CP) que indaga sobre necesidades fisiológicas, tales como dormir, desayunar y almorzar, la asistencia a consulta médica, dental, o a realizarse exámenes, junto a los traslados asociados.

Actividades de educación y aprendizaje: módulo Educación (ED), que distingue según la formalidad de la instrucción entre asistencia al establecimiento educacional (sección 1) y Otras actividades de aprendizaje (sección 2).

Ocio y vida social: presente en el módulo Ocio y vida social (VS), incluye consultas sobre temáticas como vida social, asistencia a eventos, Juegos y aficiones, deportes y uso de medios de comunicación.

4.3. Recomendaciones internacionales para medir uso del tiempo

En la elaboración de los cuestionarios fueron consideradas las recomendaciones internacionales sobre la realización de encuestas sobre uso del tiempo. Tres fuentes de información constituyen en ese sentido, la base o guía que se utilizó dentro del diseño de las preguntas, a la par de apoyo conceptual y metodológica de la medición: el primer documento es la “Guía de elaboración de estadísticas sobre el empleo del tiempo para medir el trabajo remunerado y no remunerado” (UN, 2006a). En segundo lugar “Directrices y referentes conceptuales para armonizar las encuestas sobre uso del tiempo en América Latina y el Caribe”, (Gómez, 2010) junto a la guía de UNECE *Guidelines for Harmonizing Time Use Surveys* (UNECE, 2013). Por último, se revisaron cuestionarios y documentación metodológica de diferentes encuestas sobre uso del tiempo a nivel regional, en particular se revisó la experiencia de las ENUT de Argentina 2021 (INDEC Argentina, 2021), ENUT Colombia 2020-2021 (DANE Colombia, 2022), Encuesta de Uso del Tiempo y Trabajo no Remunerado EUT Uruguay 2021 (INE Uruguay, 2023) y ENUT Costa Rica 2022 (INE Costa Rica, 2022). Todo lo anterior ayudó como insumo para revisar buenas prácticas internacionales, actualizarse en los estándares de mediciones de tiempo a nivel internacional y alinearse con listados mínimos de actividades para aplicar en el Cuestionario de Uso del Tiempo de la II ENUT.

En el área de clasificación de actividades humanas a partir del trabajo y su relación con el mercado, se consultó al manual del “Sistema de

Cuentas Nacionales 2008” (SCN, 2008), el Informe de la 19° Conferencia Internacional de Estadísticos del Trabajo (OIT, 2013), el Informe de la 20° Conferencia Internacional de Estadísticos del Trabajo (OIT, 2018) y el Manual de medición del trabajo voluntario (OIT, 2011).

4.4. Experiencias internacionales sobre Uso del Tiempo

En esta sección se abordarán algunas experiencias internacionales respecto a la medición sobre uso del tiempo.

En el 2007 se creó el Grupo de Trabajo sobre Estadísticas de Género de la Conferencia Estadística de las Américas (CEA) de la CEPAL, impulsando distintas iniciativas para la producción estadística de género y uso del tiempo en países de América Latina y el Caribe.

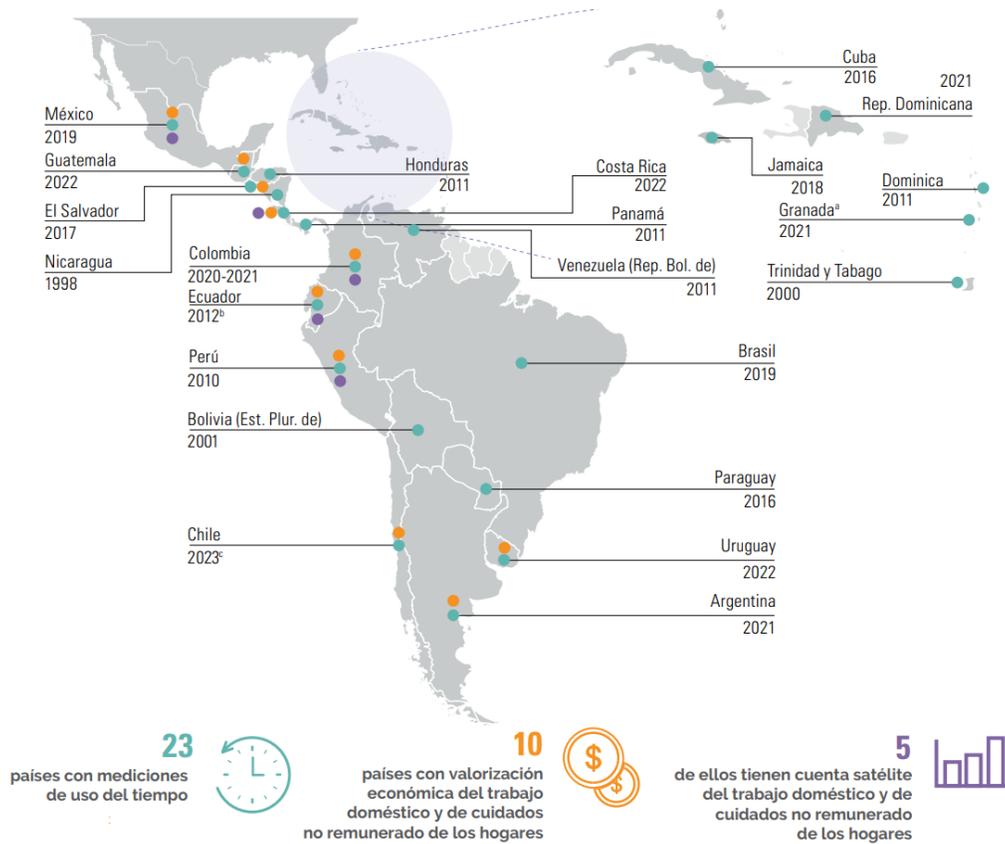
Con este grupo de trabajo se han generado variadas instancias de cooperación, que han permitido distintos avances metodológicos en la región al generar sinergias y conocer las experiencias de otras oficinas nacionales estadísticas en esta materia.

Hasta 2022, un total de 23 países de América Latina y el Caribe han realizado al menos una medición sobre uso del tiempo, enfocada en el trabajo doméstico y de cuidados, donde 10 de ellos han valorizado económicamente el trabajo no remunerado en su conjunto⁸.

La Figura 2 muestra las últimas experiencias en la región en estudios sobre uso del tiempo.

⁸ Para mayores detalles, referirse a la guía metodológica sobre mediciones de uso del tiempo en América Latina y el Caribe (Cepal, 2022)

Figura 2 América Latina y el Caribe: avances en las mediciones de uso de tiempo y del trabajo no remunerado, 1998-2021



Fuente: Comisión Económica para América Latina y el Caribe (CEPAL, 2023a). Repositorio sobre uso del tiempo de América Latina y el Caribe.

^a Hasta la finalización del documento de referencia, Granada había realizado únicamente un levantamiento piloto de información sobre uso del tiempo.

^b Los indicadores de uso del tiempo se basan en la Encuesta de Uso del Tiempo (EUT) 2012 y la valoración económica y la cuenta satélite se utilizan estimaciones en base a los datos de la EUT 2012 y los módulos de la Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEDMU) hasta 2017.

^c referente a la presente II ENUT de Chile.

Existen varios elementos metodológicos en los que los distintos países coinciden y otros en los que sigue existiendo debate. Dentro de estas dimensiones, existen dos fundamentales que se abordan a continuación.

- **Listado de actividades o Diario de registro**

La II ENUT, al igual que su primera versión 2015 utiliza listado de actividades para registrar la participación y el tiempo destinado a las actividades del quehacer diario. Esto implica que, en una entrevista, se consulta directamente por ciertas actividades concretas de interés para el estudio. Las ventajas que trae esta forma de captura, es que se busca mayor precisión en la captura de los fenómenos de mayor interés, reduciendo la subdeclaración de las personas informantes en actividades que se desean relevar y que para ellas pueden pasar desapercibidas en su día a día, como algunas tareas domésticas o de cuidados.

Por otra parte, es posible hacer una entrevista donde se va consultando por tramos horarios, qué actividades realizó en cada uno de ellos, completando un historial lineal de actividades realizadas y el tiempo destinado. Este método permite capturar mayor variedad de actividades y permite generar estudios respecto a cómo se ordenan las distintas actividades en el día consultado, pero requiere una codificación posterior de estas actividades declaradas⁹.

Como se ha indicado anteriormente, la II ENUT ha decidido utilizar un listado de actividades, manteniendo la metodología de la versión anterior, y privilegiando la mejor calidad de las estimaciones asociadas a trabajo doméstico y de cuidados no remunerados, modificando la forma de

⁹ También es posible realizar diarios pre-codificados, donde la persona encuestadora, una vez la persona informante declara una actividad, debe ir seleccionando en un listado pre-codificado aquella que corresponda con lo informado.

preguntar sobre la edad de las personas sujetos de cuidado y la participación en actividades de trabajo doméstico.

- **Períodos de referencia**

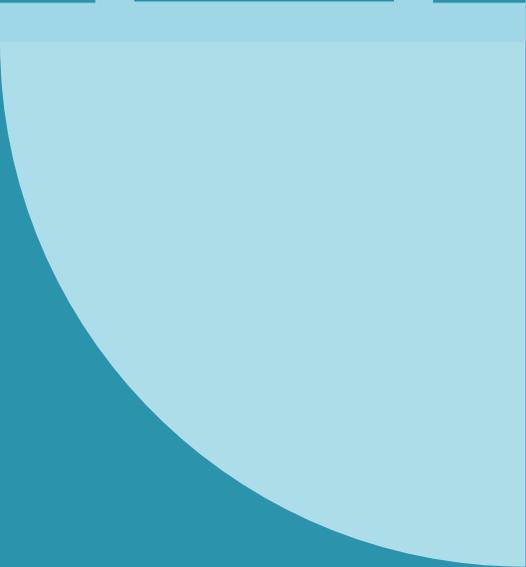
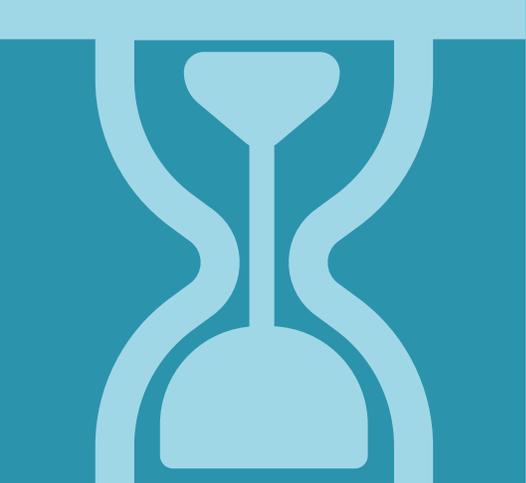
Existen dos períodos de referencia utilizados en la región¹⁰, uno de estos se enfoca en consultar por tiempo destinado durante una semana, y en otros casos se utiliza un día como referencia.

Ahora bien, existen variantes respecto a por qué día se consulta, pudiendo ser el día anterior a la entrevista, que tiene la ventaja del menor período de recuerdo posible. Por otro lado, otras experiencias preguntan respecto a un día típico, que permite a la persona informante generar un constructo general de cómo se desarrollan sus días habitualmente, dejando fuera la excepcionalidad de ocurrencias poco habituales.

En las dos mediciones realizadas en Chile, se ha optado por otro camino, asignando un día de semana y fin de semana previos a la entrevista, de tal manera de captar información balanceada de todos los días de la semana a lo largo de la muestra¹¹.

¹⁰ América Latina y el Caribe

¹¹ Para evitar periodos de recuerdo demasiado largos, se instó a los equipos de recolección a evitar entrevistar a una persona informante un día que generara distancias superiores a 4 días del día a consultar (día de referencia).



5

**DISEÑO
ESTADÍSTICO**

5. Diseño estadístico

5.1. Población objetivo

La población objetivo la constituyen todas las personas de 12 años o más, pertenecientes a 112 comunas del área urbana de todas regiones del país, que cuenten con condiciones físicas y mentales aptas para responder la encuesta y que residan habitualmente en viviendas particulares ocupadas (elegibles), durante el período de trabajo de campo de la encuesta.

5.2. Cobertura geográfica

La cobertura geográfica de la II ENUT representa el área urbana total del país, a través de la presencia de 112 comunas en la muestra, que en conjunto abarcan cerca del 80% de la muestra urbana de viviendas del país.

En la Tabla 1 se listan las comunas presentes en la encuesta.

Tabla 1. Cobertura geográfica por región

Región	Comunas presentes en la muestra
Arica y Parinacota	Arica
Tarapacá	Alto Hospicio, Iquique
Antofagasta	Antofagasta, Calama
Atacama	Copiapó, Vallenar, Caldera
Coquimbo	Coquimbo, La Serena, Ovalle
Valparaíso	La Calera, Concón, Limache, Los Andes, Quillota, Quilpué, San Antonio, San Felipe, Valparaíso, Villa Alemana, Viña del Mar
Metropolitana	Santiago, Puente Alto, Maipú, La Florida, Las Condes, Ñuñoa, San Bernardo, Peñalolén, Providencia, Pudahuel, Quilicura, Estación Central, La Pintana, Recoleta, El Bosque, Renca, Macul, San Miguel, Quinta Normal, Cerro Navia, Conchalí, Independencia, Colina, La Granja, Vitacura, San Joaquín, La Cisterna
O'Higgins	Chimbarongo, Graneros, Machalí, Mostazal, Rancagua, Rengo, San Fernando, San Vicente, Santa Cruz, Doñihue, Pichilemu
Maule	Cauquenes, Constitución, Curicó, Linares, Talca, Maule, Molina, Parral, San Javier
Ñuble	Chillán, Chillán Viejo, San Carlos, Yungay, Bulnes, Quillón
Biobío	Chiguayante, Concepción, Coronel, Curanilahue, Hualpén, Los Ángeles, Lota, Penco, San Pedro de la Paz, Talcahuano, Tomé
La Araucanía	Angol, Lautaro, Nueva Imperial, Padre las Casas, Pitrufquén, Temuco, Victoria, Villarrica, Collipulli, Pucón, Loncoche, Vilcún
Los Ríos	La Unión, Panguipulli, Río Bueno, Valdivia, Paillaco
Los Lagos	Ancud, Castro, Osorno, Puerto Montt, Puerto Varas, Quellón
Aysén	Aysén, Coyhaique
Magallanes	Punta Arenas

5.3. Unidades estadísticas

5.3.1. Unidad de observación

La unidad de información proviene de dos fuentes. La primera corresponde al informante idóneo, el cual debe responder por el Cuestionario del Hogar y es una persona de 18 años o más, con capacidades físicas y mentales que le permitan responder por la información de sí mismo y de los integrantes del hogar. La segunda fuente de información corresponde a los integrantes del hogar de 12 años o más para responder el Cuestionario sobre Uso del Tiempo, presentes en la vivienda durante el período del trabajo de campo de la encuesta, con capacidades físicas y mentales aptas para responder.

5.3.2. Unidad de análisis

La unidad de análisis corresponde a personas y hogares pertenecientes a las viviendas particulares que han sido seleccionadas en la encuesta.

5.3.3. Unidad de muestreo

Las unidades de muestreo en el diseño muestral de la II ENUT son:

- **Unidad de primera etapa:** conglomerados o unidades primarias de muestreo (UPM).
- **Unidad de segunda etapa:** Viviendas particulares ocupadas o unidades secundarias de muestreo (USM).

5.4. Definición de variables

Las variables que delimitan al producto estadístico son de dos tipos: capturadas y construidas. Por un lado, las variables capturadas se refieren a aquellas que tienen por fuente de información el cuestionario, ya sea el CH o el CUT. En el caso del CH, estas variables entregan información clave para cumplir con los objetivos del producto estadístico. Algunas variables clave son:

- Parentesco (pco):
- Sexo (sexo)
- Edad (edad)
- Asistir al establecimiento educacional (e4)
- Estar ocupado (o3, o4, o o5)
- Identificación de PSDF (d1 o d8)

En el caso del CUT, las variables capturadas permiten hacer análisis sobre uso del tiempo al nivel más desagregado de CAUTAL, así como también construir las variables que permiten hacer análisis más agregados. En general, las variables siguen una nomenclatura donde se identifica el módulo al cual pertenecen, luego el número de pregunta dentro del módulo, luego un interfijo “_t_” o “_p_” según correspondan a una variable de tiempo o participación (respectivamente) y, al final, un sufijo “_ds” o “_fds” según correspondan a variables con información de día de semana o fin de semana (respectivamente).

Por ejemplo, la pregunta *TD1. ¿Cocinó, preparó o calentó alimentos para usted o integrantes de este hogar?* es representada por las siguientes variables:

- Participación en día de semana: td1_p_ds
- Tiempo en día de semana: td1_t_ds
- Participación en día de fin de semana: td1_p_fds

- Tiempo en día de fin de semana: $td1_t_fds$ ¹²

Por otro lado, las variables construidas se refieren a aquellas que se crean a raíz de las variables capturadas y permiten sintetizar o agregar información. Estas variables se construyen mediante una rutina de código. En el caso del CH, algunas variables construidas fueron:¹³

- Condición de actividad económica (cae)
- Jornada de trabajo (jor_to)
- Clasificación internacional de la situación de empleo (ci se)

En lo que respecta a las variables del CUT, las variables construidas corresponden a los distintos niveles de agregación de CAUTAL. Es decir, hay variables a nivel de *Secciones*, *Grandes divisiones* y *Divisiones* para los distintos tipos de trabajo y para las actividades personales.

En la Figura 3 se presentan las jerarquías de clasificación sobre las actividades de uso del tiempo junto con sus nomenclaturas. Por ejemplo, a nivel de sección se presenta el *Trabajo no remunerado*, donde todas sus variables comienzan con el prefijo tnr. A nivel de gran división están los *Trabajos de cuidados no remunerados*, *Trabajo doméstico no remunerado* y *Trabajo voluntario y ayudas a otros hogares*, los cuales tienen los siguientes prefijos para sus variables: tcnr, tdnr y tvaoh. Finalmente, dentro de cada gran división se encuentran las divisiones, por ejemplo, dentro del *Trabajo doméstico no remunerado* se encuentran las distintas actividades relativas a este tipo de trabajo.

¹² Para más detalle sobre la nomenclatura en las variables, ver el capítulo 3 “Tipos de variables de la II ENUT” del Manual de Usuario II ENUT.

¹³ Para mayor detalle, ver el Diccionario de Variables de la II ENUT

Figura 3. Jerarquías de clasificación de uso del tiempo

	SECCIONES	GRANDES DIVISIONES	DIVISIONES		
Car ga Global de Trabajo	Trabajo en la Ocupación (Dentro de la frontera de la producción del SCN) - to	to Trabajo en la Ocupación	to	Trabajo en la ocupación y traslados asociados	
			tto	Traslados de Trabajo en la Ocupación	
		Trabajo No Remunerado (Fuera de la frontera de la producción del SCN) tnr	tcnr Trabajo de Cuidados No Remunerado	tcnr_ce	Cuidados Esenciales a integrantes del hogar
				tcnr_re	Cuidados relativos a la Enseñanza
				tcnr_oac	Otras Actividades de Cuidado
				tcnr_0_4	Trabajo de Cuidados en niños y niñas de 0 a 4 años
				tcnr_5_14	Trabajo de Cuidados de niños, niñas y adolescentes de 5 a 14
				tcnr_nna	Trabajo de Cuidados de niños, niñas y adolescentes de 0 a 14
				tcnr_15_65	Trabajo de Cuidados de personas de 15 a 65 años
				tcnr_66	Trabajo de Cuidados de personas de 66 años o más
tcnr_psdif	Trabajo de Cuidado para PSDF				
cgt	tdnr Trabajo Doméstico No Remunerado para el Propio Hogar			tdnr_psc	Preparación y servicio de comida
		tdnr_lv	Limpieza de la vivienda		
		tdnr_lrc	Limpieza y reparación de ropa		
		tdnr_mrm	Mantenimiento y reparaciones menores del hogar		
		tdnr_admnhog	Administración del hogar		
		tdnr_comphog	Compras para el hogar		
		tdnr_cmp	Cuidado de mascotas y plantas		
		tvaoh Trabajo voluntario y ayudas a otros hogares	tvaoh_tv	Trabajo Voluntario para ISFL y la comunidad	
			tvaoh_oh	Trabajo No Remunerado para Otros Hogares	
		ap	cpaf Cuidados Personales y Actividades Fisiológicas	cpaf_cp	Cuidados personales
cpaf_af	Actividades Fisiológicas				
ed Educación	ted		Traslados en Educación		
	vsyomcm Ocio y Medios de Comunicación Masiva		vsyo	Vida Social y Ocio	
mcm					
vsyo_csar	Convivencia Social y Actividades Recreativas				
vsyo_aa	Arte y Aficiones				
		Medios de Comunicación Masiva			

5.5. Clasificaciones estadísticas

Las clasificaciones estadísticas corresponden a herramientas que permiten organizar y categorizar ciertas características de una realidad social o económica. Las clasificaciones consideran una serie de criterios, elaborados a raíz de expertos, evidencia actualizada en la temática y su contexto de aplicación. Generalmente, las clasificaciones estadísticas son elaboradas de parte de organismos internacionales, con el fin de facilitar la comparabilidad en los análisis de distintas realidades. Así también, dependiendo del caso, las clasificaciones pueden sufrir cambios a fin de adaptarse al contexto específico de cada territorio.

Para la II ENUT, se utilizaron cuatro clasificadores:

Clasificador Internacional Uniforme de Ocupaciones (CIUO-08.CL)

Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES)

Clasificación Internacional Normalizada de la Educación (CINE) 2011

Clasificador de Actividades de Uso del Tiempo para América Latina y el Caribe (CAUTAL)

5.5.1. Clasificador Internacional Uniforme de Ocupaciones (CIUO-08.CL)

El CIUO-08.CL es una adaptación nacional de la Clasificación Internacional Uniforme de Ocupaciones (CIUO) elaborado por la OIT. El CIUO permite ordenar jerárquicamente las ocupaciones, tomando en consideración las tareas realizadas en un puesto de trabajo y las competencias requeridas para la realización de estas. La adaptación de CIUO.08.CL se efectúa con tal de reflejar de mejor manera la estructura ocupacional nacional, así como también los requerimientos de los usuarios nacionales.

El CIUO-08.CL cuenta con cuatro niveles de clasificación que van desde el mayor nivel de agregación hasta el menor nivel de agregación. Estos niveles son: Gran grupo (un dígito), Subgrupo principal (dos dígitos), Subgrupo (tres dígitos) y Grupo primario (cuatro dígitos). Por ejemplo, el grupo primario *2211. Médicos generales* se encuentra dentro del subgrupo *221. Médicos generales y especialistas*, el cual se encuentra dentro del subgrupo principal *22. Profesionales de la salud y*, a su vez, este se encuentra dentro del gran grupo *2. Profesionales, científicos e intelectuales*. En la Tabla 2 se presentan todos los grandes grupos así como también los subgrupos principales.

En la II ENUT las clasificaciones fueron realizadas a nivel de Subgrupo principal (dos dígitos), sin embargo, por razones de anonimización, la base de datos publicada contiene una versión agrupada de CIUO-08.CL. Es decir, las categorías corresponden a un nivel más agregado que el Gran grupo (un dígito) del clasificador original¹⁴.

¹⁴ Para más información, revisar la sección Metodología del procesamiento en este mismo documento

Tabla 2. Clasificador CIUO-08.CL (2 dígitos)

Gran grupo	Subgrupo principal	Glosa gran grupo y subgrupo
1		Directores, gerentes y administradores
	11	Miembros del Poder Ejecutivo y Legislativo, personal directivo de la administración pública y de otras organizaciones sociales y/o políticas, directores ejecutivos y gerentes generales.
	12	Directores y gerentes administrativos y de servicios comerciales.
	13	Directores, gerentes y administradores de producción y operaciones.
	14	Directores, gerentes y administradores de hoteles, restaurantes, comercios y de otros servicios.
2		Profesionales de las ciencias y de la ingeniería
	21	Profesionales de las ciencias y de la ingeniería.
	22	Profesionales de la salud
	23	Profesionales de la educación.
	24	Profesionales de negocios y administración.
	25	Profesionales de tecnología de la información y las comunicaciones.
	26	Profesionales en derecho, ciencias sociales y culturales.
3		Técnicos y profesionales de nivel medio
	31	Técnicos de las ciencias y la ingeniería
	32	Técnicos de la salud
	33	Técnicos en operaciones financieras y administrativas.
	34	Técnicos de servicios jurídicos, sociales, deportivos y culturales.
	35	Técnicos de la tecnología de la información y las comunicaciones.
	36	Técnicos en educación.
4		Personal de apoyo administrativo
	41	Oficinistas.
	42	Empleados en trato directo con el público.

	43	Auxiliares y ayudantes de registros contables y encargados del registro de materiales.
	44	Otro personal de apoyo administrativo.
5		Trabajadores de los servicios y vendedores de comercios y mercados
	51	Trabajadores de los servicios a las personas.
	52	Vendedores.
	53	Trabajadores de los cuidados personales.
	54	Personal de los servicios de protección y seguridad.
6		Agricultores y trabajadores calificados agropecuarios, forestales y pesqueros
	61	Agricultores y trabajadores calificados de explotaciones agropecuarias cuya producción se destina al mercado.
	62	Trabajadores forestales calificados, pescadores y cazadores cuya producción se destina al mercado
	63	Trabajadores agropecuarios, pescadores, cazadores y recolectores de subsistencia.
7		Artesanos y operarios de oficios
	71	Operarios de la construcción (no incluye electricistas).
	72	Operarios de la metalurgia y operarios de máquinas herramientas; mecánicos de vehículos, maquinarias, aviones y bicicletas.
	73	Artesanos y operarios de las artes gráficas.
	74	Trabajadores especializados en electricidad y electrónica.
	75	Operarios de procesamiento de alimentos, de la confección, ebanistas y otros oficios.
8		Operadores de instalaciones, máquinas y ensambladores
	81	Operadores de instalaciones fijas y máquinas.
	82	Ensambladores.
	83	Conductores de vehículos y operadores de equipos pesados y móviles.
9		Ocupaciones elementales
	91	Auxiliares de aseo y trabajadores de casa particular.
	92	Obreros agropecuarios, pesqueros y forestales.

93	Obreros de la minería, la construcción, la industria manufacturera y el transporte.
94	Cocineros de comida rápida y ayudantes de cocina.
95	Trabajadores ambulantes de servicios y vendedores ambulantes (excluyendo comida de consumo inmediato).
96	Recolectores de desechos y otras ocupaciones elementales.
0	Ocupaciones de las Fuerzas Armadas
01	Oficiales de las Fuerzas Armadas.
02	Suboficiales de las Fuerzas Armadas.
03	Otros miembros de las Fuerzas Armadas.

5.5.2. Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES)

CAENES corresponde a una adaptación nacional del Clasificador Chileno de Actividades Económicas CIIU4.CL 2012 para ser empleado en encuestas a hogares. CAENES facilita la clasificación de actividades económicas ya que la categoría más desagregada, puede agrupar varias categorías del CIIU4. CL 2012, o abrir otras, según sea necesario, permitiendo que el nivel de detalle requerido sea el adecuado para encuestas de hogares.

CAENES cuenta con tres niveles jerárquicos: Sección (letra), División (dos dígitos) y Clase (cuatro dígitos). Por ejemplo, la clase 4902. *Transporte de pasajeros por vía terrestre* pertenece a la división 49. *Transporte y almacenamiento* la cual, a su vez, pertenece a la sección B. *Transporte y almacenamiento*. En la Tabla 3 se presentan las clasificaciones de CAENES a nivel de sección junto con los códigos de división.

Tabla 3. Clasificador CAENES (letra)

Sección	Glosa	División
A	Agricultura, ganadería, silvicultura y pesca	01 a 03
B	Explotación de minas y canteras	04 a 09
C	Industrias manufactureras	10 a 33
D	Suministro de electricidad, gas, vapor y aire acondicionado	35
E	Suministro de agua; evacuación de aguas residuales, gestión de desechos y descontaminación	36 a 39
F	Construcción	41 a 43
G	Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas	45 a 48
H	Transporte y almacenamiento	49 a 53
I	Actividades de alojamiento y de servicio de comidas	55 a 56
J	Información y comunicaciones	58 a 63
K	Actividades financieras y de seguros	64 a 66
L	Actividades inmobiliarias	68
M	Actividades profesionales, científicas y técnicas	69 a 75
N	Actividades de servicios administrativos y de apoyo	77 a 82
O	Administración pública y defensa; planes de seguridad social de afiliación obligatoria	84
P	Enseñanza	85
Q	Actividades de atención de la salud humana y de asistencia social	86 a 88
R	Actividades artísticas, de entretenimiento y recreativas	90 a 93
S	Otras actividades de servicios	94 a 96
T	Actividades de los hogares como empleadores; actividades no diferenciadas de los hogares como productores de bienes y servicios para uso propio	97 a 98
U	Actividades de organizaciones y órganos extraterritoriales	99

5.5.3. Clasificación Internacional Normalizada de la Educación (CINE) 2011

La CINE (2011) es un marco de referencia diseñado para organizar estadísticas educativas, siguiendo un conjunto de definiciones y conceptos estandarizados a nivel mundial, la finalidad última de la clasificación es poder comparar internacionalmente niveles educativos. Este sistema emplea dos enfoques de clasificación complementarios, uno enfocado en los programas educativos (P) y otro en el nivel de logro académico (A), determinado por certificaciones reconocidas. En la II ENUT el análisis se centrará sobre los programas educativos (P).

Para efectos de la II ENUT se adaptaron las categorías de CINE 2011 a los siguientes 9 niveles (Tabla 4). Se debe mencionar que al momento de realizar la clasificación de CINE 2011 en la II ENUT, no se contaba claridad con la aplicación del nivel de educación post secundaria no terciaria en Chile, por lo cual se decidió omitir dicha categoría de la clasificación.

Tabla 4. Clasificador CINE 2011

Nivel	Glosa de nivel CINE
1	Menos que primaria
2	Primaria
3	Secundaria baja
4	Secundaria alta
5	Post secundaria no terciaria
6	Terciaria ciclo corto (técnico)
7	Terciaria (universitaria)
8	Maestría
9	Doctorado

5.5.4. Clasificación de Actividades de Uso del Tiempo para América Latina y el Caribe (CAUTAL)

La CAUTAL es un instrumento que permite la clasificación jerárquica de actividades en encuestas sobre uso del tiempo, enfocado en la región de América Latina y el Caribe. Esta clasificación permite armonizar las estadísticas de la región, ofreciendo comparabilidad internacional.

El análisis se enmarca en las actividades productivas dentro y fuera de del Sistema de Cuentas Nacionales tales como el trabajo remunerado y el trabajo no remunerado (por ejemplo, el trabajo de cuidado a miembros del hogar o el trabajo doméstico no remunerado), además de las actividades personales (tales como actividades sociales o de cuidado personal).

En cuanto a la organización del clasificador CAUTAL cuenta con la siguiente estructura: 3 secciones (letras), 9 grandes divisiones (un dígito), 34 divisiones (dos dígitos), 96 grupos (tres dígitos) y 23 subgrupos (cuatro dígitos). A efectos de la Tabla 5, sólo se detalló hasta el nivel de divisiones (dos dígitos), los cuales con en su mayoría coincidente con la clasificación realizada para la II ENUT.

Tabla 5. Clasificador CAUTAL (A nivel de sección, grandes divisiones y divisiones)

A	Trabajo en la ocupación y producción de bienes para el autoconsumo (dentro de la frontera del SCN)	
	1	Trabajo en la ocupación y actividades relacionadas
	10 *	Otras actividades productivas
	11	Trabajo en la ocupación
	12 *	Trabajo en la formación no remunerado
	13	Búsqueda de trabajo o inicio de un negocio
	14	Traslados de ida y vuelta al trabajo
	2	Trabajo para el autoconsumo de bienes
	21 *	Actividades primarias para el autoconsumo
	22 *	Actividades no primarias para el autoconsumo (se excluye la construcción)
	23 *	Construcción para uso final propio
B	Trabajo no remunerado (fuera de la frontera de la producción del SCN)	
	3	Trabajo doméstico no remunerado para el propio hogar
	31	Preparación y servicio de comida
	32	Limpieza de la vivienda
	33	Limpieza y cuidado de ropa y calzado
	34	Mantenimiento y reparaciones menores para el propio hogar
	35	Administración del hogar
	36	Compras para el hogar (incluye traslados)
	37	Cuidado de mascotas y plantas
	4	Trabajo de cuidado no remunerado a miembros del hogar
	41	Cuidado a miembros del hogar de 0 a 14 años
	42	Cuidado a miembros del hogar de 15 a 59 años
	43	Cuidado a miembros del hogar de 60 años y más
	44	Cuidado a miembros del hogar con discapacidad o dependencia permanente (todas las edades)
	5	Trabajo no remunerado para otros hogares, para la comunidad y
	51	Trabajo no remunerado para otros hogares
	52	Trabajo no remunerado para la comunidad
	53	Trabajo voluntario en instituciones sin fines de lucro
C	Actividades personales (fuera de la frontera general de la producción del SCN)	
	6	Aprendizaje y estudio
	61	Aprendizaje y estudio
	62	Traslados para actividades de estudio
	7	Convivencia social y actividades recreativas
	71	Convivencia social con familiares y otras personas
	72	Asistencia a eventos culturales, de entretenimiento y deportivos
	73	Arte y aficiones

	74	Deporte y ejercicio físico
8	Uso de medios de comunicación	
	81	Leer libros, revistas, periódicos u otro material por cualquier medio
	82	Ver televisión o videos exclusivamente
	83	Escuchar radio u otros medios de audio exclusivamente
	84	Utilizar computadora, tableta o celular (video o audio, Internet, descarga de archivos)
9	Cuidado personal	
	91	Cuidado personal
	92	Actividades fisiológicas

Fuente: Elaboración propia en base a CAUTAL (2016)

* Actividades no medidas en la II ENUT

5.6. Marco estadístico

Un marco muestral se define como una lista o mecanismo que delimita, identifica y permite acceder a los elementos de la población objetivo de una encuesta (Eurostat, 2019). La II ENUT se basa en el marco muestral más actualizado a la fecha con el que cuenta el INE: el Marco Muestral de Viviendas 2020 (MMV 2020).

5.6.1. Características del MMV 2020

Cobertura: el MMV 2020 cubre 345 comunas de las 346 que conforman el territorio nacional, continental e insular. La comuna que queda excluida corresponde a la Antártica Chilena perteneciente a la Región de Magallanes.

Estratificación geográfica: el MMV 2020 contiene variables que permiten la identificación de la división político-administrativa del país¹⁵, así como también, la identificación de las áreas urbanas y rurales de cada comuna. La definición de urbano y rural viene dada por las especificaciones del Censo de Población y Vivienda 2017.

- **Entidad urbana:** asentamiento humano con continuidad y concentración de construcciones en un amanzanamiento regular con población mayor a 2.000 habitantes, o entre 1.001 y 2.000 habitantes, donde menos de 50% de la población que declara haber trabajado se dedica a actividades primarias¹⁶.
- **Entidad rural:** asentamiento humano concentrado o disperso que posee 1.000 o menos habitantes, o entre 1.001 o 2.000 habitantes, donde más de 50% de la población que declara haber trabajado se dedica a actividades primarias.

¹⁵ Región, provincia y comuna.

¹⁶ Actividades primarias: Son todas aquellas actividades económicas que consisten en la extracción de productos naturales, tales como agricultura, pesca, caza, minería, etc.

Estratificación socioeconómica: el MMV 2020, cuenta con una estratificación socioeconómica que se realiza a nivel de Unidad Primaria de Muestreo (UPM) y se basa en un análisis de diversas variables recopiladas en el Censo 2017.

La clasificación de las UPM, de acuerdo con el nivel socioeconómico, permite aumentar la eficiencia de la inferencia en las encuestas de hogares que utilizan este marco. En línea con esto, el principal objetivo de realizar esta estratificación es inducir mayor eficiencia y precisión en las estimaciones de las estadísticas oficiales, además de servir como fuente de información para encuestas de hogares que requieren mayor análisis de carácter socioeconómico. Al estratificar se espera que las UPM dentro de un mismo estrato sean similares (disminución de la intra-varianza) y, diferentes a las de otros estratos (aumento de la inter-varianza) (Guerrero & Marín, 2020).

El detalle de la construcción de la estratificación socioeconómica y para el MMV 2020 se presenta en la sección de Anexos del presente documento.

5.6.2. Marco muestral II ENUT

El total de viviendas y UPM a nivel nacional según MMV 2020 es 6.386.059 y 35.859, respectivamente. Por su parte, el total de viviendas y UPM según el marco de selección de la II ENUT es 4.433.995 y 21.579, respectivamente. En la Tabla 6 se observan los totales de viviendas y UPM a nivel nacional y regional considerando el MMV 2020 y el marco de selección de la encuesta.

Tabla 6. Total de viviendas y UPM en el área urbana según el MMV 2020 y el marco de selección de la encuesta

Región	N° de Viviendas		N° UPM	
	MMV 2020	Marco II ENUT	MMV 2020	Marco II ENUT
Total	5.405.423	4.433.995	26.333	21.579
Arica y Parinacota	67.435	66.974	339	337
Tarapacá	105.257	99.887	514	487
Antofagasta	184.745	165.191	905	809
Atacama	93.573	75.117	464	372
Coquimbo	227.043	185.481	1.111	902
Valparaíso	617.241	498.975	3.001	2424
Metropolitana	2.296.243	1.871.485	11.168	9108
O'Higgins	247.768	200.644	1.200	963
Maule	285.388	235.786	1.394	1150
Ñuble	125.039	100.140	602	478
Biobío	490.694	395.933	2.384	1926
La Araucanía	252.543	203.758	1.225	985
Los Ríos	100.739	82.611	501	410
Los Lagos	223.697	178.954	1.090	868
Aysén	30.635	26.431	150	129
Magallanes	57.383	46.628	285	231

La cobertura nacional urbano de la II ENUT, en relación con el número total de viviendas es 82,03% y en relación con el número de UPM es de 80,67%.

Tabla 7. Porcentaje de cobertura de las 112 comunas seleccionadas sobre el total de viviendas y UPM según MMV 2020.

Región	Viviendas	UPM
Total	80,79%	80,67%
Arica y Parinacota	99,32%	99,41%
Tarapacá	94,90%	94,75%
Antofagasta	89,42%	89,39%
Atacama	80,28%	80,17%
Coquimbo	81,69%	81,19%
Valparaíso	80,84%	80,77%
Metropolitana	81,50%	81,55%
O'Higgins	80,98%	80,25%
Maule	82,62%	82,50%
Ñuble	80,09%	79,40%
Biobío	80,69%	80,79%
La Araucanía	80,68%	80,41%
Los Ríos	82,00%	81,84%
Los Lagos	80,00%	79,63%
Aysén	86,28%	86,00%
Magallanes	81,26%	81,05%

5.7. Diseño muestral

5.7.1. Dominios de estudio

Los dominios de estudio para los cuales se garantizan estimaciones con niveles de precisión adecuados incluyen: nacional desglosado por sexo y regional desglosado por sexo en el área urbana del país.

5.7.2. Estrategia muestral

El diseño muestral de ENUT 2023 implica la obtención de una muestra probabilística estratificada y bietápica. En la primera etapa, se selecciona UPM a partir de un muestreo sistemático con probabilidad proporcional al tamaño de la UPM. En la segunda etapa se selecciona la Unidad Secundaria de Muestreo (USM), la cual tiene igual probabilidad de selección al interior de cada UPM.

Al interior de cada estrato de muestreo que corresponde a la comuna, son seleccionadas las unidades de primera etapa de forma aleatoria, sistemática y con probabilidad proporcional al tamaño, según el número de viviendas que contienen.

El método de selección con probabilidad proporcional al tamaño de las UPM en encuestas de hogares es el más empleado por los países en desarrollo (Naciones Unidas, 2007). Es una técnica que utiliza información adicional para aumentar significativamente la precisión de las estimaciones de las encuestas, y, además, permite ejercer mayor control sobre el tamaño final de la muestra sin tener que incluir estratificación por tamaño de la UPM (Naciones Unidas, 2009).

En el caso de la II ENUT, la información auxiliar está definida por el tamaño de la UPM, definido por la cantidad de viviendas particulares ocupadas que contiene, por lo que las UPM con mayor cantidad de viviendas, tendrán, en general, mayor probabilidad de selección. El problema de dicho procedimiento es que las medidas verdaderas del

tamaño de las UPM casi nunca se conocen en la práctica. No obstante, muchas veces es posible obtener estimaciones válidas, como los recuentos de la población y de hogares a partir de un Censo reciente o alguna otra fuente fiable (Naciones Unidas, 2007).

Para la selección sistemática de las UPM, se emplea un ordenamiento al interior de cada comuna según el nivel socioeconómico y luego, por el número de orden geográfico de las UPM. De esta forma se introduce al diseño, una estratificación implícita por nivel socioeconómico, dando posibilidad de ser seleccionadas UPM de todos los niveles socioeconómicos, mientras que, el número de orden geográfico de las UPM permitirá que la muestra quede distribuida al interior de la comuna y nivel socioeconómico de forma balanceada.

Las unidades de segunda etapa (viviendas) son seleccionadas en forma sistemática y con igual probabilidad dentro de cada UPM seleccionada.

5.7.3. Cálculo y distribución del tamaño muestral

Los diseños muestrales complejos asociados a encuestas de hogares tienen múltiples características. Una de ellas es que la información disponible del marco muestral está sujeta a la configuración geográfica del país, hecho relevante dado que la población objetivo se encuentra en una ubicación geográfica establecida.

Para cada ubicación geográfica se pueden observar comportamientos distintos en los fenómenos de estudio y en la configuración de las unidades de primera y segunda etapa de selección, incorporando mayor complejidad al diseño muestral. Esto impacta directamente en la construcción de los factores de expansión para los distintos dominios de estudio, ya sea en las estimaciones obtenidas con la encuesta, así como en la estimación de sus errores de muestreo. Por ende, los parámetros iniciales deben considerar los elementos antes mencionados.

Para representar estas características es necesario definir un tamaño muestral que reúna la mayor parte de esta realidad, buscando que la encuesta refleje de la forma más precisa posible todos los elementos respecto a los que esta pretende dar respuesta como estudio específico. Es por este motivo que ENUT presenta un diseño muestral estratificado y bietápico.

En la II ENUT se busca estimar un parámetro desconocido de la población objetivo de modo que, esta estimación se acerque de la mejor forma al parámetro poblacional¹⁷. La distancia entre el parámetro poblacional y su estimación¹⁸ puede ser usada como una medida que cuantifica cuán lejos o cerca se está del verdadero valor. Esto se puede expresar en términos probabilísticos y tomar los márgenes de error de la estimación, como el error absoluto y relativo, con cierto nivel de confianza.

A partir de lo anterior, en esta sección se presentan los elementos considerados para dar inicio al cálculo del tamaño muestral para la II ENUT con el propósito de cumplir los objetivos propios de la encuesta, considerando factores como: el parámetro de interés, el estimador asociado a esa variable, los niveles de estimación, los errores de muestreo tolerables y disponer de una fuente de información para obtener estimaciones que servirán para dar inicio al cálculo del tamaño muestral.

Según lo anterior, fue realizado un proceso iterativo donde se consideraron distintas estrategias de muestreo (escenarios) donde se

¹⁷ Las estimaciones puntuales buscan dar el mejor valor estimado individual de un parámetro, en comparación con la estimación de intervalo, que procede de especificar un rango de valores (OCDE, 2007).

¹⁸ Esta distancia o diferencia absoluta es denominada error de estimación, error de muestreo o margen de error.

fueron ajustando los parámetros hasta obtener un tamaño muestral adecuado a los requerimientos del estudio.

5.7.4. Etapas para el cálculo del tamaño muestral para la II ENUT

El cálculo del tamaño muestral comienza con la obtención de parámetros asociados a la última aplicación de la encuesta (I ENUT) en los niveles de estimación definidos para el nuevo estudio (II ENUT). El parámetro de interés **es el tiempo medio en trabajos de cuidados no remunerados** y para efectos del cálculo del tamaño muestral se obtiene la estimación de esta variable junto con sus estadísticos asociados (error estándar y efecto del diseño).

El cálculo del tamaño muestral para la II ENUT prevé una serie de procedimientos que corresponden, en principio, al paso desde un diseño bajo muestreo aleatorio simple a un diseño complejo, seguido de otros tres ajustes, determinados por el efecto de diseño, la finitud de la población y la tasa de no logro, obteniendo así la cantidad de unidades muestrales a encuestar.

En los párrafos siguientes se describe cada una de las etapas realizadas para definir el tamaño muestral con el total de viviendas a encuestar. En principio se obtiene a nivel de región-sexo, para así determinar el tamaño muestral a nivel de región según el valor máximo entre estos y luego, a nivel de estrato de muestreo:

Etapas 1: Estimación de parámetros a nivel de región-sexo

A partir de los resultados de la ENUT 2015, se obtienen estimaciones a nivel de región-sexo, del tiempo medio en trabajos de cuidados no remunerados (t_c), junto con su error estándar y efecto del diseño.

A partir de estos parámetros se obtiene la cuasivarianza poblacional $S(t)_r^2$ según la ecuación (1).

$$S(t)_r^2 = \frac{m_r}{Def f(tc)_r} \cdot SE(t)_r^2 \quad (1)$$

Donde:

m_r : Número de viviendas logradas en el nivel de estimación r (región-sexo)

$Def f(tc)_r$: Efecto del diseño asociado a la estimación de la variable de interés en el nivel de estimación r , a nivel de vivienda

$SE(tc)$: Error estándar de la estimación de la variable de interés en el nivel de estimación r

El efecto del diseño asociado al parámetro de interés a nivel de vivienda se define en la ecuación (2).

$$Def f(tc)_r = Def f \cdot \frac{m_r}{n_r} \quad (2)$$

Donde:

$Def f$: Efecto asociado a la variable de interés a nivel de persona

m_r : Número de viviendas en el nivel de estimación r

n_r : Número de personas en el nivel de estimación r

Etapa 2: Cálculo de Tamaño bajo Muestreo Aleatorio Simple (MAS)

Se calculan los errores absolutos iniciales en cada nivel de estimación (d_{0r}) y luego, los errores relativos (e_{0r}) según la aproximación del cuantil de una distribución t de *Student* con nivel de confianza de 95% que, para efectos de la simulación, tiene un valor igual a 2. Luego, junto con la cuasivarianza, se obtiene un tamaño muestral inicial para cada nivel de estimación (m_{0r}), seleccionando el tamaño máximo entre las categorías de la variable sexo, al interior de cada región, considerando un muestreo aleatorio simple monoetápico. Esto se observa en la ecuación (3):

$$m_{0r} = \text{máx} \left(\frac{tc_{1-\frac{\alpha}{2}}^{v_a^2} \cdot S(tc)_r^2}{d_{0r}^2} \right) \quad (3)$$

Etapa 3: Ajuste por efecto del diseño

Dado que el diseño muestral definido para ENUT 2023 es un diseño complejo, el tamaño calculado bajo muestreo aleatorio simple se debe ponderar por un valor que representa el efecto provocado en la varianza al muestrear bajo un diseño complejo en lugar de un muestreo aleatorio simple. Así, el tamaño asociado al diseño complejo se obtiene desde la ecuación (4).

$$m_{1r} = m_{0r} \cdot Deff(tc)_r \quad (4)$$

Donde:

m_{0r} : Es el número de viviendas a encuestar en el nivel de estimación r bajo muestreo aleatorio simple.

m_{1r} : Es el número de viviendas a encuestar en el nivel de estimación r considerando un diseño complejo.

$Def f(tc)_r$: Es el efecto de diseño utilizado para la ENUT 2023 este valor es igual a 1.7 en todos los niveles de estimación.

Etap a 4: Corrección por población finita

El tamaño ajustado por el efecto del diseño complejo es ajustado por un ponderador que da cuenta la finitud de la población. Su forma de cálculo se observa en la ecuación (5).

$$m_{2r} = \frac{t_{1-\frac{\alpha^2}{2}}^v \cdot S(tc)_r^2 \cdot Def f(tc)_r}{d_{0r}^2 + t_{1-\frac{\alpha^2}{2}}^v \cdot Def f(tc)_r \cdot \frac{S(tc)_r^2}{M_r}} = máx \left(\frac{m_{1r}}{1 + m_{1r} + M_r} \right) \quad (5)$$

Donde:

m_{2r} : Es el total de viviendas a encuestar en el nivel de estimación r considerando un diseño complejo y asumiendo finitud de la población.

M_r : Es el total de viviendas en el nivel de estimación r , según marco de selección.

Pese a que el valor m_{2r} es calculado al interior de cada nivel de estimación, el valor que se obtiene a nivel de región es el máximo adquirido al interior de cada uno de estos y el que posteriormente se debe distribuir al interior de cada estrato de muestreo (comuna). En esta ocasión, el número de comunas incluidas en el estudio se determinó mediante un requisito de cobertura mínima del 80% en cada región, resultando en un total de 112 comunas a nivel nacional.

Etap a 5: Distribución por estrato de muestreo

Una vez determinado el tamaño muestral máximo m_{2r} en cada nivel de estimación, se distribuye este valor en forma proporcional a la cantidad de viviendas que contiene el marco de selección, en cada uno de los estratos que lo conforman. Luego, se ajusta el tamaño a una cota mínima de 30 viviendas al interior de cada estrato.

Es importante mencionar que, el tamaño inicial calculado en cada estrato de muestreo puede ajustarse en la medida que estos no cumplan con tamaños mínimos definidos, por lo que, los tamaños regionales calculados en la etapa 4 eventualmente podrían aumentar.

Etapa 6: Ajuste por no logro

Una vez definido el tamaño ajustado por la finitud del marco muestral, se debe tener en consideración que, en el trabajo de campo es posible no lograr algunas unidades por diversas razones, tales como: rechazos, moradores ausentes, entre otras. Por este motivo se hace necesario ajustar el tamaño muestral definido en la etapa 5 por un ponderador que da cuenta, a priori, de las posibles pérdidas de unidades muestrales.

Para calcular este ponderador, se utiliza información relacionada con el logro de la encuesta en levantamientos anteriores o de operaciones estadísticas de similar temática o método de recolección. El nivel de desagregación idóneo para el ponderador de ajuste es a nivel de estrato de muestreo. Obtenidos los tamaños con sobremuestreo a este nivel, los tamaños a nivel regional son calculados como la agregación de los tamaños de los estratos que las conforman.

Luego, el tamaño con sobremuestreo a nivel de estrato es calculado como se observa en la ecuación (6).

$$m_{3em} = \frac{m_{2em}}{1 - tnl_{em}} \quad (6)$$

Donde:

m_{3em} : Número total de viviendas a encuestar en el estrato de muestreo em

tnl_{em} : La tasa de no logro utilizada es el promedio a nivel comunal de las tasas de logro de la Encuesta Nacional de Empleo (ENE), en el trimestre

enero, febrero y marzo del 2022, y la IX Encuesta de Presupuestos Familiares (IX EPF).

Etapas 7: Tamaño con sobremuestreo corregido

Una vez definido el tamaño por sobremuestreo, se hace un ajuste por redondeo de la cantidad de viviendas a seleccionar dentro de cada UPM seleccionada. Esta cantidad corresponde a doce viviendas a encuestar y el redondeo se obtiene a través de la ecuación (7).

$$m'_{3r} = \text{redondear} \left(\frac{m_{3r}}{12} \right) \cdot 12 \quad (7)$$

Luego, se obtiene el tamaño ajustado por sobremuestreo a nivel de región como la suma de los tamaños con sobremuestreo a nivel de cada estrato muestral. Para el escenario escogido la muestra tiene un total nacional de 26.316 viviendas a encuestar en 2.193 UPM.

5.7.5. Resumen del proceso de simulación de tamaño muestral

Durante la etapa de simulación, previa al desarrollo de la estrategia de trabajo de campo, se estableció como hito inicial obtener un tamaño muestral óptimo para la estimación del principal parámetro de interés, con márgenes de precisión que permitiesen que la II ENUT tuviese representatividad a nivel nacional por sexo y regional por sexo. Para ello, se simularon distintos escenarios o alternativas de tamaño muestral y distribución.

Las simulaciones se realizaron considerando los parámetros del levantamiento de ENUT 2015 formulando diferentes alternativas de distribución de la muestra objetivo. Finalmente, se estableció un escenario que se ajustó a los requerimientos establecidos y consideró el presupuesto disponible para la operación.

Como base para el desarrollo de todos los escenarios trabajados se presentan los resultados del trabajo de campo de ENUT 2015, desde donde se obtiene la estimación del tiempo medio en trabajos de cuidados no remunerados junto con sus estadísticos asociados, a saber, el error estándar, el coeficiente de variación y el efecto de diseño, tanto a nivel nacional según sexo como a nivel regional según sexo.

A partir de estos resultados, fueron analizados distintos escenarios. El escenario n° 1 corresponde a un escenario base que consideró los parámetros obtenidos de ENUT 2015 y consideró la inclusión de las mismas comunas de la muestra del 2015 y que también fueron incluidas en la simulación realizada el 2020 para la segunda ENUT, que fue suspendida por la pandemia. En este escenario se consideró que cada comuna tuviese al menos 30 viviendas en la muestra. El tamaño objetivo ajustado (tamaño ajustado por finitud) fue de 15.311 viviendas a nivel nacional, y el tamaño con sobremuestreo de 26.336 viviendas.

Luego, a partir del primer escenario, se desarrolló el escenario n°2 en donde se consideró una distribución muestral con un porcentaje de inclusión del 70% de comunas para cada región y al menos un total de 30 viviendas por comuna en la muestra¹⁹. En este escenario el tamaño objetivo (tamaño ajustado por finitud) fue de 15.483 viviendas a nivel nacional, y el tamaño con sobremuestreo de 26.392 viviendas.

Finalmente, se seleccionó el escenario n° 3, que considera inclusión comunal en la distribución muestral, estableciendo abarcar al menos el

¹⁹ La elección de estos umbrales viene de la necesidad de la búsqueda de mínimos de calidad en las estimaciones. Véase el “Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares”:

<https://www.ine.gob.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>

80% de las viviendas en cada región. En este escenario el tamaño objetivo final fue de 15.490 viviendas a nivel nacional y el tamaño de sobremuestreo de 26.344 viviendas.

La Tabla 8 presenta los parámetros comunes utilizados en los escenarios considerados.

Tabla 8. Parámetros utilizados para el cálculo de tamaño muestral

Parámetro	Descripción
Parámetro asociado	Tiempo total destinado al trabajo de cuidados no remunerado a integrantes del hogar en el día tipo (tc_dt).
Estimador asociado	Media estimada del tiempo total destinado al trabajo de cuidados no remunerado a integrantes del hogar en el día tipo.
Tamaño de la población objetivo	La estimación del tamaño muestral se obtuvo a partir de un muestreo aleatorio simple en cada nivel de estimación, al cual se le aplican principalmente tres correcciones: la primera da cuenta del diseño muestral a partir de un estadígrafo denominado efecto del diseño (<i>Deff</i>); la segunda da cuenta que la población en estudio es finita; y la tercera, corrige el tamaño para compensar la falta de respuesta, pérdida usual en este tipo de estudios.
Variable de diseño	Media estimada del tiempo total destinado al trabajo de cuidados no remunerado a integrantes del hogar en el día tipo (tc_dt).
Niveles de estimación	Nacional urbano por sexo Regional urbano por sexo
Errores de muestreo	Errores absolutos de acuerdo con el estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares (INE Chile, 2020), cuya definición de aceptación se establece mediante el siguiente umbral: $\text{Máximo ee tolerable} = \begin{cases} \sqrt[3]{(r^2)/9} & ; r < 0,5 \\ \sqrt[3]{(1-r)^2}/9 & ; r \geq 0,5 \end{cases}$
Cuasi Varianza	ENUT 2015
Efecto del diseño	El efecto de diseño obtenido en ENUT 2015 ajustado por el promedio de personas en cada vivienda. Acotado a un mínimo de 1 y un máximo de 5,0 para obtener la cuasivarianza.
Tasa de sobremuestreo	Promedio de las tasas de no logro de la ENE, en el trimestre EFM 2022 y la IX EPF a nivel comunal.

La Tabla 9 resume los tres escenarios muestrales evaluados para la II ENUT donde el escenario n°3 corresponde al seleccionado.

Tabla 9. Resumen de parámetros utilizados y tamaños y errores resultantes según escenario, ENUT 2023.

N r o	Niveles de estimación	Error Absoluto	Error Relativo	Efecto Diseño	Tamaño Muestral propuesto bajo MAS	Tamaño Muestral ajustado por finitud	Tamaño con sobre-muestreo
1	Nacional	0,08	0,035	1,7	9.060	15.311	26.336
	Regional	[0,11 - 0,58]	[4,3% - 19,8%]				
2	Nacional	0,08	0,035	1,7	9.060	15.483	26.392
	Regional	[0,11 - 0,57]	[4,2% - 19,7%]				
3	Nacional	0,08	0,034	1,7	9.165	15.490	26.344
	Regional	[0,11 - 0,57]	[4,2% - 19,7%]				

Una vez seleccionado el escenario n°3 se construyeron distintos escenarios considerando distintos números de viviendas a seleccionar dentro de cada UPM. El escenario seleccionado consideró 12 viviendas a seleccionar por UPM (también se probaron 8 y 11 viviendas a seleccionar) y en dicha selección se realizó un ajuste por redondeo a la cantidad de viviendas a seleccionar al interior de cada UPM (etapa 7 en el cálculo del tamaño), por lo que el tamaño con sobremuestreo resultante fue de 26.316 viviendas considerando 2.193 UPM. En la Tabla 10 se presenta la distribución nacional y regional de la muestra con sus respectivos errores esperados (absolutos y relativos).

Tabla 10. Tamaños muestrales y errores a nivel regional y nacional.

Región	Viviendas MMV 2020	Tiempo medio en trabajos de cuidados no remunerados	Tamaño bajo MAS	Error Absoluto esperado (minutos)	Error Relativo esperado	Tamaños ajustados por el efecto del diseño	Tamaño ajustado por finitud	Tamaño con sobre-muestreo (viviendas)
Nacional	4.433.995	2,44	9.165	5,03	3,44%	15.581	15.490	26.316
Arica	66.974	2,36	420	13,15	9,29%	714	707	960
Tarapacá	99.887	3,11	475	19,57	10,50%	808	802	1.128
Antofagasta	165.191	2,43	550	11,76	8,05%	935	930	1416
Atacama	75.117	2,61	450	6,60	4,22%	765	758	1.332
Coquimbo	185.481	2,41	650	10,66	7,37%	1.105	1099	1.944
Valparaíso	498.975	2,49	730	18,90	12,67%	1.241	1.238	2.544
Metropolitana	1.871.485	2,55	1.200	18,38	12,02%	2.040	2.038	4.416
O'Higgins	200.644	2,27	520	23,52	17,31%	884	881	1.500
Maule	235.786	1,69	670	10,47	10,33%	1.139	1.134	1.872
Ñuble	100.140	1,99	415	8,47	7,09%	706	701	1.116
Biobío	395.933	2,33	720	18,27	13,09%	1.224	1.221	2.088
La Araucanía	203.758	2,15	590	22,97	17,80%	1.003	999	1.704
Los Ríos	82.611	2,91	460	34,48	19,73%	782	775	1.140
Los Lagos	178.954	2,28	525	22,36	16,38%	893	889	1.464
Aysén	26.431	2,13	390	18,13	14,19%	663	647	864
Magallanes	46.628	2,40	400	10,66	7,39%	680	671	828

5.7.6. Selección de las unidades de primera etapa

Como se ha indicado en apartados anteriores, la estrategia de muestreo de la II ENUT, sigue un diseño muestral complejo, estratificado y bietápico. En donde en la primera etapa se seleccionan unidades de primera etapa al interior de cada estrato de muestreo, de forma independiente, sistemática y con probabilidad de selección proporcional al tamaño del marco de selección. Posteriormente, se seleccionan viviendas particulares ocupadas con moradores ausentes y/o presentes (segunda etapa de muestreo) con igual probabilidad utilizando un método sistemático.

La selección de conglomerados o UPM i al interior de cada estrato de muestreo h se realiza con probabilidad proporcional al tamaño de viviendas particulares que contiene el marco muestral ajustado para la encuesta. La forma de cálculo se observa en la ecuación (8).

$$P_h(i) = \frac{n_h \cdot M_{hi}}{M_i} \quad (8)$$

Donde:

i : índice de la unidad primaria de muestreo (UPM)

h : índice del estrato de muestreo

n_h : Número de UPM seleccionadas del estrato h

M_{hi} : Número de viviendas en la UPM i del estrato h según el MMV 2020

M_h : Número de viviendas en el estrato h según el MMV2020

La selección de UPM dentro de cada estrato (comuna) es implementada con el paquete *Sampling* del *software R* que permite introducir los parámetros necesarios para seleccionar una muestra compleja. El procedimiento realiza una selección sistemática de UPM a partir de una semilla aleatoria. Previo a la selección, se ordenan las UPM según: región,

comuna, nivel socioeconómico, vector orden (variable de ordenamiento geográfico de cada UPM al interior de los estratos) y finalmente por el identificador de la UPM (ID_UPM).

5.7.7. Selección de las unidades de segunda etapa (viviendas particulares ocupadas)

Al interior de cada conglomerado, las viviendas son seleccionadas bajo un algoritmo de selección sistemático con igual probabilidad de modo que, todas las viviendas particulares ocupadas en el conglomerado, registradas en el empadronamiento, tengan la misma probabilidad de ser seleccionadas.

Luego, la probabilidad $P_{hi}(j|i)$ de incluir la j – ésima vivienda condicional a que el i – ésimo conglomerado (UPM) fue seleccionado en la muestra en el estrato h , está dada por la ecuación (9).

$$P_{hi}(j|i) = \frac{m_{hi}}{M'_{hi}} \quad (9)$$

Donde:

m_{hi} : Corresponde al número de viviendas seleccionadas en cada unidad de primera etapa i del estrato h

M'_{hi} : Corresponde al número actualizado de viviendas particulares ocupadas según el proceso de enumeración o verificación de directorios de unidades de primera etapa

Al igual que la selección de UPM, la selección de viviendas es implementada mediante el paquete *Sampling* del software R, fijando una semilla aleatoria, a fin de que dicho procedimiento pueda ser replicado en cualquier momento. Previo a la selección, se acota el directorio de viviendas según el uso de la edificación y la situación de las viviendas. El uso de la edificación debe corresponder a vivienda particular y; la situación de la vivienda, a moradores presentes o ausentes.

Para la selección de las viviendas (m_{hi}) el software ejecuta los siguientes pasos:

Paso 1

En primera instancia se ordenan geográficamente las viviendas dentro de cada UPM, en forma independiente y ascendente según el número de orden de vivienda dentro de la UPM.

Paso 2

Al interior de cada UPM se calcula el período (k) que corresponde al redondeo de $k = \frac{MM'_{hi}}{m_{hi}}$

Paso 3

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

Paso 4

Finalmente, 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: $A, A + k, A + 2K, A + 3K, \dots, A + (m_{hi} - 1)k$. La primera vivienda seleccionada es A , la segunda es el redondeo de $A + k$, la tercera es el redondeo de $A + 2k$, y así sucesivamente hasta la m_{hi} selección, dada por el redondeo de $A + (m_{hi} - 1)k$.

5.8. Factores de expansión

Como se menciona en los capítulos anteriores en este documento, el diseño muestral de la II ENUT, corresponde a un diseño probabilístico, estratificado por comunas y bietápico. La primera etapa corresponde a la selección de unidades primarias de muestreo (UPM) al interior de cada estrato muestral y, la segunda, corresponde a la selección de unidades secundarias de muestreo (viviendas) al interior de cada UPM seleccionada. Para que este diseño tenga validez en la población objetivo se debe calcular un ponderador que dé cuenta de la cantidad de personas en la población que representa cada individuo de la muestra. Este ponderador es conocido como factor de expansión. La metodología del cálculo del factor de expansión se compone de ocho etapas secuenciales que se presentan a continuación:

- i. **Ponderador de selección UPM:** corresponde al inverso de la probabilidad de selección de cada UPM al interior del estrato muestral.
- ii. **Ponderador corregido por omisión de comunas:** este ajuste es aplicado para corregir la exclusión del área urbana de las comunas que no forman parte de la muestra.
- iii. **Ponderador de selección de viviendas:** inverso de la probabilidad de selección de la vivienda condicionado a que ya fue seleccionada la UPM al interior del estrato muestral.
- iv. **Ponderador corregido por omisión de viviendas de otras muestras:** este ajuste se aplica para corregir la exclusión de las viviendas omitidas por estar seleccionadas en otras muestras.
- v. **Ajuste por elegibilidad:** viviendas clasificadas con elegibilidad desconocida durante el levantamiento se eliminan de la base de datos y sus ponderadores son distribuidos proporcionalmente en otras viviendas del mismo estrato muestral. Las viviendas

clasificadas como no elegibles son descartadas para el siguiente ajuste.

- vi. **Ajuste por no respuesta:** esta corrección tiene por objetivo reducir el sesgo asociado a la falta de respuesta de la unidad muestral, en donde las ponderaciones de las unidades que no respondieron la encuesta se distribuyen en aquellas unidades que sí respondieron la encuesta al interior de veintiles²⁰ Generados de acuerdo con la probabilidad de las unidades de responder la encuesta.
- vii. **Suavizamiento:** este procedimiento permite balancear los ponderadores con el fin de mitigar el impacto que producen valores clasificados como *outliers* permitiendo mantener el orden de la distribución inicial.
- viii. **Calibración:** en esta etapa el ponderador a nivel de vivienda es asignado a cada integrante del hogar que contestó la encuesta y se ajusta a través del método *raking* con el fin de cuadrar las estimaciones poblacionales obtenidas de la encuesta con las proyecciones de población estimadas a la fecha central del período de levantamiento.

A continuación, en cada uno de los apartados del capítulo se describe, en detalle, cada una de estas etapas.

5.8.1. Ponderador de selección de Unidades de Primera Etapa (UPM)

Este ponderador corresponde al inverso de la probabilidad de inclusión de cada UPM de la muestra al interior de cada estrato muestral. Su forma de cálculo se observa en la ecuación (10).

²⁰ Los veintiles corresponden a 20 grupos de igual tamaño generados a partir de la probabilidad de respuesta esperada resultante de un modelo de no respuesta.

$$w_{hi} = \frac{1}{P_h(i)} = \frac{M_h}{n_h \cdot M_{hi}} \quad (10)$$

Donde:

h : índice del estrato muestral correspondiente a la comuna.

i : índice del conglomerado (UPM).

M_h : Total de viviendas en el estrato h según el MMV 2020

n_h : Total de unidades de primera etapa seleccionadas al interior de cada estrato h

M_{hi} : Total de viviendas en la UPM i del estrato h

5.8.2. Ajuste por omisión de comunas

Este ajuste es aplicado para corregir la exclusión de ciertas comunas del área urbana que se realiza en el proceso de distribución de la muestra. Consiste en ponderar el inverso de la probabilidad de selección de las UPM con la razón \ddot{R}_r obtenida entre el total de viviendas del MMV 2020 y la suma total de viviendas ponderadas por el factor de selección de UPM al interior de cada región r . Esto se puede observar en la ecuación (11)

$$w'_{hi} = w_{hi} \cdot \ddot{R}_r = w_{hi} \cdot \frac{M_r}{\sum_{h \in \varphi_r} \sum_{i \in \theta_h} M_{hi} w_{hi}} \quad (11)$$

Donde:

φ_r : Listado de comunas seleccionadas en la región r

M_r : Total de viviendas en la región r según MMV 2020

5.8.3. Ponderador de selección de viviendas

Las viviendas son seleccionadas al interior de cada conglomerado bajo un algoritmo de selección sistemático con igual probabilidad de modo que, todas las viviendas particulares ocupadas en el conglomerado

tienen la misma probabilidad de ser seleccionadas. Luego, la probabilidad de incluir la j –ésima vivienda, condicional a que el i –ésimo conglomerado ya fue seleccionado en el estrato h , está dada por la ecuación (12).

$$P_h(j|i) = \frac{m_{hi}}{M'_{hi}} \quad (12)$$

Donde:

M'_{hi} : Número de viviendas particulares ocupadas según el procedo de enumeración o verificación de directorios de la UPM i seleccionada en el estrato h .

m_{hi} : Total de viviendas seleccionadas en cada UPM i del estrato h .

Como se indica en el apartado sobre la selección de viviendas para la II ENUT, en la selección de UPM se incluyeron unidades que ya habían sido seleccionadas en otras encuestas encargadas al INE. Por este motivo, previo a la selección de viviendas, es necesario realizar un ajuste que le otorgue probabilidad igual a 0 a aquellas que ya han sido seleccionadas en otras encuestas. Este ajuste afecta las probabilidades originales de selección de las viviendas reduciendo su stock. Por lo anterior, para conservar el stock original de viviendas se estima una razón entre la cantidad total de viviendas elegibles desde el directorio original y la cantidad de viviendas disponibles de dicho directorio, es decir, con todas las viviendas que contienen probabilidad de selección mayor que cero. Esta razón se observa en la ecuación (13).

$$\hat{R}_i = \frac{M'_{hi}}{\text{Total viviendas en la UPM}} \quad (13)$$

La probabilidad de selección es ponderada por esta razón, como se observa en la ecuación (14).

$$P_h(j|i)_{aj} = P_h(j|i) \cdot \hat{R}_i \quad (14)$$

Finalmente, la probabilidad condicional de selección de la vivienda j ubicada en el conglomerado i del estrato h se obtiene desde la ecuación (15).

$$P_{hi}(j) = P_h(i) \cdot P_h(j|i)_{aj} \quad (15)$$

Donde:

$P_h(j|i)_{aj}$: Probabilidad condicional ajustada al seleccionar la vivienda j del conglomerado i del estrato h .

$P_h(i)$: Probabilidad de selección del conglomerado i dentro del estrato h

Luego, el ponderador de selección de viviendas es calculado como el inverso de la probabilidad de incluir cada vivienda particular ocupada dentro de la UPM seleccionada en la muestra. Su forma de cálculo se observa en la ecuación (16).

$$w_{hij} = \frac{1}{P_h(i) \cdot P_{hi}(j|i)_{aj}} = \frac{1}{P_h(i)} \cdot \frac{1}{P_h(j|i) \cdot \hat{R}_i} = w_{hi} \cdot \frac{1}{P_{hi}(j|i)_{aj}} \quad (16)$$

Donde:

w_{hij} : Ponderador de selección de la vivienda j ubicada en el conglomerado i del estrato h

$P_h(j|i)_{aj}$: Probabilidad condicional ajustada de selección de la vivienda j del conglomerado i en el estrato h .

w_{hi} : Ponderador de selección del conglomerado i del estrato h .

$P_h(i)$: Probabilidad de selección del conglomerado i del estrato h .

\hat{R}_i : Razón de ajuste por omisión de viviendas.

Lo anterior también se puede escribir según la ecuación (17).

$$w_{hij} = \frac{M_h}{n_h \cdot M_{hi}} = \hat{R}_i \cdot \frac{M'_{hi}}{m_{hi}} \quad (17)$$

Este ponderador refleja cuantas viviendas en la población representa cada vivienda seleccionada en la muestra de la II ENUT al interior de cada conglomerado, teniendo así, la misma ponderación para cada vivienda de dicho conglomerado.

5.8.4. Ponderador corregido por elegibilidad

Este ponderador busca corregir aquellas actualizaciones de estado de las viviendas que, luego del trabajo de campo, son definidas como no elegibles al no tratarse de viviendas particulares con moradores presentes o ausentes o bien, son definidas como viviendas con elegibilidad desconocida, ya que su clasificación se desconoce. A pesar de todos los esfuerzos que se realizan por mantener marcos muestrales actualizados y validados para cubrir la población objetivo de este estudio, siempre existe la posibilidad de encontrar errores de inclusión (o de exclusión) al momento de utilizar el marco muestral.

Algunos ejemplos de estos errores pueden ser: fallas en el proceso de enumeración y/o verificación, que pueden derivar en el registro de edificaciones no elegibles (ejemplo: casas de veraneo, negocios, etc.); cambios en el uso de la edificación entre el momento de enumeración y la apertura de la entrevista, entre otros.

Al término del trabajo de campo, cuando se cierra el proceso de recolección de la información, todas las viviendas seleccionadas adquieren un código de disposición final (CDF) el que las clasifica en tres grandes grupos, éstos son:

Viviendas elegibles: corresponden a edificaciones en que el encuestador pudo determinar que se trataban de viviendas particulares ocupadas (incluye entrevistas completas, entrevistas parciales y no entrevistadas). El detalle de cada uno de los códigos se presenta en la Tabla 11.

Tabla 11. Estados de elegibilidad de la vivienda según código de disposición final

Estado de elegibilidad	Descripción	Código disposición final
Elegibles entrevistadas	Cuestionario CH con entrevista completa o parcial y todos los CUT elegibles con entrevista completa	11
	Cuestionario CH con entrevista completa o parcial y al menos un CUT elegible con entrevista completa o parcial	12
Elegibles rechazadas	Cuestionario CH con entrevista completa o parcial y ningún CUT logrado, o cuestionario CH con rechazo del hogar	21
Elegibles no contactadas	Vivienda ocupada, sin moradores presentes.	31
	Se impide el acceso a la vivienda	32
	Informante no ubicable, se contacta a otra persona	33
	Informante ausente	34
Elegibles no entrevistadas por otro motivo	Informante impedido(a) física y/o mentalmente para contestar	41
	Informante no habla idioma español	42
	Muerte del informante o algún familiar	43
	Concertó cita con informante	44
	Falseamiento del cuestionario CH	45
	Otra razón de no entrevista	49

Viviendas no elegibles: son edificaciones identificadas como negocios, viviendas colectivas, viviendas deshabitadas, viviendas de veraneo,

viviendas destruidas, etc. En síntesis, edificaciones que no pertenecen al universo de estudio. La Tabla 12 presenta el detalle de los CDF para viviendas no elegibles.

Tabla 12. Estados de no elegibilidad según código de disposición final

Estado de elegibilidad	Descripción	Código de disposición final
No elegibles	Establecimiento, oficina de gobierno u otra organización.	61
	Institución (por ejemplo: hospital, cárcel, asilo de ancianos, etc.)	62
	Dormitorio colectivo (por ejemplo: militar, de trabajo, internado, hostales, residenciales, etc.)	63
	Vivienda en demolición, incendiada, destruida o erradicada.	64
	Vivienda desocupada (por ejemplo: en venta, arriendo, etc.)	65
	Vivienda de veraneo o de uso temporal	66
	Edificación no calificable (por ejemplo: en construcción, en reparaciones, etc.)	67
	No existe informante elegible en el hogar	68
	Otra razón para ser considerada no elegible	69

Viviendas con elegibilidad desconocida: corresponden a edificaciones en que no se pudo determinar su estado. Este es el caso, por ejemplo, de unidades que nunca fueron enviadas a terreno, viviendas a las cuales no se pudo llegar o encontrar, y otros casos similares. El detalle de los CDF para las viviendas de elegibilidad desconocida se observa en la Tabla 13.

Tabla 13. Estados de elegibilidad desconocida según código de disposición final.

Estado de elegibilidad	Descripción	Código disposición final
Elegibilidad desconocida	No fue posible localizar la dirección	52
	Área peligrosa	54
	Otras razones de elegibilidad desconocida	59

La Tabla 14 muestra la clasificación de las viviendas según su estado de elegibilidad obtenida luego del trabajo de campo de la II ENUT:

Tabla 14. Distribución de viviendas seleccionadas según estado de elegibilidad

Estado de elegibilidad	Frecuencia	Porcentaje
Total	26.316	100%
Elegible	24.323	92,43%
Elegibilidad desconocida	189	0,72%
No elegible	1.804	6,85%

El ponderador de viviendas w_{hij} tiene valores válidos para todas las viviendas seleccionadas en la muestra, sin embargo, se requiere ajustar este ponderador por aquellas unidades que no se conoce su elegibilidad y luego por aquellas unidades clasificadas como no elegibles.

- **Ajuste por elegibilidad desconocida**

El procedimiento de ajuste por elegibilidad desconocida consiste en que los pesos de las edificaciones, consideradas para la selección, como viviendas con elegibilidad conocida (elegibles y no elegibles) que, luego del trabajo de campo, no ha sido posible identificar su estado, sea redistribuido proporcionalmente entre las viviendas con elegibilidad conocida, al interior de cada estrato de muestreo.

El ponderador de los pesos a redistribuir se calcula como la razón entre dos estimaciones obtenidas con base en el factor teórico (w_{hij}). En el numerador se ubica la estimación del total de viviendas de la población y en el denominador la estimación total de las viviendas con elegibilidad conocida. Esto se observa en la ecuación (18).

$$R_{h,known} = \frac{\sum_{i \in \Omega_h} \sum_{j \in \Theta_i} w_{hij}}{\sum_h \sum_{j \in \Theta_{i,known}} w_{hij}} \quad (18)$$

El factor que da cuenta de la selección de viviendas, ajustado por elegibilidad desconocida, w'_{hij} , está dado por la ecuación (19).

$$w'_{hij} = w_{hij} \cdot R_{h,known} \quad (19)$$

Finalmente, el factor de expansión asignado a las viviendas clasificadas con elegibilidad desconocida es eliminado y la suma total del factor de expansión asignado a las viviendas clasificadas con elegibilidad conocida es equivalente a la suma del factor de expansión inicial.

- **Ajuste por no elegibilidad**

Se observa que, de un total de 26.316 viviendas seleccionadas para la II ENUT, 1.804 resultaron no elegibles. A partir de este momento estas viviendas son descartadas de la muestra pues, no forman parte de la población objetivo de este estudio.

Esto se implementa asignando a las viviendas no elegibles un valor blanco (*missing value*) en el factor ajustado por elegibilidad desconocida (w'_{hij}) reduciendo los stocks de viviendas.

5.8.5. Ajuste por no respuesta

En esta etapa se consideran solo las viviendas particulares ocupadas que corresponden a viviendas elegibles, sin embargo, es posible que no se

logre conseguir respuesta en la totalidad de ellas, ya sea porque los residentes no desean participar en el estudio, o bien, porque no han estado presentes al momento de la entrevista. En estos casos, es necesario realizar un ajuste adicional al factor, considerando distribuir el peso de las viviendas donde no se logró respuesta en aquellas donde sí se logró.

Una vivienda elegible se divide entre entrevista completa, entrevista parcial y no entrevistada. Las entrevistadas parciales corresponden a aquellas viviendas en donde se logró recabar parte de la información objeto de la encuesta y esta cumple con la suficiencia y calidad mínima esperadas. Las no entrevistadas son viviendas donde existe ausencia completa de información, es decir, cuando el informante rechazó la entrevista desde el inicio o bien la vivienda se mantuvo sin moradores presentes durante todos los intentos de contacto para realizar la entrevista.

El ajuste de no respuesta aplicado al factor de expansión se realiza considerando información proveniente del levantamiento de la encuesta, y tiene como objetivo corregir las posibles fallas cuando no se logra la muestra completa de viviendas elegibles.

Para minimizar los problemas generados por la ausencia de respuesta de la unidad de muestreo, se procede a realizar un ajuste para compensar dichas pérdidas a través de la implementación del método *propensity score*. De acuerdo, a lo indicado por Valliant et al. (2013) este método consiste en modelar la probabilidad de respuesta en una encuesta como la realización de un proceso de variables latentes ($R_i^* = x_i^T \beta + u_i$) es decir, un conjunto de covariables que inciden en la “motivación” (R^*) de participar o no en la encuesta. Así, mediante un conjunto de covariables conocidas para quienes responden y quienes no responden, se busca estimar la probabilidad de responder en la encuesta ($P(R_i^* > \theta)$) y

posteriormente, se ordenan estas probabilidades subdividiéndolas en 20 partes iguales llamadas veintiles.

A partir de un set de variables que podrían contribuir a predecir la respuesta de la unidad, se realiza un primer filtro utilizando el método de bosques aleatorios, popularmente conocidos como *Random Forest* (Breiman, 2001). Con este método, se conforman conjuntos de árboles de decisión individuales, en que cada uno se prueba extrayendo una muestra aleatoria desde los datos de levantamiento mediante *bootstrapping* (o remuestreo). Esto implica que cada árbol se prueba con conjuntos de datos ligeramente distintos. En cada árbol individual, las observaciones se van distribuyendo según distintas bifurcaciones (nodos) generando la estructura del árbol hasta alcanzar un nodo terminal. La predicción de una nueva observación se obtiene agregando las predicciones de todos los árboles individuales que forman el modelo. El resultado de la aplicación de este método es un set de covariables con el que se conforman modelos *logit* con diferentes combinaciones. Las covariables que son incluidas en el modelo, por un lado, consideran el estado y entorno de la vivienda y por otro, caracterizan al encuestador en relación con la posibilidad de lograr la entrevista.

Las variables (covariables) que serán testeadas en el modelo de *Random Forest* se presentan en la Tabla 15.

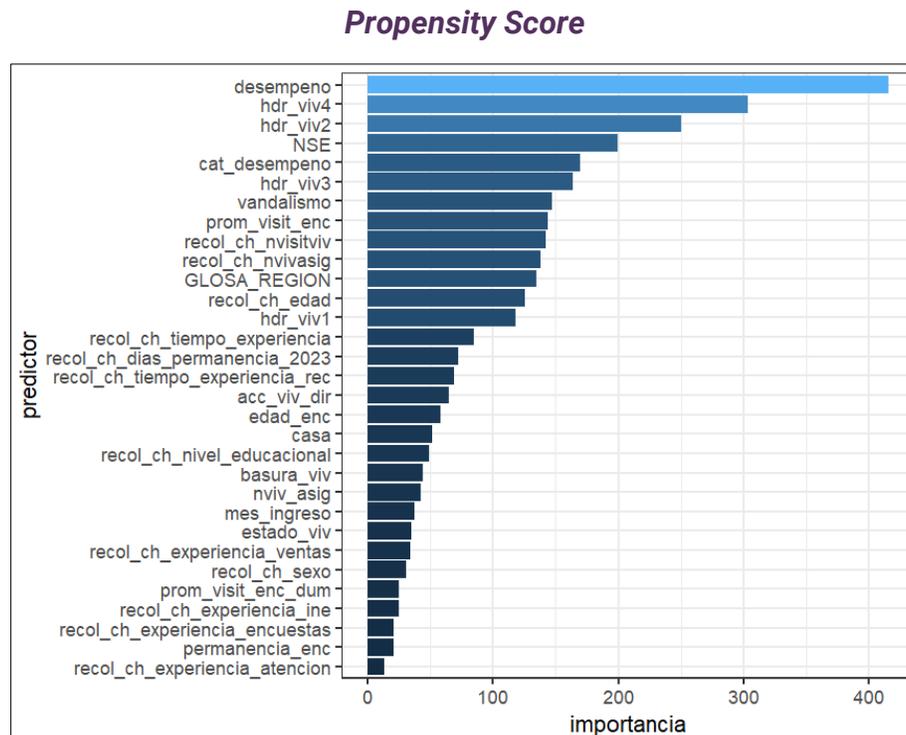
Tabla 15. Variables de la vivienda y de la persona encuestadora incluidas en el modelo

Variables de la vivienda	Variables de la persona encuestadora
Tipo de acceso a la vivienda	Sexo
Estado de la vivienda	Edad
Basura en el entorno	Nivel educacional
Vandalismo y rayados en el entorno	Tiempo de experiencia en encuestas
Tipo de vivienda	Experiencia previa en encuestas (dicotómica)
Nivel socioeconómico de la UPM	Experiencia previa en ventas
Región	Experiencia previa en atención de usuarios
	Mes de ingreso al proyecto ENUT
	N° de viviendas asignadas
	N° de viviendas logradas
	Promedio de visitas realizadas a la vivienda
	Desempeño (N° de viviendas logradas / N° de viviendas asignadas)

Así, a través del método de *Random Forest*, se observa la importancia²¹ que tienen las variables para predecir la no respuesta lo que constituye un primer filtro sobre su inclusión en el modelo *Propensity Score*. La importancia de las variables evaluadas se presenta en la Figura 4.

²¹ Es importante tener en cuenta que el *Random Forest* tiende a seleccionar variables con alta cardinalidad en la parte superior, dándole más importancia que a otras, en función de la ganancia de información. En la práctica, es una prueba de que está sobre ajustado.

Figura 4. Importancia de los predictores de no respuesta para modelo



Este gráfico permite determinar las variables que se incorporarán en los diferentes modelos *logit* con el fin de obtener el mejor modelo dentro de los propuestos, clasificándolos según el puntaje asignado a cada uno de ellos según los criterios de AIC, BIC, R^{22} . El modelo seleccionado considera las variables presentadas en la Tabla 16.

²² AIC y BIC: Criterios de información de Akaike y Schwarz. Medidas de entropía relativa que consideran la bondad de ajuste y penalizan la complejidad del modelo. Un modelo con menor AIC o BIC utiliza de mejor forma la información, mientras que R^2 corresponde a una aproximación de la medida de modelos lineales sobre el porcentaje de la varianza explicada por el modelo.

Tabla 16. Variables seleccionadas para el modelo *Propensity Score*

Variable
Región
Nivel Socioeconómico de la UPM
Vandalismo
Desempeño de la persona encuestadora
Edad de la persona encuestadora

Una vez seleccionadas las variables, se calcula la probabilidad de responder de cada una de las viviendas de la encuesta y se ordenan de menor a mayor para luego agrupar las viviendas en 20 celdas de ajuste (veintiles) de cantidades equivalentes. A continuación, se calcula el factor de expansión ajustado a aquellas viviendas que respondieron la encuesta, esto se observa en la ecuación (20).

$$w_{hij}^R = \hat{R}_v \cdot w'_{hij} \quad (20)$$

Para obtener \hat{R}_v , en el numerador se incluye la estimación del total de viviendas elegibles, y en el denominador la estimación del total de viviendas elegibles que responde. Esto se puede ver en la ecuación (21).

$$\hat{R}_v = \frac{\sum_{j \in \Phi_{eleg,v}} w'_{hij}}{\sum_{j \in \Phi_{R,v}} w'_{hij}} \quad (21)$$

Donde:

v : Índice del grupo para la corrección de no respuesta (veintil).

$\theta_{eleg,v}$: Conjunto de viviendas pertenecientes al veintil v y catalogadas como elegibles.

$\theta_{R,v}$: Conjunto de viviendas pertenecientes al veintil v , catalogadas como elegibles y que responden la encuesta.

w'_{hij} : Ponderador de selección de viviendas corregido por elegibilidad.

5.8.6. Suavizamiento del ponderador

Luego de la construcción del ponderador ajustado por no respuesta se utiliza un método denominado suavizamiento, cuya finalidad es aminorar el efecto que pudiesen ocasionar factores con valores atípicos respecto de la media (datos outliers) que podrían afectar las estimaciones de los parámetros de interés, al otorgar pesos demasiado altos (o bajos) a ciertas unidades en forma determinística.

El suavizamiento realizado al factor de expansión ajustado por no respuesta para la II ENUT, corresponde a un método mixto (MM) conformado por el método de contracción a la media (CM) y el método R-K.

- **Método de contracción a la media (CM).**

Este método contrae los factores de expansión, acercándolos a su valor promedio, haciendo que la distribución de los factores sea más homogénea. El factor suavizado o contraído a la media (o promedio) viene dado por la ecuación (22):

$$w_i^s = \beta \cdot w_i + (1 - \beta) \cdot \bar{w}; 0 \leq \beta \leq 1 \quad (22)$$

Donde β es un parámetro real que permite controlar la intensidad de la contracción, w_i es el i – *esimo* factor y \bar{w} el promedio de los factores a nivel de estrato basado en la existencia de valores atípicos o no.

- **Método r-k**

Este método se caracteriza por dar tratamiento diferenciado a los conjuntos de datos atípicos y no atípicos, generando un aumento proporcional en los factores no atípicos, mientras que, origina una disminución proporcional de los factores atípicos. Por lo que, este método, disminuye en una proporción r los valores de los factores de expansión atípicos y, por otro lado, aumenta en una proporción k los valores de los factores de expansión no atípicos, según muestra la Figura 5.

Figura 5. Factores de expansión w_i ordenados de mayor a menor

<u>Factores</u>		<u>Aumento o disminución</u>	
W_1	} Atípicos	$W_1 \cdot (1 - r)$	}
W_2		$W_2 \cdot (1 - r)$	
:		:	
:		:	
W_n		$W_n \cdot (1 - r)$	
w_1	} No atípicos	$w_1 \cdot (1 + k)$	}
w_2		$w_2 \cdot (1 + k)$	
:		:	
:		:	
w_m		$w_m \cdot (1 + k)$	

$$t_1 = \left(\sum_{i=1}^n W_i \right) - \left(\sum_{i=1}^n W_i \right) \cdot r$$

$$t_2 = \left(\sum_{j=1}^m w_j \right) + \left(\sum_{j=1}^m w_j \right) \cdot k$$

- **Método mixto**

Para aplicar el método mixto se separan los factores de expansión en dos subconjuntos (factores no atípicos y factores atípicos). Para el conjunto de los factores no atípicos se aplica el método de contracción a la media considerando β relativamente cerca de la unidad ($\beta = 0,8$). Para el conjunto de los factores atípicos, se considera $\beta = 0,2$ para que, en caso de tener un solo valor atípico al interior de un estrato, el método CM no tenga efecto sobre este valor, conservando el valor del factor atípico. Posterior a este cálculo, se aplica el método R-k, en donde se juntan ambas distribuciones de factores (no atípicos y atípicos)

considerando valores del multiplicador cercano a la unidad ($\lambda = 0,99$) para conservar una pequeña distancia entre el mayor de los factores no atípicos, y el menor de los factores atípicos, después del suavizamiento.

La Tabla 17 muestra los estadísticos asociados (recuento de viviendas, suma total y valores mínimos y máximos) para cada una de las etapas de ajuste al factor de expansión.

Tabla 17. Estadísticos asociados a cada etapa de construcción del factor de expansión

Etapas del Factor de expansión ENUT	Suma	Recuento	Mín.	Máx.
1. Teórico (Wh_{ij})	5.146.015	26.316	4,16	2.082,30
2. Ajustado por elegibilidad desconocida ($W'h_{ij}$)	5.146.015	26.127	4,19	2.132,17
3. Ajustado por no elegibilidad ($W'h_{ij}$)	4.838.205	24.323	4,19	2.132,17
4. Ajustado por no respuesta	4.838.205	16.169	7,73	3.602,70
5. Ajustado por no respuesta suavizado (W_{is})	4.838.205	16.169	34,62	2.193,18
6. Ajustado por no respuesta suavizado (W_{is_rk})	4.838.205	16.169	36,73	1.908,41

5.8.7. Calibración de los factores de expansión

En las etapas previas de este documento se consignó un factor de expansión referido a la probabilidad de seleccionar una vivienda al interior de cada estrato. No obstante, a razón de contar con estimaciones de la población actualizadas, se utilizan *stocks* poblacionales provenientes de fuentes externas, como lo son las Proyecciones de Población base censo 2017²³ referidas a la fecha correspondiente a la mitad del trabajo de campo de la encuesta que, en este caso, es el 15 de noviembre de 2023.

La calibración es un ajuste que se realiza a los factores de expansión a nivel de personas. Para esto, se utiliza el método *Raking* que permite controlar por diversas variables simultáneamente y mantener el diseño muestral original de la encuesta.

Para la II ENUT se calibran dos factores de expansión a partir del factor de expansión suavizado. El primero de ellos se asocia a todos los integrantes de los hogares que respondieron el cuestionario hogar (factor de expansión CH) y, el segundo, solo se encuentra disponible para las personas de 12 años o más que respondieron el cuestionario de uso de tiempo (factor de expansión CUT).

- **Calibración de factor de expansión cuestionario hogar (CH)**

Para obtener un factor de expansión para cada una de las personas que pertenece a las viviendas que participan en la II ENUT, se asigna a cada una de ellas el factor de expansión suavizado correspondiente a la vivienda a la que pertenece. La suma de dicho factor a nivel nacional alcanza 14.517.142 personas, cifra estimada antes de realizar el proceso de calibración que, luego de este procedimiento, debe coincidir con el

²³ <https://www.ine.gob.cl/estadisticas/sociales/demografia-y-vitales/proyecciones-de-poblacion>.

stock de personas totales provenientes de las proyecciones de población definidas para el período de levantamiento de la encuesta.

La información auxiliar se utiliza como un vector de totales poblacionales, llamados marginales de calibración, una vez suavizado el factor, se procede a revisar los coeficientes de variación de las marginales a utilizar, donde se espera obtener valores menores a 5% para todas sus categorías. Seguido de esta revisión se aplican las siguientes marginales de calibración:

- Población total por región y sexo
- Población total por región y tramo etario ([0-11], [12-14], [15-24], [25-34], [35-44], [45-54], [55-64], [65+]).
- Población total por sexo y tramo etario ([0-11], [12-14], [15-24], [25-34], [35-44], [45-54], [55-64], [65+]).

Una vez calibrados los factores se obtiene una estimación de 17.749.136 personas a nivel nacional, cifra coincidente con las Proyecciones de Población base Censo 2017 al 15 de noviembre de 2023. Cabe destacar que, para las futuras estimaciones a nivel de hogar, es necesario considerar un factor de expansión por cada hogar, mientras que, para realizar las estimaciones a nivel de personas, es necesario utilizar todas las personas con su factor de expansión asignado. También es importante señalar que, tras los resultados obtenidos a partir de la recolección de la muestra de la II ENUT, las estimaciones son representativas a nivel nacional urbano por sexo y regional urbano por sexo.

- **Calibración factor de expansión cuestionario de Uso del Tiempo (CUT)**

Por diversos motivos, algunas de las personas elegibles, esto es, de 12 o más años con capacidades físicas y mentales aptas para responder la

encuesta, no respondieron el cuestionario de Uso del Tiempo. Por lo tanto, al realizar un análisis estadístico de las variables del cuestionario de Uso del Tiempo (CUT) con el ponderador calibrado previamente, existirá una discrepancia entre el total de personas estimado a partir de las variables sociodemográficas y el universo de personas estimado que responde las preguntas del cuestionario de Uso de Tiempo.

Para subsanar esta diferencia, entre el número de personas que responde la información sociodemográfica y el CUT, se calcula un nuevo factor de expansión solo para aquellas personas de 12 años o más que respondieron el cuestionario de uso de tiempo. La calibración de este factor sigue la metodología antes expuesta y se incorpora una nueva marginal, considerando que la muestra asignada a cada par de días²⁴ debe alcanzar el mismo *stock* (como se observa en la Tabla 18). De esta manera, la calibración para el factor de expansión CUT permite cuadrar a los *stocks* poblacionales según las siguientes marginales:

- Población total por región y sexo
- Población total por región y tramo etario ([12-24], [25-34], [35-44], [45-54], [55-64], [65+]).
- Población total por sexo y tramo etario ([12-24], [25-34], [35-44], [45-54], [55-64], [65+]).
- Población total por región y par de días

²⁴ Cuando se indica par de días, nos referimos a las combinatorias posibles entre días de semana y fin de semana, que pueden ser 10 (entre 7 de semana y 2 de fin de semana).

Tabla 18. Proyecciones de población regional, para cada par de días

Región	Población para cada par de días
Nacional	1.512.631
Arica y Parinacota	20.212
Tarapacá	31.427
Antofagasta	57.947
Atacama	24.308
Coquimbo	60.580
Valparaíso	159.511
Metropolitana	690.267
O'Higgins	65.535
Maule	73.885
Ñuble	31.744
Biobío	128.559
Araucanía	62.953
Los Ríos	25.826
Los Lagos	57.930
Aysén	7.415
Magallanes	14.532

Una vez calibrados los factores se obtiene una estimación de 15.126.310 personas mayores de 12 años que responden el cuestionario del uso del tiempo.

5.9. Estimación de la varianza

Es cada vez más intensivo el uso de programas estadísticos (tales como SAS, SPSS, Stata y otros como R), que incorporan módulos en los paquetes estadísticos para el cálculo de estimaciones, varianzas y coeficientes de variación, incorporando los típicos efectos de diseño cuando se trata de muestreo complejo, por lo general polietápico.

5.9.1. Métodos convencionales para estimación de la varianza

Existen principalmente cuatro métodos convencionales para estimar las varianzas o errores muestrales para estimaciones basadas en datos de una encuesta²⁵.

- Métodos exactos.
- Métodos del conglomerado último.
- Aproximaciones por linealización.
- Técnicas de replicación.

En general se utiliza una combinación de métodos para el cálculo tanto de estimaciones como de la varianza de ellas. A continuación, se describe brevemente en qué consisten estos métodos.

i. Métodos exactos

Los métodos exactos de estimación de la varianza en diseños de muestras estándares, cuando son aplicables, son la mejor forma de estimar la varianza. Sin embargo, los diseños muestrales empleados en la mayoría de las encuestas de hogares son más complejos que el muestreo aleatorio simple o estratificado. Adicionalmente, debido a que

²⁵ Para más información ver Kish y Frankel (1974), Wolter (1985) y Lehtonen y Pahkinen (1995).

las estimaciones de interés puede que no sean funciones lineales simples de los valores observados, la varianza de muestreo no siempre puede expresarse por una fórmula de forma cerrada como la media de la muestra en el muestreo aleatorio simple o el muestreo estratificado.

ii. Métodos del conglomerado último o final

El método del conglomerado último para estimar la varianza (Hansen, Hurwitz y Madow, 1953) se emplea para calcular las estimaciones basadas en una muestra obtenida a partir de un diseño muestral complejo, generalmente polietápico o polifásico. Según este método, el conglomerado último consiste en considerar a la unidad primaria de muestreo UPM como la unidad de información última que conglojera la información de todas las unidades de muestreo de etapas posteriores dentro de ella, como por ejemplo, si la unidad primaria de muestreo corresponde a un conglomerado que contiene unidades secundarias de muestreo como las viviendas y, a su vez, estas unidades secundarias de muestreo contienen unidades terciarias como las personas, y se está interesado en alguna variable de interés como la ocupación entonces, cada UPM tendrá asociado el total de ocupados, como la sumatoria sobre todas las personas de las viviendas contenidas en la UPM. Las estimaciones de la varianza se calculan utilizando solo totales entre UPM, sin tener que calcular los componentes de la varianza en cada etapa de selección.

iii. Método de linealización

La linealización es ampliamente utilizada en la práctica porque puede aplicarse a casi todos los diseños muestrales y a cualquier estadística que pueda linealizarse o, lo que es lo mismo, expresarse como una función lineal de las estadísticas comunes como medias o totales, cuyos coeficientes se extraen de derivadas parciales necesarias para la

expansión en serie de Taylor. Se aplica la linealización generalmente para la estimación de razones o cocientes de dos variables las que se desarrollan en una serie de Taylor infinita centrada en el valor esperado (estimado) del numerador y el valor esperado (estimado) del denominador. Seguidamente, el estimador no lineal se somete a aproximación algebraica conservando solo los primeros términos de la serie infinita de Taylor, lo que produce una expresión algebraica que ya no es una función no lineal, sino una función lineal de los datos muestrales; esto es, el estimador no lineal del cociente se ha "linealizado". Ahora es posible obtener directamente la varianza estimada de la función linealizada (incluidos los términos de covarianza pertinentes). Una vez linealizada, la varianza de la estimación no lineal puede aproximarse mediante los métodos exactos descritos anteriormente (Cochran, 1997), (Lohr, 1999), para consultar información técnica detallada sobre el proceso de linealización ilustrada con ejemplos).

iv. Método de replicación

Este método consiste en tomar submuestras o réplicas de la muestra total. Se estima con la muestra total la variable de interés Y , de acuerdo con el diseño probabilístico, sea cual fuere. Enseguida se generan réplicas o se seleccionan submuestras de la muestra total, de manera que cada réplica refleje el plan de muestreo y los ajustes y procedimientos de ponderación de la muestra completa, de modo que, con cada una de ellas, se llegue a la misma estimación que con la muestra completa. Existen varias técnicas para hacerlo y en algunos *softwares* ya vienen incluidos los módulos de selección de réplicas, entre los cuales, los más utilizados son los grupos aleatorios; la replicación repetida equilibrada (BRR); la replicación *Jackknife* (JK1, JK2, y JK n) y *Bootstrap*.

Por ejemplo, suponiendo que se crean K réplicas a partir de una muestra, cada una con estimaciones $\widehat{\theta}_1, \widehat{\theta}_2, \dots, \widehat{\theta}_k$ de un parámetro $\hat{\theta}$, y supongamos que la estimación basada en la muestra completa es $\hat{\theta}_0$. La estimación de la varianza basada en la replicación viene dada por: $Var(\hat{\theta}) = \frac{1}{c} \cdot \sum_{r=1}^k (\widehat{\theta}_r - \widehat{\theta}_0)^2$, donde c es una constante que depende del método de estimación.

Estos métodos de estimación de varianzas son antiguos y pocos softwares los incorporaban o desarrollaban. Sin embargo, con el avance tecnológico y de la computación en particular, hoy en día, la mayoría de los softwares estadísticos traen incorporados módulos especiales para el tratamiento de los datos enfocados a muestras complejas, entendiendo por muestra compleja, a aquella que incorpora variables de ponderación, de estratificación y de conglomeración para las estimaciones de los principales estadísticos asociados a alguna variable de interés y también la estimación de sus varianzas.

5.9.2. Varianza con muestras complejas

Una aproximación común (WR: *With Replacement*) para la descripción de muchos planes de muestreo complejo, normalmente es utilizada en las encuestas complejas de muestreo polietápico por conglomerados. Además, es habitual la estratificación de las UPM de la población antes del muestreo en la primera etapa. Por otra parte, la estratificación de las unidades de la segunda y posteriores etapas (en una UPM seleccionada) puede tener lugar antes del muestreo en estas etapas. Sin embargo, con los métodos aproximados de estimación de varianzas, comúnmente utilizados para estos diseños complejos, no es necesario tomar en consideración todas las etapas de muestreo y estratificación. El muestreo complejo en las etapas posteriores queda automáticamente cubierto de manera adecuada con la aproximación “con reemplazo” para la primera etapa de muestreo que hemos visto en capítulos anteriores.

De hecho, pocos paquetes de software tienen la capacidad de incluir por separado todas las etapas de muestreo en la estimación de varianzas, en los casos en los que no se practica la aproximación con reemplazo en la primera etapa.

La estimación de varianzas por conglomerados finales (EVCF) puede efectuarse indistintamente por linealización mediante el desarrollo en serie de Taylor o con una técnica de replicación. Con este enfoque, las UPM se tratan como si se hubieran seleccionado con reemplazo entre los estratos de la primera etapa. En tal caso, cada elemento o individuo solo tiene que identificarse por el estrato de la primera etapa y por la UPM (dentro del estrato) de la que se haya seleccionado. Para efectos de la estimación de varianzas no se necesita información sobre las etapas de muestreo por debajo del nivel de las UPM. De este modo, la descripción del plan de muestreo se simplifica de forma que se asemeja a un muestreo por conglomerados estratificado en una etapa, esto es, una muestra estratificada de conglomerados finales completamente enumerados. Este enfoque por conglomerados finales produce una buena aproximación para la estimación de la varianza, siempre que sea razonable partir de la hipótesis de una primera etapa con reemplazo.

Esta aproximación común se denomina comúnmente WR en la bibliografía sobre encuestas. Así, cuando el plan de muestreo se describe como WR solo se necesitan tres variables del diseño de la encuesta para la estimación de varianzas:

- La variable de ponderación de la muestra o factor de expansión (necesaria también para las estimaciones).
- La variable de estratificación (o variable de pseudo-estratificación). En la II ENUT corresponde a la comuna ya que no fue necesario crear pseudo-estratos. En la base de datos designada como VarStrat.

- La variable UPM (o pseudo-conglomerado): En la II ENUT designada como VarUnit.

El empleo de WR para la aproximación al plan de muestreo complejo puede dar lugar a una ligera sobreestimación de las varianzas. Sin embargo, los analistas de datos de encuestas no tienen, por lo general, problemas en aceptar un cierto grado de sobreestimación por la relativa simplicidad de la aproximación mediante WR. No obstante, hay que tener en cuenta, que la sobreestimación puede ser apreciable si existen varios estratos en los que el muestreo en la primera etapa se hace sin reemplazo y generando grandes fracciones de muestreo. En este caso, puede ser preferible optar por un *software* que cuente con la opción de incorporar los factores de corrección por población finita o corrección por finitud para la primera etapa.

5.9.3. Algoritmo de cálculo en muestras complejas

En el cálculo de varianzas con muestras complejas, utilizando el software R, se utiliza el método de conglomerado último, en que la contribución de la segunda etapa se asume con reemplazo, lo que simplifica las fórmulas de cálculo.

Para el desarrollo de la estimación de las varianzas, es necesario contar con las ponderaciones de cada unidad de la muestra, de modo que reflejen la divergencia de las probabilidades de selección (probabilidades desiguales). La ponderación o factor de expansión de una unidad es el recíproco de su probabilidad de selección para incluirla en la muestra.

Si una vivienda se incluye en la muestra con una probabilidad $P_{hi}(j)$, su ponderación básica, viene dada por $w_{hi} = \frac{1}{P_{hi}(j)}$, es decir, el inverso de la probabilidad de selección y se denomina factor de expansión teórico, de acuerdo con el diseño probabilístico. Este ponderador es igual para todas

las viviendas dentro del conglomerado y, por lo tanto, para todas las personas dentro de ellas. En la II ENUT posteriormente, este factor se ajusta por: omisión de comunas y viviendas en otras muestras, elegibilidad y no-respuesta de las viviendas y finalmente, por la calibración para alcanzar los stocks poblacionales de personas según las proyecciones de población.

Sea y_{hij} la variable observada en la vivienda j del conglomerado i del estrato h , con ponderador base w_{hij} . Se utiliza la siguiente nomenclatura para las estimaciones y cálculo de las varianzas:

$z_{hij} = w_{hij} \cdot y_{hij}$, es el valor de la variable observada en la vivienda j a través de todos los hogares que componen la vivienda, expandido a la población.

$z_{hi} = \sum_{j=1}^{m_{hi}} z_{hij}$, es el valor de la variable observada expandida y agregada en el conglomerado i del estrato h .

$\bar{z}_h = \frac{1}{n_h} \cdot \sum_{i=1}^{n_h} z_{hi}$, es el promedio de la variable observada expandida en el estrato h .

n_h , es el número de conglomerados seleccionados en el estrato h en la muestra de la II ENUT.

N_h , es el número de conglomerados en el estrato h contenidos en el marco muestral utilizado en la II ENUT.

m_{hi} , es el número de viviendas seleccionadas y que responden dentro del conglomerado i del estrato h .

$S_h^2 = \frac{n_h}{n_h - 1} \cdot \sum_{i=1}^{n_h} (z_{hi} - \bar{z}_h)^2$, es la cuasi varianza en el estrato h , que corresponde a la variabilidad entre los conglomerados de la variable observada expandida en ellos. Asumiendo muestreo con reemplazo, se le agrega el factor de corrección por finitud $(1 - f_h) = (1 - n_h/N_h)$. Para

muestreo sin reemplazo, se utiliza una expresión más complicada, que se puede estudiar en Yates and Grundy (1953) y Sen (1953).

i. Algoritmo de cálculo de los estimadores

Para la estimación de totales, medias y razones, ya sea para cada uno de los estratos de muestreo h , como para los niveles de estimación o dominios de estudio, se utilizan los siguientes algoritmos:

El total estimado en el estrato h de la variable de interés Y es se observa en la ecuación (23).

$$\hat{Y}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_h} z_{hij} = \sum_{i=1}^{n_h} \sum_{j=1}^{m_h} w_{hij} \cdot y_{hij} \quad (23)$$

La estimación total a nivel agregado de la variable Y , es decir, la estimación agregada sobre todos los estratos y que forman algún nivel de estimación se observa en la ecuación (24).

$$\ddot{Y} = \sum_{h=1}^H \ddot{Y}_h \quad (24)$$

La estimación del número de unidades en la población en el estrato h se observa en la ecuación (25).

$$\hat{N}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \quad (25)$$

La media estimada de la variable de interés Y , en el estrato h se observa en la ecuación (26):

$$\widehat{\bar{Y}}_h = \frac{1}{\widehat{N}_h} \widehat{Y}_h \quad (26)$$

La media estimada de la variable de interés Y , en el nivel de estimación requerido se observa en la ecuación (27):

$$\widehat{Y} = \frac{1}{\widehat{N}} \widehat{Y} \quad (27)$$

El estimador de razón entre las variables Y y X , en el estrato h se observa en la ecuación (28):

$$\widehat{R}_h = \frac{\widehat{Y}_h}{\widehat{X}_h} = \frac{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij}}{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot x_{hij}} = \frac{\widehat{Y}_h}{\widehat{X}_h} \quad (28)$$

Por lo tanto,

$$R_h = \frac{Y_h}{X_h} = \frac{\bar{Y}_h}{\bar{X}_h} \quad (29)$$

El estimador de razón entre las variables Y y X , en el nivel de estimación requerido del estudio se observa en la ecuación (30):

$$\widehat{R} = \frac{\widehat{Y}}{\widehat{X}} = \frac{\sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij}}{\sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot x_{hij}} = \frac{\widehat{Y}}{\widehat{X}} \quad (30)$$

Existen también los estimadores del Total Y y la media \bar{Y} , pero a través de los estimadores de razón, esto es, si se despeja el total Y o la media \bar{Y} de la razón R se obtienen los siguientes estimadores:

Las estimaciones del total y la media de la variable Y en el estrato h , obtenidas mediante el estimador de razón se expresan en la ecuación (31):

$$\hat{Y}_{R_h} = \hat{R}_h \cdot X_h \text{ y también } \hat{\bar{Y}}_{R_h} = \hat{R}_h \cdot \hat{\bar{X}}_h \quad (31)$$

Las estimaciones del total y la media de la variable Y a nivel agregado, obtenidas mediante el estimador de razón se expresan en la ecuación (32):

$$\hat{Y}_R = \hat{R} \cdot X \text{ y también } \hat{\bar{Y}}_R = \hat{R} \cdot \hat{\bar{X}} \quad (32)$$

Para las ecuaciones (28), (29), (30), (31) y (32), se asume que el total X o la media \bar{X} son conocidas o dadas como fijas ²⁶. Sin embargo, generalmente se desconoce el valor poblacional y entonces son estimados por \hat{X} y $\hat{\bar{X}}$ en la expresión (29), o son estimadas por \hat{X}_h y $\hat{\bar{X}}_h$ en la expresión (30).

ii. Algoritmo de cálculo de las varianzas en muestras complejas

Como se ha planteado, el método de conglomerado último asume que la varianza total es debida a la variabilidad entre los conglomerados para la variable de interés (inter-varianza) y el aporte de la etapa secundaria y posterior es despreciable.

Definida la estratificación en los estratos h en la encuesta ENUT, la varianza del estimador del total se expresa en la ecuación (33):

²⁶ Por ejemplo, en la calibración a las proyecciones de población de personas, los stocks poblacionales son estimaciones obtenidas mediante modelos que son considerados fijos, como los verdaderos valores en la población. Además, si la variable es dicotómica, X coincide con N como asimismo \bar{X} coincide con \bar{N} .

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \hat{V}(\hat{Y}_h) = \sum_{h=1}^H S_h^2 = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (z_{hi} - \bar{z}_h)^2 \quad (33)$$

Desglosando z_{hi} y \bar{z}_h por las expresiones definidas previamente, queda como en la ecuación (34):

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\sum_{j=1}^{m_{hi}} z_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} z_{hi} \right)^2 \quad (34)$$

Reemplazando $z_{hij} = w_{hij} \cdot y_{hij}$ y $z_{hi} = \sum_{j=1}^{m_{hi}} z_{hij}$, estas expresiones pueden finalmente ser escritas como se presenta en la ecuación (35):

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \right)^2 \quad (35)$$

Finalmente, se muestran las estimaciones de las varianzas de los estimadores de totales, medias y razones.

La varianza o cuasivarianza estimada de la variable de interés expandida en el estrato h se presenta en la ecuación (36):

$$\hat{V}(\hat{Y}_h) = \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \right)^2 \quad (36)$$

La estimación de la varianza del Total de la variable de interés Y , en el nivel de estimación requerido se puede observar en la ecuación (37):

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \right)^2 \quad (37)$$

La estimación de la varianza en el estrato h , de la Media de la variable de interés Y se presenta en la ecuación (38):

$$\hat{V}(\hat{Y}_h) = \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left[\left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \right) / \hat{N}_h \right]^2 \quad (38)$$

La estimación de la varianza de la media de la variable de interés Y , según el nivel de estimación requerido se expresa en la ecuación (39):

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left[\left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \right) / \hat{N} \right]^2 \quad (39)$$

La varianza del estimador de razón entre las variables Y y X , en el estrato h se observa en la ecuación (40):

$$\hat{V}(\hat{R}_h) = \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left[\left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \hat{R}_h \cdot \sum_{j=1}^{m_{hi}} w_{hij} \cdot x_{hij} \right) / \hat{X}_h \right]^2 \quad (40)$$

Donde: $\hat{X}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot x_{hij}$

La varianza del estimador de razón entre las variables Y y X , en el nivel de estimación requerido se expresa en la ecuación (41):

$$\hat{V}(\hat{R}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left[\left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \hat{R} \cdot \sum_{j=1}^{m_{hi}} w_{hij} \cdot x_{hij} \right) / \hat{X} \right]^2 \quad (41)$$

La estimación de la varianza para el total y la media de la variable Y en base a la razón entre Y y X o la razón entre \bar{Y} y \bar{X} (ver expresiones (35), (36), (37) y (38) se expresa mediante las siguientes notaciones:

La estimación de la varianza del total y la media de la variable Y , en el estrato h , mediante el estimador de razón se observa en la ecuación (42):

$$\hat{V}(\hat{Y}_{R_h}) = \hat{V}(\hat{R}_h \cdot X_h) = \hat{V}(\hat{R}_h) \cdot \hat{X}_h^2 ; \hat{V}(\hat{Y}_{\hat{R}_h}) = \hat{V}(\hat{R}_h \cdot \bar{X}_h) = \hat{V}(\hat{R}_h) \cdot \hat{X}_h^2 \quad (42)$$

La estimación de la varianza del Total y la Media de la variable Y a nivel agregado, mediante el estimador de razón se expresa en la ecuación (43):

$$\hat{V}(\hat{Y}_R) = \hat{V}(\hat{R} \cdot X) = \hat{V}(\hat{R}) \cdot \hat{X}^2 ; \hat{V}(\hat{Y}_{\hat{R}}) = \hat{V}(\hat{R} \cdot \bar{X}) = \hat{V}(\hat{R}) \cdot \hat{X}^2 \quad (43)$$

5.9.4. Variables que identifican el diseño muestral complejo

Un diseño muestral complejo se origina cuando interviene más de una etapa de selección. Esta selección, en más de una etapa, suele ser generalmente, estratificada, lo que añade otro grado de complejidad al proceso. Bajo la estratificación, en un muestreo bietápico, las unidades muestrales se suelen llamar conglomerados o unidades de conglomeración que son grupos de unidades de selección de las últimas unidades de muestreo sobre las cuales se obtienen las estimaciones de las variables de interés.

La II ENUT presenta un diseño muestral que se puede caracterizar como probabilístico, estratificado y bietápico. En el muestreo estratificado la estimación de la varianza total se obtiene mediante la suma de las varianzas estimadas en cada estrato y, la varianza en cada estrato se estima a partir de la variabilidad de la característica de interés entre los conglomerados.

Para hacer factible la estimación de la varianza en un muestreo estratificado, se deben seleccionar como mínimo 2 conglomerados por estrato. Para evitar estimaciones de varianza volátiles, se recomienda seleccionar un número mínimo de unidades muestrales (ej. viviendas) por conglomerado. En consecuencia, diseños muestrales adecuados consideran al menos dos conglomerados por estrato, al momento de selección de la muestra y, un número mínimo de unidades por conglomerado. El diseño de II ENUT establece, en lo posible, un mínimo de 2 conglomerados por estrato y un mínimo de 10 viviendas por conglomerado.

Ya sea por diseño o por falta de respuesta, el problema de terminar con estratos con un solo conglomerado es recurrente en encuestas de hogares. Existen tres opciones para lidiar con el problema de estratos unitarios: (1) no hacer nada, lo que implica que no es posible estimar varianzas; (2) asumir que los conglomerados unitarios fueron seleccionados “con certeza” en cuyo caso estos estratos no contribuyen a la estimación de la varianza (se le asigna varianza igual a 0) generando una subestimación de la varianza; y (3) crear unidades que agrupen estratos unitarios junto a otros estratos (pseudo-estratos) para efectos de cálculo de varianza.

Siguiendo las mejores prácticas en la materia, se optó por esta última opción, con la finalidad de poner a disposición de los investigadores, variables que permitan identificar las características del diseño complejo de la II ENUT y producir estimaciones de varianza más estables. Las estimaciones oficiales de la varianza compleja han sido desarrolladas

utilizando pseudo-conglomerado (`varunit`) y variable de ponderación de la muestra `fe_cut` y `fe_cut` según corresponda²⁷.

5.9.5. Coeficiente de variación

El coeficiente de variación es una medida de variabilidad relativa que corresponde al cociente entre el error estándar $\hat{\sigma}_{\hat{\theta}}$ y la estimación $\hat{\theta}$, el contexto de una estimación intervalar, para la estimación de la variable de interés. Esto se expresa según la ecuación (44).

$$Cv_{\hat{\theta}} = Cv(\hat{\theta}) = \hat{\sigma}_{\hat{\theta}} / \hat{\theta} = \sqrt{\hat{V}(\hat{\theta})} / \hat{\theta} \quad (44)$$

Siendo $\hat{\theta}$ la estimación de la variable de interés θ , $\hat{\sigma}_{\hat{\theta}}$ el error estándar de la estimación, que corresponde a la raíz cuadrada de la varianza, dada por $\hat{V}(\hat{\theta})$.

En el contexto de una estimación intervalar de muestreo repetido, con una confianza dada de $(1-\alpha)\%$, el intervalo de confianza correspondiente para la variable de interés θ , está dado por la ecuación (45):

$$\theta \in [\hat{\theta} \mp t_{1-\alpha/2}^v * \hat{\sigma}_{\hat{\theta}}]_{1-\alpha} \Leftrightarrow [\hat{\theta} - t_{1-\alpha/2}^v * \hat{\sigma}_{\hat{\theta}} ; \hat{\theta} + t_{1-\alpha/2}^v * \hat{\sigma}_{\hat{\theta}}]_{1-\alpha} \quad (45)$$

Siendo $t_{1-\alpha/2}^v$ el percentil $(1 - \alpha/2)$ de la distribución *t-Student* con v grados de libertad.

El error absoluto de la estimación intervalar corresponde exactamente a la parte que se suma y resta de la estimación, esto es, $E_{abs}(\hat{\theta}) = t_{1-\alpha/2}^v * \hat{\sigma}_{\hat{\theta}}$, que, al dividirla por la misma estimación, se obtiene el error relativo, dado por la ecuación (46):

²⁷ Para la II ENUT no fue necesaria la construcción de pseudo-estratos ya que, luego del levantamiento, todos los estratos de muestreo tenían viviendas encuestadas en al menos dos conglomerados.

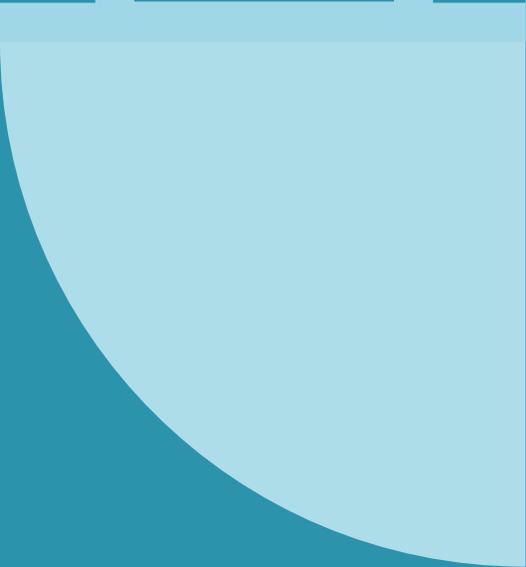
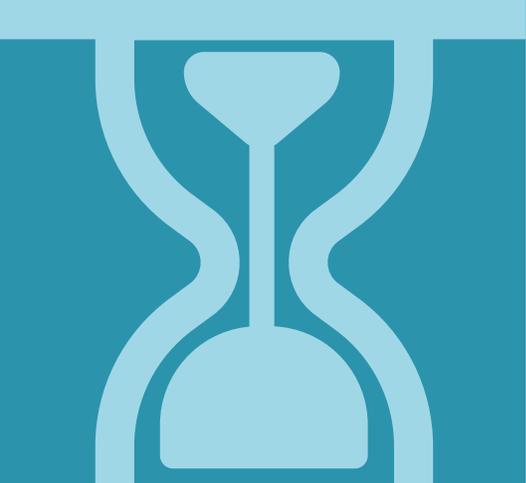
$$E_{rel}(\hat{\theta}) = t_{1-\alpha/2}^v * \hat{\sigma}_{\hat{\theta}} / \hat{\theta} = t_{1-\alpha/2}^v * Cv_{\hat{\theta}} = E_{abs}(\hat{\theta}) / \hat{\theta} \quad (46)$$

Para muestras grandes, el percentil de la distribución *t-Student* converge al percentil de la distribución Normal, es decir, $t_{1-\alpha/2}^v \approx Z_{1-\alpha/2}$, que para una confianza de 95%, corresponde al valor 1,96, esto es: $Z_{0,975} = 1,96$ cuando la confianza es $(1 - \alpha) = 95\%$.

Si solo se considera el coeficiente de variación como medida del error relativo, en el contexto de estimación intervalar, es semejante a construir intervalos de confianza de aproximadamente $(1 - \alpha) = 68\%$ de confianza, ya que $Z_{1-\alpha/2} = Z_{0,84} \approx 1$, entonces, aproximadamente $E_{rel}(\hat{\theta}) = Cv_{\hat{\theta}}$.

Se considera que el resultado de una estimación debe leerse con precaución²⁸ si su coeficiente de variación es mayor a 20%, es decir, en el contexto de una estimación intervalar de 95% de confianza, si su error relativo es mayor a 40%.

²⁸ Comúnmente, si el error relativo de una estimación es superior a 30%, es motivo de alerta, en el sentido que la estimación no está siendo tan precisa, lo que es aproximadamente equivalente a tolerar un coeficiente de variación no superior a 15%.



6

**METODOLOGÍA DE
RECOLECCIÓN
DE DATOS**

6. Metodología de recolección de datos

6.1. Características generales de la recolección de datos

Las técnicas y métodos de recolección señalan la forma en que deben recopilarse los datos de los hogares, y las personas encuestadas. Las siguientes secciones profundizan en las dimensiones de esta fase.

6.1.1. Características organizacionales de equipos participantes de la recolección de datos

La organización de los equipos del proyecto II ENUT durante el trabajo de campo fue articulada con el propósito de gestionar en los plazos estimados la totalidad de las unidades en terreno, mantener un flujo constante de encuestas recolectadas y cumplir los objetivos muestrales propuestos con altos estándares de calidad, siguiendo las recomendaciones internacionales en el desarrollo del proceso de producción estadística.

i. Organigrama del proyecto

Fue dispuesto un modelo de organización centralizado (Figura 6), que ubicó al equipo de coordinación central en la Región Metropolitana, y a nivel regional los equipos de campo.

Figura 6. Organigrama II ENUT



En relación con las funciones del equipo en la etapa de recolección, estos fueron sus objetivos centrales:

Jefatura de Proyecto²⁹: liderar los distintos equipos involucrados en el proyecto, coordinando esfuerzos de cada uno de ellos para la correcta ejecución de este. Junto a planificar y validar las ejecuciones presupuestarias asociadas a las distintas etapas del proceso.

Coordinación Nacional de Recolección³⁰: liderar el equipo de recolección realizando la planificación operativa del levantamiento, las estrategias y coordinación del avance del trabajo de campo. Debía monitorear las actividades de terreno y diseñar la proyección de la recolección.

Encargado(a) de Área Recolección: gestionar y asesorar metodológicamente a los equipos de recolección del área de recolección

²⁹ Su dependencia administrativa se encuentra en el Subdepartamento de Estadísticas Estructurales del Trabajo, ubicado dentro del Departamento de Estadísticas del Trabajo de la Subdirección técnica del INE.

³⁰ Todos los equipos relacionados con la recolección dependieron administrativamente del Departamento de Operaciones Estadísticas Sociales.

asignada (Región Metropolitana y resto de capitales regionales), a través de los encargados(as) macrozonales de recolección. Colaborar en la preparación del operativo y monitorear la ejecución de la recolección de datos.

Encargado(a) Macrozona Recolección: gestionar la programación de las actividades de recolección, coordinando y controlando a los equipos asignados a su zona. Dar soporte técnico al equipo y monitorear los avances del operativo.

Encargado(a) de Grupo: planificar y monitorear el cumplimiento de la programación, de las actividades y tareas de los equipos recolectores de datos, reportando el cumplimiento de los protocolos y la calidad de la información levantada a partir de las pautas preestablecidas en el proyecto.

Encuestadores(as): recolectar información de forma presencial en los hogares seleccionados, a través de la aplicación de técnicas de recolección, siguiendo los estándares establecidos por el estudio y la institución.

Coordinación Nacional de Supervisión: liderar al equipo de supervisión, definiendo estrategias de seguimiento a través de la coordinación de los supervisores de terreno, y el monitoreo del trabajo de campo como ente independiente del equipo de recolección.

Encargado(a) Macrozona Supervisión: gestionar la programación de las actividades de supervisión; Coordinar y monitorear el equipo de supervisión y el de soporte técnico; Vigilar el rendimiento y reportar los hallazgos del proceso al coordinador nacional de supervisión del grupo de regiones asignado a él o ella.

Supervisores(as): monitorear la calidad del proceso de recolección mediante jornadas de acompañamiento en terreno del equipo encargado,

aplicando pautas de seguimiento (supervisión directa). Además, deben de constatar información específica con los(as) respondientes (supervisión indirecta).

Coordinación del Equipo Técnico: liderar al equipo de analistas socioeconómicos, responsables de distintos procesos de la producción estadística, tales como: el diseño y construcción de este, así como del procesamiento y análisis de los datos, y la difusión de resultados.

Analistas Socioeconómicos(as): personal encargado de la ejecución del diseño del proceso de producción de la II ENUT, la construcción de los elementos requeridos para su ejecución, tales como los cuestionarios, capacitaciones, manuales, entre otros. El procesamiento de datos provenientes de la recolección, el análisis de estos datos y difusión del estudio.

Adicionalmente a estos equipos especialmente conformados para la recolección de la II ENUT, se contó con el trabajo de otros equipos de la institución que prestaron apoyo directo para el desarrollo del estudio a lo largo de toda su recolección:

Diseño de Marcos y Muestras: en la construcción de la muestra a recolectar en la encuesta y en el desarrollo de los elementos asociados al diseño complejo de la encuesta, tales como los factores de expansión y pseudo-conglomerados de varianza.

Sistemas de Procesamiento Estadístico: en el desarrollo informático del Sistema de gestión operativa (SGO) y su conexión con *Survey Solutions*.

Implementación de Diseños: en la vinculación con los equipos de tecnología de la información, y las gestiones en la implementación de sistemas.

Operaciones Logísticas: en la gestión logística de bienes y servicios a nivel central y su envío a regiones.

ii. Distribución del personal de trabajo de campo

La dotación de personal fue ajustada a la complejidad regional y la muestra asignada. En correspondencia, se estableció una proporción de un encargado(a) de grupo por cada cinco encuestadores(as) para asegurar una gestión efectiva de las labores de recolección. Esta razón arrojó a nivel nacional un total de 244 personas encuestadoras y 48 encargados(as) de grupo. Además, cada localidad contó con un equipo de supervisores de terreno, que proporcionó apoyo adicional a la coordinación directa y seguimiento de las actividades en terreno.

Los equipos operativos fueron dispuestos en cada una de las regiones. Allí los coordinaron los Encargados(as) Macrozonales desde la RM, y contaron con la colaboración directa de las **Jefaturas operativas de cada región** (ver Tabla 19).

Tabla 19. Personal de recolección en regiones

Regiones	Encuestadores(as)	Encargados (as) de grupo	Supervisores (as)
Arica y Parinacota	11	2	1
Tarapacá	13	3	2
Antofagasta	15	3	2
Atacama	12	2	1
Coquimbo	17	3	2
Valparaíso	19	4	2
Metropolitana	32	6	3
O'Higgins	14	3	2
Maule	18	4	2
Ñuble	11	2	1
Biobío	19	4	2
La Araucanía	16	3	2
Los Ríos	12	2	1
Los Lagos	14	3	2
Aysén	10	2	1
Magallanes	11	2	1
Nivel País	244	48	27

iii. Perfiles de cargo, selección y condiciones contractuales del equipo de recolección

El proceso de selección siguió el flujo habitual que desarrolla el INE para seleccionar a su personal: diseño de perfiles, publicación de la oferta laboral, evaluación de antecedentes, prueba psicométrica, entrevista personal y contratación.

La evaluación curricular se ocupó de ponderar aspectos de formación, definidos previamente en el diseño de perfil de cada cargo. La prueba psicolaboral, se encargó de aspectos tales como conocimientos generales y rasgos de personalidad. La evaluación técnica, de aptitudes

y competencias necesarias para desempeñar el rol. Y las entrevistas finales, de recursos comunicativos y experiencia laboral.

En la selección de los(as) postulantes a encuestadores(as), se privilegió la práctica previa en recolección de datos. Para el caso de los supervisores(as) y encargados(as) de grupo, se priorizó la experiencia en labores de coordinación de equipos de recolección de datos, o supervisión.

Por último, cabe señalar que en el caso de los encuestadores(as), se generó una base de datos en actualización permanente, a la par de un aviso en el portal de empleos públicos durante todo el proyecto, como una medida frente al recambio de personal por renunciadas, desvinculaciones, cambio de cargo, u otros motivos.

iv. Organigrama del equipo de supervisión

La organización del equipo de supervisión fue segmentada en dos niveles. A nivel central, se estableció la figura de una persona encargada nacional de supervisión y otras cuatro para cada macrozona del país. A nivel regional, se designaron veintisiete cupos para el cargo de supervisores(as) operativos(as), garantizando la presencia de al menos una persona supervisora por región.

Los equipos operativos de supervisión fueron dispuestos en cada una de las 16 regiones (ver Tabla 20). Allí se les coordinó desde el nivel central por las personas encargadas macrozonales de supervisión, en conjunto con las jefaturas de operaciones de cada región. La distribución de cada macrozona fue realizada de forma que cada encargado(a) macrozonal, tuviese una distribución equitativa de personas supervisoras, considerando la cantidad de unidades muestrales por monitorear. Como se observa en la Tabla 20, el detalle del personal quedó como sigue:

Tabla 20 Cantidad de personas supervisoras por región

Nacional	Macrozona	Región	Unidades Muestrales	N° de Supervisores(a s)
Coordinador/a Nacional de supervisión	Encargado/a Macrozonal 1	Valparaíso	2.544	2
		O'Higgins	1.500	2
		La Araucanía	1.704	2
		Aysén	864	1
		Total	6.612	7
	Encargado/a Macrozonal 2	Antofagasta	1.416	2
		Atacama	1.332	1
		Coquimbo	1.944	2
		Magallanes	828	1
		Ñuble	1.116	1
		Total	6.636	7
	Encargado/a Macrozonal 3	Maule	1.872	2
		Biobío	2.088	2
		Los Lagos	1.464	2
		Los Ríos	1.140	1
		Total	6.564	7
	Encargado/a Macrozonal 4	Metropolitana	4.416	3
		Tarapacá	11.28	2
		Arica y Parinacota	960	1
		Total	6.504	6
Total nacional			26.316	27

6.1.2. Modo de recolección de la información

La II ENUT empleó el registro de información a través de Dispositivos Móviles de Captura (DMC) o *tablets*, motivo por el que se dispuso de estos instrumentos a cada una de las personas encuestadoras. También fueron distribuidos cuestionarios físicos para situaciones de contingencia, a pesar de que en la práctica hubo pocas situaciones que obligaran a usar encuestas en papel (PAPI).

6.1.3. Periodo de referencia y de recolección

i. Periodicidad de los datos

En las Encuestas sobre el Uso del Tiempo realizadas en Chile, no existe una periodicidad propiamente tal. No obstante, la OIT recomienda cierta periodicidad en la recolección de estadísticas sobre el uso del tiempo, el cual en la práctica a nivel regional se ha traducido en la aplicación de encuestas con una periodicidad quinquenal.

ii. Periodos de referencia

La encuesta cuenta con distintos períodos de referencia dependiendo de las preguntas realizadas. Para las preguntas de vínculo con el mercado laboral e ingresos del trabajo se siguieron los estándares de la OIT: semana y mes anterior a la entrevista, respectivamente y, para las preguntas de otros ingresos, también se utiliza como período de referencia el mes anterior. Para las preguntas sobre uso del tiempo se asigna muestralmente a cada vivienda dos períodos de referencia: un día de semana y uno de fin de semana específicos. Esta asignación permite que los distintos días de la semana tengan una presencia balanceada al interior de la muestra, además de resguardar cierta consistencia interna

en los períodos de referencia por los cuales se consulta a todos los integrantes del hogar elegibles.

iii. Período de recolección y distribución operativa de la muestra

El proceso de recolección inició el 30 de agosto del 2023, con su marcha blanca³¹, cuya finalidad fue que los equipos de recolección y supervisión aprendieran los flujos de trabajo y se habituaran a sus labores, para reducir los errores no muestrales.

La recolección oficial comenzó el 14 de septiembre con la primera submuestra operativa, y concluyó en la séptima submuestra el 29 de diciembre. Durante este período, el objetivo fue entrevistar al total de los hogares residentes de las 26.316 viviendas seleccionadas según el diseño muestral.

La muestra objetivo según diseño fue de 15.490 viviendas, con una sobre muestra del 41,1%, véase la Tabla 21.

³¹ Las unidades recolectadas durante la marcha blanca no entran en los datos oficialmente publicados, esperando resguardar la mayor calidad posible dentro de los datos publicados.

Tabla 21. Muestra de viviendas II ENUT

Región	Muestra objetivo	Muestra total a gestionar
Arica y Parinacota	707	960
Tarapacá	802	1.128
Antofagasta	930	1.416
Atacama	758	1.332
Coquimbo	1.099	1.944
Valparaíso	1.238	2.544
Metropolitana	2.038	4.416
O'Higgins	881	1.500
Maule	1.134	1.872
Ñuble	701	1.116
Biobío	1.221	2.088
La Araucanía	999	1.704
Los Ríos	775	1.140
Los Lagos	889	1.464
Aysén	647	864
Magallanes	671	828
Total nacional	15.490	26.316

6.1.4. Ejecución operativa de la recolección de una submuestra operativa

Cada submuestra desplegada en terreno consideró las siguientes fases: Fase de preparación (4 días previos a la aplicación), Fase de aplicación (14 días de aplicación) y Fase de cierre y recuperación (5 días posteriores a la aplicación).

En la Figura 7 se muestra la estructura general de una submuestra operativa con sus fases.

Figura 7. Estructura de submuestra operativa II ENUT

Es importante indicar que cada submuestra operativa contaba con un solapamiento con la siguiente, donde durante el cierre de una, las personas encuestadoras se encontraba comenzando la fase de aplicación de la siguiente.

A continuación, se detallará en qué consistía cada una de estas fases:

i. Fase de preparación

La fase de preparación es una etapa previa a los 14 días de aplicación, cuyo propósito fue generar las condiciones óptimas de apertura de las unidades muestrales. Incluía realizar el reconocimiento de los sectores y unidades a levantar, dejando cartas en las viviendas, además de otras medidas.

Reconocimiento previo:

Con el objetivo de maximizar los resultados del trabajo de campo, las y los Encargados de Grupo, días antes del inicio de la fase de preparación, tuvieron como parte de sus tareas salir a terreno priorizando los lugares de difícil acceso (condominios, edificios, áreas lejanas de locomoción, etc.), para informar de la realización de la encuesta, sus objetivos y la fecha tentativa de visita del equipo de encuestadores(as).

Asignación de Carga:

Esta fase implicaba la distribución de las cargas de trabajo en cascada desde el Encargado Nacional, pasando por los de Área, hasta los Coordinadores/as Macrozonales, los de Grupo, y de ahí a las y los

encuestadores. Cada encuestador(a) recibió una lista de hogares específicos que debía entrevistar, según la muestra definida previamente.

Identificación de las Viviendas en Terreno:

Las personas encuestadoras estudiaron en oficina las características de las unidades muestrales que se les asignaron, considerando entre otros, las vías de acceso y la presencia de edificios o condominios. Después de la revisión, se iba a terreno a corroborar que no hubiera diferencia entre la información entregada por los encargados(as) de grupo y lo observado en campo. En caso de hallarlas, o de no encontrar las viviendas asignadas, debían reportar la situación inmediatamente a su jefatura directa.

En detalle, las instrucciones que siguieron las personas encuestadoras en la identificación de las viviendas fueron:

Previo a la salida a terreno:

Examinar la cartografía disponible y verificar rutas de acceso.

Revisar dirección y observaciones en el directorio.

En terreno:

Visitar dirección y utilizar las observaciones para hallar las viviendas.

En caso de no encontrar el domicilio, o bien identificar cambios no registrados en las observaciones que dificultaran ubicarlo, se siguieron las siguientes instrucciones:

Registrar en hoja de ruta el código pertinente: *"No fue posible localizar la dirección"*.

Dar aviso al o al Encargado(a) de Grupo (EG), quien le entregaba orientaciones o referencias para lograr encontrar la vivienda.

Una vez recibidas las orientaciones, las referencias entregadas por el Encargado(a) de Grupo se registraban en las observaciones del cuestionario.

Con la nueva información, se intentaba localizar una vez más el domicilio, registrando el código de intento respectivo de la última visita realizada.

En caso de identificarse la vivienda, la persona encuestadora se presentaba al informante y explicaba el objetivo de la encuesta, solicitando su participación. Destacando la garantía de la confidencialidad en el proceso y se resolvían posibles dudas.

ii. Fase de aplicación

La fase de aplicación es un período de 14 días continuos en cada hogar, tiempo en el que se realizó la aplicación tanto del CH, como de los CUT de las personas elegibles para este. Lo anterior se realizó por medio de un dispositivo móvil de captura (DMC).

En esta fase se genera la información fundamental para cumplir los objetivos del estudio. En ella es donde los informantes entregan la información entrevistada por las personas encuestadoras.

El esquema de entrevistas dentro de cada hogar constaba de entrevistar primero el CH en su totalidad, para luego pasar a entrevistar los CUT. Este orden es fundamental, ya que el CH entrega la información necesaria para aperturar ciertas baterías de preguntas en los CUT.

Recolección del Cuestionario del Hogar (CH)

La entrevista se llevó a cabo con un informante idóneo de 18 años o más del hogar ³². El propósito era registrar información demográfica y

³² En ocasiones, se contó con informantes idóneos menores de 18 años que respondieron el CH, estos fueron casos excepcionalmente aprobados por el equipo de recolección, dadas situaciones y contextos particulares en el hogar.

socioeconómica, además de otras características de las personas y el hogar.

Recolección de los Cuestionarios sobre Uso del Tiempo (CUT)

Posterior a la aplicación del CH, se entrevistaba con el Cuestionario sobre Uso del Tiempo, donde se recopiló información específica respecto a un listado de actividades diarias de los integrantes del hogar, de 12 años o más, realizadas en un día de semana y en otro de fin de semana específicos asignados previamente.

En promedio se visitó cada hogar cerca de 3 veces para lograr los cuestionarios requeridos y posibles de recolectar. En ocasiones cierta información del CH no era posible de obtener previamente, por lo que, si no era esencial para aplicar correctamente los CUT, era posible recuperar dicha información en visitas posteriores a la primera, para completar o ajustar dicha información.

Durante este proceso de aplicación se llevaron a cabo distintas actividades que permitieron la correcta recolección de datos:

Conseguir contacto telefónico de las personas informantes

Se recopilaban números de teléfono de los y las informantes para agendar entrevistas, para verificaciones de información y/o supervisiones. Esto resultaba crucial para asegurar la calidad y consistencia de los datos recopilados.

Revisión de la Encuesta por parte de la persona encuestadora y Encargado(a) de grupo:

La primera revisión de los cuestionarios para cerciorarse de su correcto llenado, fue responsabilidad de los y las encuestadoras. El objetivo era corregir lo más pronto posible cualquier error o discrepancia. Luego, la segunda revisión en gabinete estuvo a cargo del Encargado(a) de Grupo,

quien debía examinar las encuestas recopiladas de las personas que dependían de su gestión directa.

Un cuestionario se daba por aprobado cuando tenía el CH completo, y al menos un CUT logrado de un integrante del hogar (pero no todos), lo que se entiende como logro parcial del hogar. Por su parte, si ambos tipos de cuestionarios se encontraban terminados en todos los miembros de 12 años o más del hogar, se consideraba un logro total. Ante cualquier observación o requerimiento de datos, el cuestionario era devuelto con comentarios al encuestador o encuestadora, para que este(a) gestionara su corrección.

Supervisiones directas e Indirectas

Paralelo al trabajo de recolección en terreno, el equipo de supervisión realizó sus actividades siguiendo los criterios de verificación. Estas incluían supervisiones de encuestas seleccionadas aleatoriamente y el monitoreo directo del proceso de entrevista. Este trabajo se profundizará en secciones posteriores.

iii. Fase de Cierre y Recuperación

Esta fase se extendía por 5 días posterior al término de la aplicación de los cuestionarios del estudio. Esta fase apuntó a finalizar las tareas de revisión de las encuestas y recuperar la información faltante del CH. Cabe señalar que esta fase de cierre y recuperación se solapaba con la aplicación de la siguiente submuestra operativa.

Última revisión de Encargados(as) de Grupo

Después de la recolección de datos, las personas encuestadoras realizaron nuevamente una revisión de las encuestas recopiladas para garantizar su completitud y consistencia, atendiendo a que todos los

campos obligatorios estuviesen llenos y que las respuestas no fueran contradictorias o inconsistentes.

Recontacto.

Las encuestas con inconsistencias o falta de información se devolvieron a terreno. En esta fase, las personas encuestadoras regresaban a las viviendas específicas a aclarar dudas o completar la información faltante. La recuperación de información solo involucró al CH, dadas las definiciones metodológicas, donde no se debía recuperar información del CUT en un momento distinto a la entrevista original.

6.1.5. Indicadores de desempeño operativo

La información registrada durante el proceso de recolección en las Hojas de Ruta tanto de CH como de CUT, permitió dar seguimiento al trabajo de campo a través de indicadores de desempeño que se estimaron a partir de ellas. Al revisar esta data por submuestra y región se accedía a contenido valioso y oportuno para la toma de decisiones en mejora de los resultados de la encuesta.

Los códigos de disposición final de las encuestas permitían calcular los indicadores de desempeño operativo. En la Tabla 22 se encuentran desagregados los códigos de disposición finales (CDF)³³ con los que se calcularon las tasas e indicadores.

³³ El código de disposición final (CDF), resulta de todas las visitas realizadas por una persona encuestadora, dando cuenta del estado final de la unidad encuestada, indicando el código de máximo contacto logrado con esta.

Tabla 22. Clasificación de viviendas según tipo de código de disposición final

Categoría según pertenencia a población objetivo	Categoría según nivel de contacto	Clasificación	Sigla	Código de disposición CH (Hogar)	Códigos de disposición CUTs (Personas de 12 años o más)	Código de disposición hogar (CDF)
Elegibles	Entrevistadas	Entrevista completa	E	CH = 11	Todos los CUT elegible logrados completos = 11	11
		Entrevista parcial	EP	CH = 11 12	Al menos un CUT elegible logrado, pero no todos = 11 12	12
	Rechazadas	Elegibles, break-off	BO	CH = 11 12	Sin ningún CUT logrado	21
		Rechazo inicial	RI	CH = 21	-	21
	No contactadas	No contactadas	NC	CH = 31 32 33 34	-	31 32 33 34
	No entrevistadas por otros motivos	Otras no entrevistadas	O	CH = 41 42 43 44 45 49	-	41 42 43 44 45 49
Unidades de elegibilidad desconocida			ED	CH = 51 52 53 54 59	-	51 52 53 54 59
Unidades no elegibles			NE	CH = 61 62 63 64 65 66 67 68 69	-	61 62 63 64 65 66 67 68 69

Nota: Para mayores detalles del significado de cada código de disposición, remitirse a la Hoja de ruta presente en el Cuestionario del Hogar de la II ENUT, página 3.

|: este símbolo se utiliza para establecer una relación lógica de "o", donde se puede cumplir una u otra situación o código en este caso.

La construcción de los indicadores operativos sigue las recomendaciones de cálculo entregadas por los equipos de Calidad y Estándares del INE. Las fórmulas de estimación utilizaron las siguientes nomenclaturas:

- **E** = Número de unidades con entrevista completa.
- **P** = Número de unidades con entrevista parcial.
- **R** = Número de unidades rechazadas (incluye unidades con entrevista parcial pero baja calidad respecto a los datos – break off).
- **NC** = Número de unidades no contactadas.
- **O** = Número de unidades no entrevistada por otros motivos.
- **ED** = Número de unidades de elegibilidad desconocida.
- **NE** = Número de unidades no elegibles
- **M (Total de unidades muestrales)** = total de unidades seleccionadas en la muestra, incluyendo unidades asignadas como sobremuestra que buscan mitigar el no logro esperado.
- **Mo (Total de unidades muestra objetivo)** = total de unidades seleccionadas en la muestra para garantizar los niveles de precisión diseñados para el estudio. Se excluyen las unidades asignadas como sobremuestra.

La Tabla 23 muestra las fórmulas de cálculo de los principales indicadores operativos.

Tabla 23. Cálculo de Indicadores

Nombre	Descripción del Indicador	Formula
Tasa de No Elegibilidad	Se expresa como la razón entre el número de unidades que constituyen casos no elegibles sobre el total de unidades muestrales a encuestar en un período de tiempo determinado.	$TNE = \left(\frac{NE}{M}\right) * 100$
Tasa de Elegibilidad Desconocida	Se expresa como la razón entre el número de unidades que constituyen casos con elegibilidad desconocida sobre el total de unidades muestrales a encuestar en un período de tiempo determinado.	$TED = \left(\frac{ED}{M}\right) * 100$
Tasa de Contacto	Describe el número de unidades donde se estableció algún tipo de contacto (se entrevistó, rechaza o no por otros motivos), sobre el total de unidades muestrales elegibles y de elegibilidad desconocida, en un tiempo determinado.	$TCON = \left(\frac{E + P + R + O}{E + P + R + NC + O + ED}\right) * 100$
Tasa de Cooperación	Describe el número de unidades donde se realizó una entrevista completa y/o parcial, sobre el total de unidades muestrales elegibles y contactadas de la muestra (incluyendo entrevistadas, rechazos y otras unidades no entrevistadas que pueden haber sido contactadas) en un período de tiempo determinado.	$TCOOP = \left(\frac{E + P}{E + P + R + O}\right) * 100$
Tasa de Rechazo	Describe el número de unidades donde se obtiene un rechazo, sobre el total de unidades muestrales a encuestar que son potencialmente elegibles (entrevistadas, rechazadas, no contactadas y no entrevistadas por otros motivos) y aquellas de elegibilidad desconocida, en un período de tiempo determinado.	$TREC = \left(\frac{R}{E + P + R + NC + O + ED}\right) * 100$

Tasa de Respuesta (no ponderada)	Describe el número de unidades donde se ha obtenido una entrevista completa o parcial, sobre el total de unidades muestrales a encuestar que son potencialmente elegibles (entrevistadas, rechazadas, no contactadas y no entrevistadas por otros motivos) y aquellas de elegibilidad desconocida, en un período de tiempo determinado.	$TRnp = \left(\frac{E + P}{E + P + R + NC + O + ED} \right) * 100$
Tasa de Logro Muestra Objetivo	Describe el número de unidades muestrales a encuestar en un período determinado y donde se obtiene una entrevista (completa o parcial), en relación con las unidades de la "muestra objetivo" definidas en la encuesta. A diferencia de la Tasa de respuesta, este indicador no distingue en su cálculo a aquellas unidades que no forman parte de la población objetivo del estudio (no elegibles).	$TLO = \left(\frac{E + P}{Mo} \right) * 100$
Tasa de Logro Muestral Total	Describe el número de unidades muestrales a encuestar en un período determinado y donde se obtiene una entrevista (completa o parcial), en relación con las unidades de la muestra total a gestionar, considerando la sobremuestra.	$TLT = \left(\frac{E + P}{M} \right) * 100$

Las tasas calculadas en base a los estados finales de cada unidad encuestada de la II ENUT se pueden encontrar en la Tabla 24, en base a las fórmulas presentadas previamente.

Tabla 24. Indicadores de Desempeño Operativo

Indicadores de desempeño operativo	Tasa
Tasa de No Elegibilidad	6,9%
Tasa de Elegibilidad desconocida	0,7%
Tasa de Contacto	82,6%
Tasa de Cooperación	79,9%
Tasa de Rechazo	13,7%
Tasa de Respuesta	66,0%
Tasa de Logro Muestra Objetivo	104,4%
Tasa de Logro Total	61,4%

6.1.6. Estrategias de sensibilización

El trabajo de sensibilización de las personas informantes de los hogares estuvo centrado en fomentar su involucramiento a través de dos elementos centrales: confianza en la institución y relevancia en la medición del uso del tiempo.

Esta labor fue realizada por los equipos de recolección (encuestadores y supervisores) de cada región. Para ello, se les capacitó en aspectos como la Ley 17.374 de resguardo del secreto estadístico, el enfoque de género de la encuesta y la importancia para el país de la visibilización del tiempo destinado al trabajo doméstico y a cuidados no remunerados en el hogar.

El personal de trabajo de campo contó con material de apoyo, como cartas de presentación para las viviendas seleccionadas, dípticos informativos sobre las características centrales de la encuesta y afiches que se ubicaron en negocios o lugares concurridos de los sectores donde se tuvo presencia.

6.2. Diseño de los instrumentos de captura de información

6.2.1. Tecnología de producción de la información

Los instrumentos de recolección de la II ENUT se aplicaron por medio de un dispositivo móvil de captura (DMC). En este dispositivo cada persona encuestadora registró las respuestas dadas por las y los informantes para los dos cuestionarios presentes en el estudio.

En relación con el diseño y construcción de los cuestionarios, este se realizó a través de *Survey Solutions (SuSo)*, programa desarrollado por el Banco Mundial.

Este programa entrega tres herramientas (ver Tabla 25) que permiten desarrollar desde la construcción de los cuestionarios digitales (*Designer*), pasando por el aplicativo destinado a que las personas encuestadoras aplicaran los cuestionarios (*Interviewer*), hasta la plataforma que permite gestionar de forma centralizada las asignaciones de carga de encuestas, personal y verificación de estado (*Headquarters*).

Tabla 25. Herramientas de SuSo en la II ENUT

Herramienta de SuSo	Características	Usos en II ENUT
Designer	Servicio en plataforma web que permite desarrollar cuestionarios complejos, colaborar entre diseñadores en una misma versión, identificar problemas de compilación de la programación del cuestionario, posibilidad de testear los cuestionarios en línea y DMC, y generar documentos en PDF detallando todos los elementos del cuestionario construido.	Herramienta que permitió el diseño y construcción del cuestionario de la II ENUT. Proceso colaborativo e iterativo que incluyó la implementación de la estructura general del cuestionario, preguntas, saltos, alertas y validaciones iniciales.
Interviewer	Aplicación que permite su utilización vía web o en DMC, la que permite gestionar a las personas encuestadoras la carga de unidades que le son asignadas, y realizar las entrevistas asociadas.	La II ENUT utilizó la aplicación que fue posible instalar en los DMC de cada persona encuestadora, permitiendo aplicar cada entrevista, y gestionar el estado de sus cargas individuales, identificando unidades asignadas, iniciadas y completadas. Para luego realizar la sincronización vía internet de los datos a la plataforma de <i>Headquarters</i> .
Headquarters	Herramienta web que permite realizar seguimiento del progreso general del estudio, realización de revisiones, aprobaciones y rechazos de entrevistas terminadas, gestión de los equipos y roles, configuración de la encuesta que permite importar instrumentos de recolección construidos por medio de <i>Designer</i> y exportar los datos recolectados.	El servidor de <i>Headquarters</i> para la II ENUT, fue instalado por los equipos de tecnología de la información dentro de los servidores institucionales, resguardando la seguridad de la información recolectada, manteniéndose en una red interna. Dado el desarrollo interno INE, del Sistema de Gestión Operativa (SGO), se integraron estos dos sistemas para la gestión central de la recolección.

El cambio tecnológico de papel a una encuesta digital implicó ciertas ventajas como, por ejemplo:

El no requerir de una fase de digitación permitió una reducción de los tiempos entre la recolección y la obtención de los datos para su procesamiento, posibilitando reaccionar y tomar decisiones con mayor celeridad frente al despliegue del trabajo de campo. Asimismo, fue posible articular de forma más temprana la programación destinada a la limpieza y tratamiento de la información. Por último, esta nueva modalidad dio acceso a una forma automática para obtener datos contextuales de la recolección, llamada parados³⁴, resultando claves en la supervisión y análisis del proceso de recolección.

Cabe mencionar que la plataforma digital admitía también la programación de aclaraciones y alertas frente a determinadas respuestas; duraciones muy extensas o bajas dependiendo de la actividad. Ante estos escenarios se activaba un mensaje que pedía a la persona encuestadora corroborar inmediatamente con el(la) encuestado(a) la respuesta entregada.

Para aquellos casos donde no fuera posible capturar los datos con los DMC, se distribuyeron cuestionarios físicos para situaciones de contingencia, a pesar de que en la práctica hubo pocas situaciones que obligaran a usar encuestas en papel (PAPI).

A continuación, se describen estos instrumentos que fueron recolectados en PAPI y DMC.

³⁴ Más información sobre los parados y su uso en la II ENUT se puede encontrar en la sección 9.2.3.

6.2.2. Cuestionario del Hogar (CH)

El objetivo de este cuestionario es caracterizar al hogar entrevistado y a las distintas personas que lo componen, ayudando a distinguir a quienes pertenecen al hogar. Además, indaga en sus características sociodemográficas tales como sexo, edad, relaciones de parentesco, etc., a la par de las cualidades del propio hogar y vivienda; presencia de servicio doméstico o de personas que ayuden a las labores del hogar; ocupación y sus ingresos, entre otras dimensiones.

El CH se aplicó a un informante idóneo miembro del hogar seleccionado, de 18 años o más, con los conocimientos necesarios para responder por el resto de los integrantes. Adicionalmente a la propia caracterización lograda con esta entrevista, permite identificar a las personas a las que debía aplicarse el segundo instrumento llamado Cuestionario sobre Uso del Tiempo (CUT), como también los módulos que tendrán que responder. Por último, en caso de que en una misma vivienda seleccionada se detectara más de un hogar, cada uno debía responder un CH independiente.

Sobre los ámbitos que indagó el CH, este cuenta con siete módulos temáticos que incluyeron: “Identificación del hogar”, “Características sociodemográficas”, “Educación”, “Trabajo en la ocupación”, “Ingresos”, “Personas en situación de dependencia funcional” y “Ayudas que recibe el hogar”. A continuación, se presenta un breve resumen de cada uno.

Módulo Identificación del Hogar (H): su objetivo es identificar el número de personas y hogares en la vivienda. A partir de aquí, cada módulo registró los datos de todos los integrantes de 12 años o más que componen un hogar, salvo excepciones indicadas en las mismas preguntas.

Módulo Características sociodemográficas (C): busca caracterizar sociodemográficamente a las personas a través de preguntas como la edad, sexo, género, estado conyugal y nacionalidad.

Módulo Educación (E): su objetivo es registrar la situación educacional de todos los integrantes. Se consultan aspectos tales como el nivel educativo más alto alcanzado, y si asiste actualmente a algún establecimiento educativo, entre otros.

Módulo Ocupación (O): su propósito es indagar en la situación ocupacional de cada persona de 15 años o más, considerando su vínculo con el mercado laboral, razones que explican estar fuera de la fuerza de trabajo, situación ocupacional, informalidad laboral y horas de trabajo.

Módulo Ingresos (Y): la finalidad es establecer los ingresos monetarios que tiene cada persona de 15 años o más, según su situación ocupacional descrita en el Módulo O. Para estimar este monto se consideran los principales ingresos de la ocupación, jubilaciones, pensiones, arriendos y transferencias.

Módulo Personas en situación de dependencia funcional - PSDF (D): identifica si en el hogar existen personas en situación de discapacidad, que requieran de la asistencia de un tercero para realizar sus actividades cotidianas. También indaga quién es el(a) cuidador(a) principal (en términos de tiempo destinado al cuidado) de las personas en situación de dependencia funcional.

Módulo Ayudas que recibe el hogar (A): su objetivo es comprobar si el hogar recibe apoyo de una o más personas externas en el trabajo doméstico y de cuidados, distinguiendo si esta ayuda es remunerada o no.

Otras secciones son:

Hoja de ruta II ENUT: completada por el encuestador, permite el registro de los esfuerzos realizados para contactar a los distintos informantes, ingresando la elegibilidad de la vivienda, las visitas de apertura y los códigos de disposición de cada visita, junto al estado final de la unidad entrevistada.

Observaciones: espacio destinado a que la persona encuestadora registrara la información relevante del hogar o sobre alguna situación anormal que ayudara a entender respuestas particulares en el cuestionario.

6.2.3. Cuestionario sobre Uso del Tiempo (CUT)

El Cuestionario sobre Uso del Tiempo (CUT) tiene como objetivo registrar las actividades realizadas por los distintos integrantes del hogar, de **12 años o más**, dentro del período de referencia (que considera un día de semana y otro de fin de semana), junto al tiempo que le dedicaron. Para cumplir con ello, la estructura del cuestionario consideró 11 módulos. En cada uno se incluye un listado de actividades asociadas a la temática abordada. En su mayoría, la primera pregunta por la que se consulta es sobre **participación**, es decir, si el informante efectuó o no la actividad durante el día asignado. Cuando la respuesta era afirmativa, se interrogaba luego por la duración, **cuánto tiempo** invirtió en ella considerando todas las veces que desarrolló la acción durante el día.

Algunos tipos de actividades incluyen a terceras personas, como aquellas del Módulo de Trabajos de Cuidados (TC). Por ejemplo, para saber si para alguna Persona en Situación de Dependencia Funcional (PSDF) o persona de 0 a 4 años del hogar, el respondiente lo/la mudó o llevó al baño. Cuando se contesta afirmativamente (participación), se indagaba a quiénes, individualizando de ese modo a las personas sujetas de cuidado. Otro tipo de preguntas, abordan los traslados, consultando cuánto tiempo se dedicó a movilizar a algún integrante del hogar,

pudiendo responder de forma desagregada por persona, o bien un tiempo total para todos los participantes del traslado.

Sobre la aplicación del CUT, es importante señalar que debía ser respondido de manera **directa y presencial por cada integrante del hogar de 12 años o más**, y que cada CUT tenía que ser aplicado **dentro del mismo día**, por lo que no debía recuperarse información en entrevistas posteriores, ya que esto podía provocar que las respuestas de una segunda instancia fueran menos precisas, dado el mayor período de recuerdo e inconsistencia interna con las respuestas dadas previamente.

Referente a los módulos del CUT, estos son:

Módulo Contextualización de los días asignados (CD): apunta a facilitar el recuerdo de las actividades realizadas en los días asignados (tanto de semana, como para el de fin de semana), formulando preguntas de control sobre el itinerario de dichas fechas. Por ejemplo: ¿Almorzó?, ¿Estuvo enfermo?, ¿Estuvo fuera de su casa por más de 12 horas?, ¿Hizo tareas domésticas como cocinar o hacer aseo? entre otras. Esto permitía situar a la persona informante antes de comenzar a consultarle sobre la participación y tiempo destinado a cada actividad.

Módulo Trabajo en la ocupación (TO): examina la participación y tiempo dedicado al trabajo en la ocupación, traslados de ida y vuelta, y si el individuo buscó trabajo o inició un negocio durante los días de referencia. Al mismo tiempo indaga, en los casos que corresponde, la interrupción de la jornada laboral por motivos de cuidado de un niño o niña de 0 a 4 años.

Módulo Trabajo de cuidados no remunerado (TC): registra la participación y tiempo dedicado a actividades de cuidado no remunerado de integrantes del hogar. En esta versión, y con el fin de registrar para quiénes se realizó trabajo de cuidados, se utilizó como identificador el

número de línea de la persona sujeta de cuidado, de esta forma es posible identificar tanto la edad del integrante, como si este está en situación de dependencia funcional.

Módulo Trabajo doméstico no remunerado (TD): indaga sobre las actividades de trabajo doméstico no remunerado. Considera aspectos como la preparación y servicio de comida, limpieza de la vivienda, limpieza de ropa y calzado, mantenimiento y reparaciones menores del hogar, y por último administración del hogar y abastecimiento.

Módulo Trabajo voluntario y ayudas a otros hogares (TV): busca caracterizar, durante los días asignados, las transferencias de trabajo a otros hogares y a la comunidad que no involucran remuneraciones.

Módulo Cuidados personales (CP): registra los tiempos que se destinan a actividades personales. Incluye por ejemplo desayunar, almorzar e ir a una consulta médica.

Módulo Educación (ED): examina los tiempos que la persona invierte en educación formal (para personas que se encuentran estudiando) y no formal. Considera las duraciones de los traslados, estadía en clases y si se recibieron clases particulares de apoyo escolar.

Módulo Ocio y vida social (VS): identifica los tiempos que los integrantes destinan a la vida social y a actividades de ocio. Incluye, por ejemplo, conversar o compartir con familiares, pareja o amistades durante el tiempo libre; asistir a eventos culturales, de entretenimiento o deportivos; practicar algún deporte y hacer ejercicio físico; y actividades tales como ver televisión o utilizar medios de comunicación masiva.

Módulo Simultaneidad (SIM): registra las actividades que la persona informante efectuó de forma simultánea, permitiendo identificar distintos pares de actividades realizadas al mismo tiempo durante los días de referencia.

Este módulo es una innovación respecto a la ENUT 2015, posibilitando contar con información respecto a cantidad de actividades realizadas en simultáneo e indagar en aquellas que más se realizan en paralelo junto a otras, abordando la dimensión de la intensidad en el uso del tiempo.

Módulo Bienestar Subjetivo (BS): apunta a caracterizar el bienestar subjetivo de los integrantes respondientes del CUT, a través de la consulta del nivel de satisfacción del(a) entrevistado(a) con su vida en general, respecto a cómo se reparten las tareas domésticas en el hogar, y las actividades de cuidado hacia otras personas. Adicionalmente, indaga que tan conforme se encuentra con el tiempo que dedica en la actualidad a diversas actividades.

Módulo Características específicas de la persona (ID): se consulta sobre el Rut de quienes responden el CUT y la orientación sexual (esta última solo para mayores de 15 años o más).

6.3. Prueba y evaluación de los instrumentos de captura de información

6.3.1. Prueba Piloto II ENUT

La Prueba Piloto de la II ENUT tuvo lugar entre el 6 de marzo y el 3 de mayo de 2023 en las zonas urbanas de las regiones de Valparaíso, Metropolitana y Biobío, considerando una gestión total de 1.440 viviendas.

El objetivo general de la Prueba Piloto fue:

- *Evaluar la presencia de problemas de construcción, medición y respuesta en los cuestionarios piloto de la II ENUT 2023.*

Mientras que los objetivos específicos fueron:

- Analizar la duración de las entrevistas para la aplicación de los cuestionarios CH y CUT, incluyendo la comparación entre diferentes grupos de interés.
- Examinar las respuestas, distribución, magnitud y comportamiento en módulos o preguntas de importancia crítica para el instrumento, incluyendo comparación entre diferentes grupos de interés.
- Determinar la eficacia de las categorías de respuesta para el registro correcto de datos provenientes de las y los participantes en el estudio.
- Identificar errores en la medición mediante el uso de rejillas de validación.
- Evaluar el logro en la aplicación de la encuesta ENUT vía el análisis de los códigos de disposición obtenidos para cada cuestionario.
- Evaluar la aplicación del diseño experimental seleccionado para comparar la recolección de datos utilizando formularios en papel frente a formularios en Tablet.
- Aplicar cambios al instrumento en función del análisis de los resultados obtenidos.

Entre los puntos anteriores, destaca la comparación de experiencias en la aplicación de formularios en versión papel, frente a los formularios versión *tablet*. Para esto, se realizó un diseño experimental por muestra dividida aleatoriamente, distribuyendo la muestra cómo se observa en la Tabla 26.

Tabla 26 Muestra de viviendas a gestionar, Prueba Piloto II ENUT

Región	CAPI	PAPI	Total
Región Metropolitana	240	240	480
Biobío	240	240	480
Valparaíso	240	240	480
Total	720	720	1.440

La mitad de la muestra fue el grupo control, aplicando los cuestionarios por medio de papel, mientras la otra mitad (grupo de tratamiento) fue aplicada por medio de DMC.

De estas gestiones se lograron encuestar a 964 hogares, de los cuales 490 fueron en CAPI y 474 en PAPI.

Si bien la Prueba Piloto tenía dentro de sus objetivos específicos la comparativa CAPI vs PAPI, la experiencia fue tomada para realizar una serie de ajustes adicionales los flujos de trabajo técnico-operativos.

Los análisis y observaciones adicionales se realizaron desde la perspectiva de varios ejes catalogados como críticos para el funcionamiento de las entrevistas. Entre ellos:

- Evaluación de la **no respuesta y frecuencia de los ítems nuevos o más complejos**, con el objetivo de revisar el volumen de uso de las alternativas de cada una de estas preguntas.
- **Tiempo de aplicación de las entrevistas**, evaluando la pertinencia de la reducción de alguno de los módulos propuestos en la Prueba Piloto.

- Evaluar la distribución, magnitud y el **comportamiento de las principales variables del estudio**, incluyendo comparación entre los principales grupos de interés.
- **Flujo de trabajo general** de los equipos de recolección y supervisión.

En los principales indicadores que se evaluaron de la encuesta, como lo es el tiempo dedicado a trabajo no remunerado, tiempo destinado a la ocupación y carga global de trabajo, no se encontraron diferencias estadísticamente significativas entre las aplicaciones en PAPI y aquellas realizadas en el grupo tratamiento con CAPI.

6.3.2. Hitos previos a la Prueba Piloto II ENUT

La elaboración del instrumento testado en la Prueba Piloto II ENUT, así como los análisis posteriores realizados, estuvieron informados por una serie de pruebas que data de dos hitos de recolección de datos sobre uso del tiempo del INE: el proyecto original de II ENUT de 2020 y el proyecto de Módulo rotativo sobre Uso del Tiempo (MUT) en la Encuesta Nacional de EMPLEO (ENE).

La primera iniciativa consistía en una II ENUT para ser aplicada durante el primer semestre de 2020, la cual fue suspendida por la emergencia sanitaria y las restricciones presupuestarias derivadas de la pandemia de COVID-19. La ejecución del proyecto fue interrumpida durante la semana de capacitaciones previa al lanzamiento del instrumento piloto. Más tarde, gracias al trabajo conjunto del Subdepartamento de Estadísticas Estructurales del Trabajo (SEET) y el Subdepartamento de Estadísticas Continuas del Trabajo (SECT) se elaboró durante 2021 el MUT con miras a proporcionar la información necesaria para responder al indicador 5.4.1 de los Objetivos de Desarrollo Sostenible, pero el módulo diseñado no llegó a ser aplicado en la ENE.

Para efectos de la elaboración y el análisis de la Prueba Piloto II ENUT, la experiencia de estos dos hitos (La II ENUT no aplicada el 2020 y el MUT) destacan la elaboración del cuestionario 2020; el primer diseño de una encuesta ENUT en la plataforma *Survey Solutions* (SUSO) para aplicación de un cuestionario en su versión en *Tablet*; un análisis QAS (*Questionnaire Appraisal System* o Sistemas de Evaluación de Cuestionarios de Willis y Lessler, 1999) en el marco de una serie de pruebas cognitivas al instrumento elaborado; y el trabajo de mesas interministeriales para recoger demandas sobre organismos usuarios de los datos sobre uso del tiempo.

6.3.3. Ajustes para versión oficial

A continuación, se describen las principales modificaciones que se realizaron al cuestionario, a raíz de las observaciones y análisis técnico-operativos de la Prueba Piloto.

Debido al buen desempeño de los Dispositivos Móviles de Captura en la prueba Piloto, se decidió utilizar estos dispositivos para la versión oficial, incluyendo los ajustes que se detallan a continuación. Se mejoraron saltos y filtros de preguntas, que no implicaban un exceso de carga de procesamiento para el dispositivo móvil de captura, y permitió mejorar la experiencia de la entrevista.

Junto a esto se incorporaron nuevas alertas y aclaraciones en algunas preguntas, dados ciertos errores sistemáticos de aplicación identificados durante la Prueba Piloto. Un caso de este tipo ocurrió en la pregunta c1, donde las personas encuestadoras registraban en el número de línea 1 al informante idóneo, en lugar de al proveedor principal.

Se redujo la extensión del cuestionario para disminuir la duración de la entrevista. En el módulo de Trabajo en la ocupación (O), las personas informantes y el personal de trabajo de campo, en grupos focales,

hicieron énfasis en la necesidad de reducir esta batería de preguntas, por lo que se priorizaron las preguntas necesarias para mantener cierto nivel de armonización con las otras mediciones oficiales respecto a ocupación de la institución, en particular, con respecto a la Encuesta Nacional de Empleo (ENE).

Por otro lado, durante la prueba Piloto, la forma de preguntar los trabajos de cuidados y el tiempo destinado a ellos se consultó de dos formas diferentes. Para algunas secciones se consultaron según cada subpoblación de relevancia (formato ENUT 2015), mientras otras se abordaron consultando por la tarea realizada e identificando a la persona sujeta de cuidado directamente, por medio del identificador dentro del hogar (número de línea). Esta segunda modalidad mostró mejor desempeño en las pruebas, con una baja no respuesta, lo que se tradujo en que la versión oficial de la II ENUT se incorporara esta modalidad de consulta en la totalidad de las secciones del módulo de trabajos de cuidados (TC). Este cambio permite construir las subpoblaciones que el usuario(a) requiera para sus análisis respecto a las personas que requieren cuidados.

6.4. Método de capacitación

A continuación, se presentan detalles respecto a los procesos de capacitación para la II Encuesta Nacional sobre Uso del Tiempo.

6.4.1. Capacitación del equipo de recolección

El proceso de capacitación entregó herramientas y detalles sobre la recolección y aplicación de la encuesta, tanto a los(as) encuestadores(as) como a quienes debían desempeñar un cargo de coordinación. Su modalidad fue telemática en regiones y presencial en la Región Metropolitana, para ambos casos en jornadas desde las 9:30 horas hasta las 18:00 horas durante 5 días. Cada una de ellas, pasa a describirse resumidamente a continuación:

Día 1

- Exposición de Antecedentes generales de la ENUT.
- Organización del trabajo de Campo y conceptos operativos, por ejemplo, qué es una UPM, una vivienda, un hogar, entre otros.
- Metodología de trabajo de Campo, esquema de visitas, qué es una submuestra y sus fases, etc. Introducción a la Hoja de Ruta y cuestionarios a utilizar.

Día 2

- Presentación de las Hojas de Ruta en formato papel (PAPI) y formato digital (SuSo)
- Exposición del Cuestionario del Hogar (CH) módulo por módulo en formato PAPI y CAPI. Primera parte.

Día 3

- Exposición del Cuestionario del Hogar (CH) módulo por módulo en formato PAPI y CAPI. Segunda parte y final.

Día 4

- Exposición del Cuestionario sobre Uso del Tiempo módulo por módulo. en formato PAPI y CAPI. Primera parte.

Día 5

- Exposición del Cuestionario sobre Uso del Tiempo módulo por módulo. en formato PAPI y CAPI. Segunda parte y final.
- Ejercicios prácticos de llenado del cuestionario completo y solución a situaciones hipotéticas de registro.
- Parte de esta última jornada se dedicó a capacitar en exclusivo a las personas encargadas de grupo, sobre el flujo de información que debían monitorear junto a sus atribuciones particulares.

La capacitación fue acompañada de material de consulta y para facilitar la comprensión y funcionamiento de la encuesta, y en particular de sus distintos módulos. Para ello se contó con un Manual de trabajo de campo, el cual fue distribuido entre los distintos perfiles capacitados. Este Manual contenía toda la información pertinente al trabajo de recolección y a la forma de organización del trabajo de campo.

Esta capacitación fue central, por lo que se llevó a cabo para todos los perfiles que desempeñaron labores directas y continuas en el proyecto.

Adicionalmente, se mantuvo habilitada una modalidad de capacitación continua a cargo de las personas encargadas de grupo, para los encuestadores(ras) que fueron incorporándose al proceso durante el avance del trabajo de campo.

Contar con una capacitación completa de una encuesta compleja como lo es la II ENUT, demostró ser necesaria, junto a la incorporación de mayor cantidad de ejercicios prácticos. En el caso del presente estudio, una semana de capacitación fue lo requerido para poder abarcar toda la información teórica y práctica para su posterior aplicación por el personal de trabajo de campo.

El equipo de supervisión, adicional a la capacitación general de recolección, realizó una capacitación adicional, centrada en las tareas concretas asociadas al perfil de supervisores(as) operativos(as), llevándose a cabo de forma telemática para incorporar al personal de regiones diferentes a la metropolitana.

6.5. Supervisión de la recolección

Se aplicaron distintos métodos de supervisión para garantizar la calidad de la información y asegurar el cumplimiento de los procedimientos establecidos en la operación estadística de la II ENUT.

El equipo de supervisión tuvo un rol independiente de la recolección, con personal diferenciado, dependiendo directamente de la jefatura de proyecto, lo que permite contar con una supervisión externa a las tareas del equipo de recolección de datos.

Una serie de medidas fueron implementadas durante el trabajo de campo, con el propósito de identificar y corregir errores que pudiesen invalidar total o parcialmente los datos, asegurando, que el proceso de recolección fuese riguroso y representativo.

Dentro de los tipos de errores que afectan a una encuesta, es posible distinguir entre los que provienen del diseño muestral, y aquellos ajenos a ese factor, como resultado de problemas de cobertura, la no respuesta, la medición o el procesamiento. La estrategia de supervisión utilizada resultó útil sobre todo para gestionar errores de medición, ya que prestó

atención a sus factores causales: las personas encuestadoras, los informantes, los problemas con el cuestionario y equivocaciones producidas en el método de recolección de datos (Díaz de Rada, 2019).

De acuerdo con las recomendaciones internacionales y los estándares establecidos en otras operaciones estadísticas del INE, se definió para la II ENUT una cobertura de supervisión de al menos un 20% de la muestra gestionada por recolección. Ese porcentaje corresponde a cerca de un 11,3% de supervisión directa, mientras que las supervisiones indirectas representaron un 23,9% de las encuestas a nivel nacional. Estos dos tipos de supervisión se explicarán con mayor detalle dentro de esta misma sección.

6.5.1. Tareas del equipo de supervisión

El equipo de supervisión fue responsable de distintas tareas y funciones durante las etapas de diseño y recolección del proyecto. Respecto de la primera etapa, se identificaron cuatro áreas esenciales en el control de calidad de la encuesta: Planificación y gestión administrativa, Sistema informático, Cuestionarios de supervisión y Capacitación de los supervisores.

En relación con el subproceso de ejecución de la recolección, resulta útil describir y diferenciar los roles que asumieron los involucrados en el proceso.

Comenzando por el **coordinador o coordinadora nacional de supervisión**, encargado(a) de dirigir la ejecución del proceso a nivel nacional de supervisión de la II ENUT, que debió planificar y monitorear las actividades de las personas encargadas macrozonales de supervisión y de las personas supervisoras operativas, anticipando riesgos de incumplimiento e implementando ajustes oportunos en la planificación y desarrollo de las acciones de equipo.

Respecto de la persona responsable de la implementación del plan operativo; **encargada de cada macrozona de supervisión**, sus responsabilidades fueron, planificar las actividades de las personas supervisoras, analizar los resultados de la aplicación de los instrumentos de recolección con el propósito de asignar y controlar las cargas de trabajo de cada supervisor(a) operativo(a), y revisar los datos y consistencia de los instrumentos de monitoreo gestionados por las personas supervisoras operativas.

Por su parte, el rol de **la persona supervisora operativa** responsable de certificar la labor realizada por las y los encuestadores, tuvo dentro de sus tareas:

- Monitorear el correcto cumplimiento de los procesos de recolección, aplicando formularios de supervisión directa y/o indirecta.
- Elaborar y remitir a la coordinación de supervisión, informes de cumplimiento de los protocolos y calidad de los datos recolectados.
- Detectar e informar a la coordinación de supervisión las debilidades en el proceso de recolección, asegurando que todas las actividades se llevasen a cabo en conformidad con los procedimientos establecidos.
- Comunicar oportunamente problemáticas frecuentes y de excepción en el transcurso del proyecto, identificando soluciones adecuadas.

A continuación, se describen los dos tipos de supervisiones llevadas a cabo por este equipo.

6.5.2. Supervisión directa

Este método de supervisión consiste en el acompañamiento en el momento de la entrevista a cada una de las personas encuestadoras por

parte de un supervisor, para monitorear sus actividades en terreno durante su jornada de trabajo. La persona supervisora debía adoptar una actitud de observador; prestando atención al flujo de preguntas diseñado en el dispositivo y si la persona encuestadora lo seguía, aplicaba y fraseaba correctamente las preguntas. Además de verificar la gestión de la apertura y cierre de la entrevista. La cordialidad de las personas encuestadoras y el temple ante rechazos o una disposición negativa hacia la entrevista de las personas informantes, también fueron evaluados.

Toda la información anterior era recopilada en un Cuestionario de Supervisión Directa, que al igual que la encuesta, fue realizada en dispositivo móvil de captura, por medio del programa *Survey Solutions*. Este cuestionario consideraba los siguientes módulos en términos de supervisar ciertos elementos de la entrevista realizada por la persona encuestadora: contacto inicial, aplicación del Cuestionario del Hogar, aplicación del Cuestionario de Uso del Tiempo, aplicación de protocolos y cierre de la entrevista.

Posterior a la visita de supervisión directa, el supervisor o supervisora operativa realizaba una retroalimentación a la persona encuestadora, en base a un informe cualitativo estandarizado, donde registraba el detalle de lo observado en el transcurso de la jornada de acompañamiento. La retroalimentación se efectuaba directamente al encuestador o encuestadora supervisada, una vez concluida cada encuesta.

En la retroalimentación se hacía énfasis en la evaluación de tres dimensiones: la apertura de la vivienda, la aplicación de cuestionarios y el cierre de la entrevista.

Al inicio de la recolección de la encuesta se dio prioridad a este tipo de supervisión, para retroalimentar al equipo de recolección y así tomar medidas de mejora continua a tiempo.

6.5.3. Supervisión indirecta

El formulario de supervisión indirecta fue una medida de control cuyo propósito es corroborar que la información recopilada por las personas encuestadoras fuera fidedigna, la correcta aplicación de los protocolos correspondientes y que no se hubiese incurrido en la falsificación de datos. Para lograr lo anterior, los supervisores(as) operativos(as) aplicaron un cuestionario de supervisión indirecta a las unidades encuestadas, que al igual que el de supervisión directa, fue realizado en dispositivo móvil de captura.

El cuestionario de supervisión indirecta se aplicó tanto de manera presencial como telefónica, aunque se dio preferencia a la modalidad telefónica. Para las unidades que no fueron entrevistadas, el contacto se realizó exclusivamente de forma presencial.

La información de la supervisión indirecta se contrastó con la información levantada en la etapa de recolección donde se reaplicaron algunas preguntas del cuestionario CH y CUT, para realizar una comparativa entre las respuestas registradas y las obtenidas durante la supervisión.

Para aquellas encuestas cuyos códigos de disposición eran clasificadas como “entrevistas completas o parciales” (11 o 12), se verificó, además, que tuviera registros tanto en el CH como en el CUT.

La pauta contaba, además, con un módulo para la supervisión de unidades de código de disposición final no elegibles, permitiendo verificar en los hogares que era pertinente si su uso había sido bien

empleado o no. El disponer de una única pauta de monitoreo contribuyó con agilizar el proceso en tanto hizo más fácil la asignación de casos.

Este método de supervisión se implementó al concluir la fase de recolección realizada por las personas encuestadoras en cada submuestra operativa. Los criterios presentes en la selección de los hogares a monitorear de forma indirecta incluyeron:

- Viviendas con 3 o más hogares.
- Hogar con 9 o más integrantes.
- Cuestionarios con tiempos de aplicación inusuales.
- Casos de personas encuestadoras con el 50% de hogares unipersonales.
- Casos de personas encuestadoras con el 90% de sus hogares con logro completo.
- Selección propuesta por encargado/a de grupo de recolección o equipo técnico.
- Supervisión aleatoria por sistema.

Sobre la retroalimentación al equipo de recolección, en casos donde las personas supervisoras identificaban algún hallazgo, se notificaba de inmediato al(a) encargado(a) de la macrozona de supervisión. Este accionar activaba el protocolo de informar al equipo de recolección y dar seguimiento al caso.

Sobre la retroalimentación al equipo de recolección, en casos donde las personas supervisoras identificaban algún hallazgo, se notificaba de inmediato al(a) encargado(a) de la macrozona de supervisión. Este accionar activaba el protocolo de informar al equipo de recolección y dar seguimiento al caso.

6.5.4. Indicadores de supervisión

Los indicadores de supervisión (ver Tabla 27) fueron utilizados durante todo el proceso de recolección oficial de datos, permitiendo monitorear la aplicación de la encuesta, para evaluar si se debían o no realizar ajustes al proceso de control de calidad. En su construcción se empleó la información de los estados de logro de la Hoja de ruta de supervisión, que utilizaron códigos adaptados por el INE desde el estándar de la *American Association for Public Opinion Research (AAPOR, 2023)*.

Tabla 27. Descripción de los Indicadores de Supervisión

Nombre	Descripción	Fórmula
Porcentaje de gestión de supervisión (GS)	Este indicador corresponde a un porcentaje determinado por el número de unidades de supervisión directa e indirecta gestionadas aprobadas por supervisión, en un período de tiempo determinado, sobre el total de unidades gestionadas aprobadas por recolección.	$\frac{\text{Número de unidades de supervisión directa e indirecta gestionadas aprobadas por supervisión en un periodo } t}{\text{Total de unidades gestionadas aprobadas por recolección en un periodo } t} \times 100$
Porcentaje de gestión de supervisión directa (GSD)	Este indicador corresponde a un porcentaje determinado por el número de unidades de supervisión directa gestionadas aprobadas por supervisión, en un período de tiempo determinado, sobre el total de unidades gestionadas aprobadas por recolección.	$\frac{\text{Número de unidades de supervisión directa gestionadas aprobadas por supervisión en un periodo } t}{\text{Total de unidades gestionadas aprobadas por recolección en un periodo } t} \times 100$
Porcentaje de gestión de supervisión indirecta corroboración de datos (GSI)	Este indicador corresponde a un porcentaje determinado por el número de unidades de supervisión indirectas, gestionadas aprobadas por supervisión, en un período de tiempo determinado sobre el total de unidades gestionadas aprobadas por recolección.	$\frac{\text{Número de unidades de supervisión indirecta gestionadas aprobadas por supervisión en un periodo } t}{\text{Total de unidades gestionadas aprobadas por recolección en un periodo } t} \times 100$
Porcentaje de efectividad en el logro de supervisión (ELS)	Este indicador corresponde a un porcentaje determinado por el número de unidades de supervisión directas e indirectas, que resultaron efectivas, logradas aprobadas por supervisión, en un período de tiempo determinado sobre el total de unidades logradas y parciales aprobadas por recolección.	$\frac{\text{Número de unidades efectivas logradas por supervisión aprobadas en un periodo } t}{\text{Total de unidades logradas y parciales aprobadas en recolección en un periodo } t} \times 100$
Porcentaje de gestión de supervisión	Este indicador corresponde a un porcentaje determinado por el número de unidades de supervisión directa e indirecta, denominadas no elegibles, gestionadas aprobadas por supervisión,	

indirecta no elegibles (GSNE)	en un período de tiempo determinado sobre el total de unidades no elegibles gestionadas aprobadas por recolección.	$\frac{\text{Número de unidades no elegibles gestionadas por supervisión aprobadas en un periodo } t}{\text{Total de unidades no elegibles aprobadas en recolección en un periodo } t} \times 100$
--------------------------------------	--	--

6.6. Sistema informático

Para efectos de la gestión operativa de la encuesta se empleó una plataforma de recolección de estadísticas sociales desarrollada por el INE, llamada Sistema de Gestión Operativa (SGO). Este sistema, enlazado al ambiente *Headquarters* de *Survey Solutions* (SuSo) en el que se programaron los cuestionarios, colaboró, tras añadir los siguientes componentes, a una mejor experiencia de los usuarios(as):

- Asignación en cascada de los distintos perfiles operativos.
- Asignación por áreas de levantamiento (Regiones, comunas, UPM, viviendas).
- Precarga de datos para supervisión desde cuestionarios recolectados.
- Gestión de la supervisión.

6.7.1. Funcionalidades asociadas a la ejecución de recolección

El sistema informático desempeñó un papel clave en las tareas de monitoreo realizadas por la coordinación nacional de recolección, facilitando el acceso a información actualizada. Entre las funciones más importantes del sistema se destacaron las siguientes:

- **Gestión muestral**

Tras la carga de las bases de datos de las viviendas seleccionadas, el sistema permitía asignar las cargas de trabajo a las personas encuestadoras, habilitar la creación de las claves de entrevistas en los dispositivos móviles, acceder a los módulos digitales, supervisar la gestión de las personas encuestadoras mediante la hoja de ruta y generar nuevos hogares en las viviendas seleccionadas.

- **Revisión de cuestionarios**

El sistema permitió acceder al cuestionario digital, el cual podía ser revisado por la persona encargada de grupo de recolección, quien tenía la posibilidad de agregar observaciones y comentarios, señalando la presencia de errores, omisiones o inconsistencias en la información de cada módulo. Cabe señalar que los encargados(as) de grupo no podían editar la información registrada por la persona encuestadora, dejando una observación para una posterior corrección o justificación del o la encuestadora.

- **Consultas y solicitudes de autorización de medidas especiales**

El sistema permitía brindar apoyo a los equipos de recolección cuando surgían dificultades durante el proceso como, por ejemplo, problemas relacionados con la identificación de unidades muestrales, la necesidad de actualizar información domiciliaria o ajustes en el número de visitas solicitadas por el hogar. Estas peticiones eran revisadas por el equipo de coordinación nacional de recolección, que evaluaba las posibles soluciones y decidía aprobar o rechazar cada solicitud. Un aspecto clave de este módulo era que se generaba un registro de los ajustes realizados y sus motivos, proporcionando mayor contexto sobre el proceso de recolección.

- **Descarga de bases**

Con la información proporcionada por el SGO y *SuSo*, era posible calcular los indicadores de desempeño operativo y generar informes del proceso de recolección. Estos informes incluían resultados a nivel nacional y desglosados por región, personas encargadas de grupo y encuestadores(as). La implementación de la aplicación de la encuesta en DMC permitió que, al sincronizar los datos, estos pudieran ser visualizados y gestionados de manera centralizada.

6.7.2. Funcionalidades asociadas a la supervisión de recolección

La supervisión del proceso de recolección de datos requirió herramientas eficaces para monitorear las actividades operativas, empleando recursos como el sistema de gestión, la hoja de ruta de supervisión, y varios indicadores de las actividades de control de la encuesta. Gracias a estas herramientas, se contó con datos clave para realizar ajustes y planificar acciones del control de calidad.

La gestión del proceso de supervisión se realizó mediante el Sistema de Gestión Operativa SGO, el cual interactúa directamente con la plataforma *Survey Solutions* (*SuSo*). Por su parte, las secciones de supervisión del SGO permitían monitorear y aprobar las unidades seleccionadas, y conferían acceso a los datos y administración operativa de la encuesta. Entre los principales módulos utilizados se encontraban:

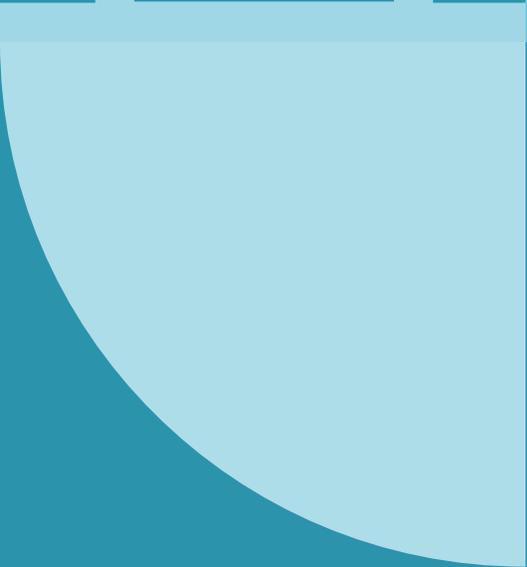
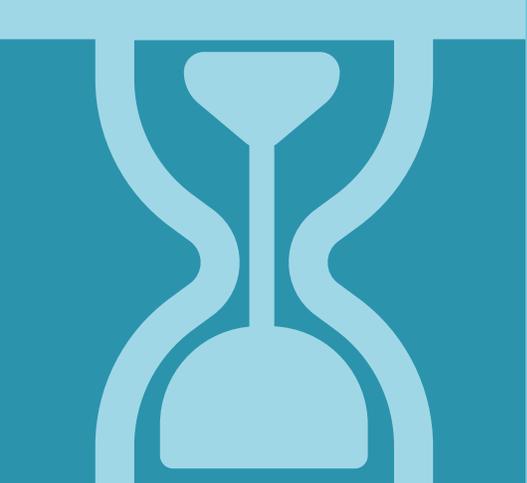
Administración de usuarios: orientado a la creación, gestión y edición de los perfiles del equipo de supervisión.

Asignación de área de levantamiento: indica la distribución de macrozonas y equipos de trabajo correspondientes a los colaboradores,

tanto para los perfiles de las personas supervisoras operativas, como para los encargados/as macrozonales.

Asignación de folios: que permitió monitorear el llenado de los formularios de supervisión en curso y la asignación y reasignación de folios a supervisar para cada persona encargada de esta tarea.

Hoja de ruta de supervisión: cuyo objetivo fue caracterizar la gestión realizada en cada supervisión, representando un insumo para la construcción de indicadores de eficiencia, que fueron de ayuda luego en el seguimiento del proceso de supervisión de la encuesta.



7

**METODOLOGÍA DE
PROCESAMIENTO
DE DATOS**

7. Metodología de procesamiento de datos

El procesamiento de datos de la II ENUT es el proceso donde se transformaron los datos de origen (brutos), sin organizar ni tratar, provenientes directamente del registro de la información recolectada por medio del programa informático *Survey Solutions* en dispositivos móviles de captura, en datos definitivos para fines de análisis y difusión.

7.1. Estructura de trabajo

Los ocho subprocesos que se abordan en la siguiente sección, fueron realizados a través del programa estadístico R, repositorios de *Gitlab* alojados en servidores internos INE. La utilización de esta herramienta permitió:

- i. Desarrollo de tareas de procesamiento en un entorno colaborativo de programación por los distintos miembros del equipo técnico del estudio.
- ii. Contar con control de versiones para el seguimiento de cambios y ajustes realizados en las distintas actividades requeridas por el proceso.
- iii. Disponer de reproducibilidad de las distintas etapas del proceso.

7.1.1. Entregas por submuestra

Tal como se detalló en la sección sobre Diseño muestral, si bien la muestra es una sola en términos de diseño, esta fue dividida en siete submuestras operativas, para organizar el trabajo de recolección, esto permitía contar con entregas periódicas durante los cuatro meses que duró el subproceso de recolección de datos.

Lo anterior facilitó la construcción del procesamiento desde un inicio. Esta construcción fue un trabajo iterativo, que lleva a mejoras continuas durante todo el proceso. Al terminar la recolección de datos se contó con la totalidad de la muestra, por lo que, desde ese momento, las actividades de procesamiento fueron realizadas con la totalidad de los datos, no particionado por submuestra operativa.

7.1.2. Nomenclatura de las bases de datos

A continuación, se detallan los lineamientos establecidos en el estudio con el objetivo de estandarizar la denominación e identificación de las bases de datos, alcanzando la trazabilidad de los insumos de salida de los procesos de producción estadística.

Para la II ENUT se generó una estructura estándar en la nomenclatura de los datos resultantes del procesamiento de datos, que incluye la información en base a un patrón establecido y reconocible, independiente de las características y temporalidad de la operación estadística.

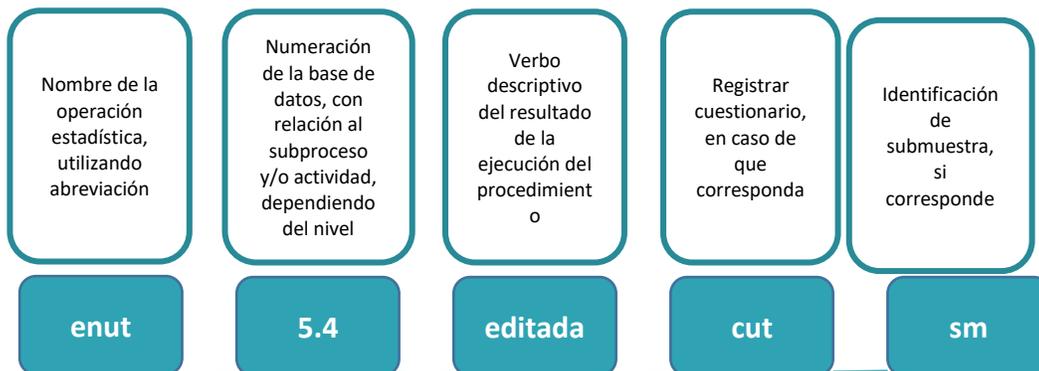
1. Se utilizó solo minúsculas y los espacios fueron representados por guiones bajos (“_”).
2. Nombre de la operación estadística (OE), utilizando la abreviación en minúscula (ej.: epf, enut, encla, eme, etc.).
3. Número del subproceso (segundo nivel) o actividad ejecutada (tercer nivel) (ej.: 5.4 o 5.4.1), dado el GSBPM.
4. En el caso de la numeración a nivel inferior a subproceso, como la de actividades (tercer nivel), se mantuvo el número del segundo nivel al que pertenece la acción que se está realizando, y se incorporó numeración consecutiva en el tercer nivel, aun cuando sean bases de datos intermedias previas a la que termina con el subproceso (ej.: la base de datos 5.4.1 y 5.4.2 serían intermedias del subproceso 5.4). De esta forma, se facilita el identificar la

etapa (subproceso) a la que pertenece el objeto intermedio (tercer nivel), en término de las acciones necesarias para llegar al objeto final del mismo (ej.: 5.4).

5. Se utilizó un verbo descriptivo de lo que se logra al crear el objeto. Utilizando el participio pasado femenino singular de la acción realizada, como adjetivo (ej.: codificada, editada, etc.).
6. Cuando corresponde, se agregó el cuestionario o sección específica, utilizando la abreviación del instrumento, cuestionario, módulo o sección de recolección.
7. Por último, para las bases de datos intermedias que requerían identificar la submuestra operativa, se utilizó la abreviación “sm” seguida del identificador numérico específico (ej.: sm1, sm2, sm99)

A continuación, se presenta un ejemplo de esta nomenclatura aplicado a una base de datos, enut_5.4.4_editada_cut_sm2.

Figura 8. Ejemplo de Nomenclatura de objetos bases de datos



7.2. Subprocesos del procesamiento

El procesamiento de datos al igual como otras partes del proceso de producción estadística tuvo como marco el modelo GSBPM, explicado en mayor detalle en el Anexo 13.1 de este documento.

A continuación, se detallan las principales actividades realizadas por subproceso. Es importante recordar que, si bien estos se describen de manera secuencial, muchas de estas actividades son iterativas y desde un subproceso posterior es posible volver a uno anterior si así se requiere.

7.2.1. Integración

Los datos exportados directamente desde el sistema informático de *Survey Solutions* vienen en distintas tablas separadas dada una lógica interna de la plataforma, que no cumple lo requerido para trabajar los datos con una estructura eficiente para las actividades de procesamiento. Por lo anterior, la integración de datos organiza y estructura los datos según los requerimientos del equipo técnico.

i. **Descarga de datos originales (brutos)**

El encargado nacional de recolección informaba al encargado del equipo técnico el término de la recolección de cada submuestra operativa, lo que permitía la descarga de los datos recolectados por parte del equipo técnico directamente desde el *Headquarters* (HQ) de *Survey Solutions* ³⁵.

ii. **Verificación de existencia de datos**

Los analistas socioeconómicos encargados de la integración verificaban la existencia de los elementos esperados en la descarga de datos.

En la práctica no se encontraron errores en la descarga de datos durante el período de recolección desde el HQ.

iii. **Unión de tablas CH y CUT**

³⁵ *Headquarters*: plataforma que permite ver reportes para evaluar el progreso general de la encuesta, herramientas de aprobación o rechazo de entrevistas, administrador de roles y equipos de trabajo (personas encuestadoras, supervisoras), carga y cuestionarios y descarga de datos recolectados.

Se unificaron las distintas tablas que contenían información del CH y CUT, además de sus hojas de ruta, construyendo un formato trabajable en términos del procesamiento que llevó a cabo el equipo técnico.

Esta base de datos unificada solo contiene los hogares donde se comenzó una entrevista, puesto que contiene la información de ambos cuestionarios.

iv. **Llenado de campos vacíos por flujo**

Una de las ventajas identificadas al utilizar DMC es que permite construir saltos/flujo más complejos, que ayudan a reducir las preguntas realizadas a los informantes si no son necesarias³⁶. Por esta razón, en integración se completó esta información faltante.

v. **Consolidación**

Luego, se unifican los datos que se encontraban separados por submuestras operativas, en una sola base de datos. En esta, se realiza el ordenamiento de las variables, para su uso más eficiente durante las siguientes actividades del procesamiento.

vi. **Revisión de unidades válidamente recolectadas**

Por último, se realiza una revisión de los datos consolidados, para evitar la presencia de casos duplicados u hogares erróneamente gestionados.

vii. **Construcción de base de Códigos de disposición**

En paralelo, se construye una base de datos que contiene información respecto a la gestión de las unidades muestrales, incluyendo aquellas viviendas no entrevistadas, que permite

³⁶ En ocasiones, la información ya dada por los informantes posibilitaba deducir las respuestas de algunas preguntas, lo que permite generar un salto que evite preguntar por estas durante la entrevista, reduciendo la carga a la persona entrevistada.

llevar control del estado de cada unidad, y el cálculo de los indicadores de trabajo de campo, tales como las tasas de respuesta, logro, elegibilidad, contacto, entre otras.

El elemento resultante de la integración de datos entrevistados es una base de datos llamado enut_5.1_integrada.RDS, que contiene una fila/observación por persona integrante del hogar. Tal como se indicó en las actividades del subproceso de integración, este elemento contiene solo las entrevistas logradas (tanto parciales como completas).

7.2.2. Clasificar y codificar

La codificación de datos obedece principalmente a dos baterías de preguntas que capturan información que debe ser clasificada en categorías estandarizadas.

- **Ocupación:** para esto se utiliza el Clasificador Chileno de Ocupaciones (CIUO 08.CL³⁷). Este es el estándar oficial para el procesamiento y análisis de las ocupaciones, permitiendo organizar estas en relación con las tareas que realizan los informantes ocupados y las competencias que requiere para realizarlas.
- **Rama de Actividad Económica:** para esto se utiliza el Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES³⁸). Este clasificador nos permitió codificar las actividades económicas de las empresas o

³⁷ Para más información respecto a CIUO 08.CL, véase: www.ine.gob.cl/docs/default-source/buenas-practicas/clasificaciones/ciuo/clasificador/ciuo-08-cl.pdf.

³⁸ Para más información del CAENES, véase: [www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/otros/documento-clasificador-de-rama-de-actividad-econ%C3%B3mica-caenes-\(abril-2016\).pdf?sfvrsn=1b1cc016_3](http://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/otros/documento-clasificador-de-rama-de-actividad-econ%C3%B3mica-caenes-(abril-2016).pdf?sfvrsn=1b1cc016_3).

instituciones en las que se desempeñan los informantes respecto a su ocupación principal.

Luego de la integración de datos, los datos resultantes (enut_5.1_integrada) se derivan tanto a Clasificación y Codificación, como a Revisar y Validar en paralelo.

La clasificación se realizó por medio de algoritmos de codificación automatizada por parte del equipo técnico, además de distintas etapas de control de calidad que se describen a continuación.

i. Codificación automatizada

Se realizó la codificación de ambos elementos por medio de dos modelos de clasificación diferentes. El primero fue por medio del mecanismo llamado *Gated Recurrent Units* (GRU) de *deep learning* y, en segundo lugar, con algoritmos de *Support Vector Machine*, ambos modelos fueron entrenados con información proveniente de la Encuesta Nacional de Empleo (ENE), y en el caso del primer modelo también toma algunos casos provenientes de la prueba piloto 2020 en DMC de la Encuesta de Presupuestos Familiares (EPF).

ii. Revisión y codificación manual

Por parte de los analistas socioeconómicos encargados del CH se revisaron y codificaron aquellos casos que ninguno de los modelos pudo clasificar.

iii. Auditoría interna (equipo técnico)

Se revisaron y recodificaron los casos con probabilidad predicha según los modelos más baja (bajo 80%). Además, se seleccionaron de forma aleatoria 6% de los casos codificados para su revisión manual. Esta última acción sirvió como medio de control interno de calidad para contrastar la codificación automática con la codificación manual.

iv. Auditoría por equipo externo al proyecto

Se realizó una auditoría por la Sección de Nomenclaturas del INE, los casos auditados fueron de 2.500 respuestas considerando CAENES y CIUO, lo que equivale a cerca del 6% del total de casos. Esta auditoría reveló algunos puntos críticos con respecto al proceso de clasificación y codificación de la II ENUT, los que se pueden resumir en los siguientes puntos: desempeño de los codificadores y/o sistemas de codificación, instrumentos de clasificación y calidad de las respuestas abiertas.

En cuanto al sistema de clasificación, se sugiere entrenar mejor los modelos automatizados de GRU y SVM para corregir posibles desviaciones de los modelos. En cuanto a la codificación manual, se sugiere realizar retroalimentación para reforzar buenas prácticas y poder identificar áreas de mejora. También, para próximos procesos de la encuesta, se sugiere establecer o dar énfasis a protocolos de la recolección de información en terreno, ya que en algunos casos la información recolectada en las glosas era poco precisa o incompleta.

Las recomendaciones anteriores derivaron en dos acciones:

- a. Reemplazar códigos enviados a auditar por aquellos ajustados por el equipo auditor. En este paso se incluyeron codificaciones en los cuales no había conformidad con la clasificación realizada del equipo ENUT
- b. En base a las recomendaciones basados en las codificaciones con estado No conforme y con el estado Conforme con observaciones, se tomaron los criterios dados por la auditoría para realizar una recodificación por parte del equipo técnico del proyecto, de los casos similares de la base de datos total, que no fueron

auditados. La auditoría realizada por el equipo externo detalló las principales secciones y áreas con falencias en la codificación, pudiendo aplicar estas recomendaciones en otros registros no auditados en esta oportunidad.

v. Revisión de casos codificados automáticamente, no coincidentes

Por último, los casos donde ambos modelos clasificaron con distintos códigos se revisaron manualmente para decidir el código final asignado.

7.2.3. Revisar y validar

Los subprocesos de “Revisar y validar” junto a “Editar e imputar” son los de mayor interacción en un flujo iterativo, ya que luego de cada actividad importante de edición de datos, le sigue una actividad de validación de los datos editados.

Por lo anterior, para facilitar la explicación de las actividades clave de este subproceso se describirán las validaciones en términos globales.

i. Malla de validaciones

a. Validaciones básicas de primer nivel

Se construyó una malla automática de validaciones básicas, buscando inconsistencias en términos de **rango, campos en blanco y saltos**.

b. Validaciones de segundo nivel

Estas validaciones buscan identificar **inconsistencias** y **advertencias**. Las inconsistencias se tipificaron como aquellas situaciones en base a interacción de distintas variables recolectadas que a priori no deberían encontrarse. Por su parte, las advertencias se entendieron como situaciones de interacción

entre variables que son poco plausibles, pero se estimó necesario revisar para corroborar en conjunto a la entrevista en su totalidad. La evaluación de la malla de validaciones, tanto de primer como segundo nivel, se aplicó luego de cada actividad importante de edición de datos, de forma iterativa, de tal manera de evaluar que las ediciones corrigiesen distintas situaciones identificadas por las validaciones y, en segundo lugar, permitieron identificar inconsistencias que pudiesen producir de forma involuntaria cualquier edición realizada para su posterior corrección.

ii. Grilla técnica de mínimos de suficiencia y calidad

Si bien durante la recolección se cuenta con distintas herramientas de control de calidad y supervisión, durante el procesamiento se realiza una evaluación con mayor profundidad en cuanto a dos factores, **suficiencia**³⁹ y **calidad temática**⁴⁰.

Esta Grilla de técnica se divide en la evaluación respecto a la información del Cuestionario del Hogar (CH) y para el Cuestionario sobre Uso del Tiempo (CUT).

a. Grilla técnica CH

El CH evalúa la presencia de información en las preguntas que se consideraron clave para poder entrevistar correctamente el cuestionario temático (CUT), estas variables fueron:

- Parentesco: c2
- Edad: c3
- Sexo: c4

³⁹ Suficiencia: entenderemos suficiencia como la evaluación respecto a la completitud del cuestionario evaluado en términos de las preguntas habilitadas por flujo para cada entrevista.

⁴⁰ Calidad temática: en este contexto, entenderemos calidad como una evaluación temática respecto a la información registrada, donde se cuente con mínimos respecto a cierta información específica y crítica.

- Asistencia a un establecimiento educacional: e4
- Ocupación: o3, o4 y o5
- Personas en Situación de Dependencia Funcional (PSDF): d1 a d9⁴¹

Los CH que no contuvieran la información válida en alguna de las preguntas previamente indicadas, se reclasificó a no logro⁴².

b. Grilla técnica CUT

En el caso de la grilla técnica asociada al CUT, cuenta con tres evaluaciones realizadas en momentos diferentes del procesamiento de datos:

1. **Suficiencia 50% centrales y 50% esenciales:** se evalúa que los CUT cuenten con al menos el 50% de información válida desde el Módulo TO a SIM, es decir, variables de participación y tiempo e información asociada a ella que afecte el flujo de su aplicación. En segundo lugar, se evalúa que exista información válida en el 50% de los campos identificados como esenciales, correspondiente a aquellos de mayor relevancia para los indicadores principales de la encuesta, asociadas a Trabajo en la ocupación, Trabajo doméstico y de cuidados no remunerado (TCDNR).
2. **Suficiencia actividades por día:** se le exige 4 actividades mínimo por día con información de participación y tiempo. Esta evaluación se aplica posterior a un proceso de edición/imputación por sustitución respecto a la información

⁴¹ En la base de datos se podrán encontrar algunos casos que no cuenten con información en d8 (con valor 96, de dato faltante), que fueron justificados al ser hogares unipersonales a los que no les afecta en la aplicación del CUT.

⁴² Es importante recordar que, si por flujo no se requiere la contestación de una pregunta concreta, esto sigue siendo un valor válido. Como sucede con la pregunta d9, que no todos deben contestar.

que se pueda encontrar en el cuestionario y sus observaciones de campo.

3. **Evaluación de tiempos mínimos y máximos por día:** por último, se evalúan los tiempos mínimos y máximos en los días de referencia entrevistados. Se exige un mínimo de 10 horas y un máximo de 48 horas en esos días. Esta evaluación se realiza posterior a los procesos de identificación y edición de casos atípicos de tiempo en el CUT.

iii. Revisión de parادات

Un elemento importante en el desarrollo de la II ENUT fue la utilización de los parادات como fuente de información para la validación y edición durante el procesamiento de datos.

Los parادات, son datos respecto al proceso de respuesta de la encuesta en sí mismo (Callegaro, 2013). En el caso de la II ENUT, que se desarrolló por medio de *Survey Solutions*, los parادات que se pueden descargar del sistema entregan información sobre el orden en que se responde una pregunta, la fecha y hora de registro, además del momento en que se inicia o sincronizan las respuestas para enviar la información hacia los servidores (Olson y Parkhurst, 2013). Pudiendo rastrear cada registro realizado y sus modificaciones, lo que permite identificar información con la que no se cuenta en los datos finales registrados en la descarga de la encuesta.

Lo anterior dio acceso a los registros previos de las distintas preguntas, pudiendo generar un historial que ayudó en las decisiones de edición respecto a casos complejos de información inconsistente o poco plausible.

7.2.4. Editar e imputar

Las ediciones e imputaciones realizadas durante el procesamiento de datos de la II ENUT buscan **justificar, corregir, reemplazar o validar** información que se identifica como inválida en base a los datos. Para esto se utiliza información adicional, tal como observaciones y comentarios de las personas encuestadoras, la coherencia con otras variables y módulos, además del uso de parados recopilados durante la recolección.

A continuación, se listan las distintas actividades que se realizaron en términos de ediciones e imputaciones.

i. Ediciones por validaciones primer nivel CH y CUT

Se realizaron las ediciones en base a las validaciones de primer nivel descritas en la sección anterior. Estas ediciones en su mayoría obedecen a llevar el dato a inválido o vacío (dependiendo el flujo de cuestionario), con excepción de casos justificados por medio de información contextual, como observaciones de los equipos de terreno.

ii. Reclasificación de la respuesta de CH por suficiencia

Estas ediciones generan la reclasificación a no respuesta (no logro) de los Cuestionario del Hogar que no pasan los criterios de mínimos de suficiencia y calidad evaluados en la grilla técnica.

iii. Ediciones por validaciones segundo nivel CH

Estas ediciones corresponden a cambios en los datos inválidos del CH dada su interacción con otras partes del cuestionario, utilizando la información adicional de la entrevista, tal como para aquellas ediciones por validaciones de primer nivel.

iv. Ediciones por revisiones de parados CUT

El cuestionario temático sobre uso del tiempo debía levantarse en su completitud durante una sola entrevista⁴³, sobre todo por la información referente a participación y tiempo en la realización de actividades. Si bien en la práctica se permitió registros en otro momento, estos debían ser en su mayoría en módulo que no registran tiempo ni participación en actividades, tales como el de Bienestar Subjetivo (BS) o Características específicas de la persona (ID). Para corroborar esto, se realizaron revisiones de los parados del CUT, pudiendo evaluar los momentos concretos de registro de la información en cada pregunta.

Con la evaluación anterior se invalidaron los campos del CUT registrados en una fecha diferente a la entrevista mayoritaria (en la que se registró más información central del cuestionario).

v. Ediciones por validaciones segundo nivel CUT

Estas ediciones corresponden a cambios en los datos inválidos del CUT dada su interacción con otras partes del cuestionario y con características de los encuestados provenientes del CH. Para esto se utilizó información contextual de la entrevista, al igual que para las ediciones del CH.

vi. Reclasificación de la respuesta de CUT - por suficiencia 50%

En base a lo descrito en la sección Revisar y validar 9.2.3, en el apartado de grilla técnica, se evalúa los mínimos de suficiencia y calidad, para en esta etapa de edición reclasificar como no respuesta los CUT que no cumplen con un mínimo de información central y esencial del cuestionario (50% de estos campos).

vii. Ediciones de casos atípicos de ingresos

⁴³ Esto para resguardar la coherencia interna de la información registrada con el mismo periodo de recuerdo.

Luego de contar con las ediciones en base a las validaciones de primer y segundo nivel, se realizaron los análisis de casos atípicos de los ingresos registrados en el CH para personas de 15 años o más.

El método utilizado para detectar ingresos atípicos fue *adjusted outlyingness* (AO), el cual tiene en cuenta la distancia y dirección de los ingresos en el espacio de datos y realiza un ajuste basado en la escala y dispersión de estos. Este método se utilizó para los ingresos de trabajadores dependientes e independientes.

Si bien se utilizó el método AO univariado, se generaron perfiles para determinar los casos atípicos dentro del contexto propio de cada persona trabajadora. En este sentido, se evalúa la distribución de ingresos para personas que comparten las mismas características y se le asignó un puntaje AO a cada observación según su ubicación en la distribución.

Las **variables utilizadas** para la generación de los perfiles fueron: nivel educacional, situación en el empleo, clasificación de la ocupación, sexo, tramos de edad, actividad económica del lugar de trabajo (rubro) y región. Se estableció que cada grupo constara de al menos 20 casos para tener una distribución de los datos suficiente para aplicar el método. En caso de no cumplir con este criterio, se fue relajando la inclusión de variables de caracterización privilegiando las que tienen mayor correlación con los ingresos, así se aseguró que todas las observaciones tengan un puntaje de AO según su tipo de ingreso.

Se definió el puntaje de corte a partir del cual fue catalogado como atípico según cada perfil utilizando la metodología descrita en Hubert & Van der Veeken (2008). Las observaciones que fueron catalogadas como atípicos se revisaron de forma manual para evaluar si existía justificación para tales valores, dado el contexto del informante y comentarios dejados por la persona encuestadora. En caso de tener fundamentos

para considerar que existe un error en el monto, estos valores se pasaron a perdidos.

viii. Imputación de ingresos

Las preguntas que recopilan información respecto a ingresos de los integrantes del hogar, en este caso de 15 años o más, presentan cierto nivel de no respuesta, por lo que se decidió tratar estos casos por medio de imputación de los datos faltantes. Incluyéndose aquellos casos pasados a perdidos en la actividad anterior de edición de casos atípicos en ingresos.

El primer paso corresponde a la evaluación del nivel de no respuesta presente en cada partida de ingresos.

Tabla 28. No respuesta en ingresos

Ingresos	Observaciones con respuesta	Observaciones con datos imputados	% Datos imputados	N total
Dependientes	13.006	3.239	19,9%	16.245
Independientes	4.183	1.160	21,7%	5.343
Jubilados	6.191	623	9,1%	6.814

Con estos niveles de respuestas faltantes se realizaron pruebas por medio de simulaciones para evaluar los resultados de la imputación de estos datos. Para estas simulaciones se tomaron todos los casos con información válida en el módulo de ingresos y artificialmente se eliminaron datos en base a un modelo de no respuesta⁴⁴. Luego de este

⁴⁴ Estas eliminaciones solo fueron realizadas para efecto de los ejercicios de simulación, en ningún caso se eliminaron de forma permanente respuestas de los respondientes de forma arbitraria.

borrado momentáneo de datos, se procedió a probar dos métodos de imputación, **hot deck condicional**⁴⁵ y **mediana condicional**⁴⁶.

Se realizaron 1.000 iteraciones (simulaciones) para cada método, con el fin de evaluar la distribución de las medias de ingresos en las partidas de ingresos imputada. Este ejercicio permitió comparar estas medias imputadas de ingresos con los datos originales, previo a la eliminación de datos para probar las imputaciones. El rendimiento del método *hot deck* fue mejor que el de mediana condicional bajo dos parámetros. (i) En términos de su acercamiento a la media y distribución de ingresos previo a la eliminación artificial, *hot deck* presentó mejor comportamiento. (ii) Adicionalmente, al evaluar los cambios en la variabilidad de los datos originales, *hot deck* presentó menos cambios en esta, lo que tiene sentido, dado que la media condicional utiliza una medida de tendencia central dentro de cada *cluster* para realizar su imputación/elegir donante, lo que lleva a reducir de forma artificial la varianza de los datos.

Para la realización de la imputación fue necesario construir perfiles por características sociodemográficas de los informantes, respecto a variables que teóricamente correlacionan con los ingresos, de tal manera de buscar donantes dentro de los datos de la misma encuesta que contaran con características similares a quien recibiría la imputación. Las variables consideradas para esto fueron las contenidas en la Tabla 29:

⁴⁵ *Hot deck* condicional: este método, condicionado a ciertas variables sociodemográficas de los respondientes, genera agrupaciones o *clusters* de personas que comparten estas características dentro de la misma encuesta, seleccionando de forma aleatoria un donante dentro de estos grupos.

⁴⁶ media condicional: método que al igual que *hot deck*, condiciona la agrupación de las personas en base a características explicativas de los ingresos, pero en este caso selecciona la mediana dentro de cada *cluster* para elegir el dato donado.

Tabla 29. Variables de agrupación para búsqueda de donante en la imputación de ingresos

Variable	Tipo Variable	Tipo		
		Dependientes	Independientes	Jubilados
Edad	Teórica	✓	✓	✓
Escolaridad	Teórica	✓	✓	✓
Sexo	Teórica	✓	✓	✓
Nivel socioeconómico	Teórica	✓	✓	✓
Región	Teórica	✓	✓	✓
Nacionalidad	Empírica	✓	✓	✓
Sustentador/a principal	Empírica	✓	✓	✓
Presencia menores 15 años	Empírica	✓	✓	-
Situación en el empleo	Empírica	✓	✓	-
Clasificación de la ocupación	Empírica	✓	✓	-
Actividad económica del lugar de trabajo	Empírica	✓	✓	-

Se entiende como **variables teóricas**, aquellas que, dados distintos trabajos y modelos académicos, presentan cierta correlación con los ingresos de la ocupación. La teoría del capital humano, la teoría de señalización (Weiss, 1995) y la de ciclo de vida (Mincer, 1974). Por su parte, las **variables empíricas** se entienden como aquellas que caracterizan a quien recibe ingresos adicionales a las teóricas y que dados distintos análisis estadísticos se observa una correlación con los ingresos.

Por último, debido a las exigencias en la selección de donantes, en términos de la especificidad de las características buscadas, se fueron flexibilizando los criterios de búsqueda, utilizando niveles más agregados de alguna de las variables de características de búsqueda o eliminándolas. Esta flexibilización en la búsqueda de donante fue paulatina, solo se dio si no era posible encontrar donante con los criterios

iniciales. En tanto, la selección de las variables a flexibilizar fue dada por la correlación con la variable de interés (ingreso).

ix. Reclasificación de la respuesta de CUT - por cantidad de actividades por día

El siguiente criterio por el cual se evaluó la calidad de los CUT fue la cantidad de actividades registradas en un día, donde se esperan al menos 4 actividades diarias, considerando dormir y comer. Aquellos casos donde en un día contara con 3 o menos actividades sin justificación fue reclasificado como no logrado (rechazo implícito).

La exigencia de un mínimo de actividades diarias busca limpiar los datos de casos donde la persona informante deja de tener interés en responder correctamente y comienza a dar respuestas negativas a cada actividad por la que se le consulta.

Es importante mencionar que tal como se indicó previamente, los casos invalidados solo fueron aquellos que no contaban con una justificación clara para declarar menos de 4 actividades. Por ello, se realizó una revisión exhaustiva de los casos identificados a través de la grilla técnica. Donde los casos que declararon una situación excepcional en dicho día, o que, dadas sus características sociodemográficas, tal como su ocupación, permitieran entender la situación se mantuvieron.

x. Ediciones por casos atípicos de tiempo

El listado de actividades con que fue medido el uso del tiempo permitió que cada persona declarase la duración en minutos, considerando como rango de respuesta válida desde un minuto hasta las 24 horas del día. No obstante, la experiencia de la ENUT 2015 contribuyó a programar avisos de alerta dentro del cuestionario en *Survey Solutions*, en circunstancias donde la respuesta calificaba como posible valor atípico. En dichas situaciones la persona encuestadora, debió corroborar nuevamente con

el entrevistado o entrevistada la duración que se había señalado inicialmente.

Tras esto, y concluida la recolección de la información en terreno, fue necesario evaluar el comportamiento de los datos. Los resultados de las medidas de tendencia central y dispersión constataron valores anómalos para las distintas actividades consultadas. A partir de ese análisis, se procedió a definir un modelo estadístico de identificación de casos atípicos, optando por el de Hadi, dada su capacidad de incorporar diversas variables del informante o de contextualización en la estimación de los casos. Siendo el mismo método utilizado durante la ENUT 2015 con buenos resultados.

Los perfiles de informantes utilizados se obtuvieron luego de testear distintas variables de clasificación, en perspectiva a las actividades evaluadas. Sexo y tramos de edad, aportaron en ese sentido los resultados más consistentes tanto en la cantidad de casos atípicos detectados, como en los umbrales de tiempo que separaban los casos típicos de los atípicos. Solo en las actividades de cuidado no remunerado, el modelo más eficiente incorporó sexo, tramos de edad y estrato económico.

Definidas las variables de agrupamiento de las respuestas de tiempo, el modelo identificó los casos atípicos en función de la distancia de las observaciones al centroide de cada perfil del informante. Desde ese punto, y con la finalidad de restringir la cantidad de casos totales a editar, se generó un segundo corte dentro de los casos atípicos identificados inicialmente, estandarizando las distancias calculadas por el modelo de 0 a 1, y focalizando la revisión en los casos donde la distancia estandarizada resultara mayor a 0,5.

Únicamente en aquellas variables con una cantidad de casos atípicos igual o inferior a 200 casos, se examinó la totalidad de los valores indicados como extremos por los modelos.

Respecto de la revisión final de las respuestas para decidir si se mantenía o no el valor atípico, la comprobación se realizó en base a los datos del módulo de contextualización, las características del(a) encuestado(a), otras respuestas entregadas por él (ella) y los registros del módulo de observaciones, donde el personal de campo tenía instruido registrar cualquier hecho complementario que explicara situaciones atípicas.

xi. Reclasificación de la respuesta de CUT - por suficiencia de tiempo mínimo y máximo

El último paso dentro de la edición e imputación en el procesamiento fue la reclasificación a no respuesta de los CUT que no cumplieran con un mínimo de 10 horas por día⁴⁷, o superaran un máximo de 48 horas diarias de realización de actividades⁴⁸.

Es importante recordar que la ENUT no es exhaustiva, es decir, no pregunta por todas las actividades posibles de realizar por las personas, por esta razón es posible encontrar días de menos de 24 horas. Por otro lado, esta es una encuesta declarativa, por lo que es esperable que los tiempos estimados por cada persona entrevistada no sean exactos.

Otra razón que ayuda a entender los tiempos superiores a 24 horas es la existencia de la realización de actividades simultáneas y casos de personas con jornadas laborales extensas, donde si bien se declara

⁴⁷ Tal como en la ENUT 2015 se evaluó la distribución de tiempo total por día, donde se permitieron tiempos diarios no menores al percentil 1 de lo recolectado en la encuesta.

⁴⁸ Existen casos justificados dado su contexto y observaciones en terreno, que pueden superar las 48 horas.

trabajar, por ejemplo, 24 horas, dentro de ese periodo se suele comer, dormir o incluso realizar trámites entre otras actividades.

7.2.5. Derivar nuevas variables y unidades

El subproceso de Derivación de nuevas variables construye variables que toman como base aquellas provenientes de la recolección y el procesamiento de datos, y que son requeridas para los distintos análisis a realizar por parte del propio equipo técnico, como también de los usuarios externos.

Las distintas variables creadas se pueden agrupar en tres categorías:

- i. **Sociodemográficas:** tales como aquellas variables de nivel educacional derivadas de las preguntas del módulo E del Cuestionario del Hogar (CH), Codificación de ocupación, actividad económica del lugar de trabajo, tramificación de la edad de los informantes y tipologías de hogar.
- ii. **Ingresos:** se construyeron las variables de ingresos, en base a las variables recolectadas del módulo y a nivel de persona, tales como ingresos del trabajo, ingreso de la ocupación principal y jubilaciones. Por otra parte, se construye aquella variable agregada de ingresos totales del hogar, per cápita del hogar y la variable que identifica los casos que fueron imputados por datos faltantes, véase sección 9.2.4 .
- iii. **Variables de Tiempo y Participación:** los distintos tiempos y participaciones en actividades diarias capturados en el CUT, permiten generar distintos constructos de actividades en temáticas que es posible agregar para su posterior análisis.
En este sentido, se construyen agregaciones como la suma de los tiempos destinados a trabajos de cuidados a niños y niñas, de cuidados a adultos mayores, cuidados a personas en situación de

dependencia funcional (PSDF), tiempos de traslados por educación, traslados por trabajo en la ocupación, tiempo en actividades educacionales, trabajo doméstico en el hogar, entre otros. Junto a si destinó al menos un minuto en alguna de estas actividades o no (participación).

Todas estas estimaciones se realizan tanto para el día de semana y fin de semana, por los que se consultan en el cuestionario. Adicionalmente, se construyen estimaciones para el **día tipo**. Este concepto es un constructo en base a la información del día de semana y fin de semana declarados por los informantes, realizando una suma ponderada entre ambos, entregándole un peso de 2/7 al tiempo declarado en el día de fin de semana y de 5/7 al día de semana.

$$\begin{aligned} & \left(tiempo\ dia\ semana_i \times \frac{5}{7} \right) + \left(tiempo\ dia\ fin\ de\ semana_i \times \frac{2}{7} \right) \\ & = tiempo\ día\ tipo_i \\ & i = actividad\ i\ realizada\ por\ la\ persona\ encuestada \end{aligned}$$

Es importante indicar que el **día tipo** que se utiliza en la II ENUT, no refiere a lo que se conoce como “día típico”, que es un concepto que refiere a habitualidad⁴⁹, y no se utiliza en este estudio.

⁴⁹ Este concepto se utiliza en algunas mediciones ajenas a este estudio, donde se consulta por actividades realizadas en un día habitual (promedio), y no se utiliza un día de referencia específico.

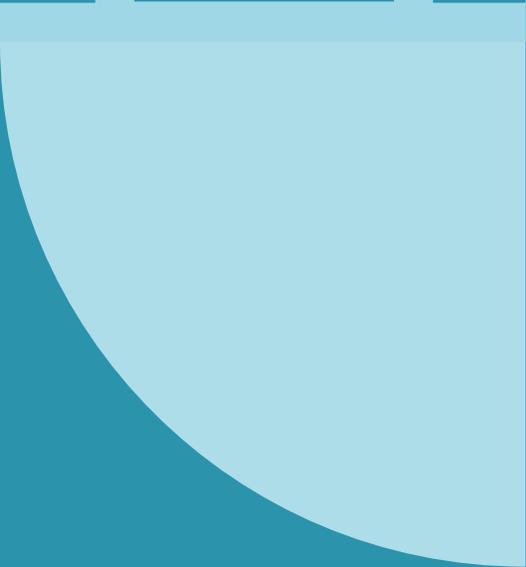
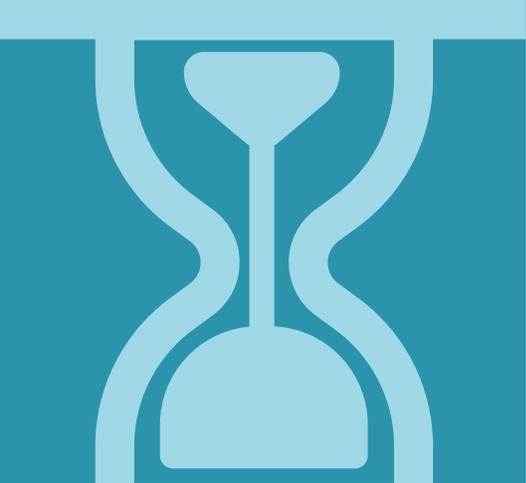
7.2.6. Calcular ponderadores

El equipo técnico del proyecto, además de trabajar los datos centrales de los cuestionarios recolectados, procesó los datos de las hojas de ruta del cuestionario, que recopilaban información del estado final de las unidades encuestadas, lo que permitió generar el insumo necesario para que se construyesen las variables requeridas para poder realizar las estimaciones considerando el diseño muestral complejo del estudio, que se explica en detalle en la sección 7.4.

En este subproceso se calcularon los factores de expansión (ponderadores o pesos) requeridos por el estudio, además de los pseudo-conglomerados de varianza requeridos.

7.2.7. Finalizar los archivos de datos

Finalmente, se consolida la base de datos que da término al proceso de procesamiento (enut_5.4_editada), verificando que sus variables se encuentren en los formatos correctos y con los valores esperados luego del proceso, resguardando que los últimos pasos de edición no hayan afectado otros datos que no debían ser modificados.



8

METODOLOGÍA DE ANÁLISIS DE RESULTADOS

8. Metodología de análisis de resultados

8.1. Preparar borradores de resultados

Una vez se cuenta con la base de datos validada y editada, junto con los factores de expansión y los pseudo-conglomerados de varianza, se procede a generar las estimaciones planificadas en el diseño del análisis de resultados considerando el diseño complejo de la encuesta.

En esta etapa se generaron cálculos espejo por distintos analistas del equipo técnico que permitieron corroborar el correcto cálculo de las estimaciones.

Junto a lo anterior, esta etapa permite identificar errores en las variables creadas/calculadas en el subproceso de procesamiento. Si se identificaba alguna variable con comportamiento no esperado se revisaron tanto los cálculos como la propia creación de variables en el subproceso anterior. Si el error provenía de la creación de la variable utilizada, se realizaron los ajustes pertinentes, siendo un proceso iterativo.

En esta etapa es donde se comienza a construir el informe de principales resultados que acompaña la publicación de los microdatos del estudio.

8.2. Validar los resultados

En paralelo a la preparación de resultados, el equipo técnico comparó en ciertas características sociodemográficas e ingresos los datos obtenidos por la II ENUT, con otras fuentes, tanto producidas por la propia institución, como con operaciones estadísticas externas a ella.

El primer paso fue identificar productos clave que pudieran dar cuenta de que la realidad sociodemográfica que se está caracterizando estuviese en línea con lo encontrado en otros estudios.

- **Encuesta CASEN 2022**

Con la encuesta CASEN se realizaron comparaciones respecto a nivel educacional, población activa e ingresos del trabajo (dependientes, independientes y otros ingresos del trabajo), en las que no se encontraron diferencias importantes en estas estimaciones. Para llegar a estas conclusiones se analizaron dominios similares (sector urbano) y se analizó las estimaciones de la media poblacional en conjunto con los intervalos de confianza. En particular, para el caso de comparar ingresos del trabajo, se utilizó la calculadora IPC para llevar los precios del periodo de recolección de la CASEN 2022 al mes central, Octubre, de la II ENUT.

- **Encuesta Suplementaria de Ingresos (ESI) 2023**

Con la Encuesta Suplementaria de Ingresos se compararon los ingresos provenientes del trabajo dependiente, independiente y otros ingresos del trabajo. En esta comparación a nivel de estimaciones medias a nivel poblacional, tampoco se observaron diferencias relevantes a nivel nacional respecto a lo encontrado por la II ENUT. Para realizar de forma más acertada la comparación, se seleccionaron las mismas comunas en las que la II ENUT y la ESI tienen presencia. Cabe mencionar que el periodo de referencia de la ESI es del último trimestre del año 2023, periodo de recolección similar al que la II ENUT.

- **Encuesta Nacional de Empleo (ENE)**

Por último, para realizar las comparaciones con la ENE, se consideraron los trimestres móviles de SON y OND⁵⁰, analizando

⁵⁰ SON: trimestre móvil de septiembre, octubre y noviembre. OND: trimestre móvil de octubre, noviembre y diciembre 2023, que abarcan el periodo de recolección de la II ENUT.

las diferencias en la proporción de la fuerza de trabajo de la población ocupada, desocupada y fuera de la fuerza de trabajo. Las comparaciones se realizaron a nivel nacional. Al respecto, se encuentran diferencias aceptables en contraste con los provenientes de la II ENUT.

Estos análisis de contraste consideran que cada producto estadístico cuenta con sus propias características de diseño muestral y de factores de expansión, por lo que, si bien no se espera que los resultados de estas comparaciones sean idénticos, si permite cerciorarse que no existen diferencias relevantes en características básicas de la población a la que se busca representar. Lo anterior se verificó revisando los intervalos de confianza de cada una de las estimaciones poblacionales.

8.3. Interpretar y explicar los resultados

En este subproceso se tomó como base los análisis realizados en la preparación de los borradores de resultados, buscando interpretar los estadísticos resultantes obtenidos en términos temáticos para, de esta manera, generar un relato lógico alrededor de estos.

Adicionalmente, esto permite evaluar la pertinencia en la presentación de algunos resultados en cuanto a si todos aquellos análisis planificados, son pertinentes en términos de relevancia y a aportar información adicional a la que otros cruces o análisis ya están mostrando.

Otro paso importante en paralelo a esta revisión de resultados es la evaluación de cada estimación presente en el plan de análisis en cuanto a los criterios de **calidad estadística para las estimaciones establecidos por la institución** en el “Estándar para la evaluación de la calidad de las

estimaciones en encuestas de hogares”, publicado en febrero del año 2020⁵¹.

El paso anterior es fundamental para asegurar que los datos presentados como resultantes de los análisis cumplen cierto mínimo de calidad estadística, y que las conclusiones que se puedan obtener se generen en base a datos que cumplan estos criterios. La aplicación del estándar no condicionó la publicación de las principales estimaciones asociadas a trabajo no remunerado a nivel nacional, regional por sexo.

8.4. Aplicar control a la divulgación

El subproceso de aplicación del control a la divulgación de datos busca que los datos publicados cumplan con el marco normativo de confidencialidad de la información entregada por los informantes, los cuales se encuentran amparados por la **ley de secreto estadístico N° 17.374 y de la ley N°19.628 sobre protección de datos**.

El INE cuenta con una guía para el control de la divulgación, que sistematiza los pasos que se deben seguir para asegurar un análisis de los riesgos de re-identificación de las personas informantes y las posibles mitigaciones que se pueden implementar para reducir este riesgo, cuidando la utilidad de los microdatos a publicar. Lo anterior, implica un trabajo iterativo de evaluación del riesgo, junto con la utilidad de los datos al aplicar métodos de anonimización.⁵²

⁵¹ Para más antecedentes de dicho estándar, este se puede encontrar en el siguiente enlace: <https://www.ine.gob.cl/docs/default-source/buenas-practicas/estandares/estandar/documento/estandar-evaluacion-de-calidad-de-estimaciones.pdf>

⁵² Guía para el control de la divulgación estadística en microdatos: https://www.ine.gob.cl/docs/default-source/buenas-practicas/estandares/estandar/documento/qu%C3%ADa-control-divulgaci%C3%B3n-estad%C3%ADstica-microdatos.pdf?sfvrsn=fb568638_2

Este análisis asume que puede existir una persona intrusa, mal intencionada que intente cruzar la información publicada por este estudio con otras fuentes, por medio de identificadores directos o indirectos/clave.

Los pasos para realizar este subproceso son los siguiente:

- **Definiciones previas al control a la divulgación**

Aquí se realizó una revisión de los insumos que se esperan publicar, las necesidades potenciales de las personas usuarias de los microdatos y las posibles fuentes adicionales que podrían utilizar usuarios externos para re-identificar a las personas informantes.

Es en esta etapa que se identificaron las variables clave o de identificación indirecta ⁵³.

- **Preparación y exploración de datos originales**

Aquí se eliminaron los identificadores directos, que refiere a aquellos que de forma evidente permiten individualizar a un informante, tales como nombres, números de identificación personal y direcciones. Esta supresión de datos evita que aquellos a publicar se vinculen directa e inequívocamente con una persona informante.

Luego se seleccionaron las variables relevantes (indirectas) para las personas usuarias finales, que debían, inicialmente ser incluidas en el conjunto de datos a publicar.

Habiendo seleccionado los datos relevantes a publicar, se identifican las variables que pueden servir de identificadores

⁵³ Variables de identificación indirecta: son aquellas variables que, si bien no identifican inequívocamente unidades estadísticas, en combinación se pueden vincular a información externa para reidentificar a los informantes en el conjunto de datos publicado. También se les denomina "cuasi-identificadores" o variables clave.

indirectos, que habitualmente son características sociodemográficas estructurales de los informantes, tales como sexo, edad, educación, ocupación, zona geográfica de residencia (región), entre otras.

- **Medir y evaluar riesgos**

Luego de los pasos previos, se evaluó el riesgo inicial, con los datos seleccionados como relevantes a publicar.

Para esta evaluación se utiliza el paquete del programa estadístico R *SDCMicro*, calculándose las siguientes métricas: k-anonimato, riesgo global y riesgos individuales.

- **Seleccionar, aplicar y evaluar**

Es en este punto donde comienza un trabajo iterativo, donde dado el riesgo inicialmente calculado, se aplicaron técnicas de anonimización sobre las variables de identificación indirecta para reducir las probabilidades de que una persona informante sea reidentificada por medio de estas.

El riesgo calculado se ve afectado por la cantidad de personas que presenten una combinatoria similar en las variables de identificación indirecta. Si solo existe una persona con ciertos valores combinados en estas variables, existe un mayor riesgo de reidentificarla. Por otro lado, mientras más personas presenten el mismo patrón, este riesgo se reduce, por no contar con características únicas que la individualicen.

Por lo anterior, para reducir el riesgo se aplicaron *técnicas no perturbativas*⁵⁴, recategorizando algunas variables a versiones

⁵⁴ Una técnica perturbativa es aquella que altera algunos de los datos de cierta variable, como la supresión local (eliminación de algunas celdas en una variable), o el *swapping*, que intercambia un dato de una fila a otra dentro de un grupo, para no alterar la estimación global resultante del uso de dichos datos. En el caso de la II ENUT, no se utilizaron este tipo de métodos.

menos desagregadas, tales como nivel educacional, clasificación de educación, entre otras. Además de no incluir en la base de datos a publicar algunas variables muy específicas, como en nacionalidad dar cuenta de quién es nacional o extranjero (pregunta c11a), pero no el detalle de la nacionalidad de dicha persona extranjera (pregunta c11b).

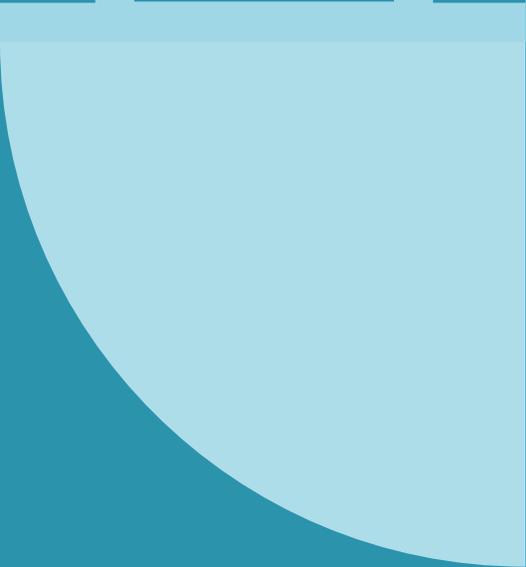
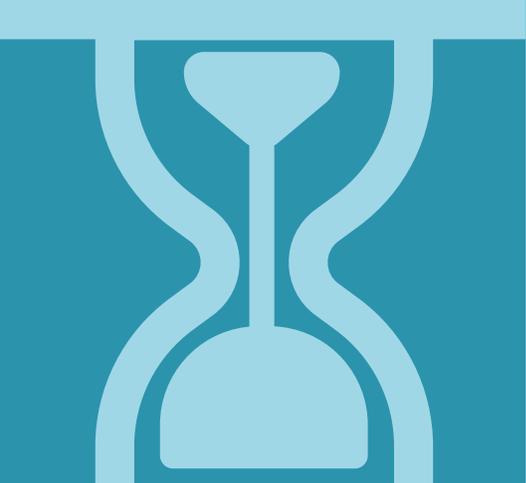
Esta etapa requirió de varias iteraciones, como se indicó al inicio de este apartado, de tal manera de reducir el riesgo lo máximo posible sin reducir la utilidad de los datos en sus variables más críticas para los posteriores análisis de personas usuarias. Para esto se trabajó a través del *método no perturbativo* de recodificación global.

- **Liberación de datos**

Por último, se crea la base de datos a publicar, que contiene el menor riesgo de identificación tanto global como individual y en la medida de k-anonimato usando los umbrales de 2k anonimato, 3k anonimato y 5k anonimato, todo esto con el fin de resguardar la utilidad de los microdatos presentes en el estudio.

8.5. Finalizar resultados

Luego de la aplicación del control de la divulgación, al contar finalmente con los datos que son posibles publicar, se ejecuta una revisión de los análisis realizados en el informe de principales resultados, para que se encuentren en concordancia con los datos seleccionados y validados.



9

DIFUSIÓN DE RESULTADOS

9. Difusión de resultados

9.1. Productos estadísticos para publicar

La publicación de la II ENUT consiste en los siguientes productos:

i. Base de datos II ENUT

Este elemento corresponde a los microdatos que contienen toda la información validada y editada, que se recopiló por medio de los cuestionarios del estudio, además de las variables derivadas creadas a partir de esta información inicial. Esta información es de carácter sociodemográfica, sobre los ingresos de las personas, el uso del tiempo que destinan a distintas actividades y su bienestar subjetivo en distintas dimensiones.

Esta base de datos se encuentra a nivel de personas integrantes de los hogares entrevistados. Cada observación o fila de esta es una persona.

La base de datos de la II ENUT estará disponible en formato de 'R', Stata y CSV.

ii. Informe de principales resultados II ENUT

Este documento presenta las estimaciones centrales del estudio, en cuanto los tiempos promedio destinados a grupos de actividades y la proporción de hombres y mujeres que las realizan, con desagregaciones a nivel regional para los principales indicadores. Adicionalmente, se presentan análisis respecto al bienestar de las y los respondientes, en base a preguntas de percepción.

iii. Manual de usuarios de base de datos II ENUT

El objetivo de este documento es entregar una guía para las personas usuarias de la base de datos, en base al programa estadístico R, de tal

manera de entregar una explicación de los datos, sus variables y de cómo se debe utilizar el diseño complejo del estudio para el correcto cálculo de estimaciones.

iv. Manual metodológico II ENUT

El presente documento detalla los procedimientos y decisiones metodológicas adoptados para garantizar el correcto uso de los datos, incluye antecedentes conceptuales del estudio y del diseño muestral, aspectos relevantes de la recolección, información acerca del procesamiento de los datos, así como también los métodos de análisis utilizados en la II ENUT.

v. Síntesis de resultados II ENUT

Documento que presenta en forma resumida los principales resultados del estudio.

vi. Cuestionarios II ENUT

Incluye dos cuestionarios, el Cuestionario del Hogar y el Cuestionario sobre Uso del Tiempo en su formato papel. Además, se pone a disposición la versión de *Survey Solutions* aplicada por medio de dispositivo móvil de captura.

9.2. Accesibilidad a los datos y marco legal

La publicación de microdatos y documentos oficiales del estudio tiene como requisito la anonimización previa de los datos recolectados, así como la indeterminación de estos. La II ENUT, como el resto de los productos estadísticos elaborados en el INE, tiene como fundamento la Política de Protección de Datos Personales institucional en el aseguramiento de la confidencialidad de la información producida. En ella, se establece que los datos producidos por el organismo no pueden hacer referencia expresa a personas que participen en sus estudios. Por

ese motivo, las bases definitivas excluyeron variables de identificación de las unidades de análisis que participaron en el estudio (personas, hogares, viviendas) o variables que aumentarían el riesgo de identificación de estas. Adicionalmente, las bases de datos se estructuran de acuerdo con las áreas de estimación determinadas en el diseño muestral.

Las normativas vigentes, que enmarcan el quehacer de la II ENUT e INE en general sobre esta materia son las siguientes:

- Compromiso de confidencialidad y reserva de datos protegidos por el secreto estadístico e información sensible y/o sujeta a la ley 19.628 sobre la protección de datos personales.
- Ley 20.285 regula el principio de transparencia de la función pública y el derecho de acceso a la información de los órganos de la Administración del Estado.
- Artículo 29° de la Ley 17.374. 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.
- Artículo 30° de la Ley 17.374. 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” (Extractos de la Ley

Orgánica 17.374 de creación del Instituto Nacional de Estadísticas, publicada el 10-12-1970).

- 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”.
- La difusión de información y accesibilidad a los datos involucra la tensión entre dos principios. En primer lugar, uno referido a la importancia de la libre difusión de los datos. Al respecto Naciones Unidas menciona que las estadísticas oficiales constituyen un elemento indispensable en el sistema de información de una sociedad democrática, de manera tal que los organismos nacionales en la materia deben facilitar que la población ejerza su derecho a información mediante un acceso imparcial a las estadísticas producidas. En segundo lugar, otro referido a la necesidad de asegurar esta información, para prevenir su mal uso y mantener la confianza de las poblaciones a las que sirven.

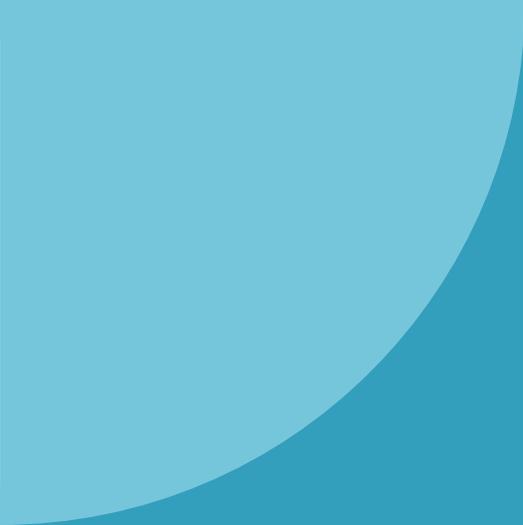
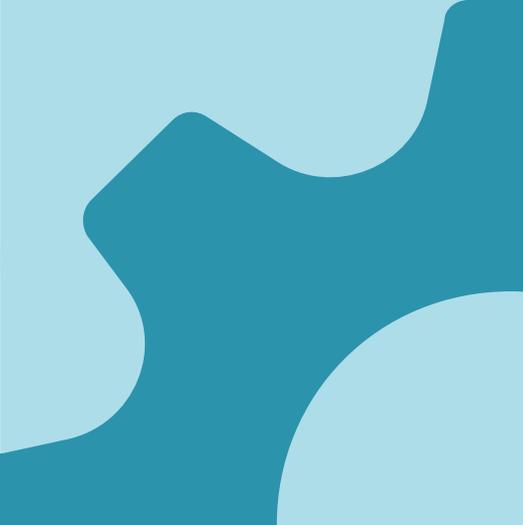
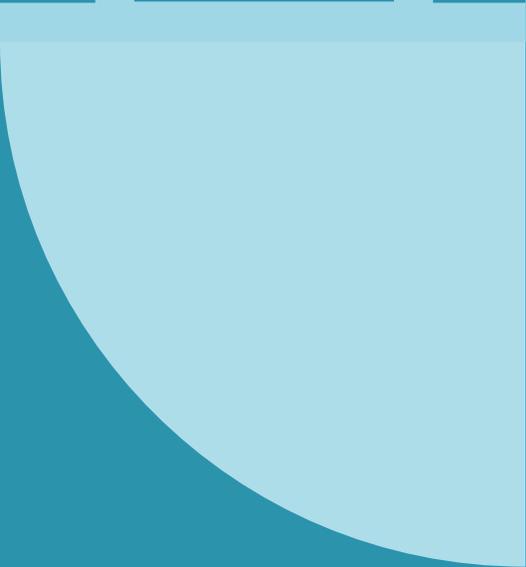
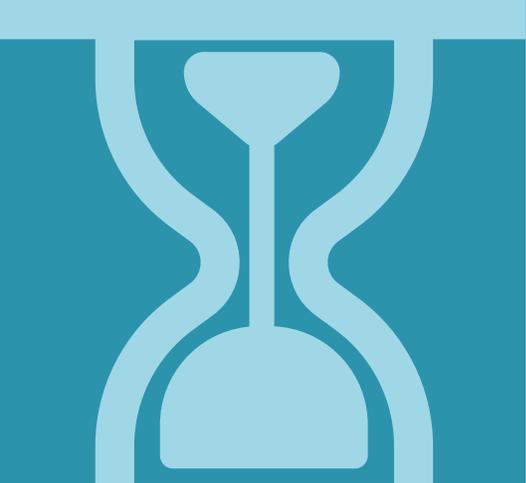
Es posible señalar, con relación a la seguridad de datos, que el INE se adhiere al Código Regional de Buenas Prácticas en Estadísticas para América Latina y El Caribe (CEPAL, 2023b) que en su principio 4 sobre confidencialidad establece que “la oficina nacional estadística y los demás miembros del sistema estadístico nacional (garantizan la protección y confidencialidad de la información con la que se producen las estadísticas oficiales y, evitan la identificación de las fuentes” (INE, 2015a, p. 8).

Además de la exclusión de las variables de identificación directa, la labor del Instituto Nacional de Estadísticas y de la II ENUT en particular

consideró dentro del proceso de difusión de información la evaluación de métodos de Control a la Divulgación Estadística (SDC según su sigla en inglés: *Statistical Disclosure Control*), también llamados de “anonimización”, que buscan establecer un equilibrio entre los principios de libre acceso a la información y aseguramiento de la confidencialidad de la información, entendiendo que existe una tensión o *trade-off* entre la protección de la seguridad de los datos y la utilidad de su libre difusión.

El objetivo de un proceso SDC bien implementado es encontrar el punto óptimo en el que la utilidad para los usuarios finales se maximice a un nivel de riesgo aceptable.

Mayores detalles al respecto formarán parte de un documento específico sobre el proceso de anonimización de la II ENUT.



10

GLOSARIO

10. Glosario

Actividades personales: actividades no productivas que involucran cuidados personales y necesidades fisiológicas, actividades de educación y aprendizaje, y ocio y vida social. Estas actividades se encuentran fuera de la frontera general de la producción (CEPAL, 2016).

Actividades educativas: actividades dedicadas a adquirir conocimientos y que distingue por la formalidad del establecimiento (CEPAL, 2016).

Carga global de trabajo: también conocido como tiempo total de trabajo. Corresponde a la suma de tiempo destinado al trabajo en la ocupación (y sus traslados) y al trabajo no remunerado. Permite dar cuenta del trabajo total que realizan las personas, independiente de si se realiza en el mercado, en el propio hogar o para otros hogares.

Cuidador principal: aquella persona dentro del hogar que destina más horas al cuidado de otra persona. Para el caso de las Personas en Situación de Dependencia Funcional (PSDF), se identifica la persona dentro del hogar que habitualmente dedica mayor tiempo al cuidado de forma declarativa (pregunta d9 del CH relativa a cuidador/a principal habitual), mientras que para Niños, Niñas y Adolescentes (NNA), este se identifica a partir de las horas registradas en el CUT, como aquella persona que destina más tiempo a tareas de cuidado en el hogar (cuidador/a principal en el período de referencia).

Cuidados personales: comprenden las actividades fisiológicas necesarias para la vida (comer y dormir) y otras actividades como aseo personal y cuidados de salud (CEPAL, 2016).

Personas en Situación de Discapacidad (PESD): son aquellas personas que, teniendo una o más deficiencias físicas, mentales y/o de movilidad, de

carácter temporal o permanente, al interactuar con diversas barreras presentes en el entorno, ve impedida o restringida su participación plena y efectiva en la sociedad, en igualdad de condiciones con los demás (Ley 20.422).

Personas en Situación de Dependencia Funcional (PSDF): son aquellas personas que, considerando su capacidad mental, física y/o de movilidad, requieren apoyo de un tercero para realizar ciertas tareas (SENADIS, 2017).

Responsable principal y secundario: son aquellas personas que toman las decisiones acerca del cuidado de la persona. Para el caso de Niños, niñas y adolescentes (NNA), se identifica la persona responsable principal de forma declarativa (pregunta c13a del CH). Además, en los casos que corresponda, se identifica un responsable secundario en el hogar (a partir de la pregunta c13b).

Tasa de participación: corresponde al porcentaje de personas que realizaron una determinada actividad. Es la relación entre las personas que declararon realizar una actividad y el total de personas consultadas elegibles.

Tiempo de participantes: es el tiempo destinado a una determinada actividad, expresado en horas o minutos, entre las personas que efectivamente la realizaron (destinan al menos un minuto declarado en la encuesta a la actividad).

Tiempo social: también llamado tiempo poblacional. Es el tiempo destinado a una determinada actividad, expresado en horas o minutos, entre toda la subpoblación evaluada, independiente si realizaron la actividad o no (cuando no se realizó la actividad, se considera el tiempo con valor cero en el numerador).

Trabajo: comprende todas las actividades realizadas por personas de cualquier sexo y edad con el fin de producir bienes o prestar servicios para el consumo de terceros o para uso final propio, independiente de si es legal o informal, y si es pagado o no pagado (OIT, 2013).

Trabajo en la ocupación: comprende todo el trabajo para la producción de bienes o prestación de servicios realizado para terceros, a cambio de una remuneración o beneficio. Incluye el trabajo independiente (empleadores y trabajadores por cuenta propia), el trabajo asalariado (asalariados y servicio doméstico remunerado), el trabajo familiar auxiliar en actividades de mercado y el trabajo de los contratistas dependientes (OIT, 2018). Estas actividades se encuentran dentro de la frontera de producción del Sistema de Cuentas Nacionales (SCN), y se agrupan todas las actividades que hacen posible, facilitan o fomentan la producción, incluyendo los traslados o viajes requeridos para realizar dichas tareas, además de la actividad de búsqueda de trabajo, como actividad asociada a la ocupación.

Trabajo no remunerado: son las actividades productivas que no son pagadas ni transadas por un bien en el mercado, que realizan las personas vinculadas a la prestación de servicios para los propios integrantes del hogar, para otros hogares o para la comunidad. Se compone de los trabajos domésticos y de cuidados no remunerados del propio hogar y el trabajo doméstico y de cuidados para otros hogares, para la comunidad y voluntario. Esta forma de trabajo está fuera de la frontera de producción económica, pero dentro de la frontera general de producción del Sistema de Cuentas Nacionales (CEPAL, 2016).

Trabajo doméstico no remunerado: actividades relativas a tareas domésticas desempeñadas por integrantes del hogar, para su propio consumo o beneficio del hogar, sin que exista retribución monetaria

alguna. “Son las actividades relacionadas con el concepto de reproducción social, es decir, de mantenimiento y reproducción de la fuerza de trabajo. Se excluyen todas las actividades realizadas en los hogares cuyo resultado esté orientado a su venta en el mercado y por las cuales se obtiene una remuneración, como puede ser el caso de la elaboración de comida para vender, el lavado y planchado de ropa ajena y la confección y reparación de ropa realizadas como actividad económica para el mercado y remunerada, entre otras.” (CEPAL, 2016). Incluye los traslados asociados a estas actividades.

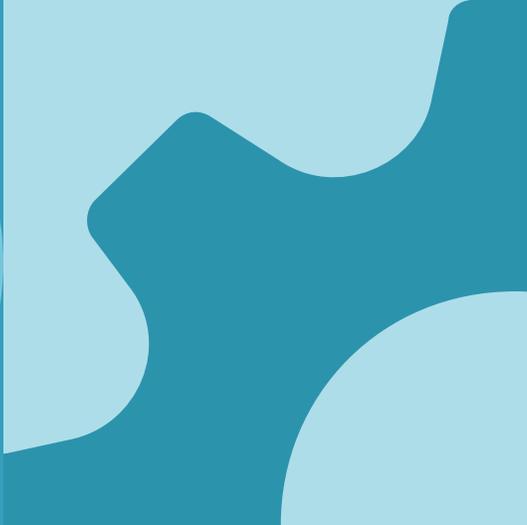
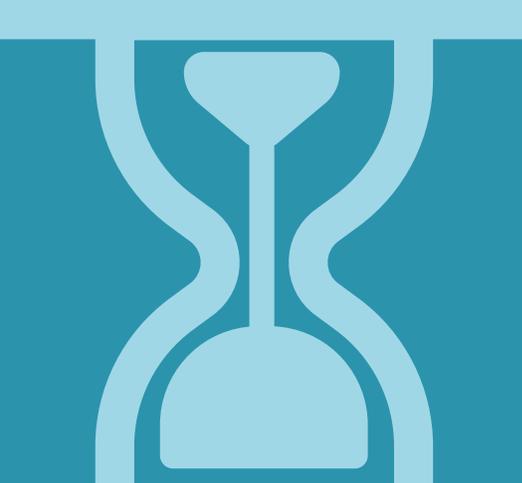
Trabajo de cuidados no remunerado: refiere a aquellas actividades de cuidado que se realizan a otra persona integrante del hogar, sin recibir un pago a cambio. Este trabajo reproduce y mantiene la fuerza de trabajo. Forma parte del trabajo de producción de servicios para el autoconsumo junto con el trabajo doméstico no remunerado (CEPAL, 2016).

Trabajo no remunerado para otros hogares, para la comunidad y voluntario: actividades de trabajo cuyo destino está fuera del propio hogar. Comprende el trabajo doméstico y de cuidados no remunerado dirigido a personas de otros hogares, los servicios voluntarios que se hacen a través de instituciones sin fines de lucro, y las actividades, ya sea individuales o de pequeños grupos de personas, que no están organizadas formalmente y que se realizan para dar un servicio a individuos o a la comunidad (CEPAL, 2016).

Ocio y vida social: considera actividades de carácter recreativo y de convivencia social, fuera de frontera general de producción.

Uso de Medios de Comunicación Masiva (MCM): actividades que involucran leer diarios, revistas o libros, ver televisión, escuchar radio, y utilizar redes sociales o correo electrónico, fuera de frontera general de producción (CEPAL, 2016).

Simultaneidad: actividades indicadas en el CUT que fueron realizadas al mismo tiempo, es decir, en un mismo lapso de tiempo en el día asignado. También incluye el registro de traslados multipropósitos, es decir, desplazamientos asociados a más de una actividad (un desplazamiento que fue registrado como viaje de vuelta en una actividad, y como viaje de ida hacia otra).



11

REFERENCIAS

11. Referencias

Aguirre, R. (2024). *Cuidados en agenda. Género, trabajo y uso del tiempo.*

<https://libreria.clacso.org/publicacion.php?p=3224&c=24>

Alvarado, M., & Pizarro, M. (2019). *Fundamentos de la nueva metodología de calibración de los factores de expansión de la Encuesta Nacional de Empleo.* Documento de trabajo n° 12. Instituto Nacional de Estadísticas de Chile.

<https://www.ine.gob.cl/inicio/documentos-de-trabajo/documento/documento-de-trabajo-fundamentos-de-la-nueva-metodolog%C3%ADa-de-calibraci%C3%B3n-de-los-factores-de-expansi%C3%B3n-de-la-encuesta-nacional-de-empleo>

American Association for Public Opinion Research (AAPOR) (2023). *Standard Definitions: Final dispositions of case codes and outcomes rates for surveys.* Tenth Edition.

<https://aapor.org/wp-content/uploads/2024/03/Standards-Definitions-10th-edition.pdf>

Bustelo, M., Suaya, A., Veza, E. (2021). *Hacia una nueva realidad laboral para las mujeres. Soluciones para recuperar el empleo femenino en ALC.* Banco Interamericano de Desarrollo.

<https://publications.iadb.org/es/publications/spanish/viewer/Hacia-una-nueva-realidad-laboral-para-las-mujeres-soluciones-para-recuperar-el-empleo-femenino-en-ALC.pdf>

Breiman, L. (2001). *Random forests.* Machine learning, 45, 5-32.

Callegaro, M. (2013). *Paradata in web surveys.* F. Kreuter, Improving surveys with paradata, 339-359. New Jersey: John Wiley & Sons, Inc.

Cochran, W. G. (1997). *Sampling techniques*. Hoboken, NJ: John Wiley & Sons, Inc.

CEPAL (2016). *Clasificación de Actividades de Uso del Tiempo para América Latina y el Caribe* (CAUTAL). <https://www.cepal.org/es/publicaciones/39624-clasificacion-actividades-uso-tiempo-america-latina-caribe-cautal>

CEPAL (2022). *Guía metodológica sobre las mediciones de uso del tiempo en América Latina y el Caribe* (LC/CEA. 11/17), Santiago, Comisión Económica para América Latina y el Caribe.

CEPAL (2023a). Repositorio sobre uso del tiempo de América Latina y Caribe. Observatorio de Igualdad de Género. https://www.cepal.org/sites/default/files/infographic/files/repositorio_sobre_uso_del_tiempo_de_america_latina_y_el_caribe_-_sep_2023.pdf

CEPAL (2023b). *Código Regional de Buenas Prácticas en Estadísticas para América Latina y el Caribe*. Conferencia Estadística de las Américas. <https://repositorio.cepal.org/server/api/core/bitstreams/c55d4f3f-4ac0-4b7b-8c96-4df45f251fe8/content>

Departamento Administrativo Nacional de Estadística de Colombia (DANE) (2022). Encuesta nacional de Uso del Tiempo 2020-2021: Resultados. https://www.dane.gov.co/files/investigaciones/boletines/ENUT/Presentacion_ENUT_2020_2021.pdf

Díaz de Rada, V. (2019). *Calidad de respuesta y modo de administración del cuestionario*. *Revista Internacional de Sociología* 77(1): e122. <https://doi.org/10.3989/ris.2019.77.1.17.135>

- Díaz Jiménez, R.M.** (2007). *Feminización de la dependencia. Reflexiones sobre el sistema para la autonomía y la atención a personas en situación de dependencia*. Portularia, 7(1-2).
- European Commission, Eurostat.** (2019). *Quality Guidelines for Frames in Social Statistics (QGFSS) (Version 1.51)*. Luxembourg: European Commission. http://ec.europa.eu/eurostat/cros/content/essnet-quality-multisource-statistics_en
- Feres, J. C.** (mayo y junio de 2008). *Pobreza y Uso del Tiempo*. Curso Internacional: Redistribución del Tiempo, un indicador de igualdad. CEPAL. Seminario
- Galván, M., Medina, F.** (2007). *Imputación de datos: teoría y práctica*. Serie Estudios Estadísticos y Prospectivos. División de Estadísticas y Proyecciones Económicas. <https://repositorio.cepal.org/server/api/core/bitstreams/02dd479f-fae2-43c4-b5ec-5419fa7f6190/content>
- Gerber, A., Arceneaux, K., Boudreau, Ch., Dowling, C., Hillygus, S., Palfrey, T., Biggers, D. R., Hendry, D. J.,** (2014) *Reporting Guidelines for Experimental Research: A Report from the Experimental Research Section Standards Committee*. Journal of Experimental Political Science, Volume 1, Issue 1, Spring 2014, pp. 81 – 98. <https://doi.org/10.1017/xps.2014.11>
- Gómez Luna, M. E.** (2010). *Directrices y referentes conceptuales para armonizar las encuestas sobre uso del tiempo en América Latina y el Caribe*. Documento DE Trabajo para discusión. México. https://inec.cr/wwwisis/documentos/Otros_doc/Directrices_C_EUT_AL.pdf

Guerrero, M, Guerrero, J, Marín, A & Pizarro, M. (2019). *Estratificación socioeconómica del marco muestral de viviendas 2017 (MMV 2017)* (Documento de trabajo No. 11). Instituto Nacional de Estadísticas. <https://www.ine.gob.cl/docs/default-source/documentos-de-trabajo/estratificación-socioeconómica-del-marco-muestral-de-viviendas-2017.pdf>

Guerrero, J., & Marín, A. (2020). *Estratificación socioeconómica del marco muestral de viviendas 2017 (MMV 2017): Implementación en R* (Documento de trabajo No. 17). Instituto Nacional de Estadísticas. [https://www.ine.gob.cl/docs/default-source/documentos-de-trabajo/estratificación-socioeconómica-del-marco-muestral-de-viviendas-2017-\(mmv-2017\)-implementación-en-r.pdf](https://www.ine.gob.cl/docs/default-source/documentos-de-trabajo/estratificación-socioeconómica-del-marco-muestral-de-viviendas-2017-(mmv-2017)-implementación-en-r.pdf)

Hansen, M. H., Hurwitz, W. N., & Madow, W. G. (1953). *Sample survey methods and theory.*

Hubert, M., & Van der Veeken, S. (2008). *Outlier detection for skewed data. Journal of Chemometrics: A Journal of the Chemometrics Society*, 22(3-4), 235-246. <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.1123>

International Labour Organization (ILO) (2023) *Spotlight on Work Statistics n°12 New data shine light on gender gaps in the labour market.* <https://www.ilo.org/publications/new-data-shine-light-gender-gaps-labour-market>

Instituto Nacional de Estadísticas y Censos de Argentina (INDEC Argentina) (2021). *Cuestionario de la Encuesta Nacional de Uso del Tiempo* 2021.

https://www.indec.gov.ar/ftp/cuadros/sociedad/cuestionario_enut_2021.pdf

Instituto Nacional de Estadísticas de Chile (INE) (2006). *Encuesta Nacional De Empleo Diseño Muestral*.
http://www.ine.cl/canales/chile_estadistico/mercado_del_trabajo/empleo/metodologia/pdf/ene_espanol.pdf

Instituto Nacional de Estadísticas de Chile (INE). (2015). *Actualización de límites urbanos censales*. Serie Boletines No. 1.
<https://geoarchivos.ine.cl/File/boletines/Actualización%20Límites%20Urbanos%20Censales.pdf>

Instituto Nacional de Estadísticas de Chile (INE) (2016). *Documento Metodológico ENUT 2015*. Santiago.
<https://www.ine.gob.cl/docs/default-source/uso-del-tiempo-tiempo-libre/metodologias/enut-2015/documento-metodol%C3%B3gico-enut-2015.pdf?sfvrs->

Instituto Nacional de Estadísticas de Chile (INE) (2020). *Estándar para la evaluación de calidad de estimaciones*.
<https://www.ine.gob.cl/docs/default-source/buenas-practicas/estandares/estandar/documento/estandar-evaluacion-de-calidad-de-estimaciones.pdf>

Instituto Nacional de Estadísticas y Censos de Colombia (INE Costa Rica) (2022). *Encuesta Nacional de Uso del Tiempo: Principales resultados*. San José, Costa Rica.
https://admin.inec.cr/sites/default/files/2023-06/reENUT2022_0.pdf

Instituto Nacional de Estadísticas de Uruguay (INE Uruguay) (2023). *Cuestionario de la Encuesta de Uso del Tiempo y Trabajo no Remunerado*

2021.

<https://www4.ine.gub.uy/Anda5/index.php/catalog/732/download/1214>

Kish, L., & Frankel, M. R. (1974). Inference from complex samples. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 36(1), 1-22.

Laurie, H. (2003). *From PAPI to CAPI: consequences for data quality on the British Household Panel Study*. ISER Working Paper Series, No. 2003-14, University of Essex, Institute for Social and Economic Research (ISER), Colchester. <https://www.econstor.eu/bitstream/10419/92209/1/2003-14.pdf>

Lehtonen, R., & Pahkinen, E. (1995). *Practical methods for design and analysis of complex surveys*. John Wiley & Sons.

Lohr, S.L. (1999) *Sampling: Design and Analysis*. Duxbury Press, Pacific Grove, CA.

Mincer, J. (1974). *Schooling, experience, and earnings*. New York: National Bureau of Economic Research; distributed by Columbia University Press.

Naciones Unidas (2006a). *Guía de elaboración de estadísticas sobre el empleo del tiempo para medir el trabajo remunerado y no remunerado*. Departamento de Asuntos Económicos y Sociales. División Estadística. https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/time-use/SeriesF_93-S.pdf

Naciones Unidas (2006b). *Convención sobre los Derechos de las Personas con Discapacidad y Protocolo Facultativo*. Naciones Unidas.

<https://igualdad.cepal.org/es/digital-library/convencion-internacional-sobre-los-derechos-de-las-personas-con-discapacidad>

Naciones Unidas. (2007). *Encuestas de hogares en los países en desarrollo y en transición* (Vol. Serie F N°96). Nueva York. Obtenido de https://unstats.un.org/unsd/demographicsocial/Standards-and-Methods/files/Handbooks/surveys/seriesf_96-S.pdf

Naciones Unidas. (2009). *Diseño de muestras para encuestas de hogares. Directrices prácticas Serie F N°98*. Nueva York. https://unstats.un.org/unsd/publication/seriesf/Seriesf_98s.pdf

Naciones Unidas. (2019). *Recopilación y gestión de datos como medio para fomentar la sensibilización acerca de la violencia y la discriminación por motivos de orientación sexual o identidad de género*. Asamblea General de las Naciones Unidas: Consejo de Derechos Humanos. A/HRC/41/45.

Nock, S. L., Guterbock, T. M. (2010). Survey experiments. En *Handbook of Survey Research*, editado por Peter V Marsden y James D Wright, 2. ed., 837-65. Editorial Emerald, Bingley, UK, 2010. <http://books.emeraldinsight.com/display.asp?K=9781848552241>.

OCDE. (2007). *Glossary of Statistical Terms*. Organisation for Economic Cooperation and Development, Nueva York.

OIT (2008). *Resolution I, Resolution concerning the measurement of working time*. Conferencia Internacional de Estadísticos del Trabajo Ginebra, 24 de noviembre - 5 de diciembre de 2008. https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@dgregperts/@stat/documents/normativeinstrument/wcms_112455.pdf

OIT (2011). *Manual de medición del trabajo voluntario*. STATISTICS Departamento de Estadística. Oficina Internacional del Trabajo. http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_167833.pdf

OIT (2013). *19° Conferencia Internacional de Estadísticos del Trabajo, CIET*. Conferencia Internacional de Estadísticos del Trabajo Ginebra, 2 a 11 de octubre de 2013. <https://www.ilo.org/es/media/456071/download>

OIT (2018). *Resolución de la 20° Conferencia Internacional de Estadísticos del Trabajo (CIET) sobre las estadísticas de las relaciones de trabajo*. Conferencia Internacional de Estadísticos del Trabajo Ginebra, 10 a 19 de octubre de 2018. <https://www.ilo.org/es/media/216761/download>

OIT (2023a). *Resolución de modificación de la resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo*. 21° Conferencia Internacional de Estadísticos del Trabajo. 11-20 de octubre de 2023. <https://www.ilo.org/es/resource/resolucion-de-modificacion-de-la-resolucion-sobre-las-estadisticas-del-1>

OIT (2023b). *Panorama laboral 2022. América Latina y el Caribe*. Oficina Regional de la OIT para América Latina y el Caribe. https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@americas/@ro-lima/documents/publication/wcms_867497.pdf

Rogero-García, J. (2009). *Distribución en España del cuidado formal e informal a las personas de 65 y más años en situación de dependencia*. *Revista española de salud pública*, 83(3), 393-405.

Olson, K., & Parkhurst, B. (2013). *Collecting paradata for measurement error evaluations*. F. Kreuter, Improving surveys with paradata. 43-72. New Jersey: John Wiley & Sons, Inc.

Shadish, W. R., Thomas, D. C., Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton, Mifflin and Company. Boston, New York.
<https://psycnet.apa.org/record/2002-17373-000>.

Smith, T. W., Jibum K. (2015). *A Review of Survey Data Collection Modes*. *Sociological Theory and Methods* 30 (2): 185-200.
<https://doi.org/10.11218/ojjams.30.185>.

Sen, A. R. (1953). *On the estimate of the variance in sampling with varying probabilities*. *Journal of the Indian Society of Agricultural Statistics*, 5, 119–127.

Servicio Nacional de la Discapacidad (2017). *Discapacidad y Dependencia: Caracterización de la dependencia en las personas en situación de discapacidad a partir del II Estudio Nacional de la Discapacidad*. Departamento de Estudios.

Sistema de Cuentas Nacionales (SCN) (2008). *Sistema de Cuentas nacionales* 2008.
https://www.cepal.org/sites/default/files/document/files/sna2008_web.pdf

UNECE (2013). *Guidelines for Harmonizing Time Use Surveys*. Naciones Unidas. Génova.
<https://digitallibrary.un.org/record/787787?ln=es&v=pdf>

UNECE (2019). *Generic Statistical Business Process Model GSBPM*.
<https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1>

UNSD. (2017). *International Classification of Activities for Time-Use Statistics 2016 (ICATUS 2016)*. Economics and Social Affairs. Naciones Unidas. <https://unstats.un.org/unsd/classifications/Family/Detail/2083>.

Urquidi, M., Chalup, M. (2023). *Brecha de ingresos laborales por género en América Latina y el Caribe: Un análisis de sus componentes*. Banco Interamericano de Desarrollo. <http://dx.doi.org/10.18235/0004785>

<https://publications.iadb.org/es/brecha-de-ingresos-laborales-por-genero-en-america-latina-y-el-caribe-un-analisis-de-sus-diferentes>

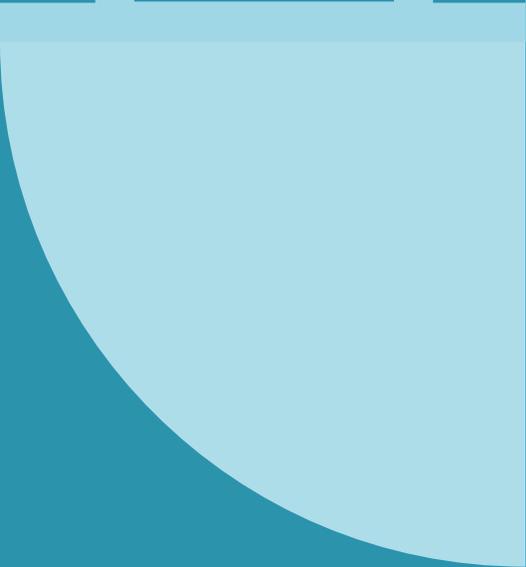
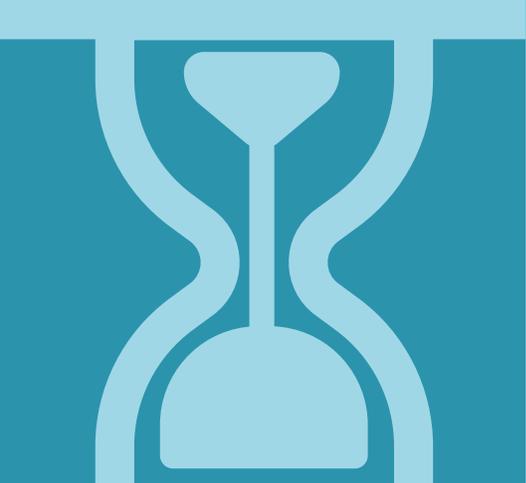
Valliant, R., Dever, J. A., & Kreuter, F. (2013). *Practical tools for designing and weighting survey samples* (Vol. 1). New York: Springer.

Weiss, A. (1995). *Human Capital vs. Signalling Explanations of Wages*. *Journal of Economic Perspectives*, 9(4), 133–154. <https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.9.4.133>

Willis, G. B., Lessler, J. T. (1999). *Question Appraisal System QAS-99*. Research Triangle Institute.
http://www.websm.org/uploadi/editor/1364216022Willis_Lessler_1999_QAS_99.pdf

Wolter, K. M. (1985). *Introduction to variance estimation*. New York: Springer-Verlag.

Yates, F., & Grundy, P. M. (1953). *Selection without replacement from within strata with probability proportional to size*. Journal of the Royal Statistical Society: Series B (Methodological), 15(2), 253-261.



12

ANEXOS

12. Anexos

12.1. GSBPM en la producción estadística

El *Generic Statistical Business Process Model*, o GSBPM por sus siglas en inglés, es un estándar internacional consolidado que propone una estructura de 8 fases estandarizadas para el proceso de producción estadística, en el que se utiliza un enfoque principalmente operacional.

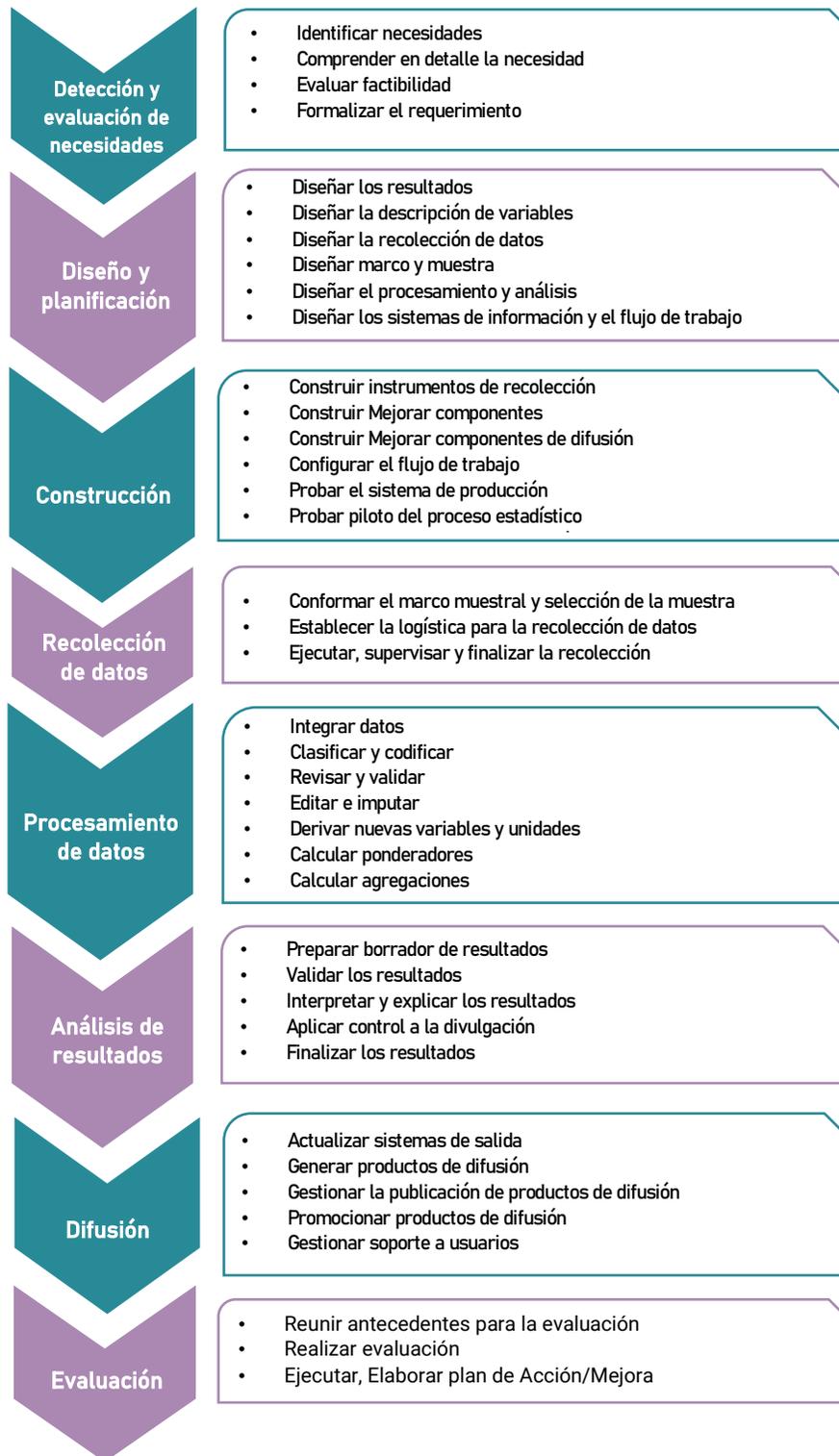
En base a este marco durante los últimos años, el INE ha establecido directrices e implementado una adaptación de este modelo en su segmento de negocios (procesos de producción estadística).

Algunas ventajas de su uso, a la par de la guía al diseño y modernización del proceso, son las siguientes (UNECE, 2019):

- Proporciona una **estructura para organizar cada fase de la producción estadística**, contribuyendo así a la documentación y estandarización del proceso.
- Entrega un marco de referencia para **evaluar la mejora y calidad de la producción**, al mismo tiempo que ayuda a reconocer más fácilmente y de forma ordenada las buenas prácticas (fase en la que aportaron y de qué forma).
- Ayuda a **identificar componentes que no están siendo ejecutados correctamente**, y/o visibiliza limitaciones del proceso actual, que llevan más rápidamente al diseño de nuevas actividades o pasos necesarios para la producción estadística.

Por otro lado, las 8 etapas que componen el GSBPM adaptado en la institución son las siguientes (ver Figura 9):

Figura 9. GSBPM Modelo de Producción Estadística Adaptado



1) Detección y evaluación de necesidades

En esta fase del GSBPM, se identifica la necesidad de nuevas estadísticas y define la información necesaria para darle respuesta, evaluando factibilidad y disponibilidad de datos. El resultado debe concretar un plan general de cómo lograr ese propósito. Por otra parte, los 4 subprocesos que la fase considera son: 1.1 Identificación de necesidades, 1.2 Comprender en detalle la necesidad, 1.3 Evaluar factibilidad y 1.4 Formalizar el requerimiento.

2) Diseño y Planificación

Este proceso como su nombre lo indica, corresponde al diseño y planificación de actividades del proceso enunciadas en las subetapas de: 2.1 Diseñar los resultados, 2.2 Diseñar la descripción de variables, 2.3 Diseñar la recolección de datos, 2.4 Diseñar marco y muestra, 2.5 Diseño del procesamiento y análisis y 2.6 Diseñar los sistemas de información y el flujo de trabajo.

3) Construcción

En esta etapa se elabora y evalúa la propuesta de producción de la fase de “diseño”. En cuanto a los subprocesos que la componen, se encuentran: 3.1 Construcción de instrumentos de recolección, 3.2 Construcción o mejora de componentes, 3.3 Construcción o mejora componentes de difusión, 3.4 Configuración de flujos de trabajo, 3.5 Probar sistema de producción, 3.6 Probar piloto del proceso estadístico y 3.7 Finalizar los sistemas de producción.

4) Recolección de datos

Esta etapa abarca la selección muestral, y la puesta en marcha de los métodos de recopilación de datos para la elaboración estadística. En ese sentido considera los subprocesos de: 4.1 Conformar el marco muestral y selección de la muestra, 4.2

Establecer la logística para la recolección de datos, 4.3 Ejecutar, supervisar y finalizar la recolección.

5) Procesamiento

Procesamiento incluye la validación, depuración de los datos y su preparación para el análisis. En esta fase se desarrollan una serie de subprocesos en que se revisan, editan/imputan los insumos, para que adquieran un nivel de calidad que permita su análisis y difusión como productos estadísticos confiables. Los subprocesos de esta etapa involucran: 5.1 Integrar datos, 5.2 Clasificar y codificar, 5.3 Revisar y validar, 5.4 Editar e imputar, 5.5 Derivar nuevas variables y unidades, 5.6 Calcular ponderadores, 5.7 Calcular agregaciones y 5.8 Finalizar los archivos de datos.

6) Análisis de resultados

Consiste en la obtención de los resultados estadísticos, su examen y validación de forma previa a la difusión. Supone al mismo tiempo la construcción de una reseña referente a los distintos procesos estadísticos realizados, y el cumplimiento de los protocolos de calidad. En cuanto a los subprocesos de la fase, estos se componen de: 6.1 Preparar borrador de resultados, 6.2 Validar los resultados, 6.3 Interpretar y explicar los resultados, 6.4 Aplicar control a la divulgación y 6.5 Finalizar los resultados.

7) Difusión

Abarca la publicación de los productos estadísticos y la documentación que lo acompaña. Considera los subprocesos de 7.1 Actualizar sistemas de salida, 7.2 Generar productos de difusión, 7.3 Gestionar la publicación de productos de difusión, 7.4 Promocionar productos de difusión y 7.5 Gestionar soporte a usuarios.

8) Evaluación y retroalimentación

Esta fase persigue evaluar el éxito de instancias específicas del proceso hacia el objetivo de identificar mejoras potenciales, empleando información cualitativa y cuantitativa. Los subprocesos que la componen son: 8.1 Reunir antecedentes para la evaluación, 8.2 Realizar evaluación y 8.3 Elaborar plan de acción/mejora.

Al utilizar este modelo, es importante entender dos puntos relevantes:

- **No linealidad:** si bien el modelo presenta estas ocho etapas con sus subprocesos ordenados secuencialmente, al diseñar los flujos de actividad, las relaciones entre las fases y subetapas no necesariamente ocurren siempre en el orden señalado, dado que en ocasiones un cierto tramo del proceso se desarrolla más de una vez. O bien, los resultados de cierta fase impactan el diseño o construcción original de una etapa anterior. Por ejemplo, de acuerdo con el GSBPM adaptado, dentro del Procesamiento, primero se revisa y valida una base de datos (5.3), y luego se edita e imputa según estas revisiones (5.4). Sin embargo, es posible y hasta recomendable, que luego de dicha edición, esa misma base se someta nuevamente a validación (un subproceso anterior) para revisar el efecto de los cambios en ella. Rompiendo así el flujo lineal expuesto con anterioridad.
- **Marco de referencia:** el modelo propone etapas y subprocesos, buscando ser lo suficientemente flexible para ser adaptado con éxito a la producción de estadísticas de diversa clase. Pero finalmente son las particularidades del producto las que definirán en última instancia las fases y actividades más pertinentes a realizar, además del cómo serán implementadas. 5.3 Revisar y validar, subproceso del Procesamiento, puede implicar en un

estudio comparar ciertas respuestas entregadas por el entrevistado, mientras que en otra investigación pueden asumir un criterio de comprobación diferente, dando en ambas situaciones una respuesta igualmente satisfactoria al requerimiento de “validar” atinente a su contexto particular.

12.2. Construcción del marco estadístico

La construcción de los marcos estadísticos convive con limitaciones, entre ellas, la sub o sobreestimación de la población objetivo. Este tipo de limitaciones derivan de problemas de actualización del *stock* de viviendas, así como también de dinámicas territoriales que se insertan a la producción con rezago, generando errores de medición. Además, si la falta de cobertura no tiene un comportamiento aleatorio, se genera una subrepresentación de determinados grupos de población y eventualmente, la sobrerrepresentación de otros. En general, estas situaciones suelen darse cuando los marcos muestrales están obsoletos, por ende, una solución a este problema es considerar marcos muestrales lo más actualizados posible.

En atención a lo anterior, desde el segundo semestre del 2019, el INE pone a disposición el Marco Muestral de Viviendas (MMV 2017) construido con base en el Censo de Población y Vivienda de 2017 que, 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 2021 se encuentra disponible el MMV 2020 que es una versión actualizada del MMV 2017, con base en enumeraciones en terreno de unidades geográficas priorizadas y verificaciones de gabinete (en oficina), las cuales hacen uso de los registros administrativos con los que cuenta el INE para tal fin. En este sentido, la encuesta ENUT utilizó el MMV 2020 para fines del diseño y selección de la muestra.

12.2.1. Estratificación socioeconómica de las UPM

Para la determinación de los niveles o estratos socioeconómicos, se consideraron diversos métodos y niveles de agregación de la información. Los métodos considerados fueron algoritmos de clasificación multivariados, tales como *Principal Component Analysis* (PCA), *K-medias*, y *Principal Component Analysis by means of Alternating Least Squares* (PRINCALS).

A partir de las distintas combinaciones de técnicas, agrupaciones, cantidades de estratos y variables, se obtuvieron 180 vectores de estratificación, que consideraban entre 3 y 5 estratos. Para evaluar el desempeño de estos, se calculó el efecto de diseño de 24 indicadores, con lo cual fue posible escoger la estratificación que redujera la varianza de estos. Los indicadores seleccionados fueron:

- Tasa de ocupación por sexo.
- Tasa de desocupación por sexo.
- Tasa de inactividad por sexo.
- Porcentaje de extranjeros.
- Porcentaje de personas dentro de la fuerza de trabajo primaria respecto de la población económicamente activa (PEA) por sexo.
- Porcentaje de hogares unipersonales.
- Porcentaje de población femenina ocupada por rama.
- Porcentaje de población masculina ocupada por rama.
- Porcentaje de personas según nivel educacional por sexo.

Luego de la evaluación, el método utilizado para la estratificación del marco es el de estratificación óptima de la primera componente principal del *PCA*, que considera los siguientes indicadores:

- Porcentaje de personas en la educación superior.
- Tasa de ocupación.

- Porcentaje de viviendas con índice de materialidad alto.
- Indicador sobre total de hijos nacidos vivos.

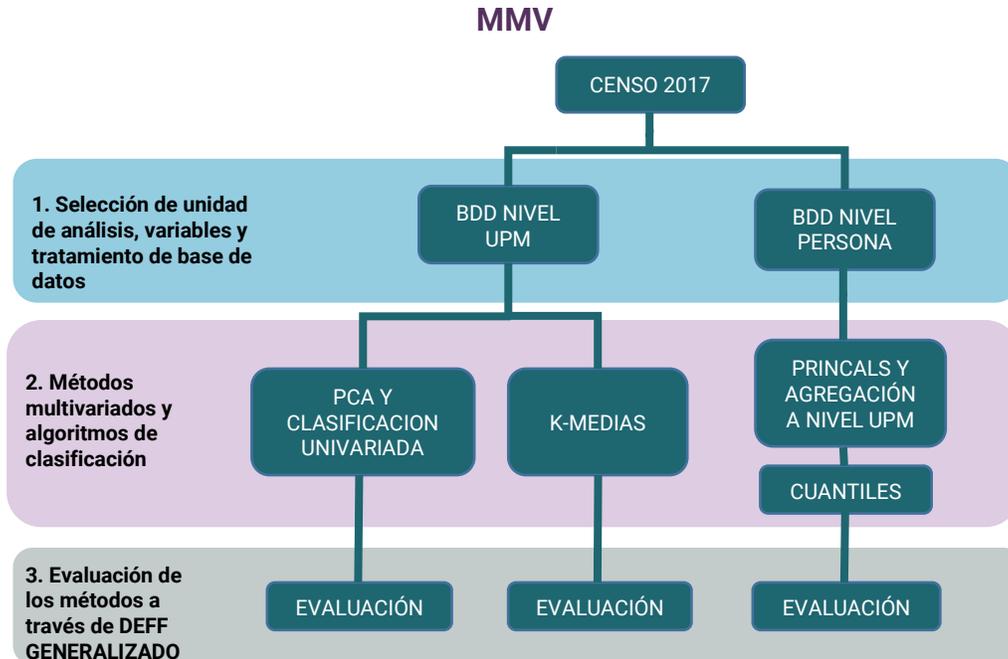
La elección del número de estratos (3) se basa en el análisis de las distribuciones de las UPM, según el cruce de comuna y área, ya que la baja prevalencia de UPM en regiones extremas, particularmente en el área rural, hace que sea más conveniente considerar solo 3 estratos. Finalmente, se obtiene como resultado una variable de estratificación con tres categorías de nivel socioeconómico, señalada en el MMV 2020 como clasificación socioeconómica de la UPM ⁵⁵.

⁵⁵ Para mayores antecedentes consultar:

[https://www.ine.cl/inicio/documentos-de-trabajo/documento/estratificaci%C3%B3n-socioecon%C3%B3mica-del-marco-muestral-de-viviendas-2017-\(mmv-2017\)-implementaci%C3%B3n-en-r](https://www.ine.cl/inicio/documentos-de-trabajo/documento/estratificaci%C3%B3n-socioecon%C3%B3mica-del-marco-muestral-de-viviendas-2017-(mmv-2017)-implementaci%C3%B3n-en-r)

En la Figura 10 se presentan gráficamente los pasos seguidos para generar la estratificación socioeconómica del MMV.

Figura 10. Esquema general de la metodología de estratificación del



Fuente: Guerrero et al., 2019, pág. 10

12.2.2. Áreas especiales

Dentro de los atributos del nuevo MMV 2020, al igual que su antecesor (MMV 2017), destaca la incorporación de la identificación de las áreas de tratamiento o áreas especiales, aportando información necesaria para identificar áreas susceptibles de exclusión, a nivel de UPM, en sustitución de lo que los marcos anteriores definían como Áreas de Dificil Acceso (ADA) y Áreas de Alto Costo (ACO)⁵⁶. Estas áreas especiales suelen estar afectadas por factores que dificultan su levantamiento, los cuales pueden ser climáticos, de transporte, costo de acceso y de acceso

⁵⁶ Las áreas conocidas como Áreas de Dificil Acceso (ADA) y Áreas de Alto Costo (ACO) no eran parte de los marcos de muestreo del INE, ya que al ser clasificadas de esta manera a partir de la información del Censo 2002, no existía cartografía asociada. En general, en todas las antiguas ADA existe un alto porcentaje de UPM especiales (un 76,9% en promedio), a excepción de la comuna Laguna Blanca, que no tiene ninguna.

restringido, entre otros. El MMV 2020 aporta información mediante ocho tipologías que dan cuenta de las complejidades presentes en la realización de trabajos operativos, información que resulta de gran utilidad para la coordinación de las actividades en terreno, tales como:

- Tiempo de trayecto
- Clima
- Altitud
- Transporte especial
- Insularidad
- Acceso pedestre
- Estado de los caminos
- Ingreso restringido

12.2.3. Actualización de los límites urbanos censales (LUC)

Una de las tareas realizadas para preparar el Precenso 2016 fue revisar la base territorial, es decir, la información geográfica existente respecto a los límites del territorio nacional. Esto buscaba segmentar el territorio y, posteriormente, distribuir las cargas de trabajo además de controlar los avances del operativo, e identificar las modificaciones geográficas. Para ello, fue utilizada como punto de partida la división político-administrativa del país y la división del territorio comunal en distritos y zonas censales. Luego se revisa el Límite Urbano Censal (LUC) que establece la división entre áreas urbanas y rurales, para lo cual fue necesario realizar un trabajo en terreno y en oficina, en cada una de las regiones del país. A partir de estos criterios se establecieron los límites que debían ser utilizados en el Precenso 2016, los cuales constituyeron la base para la realización del levantamiento censal del 2017.

En particular, la actualización del LUC se realiza producto del dinamismo que experimentan ciudades y pueblos, tanto por el crecimiento urbano como por la modificación de los elementos geográficos de referencia para su trazado. Para una correcta georreferenciación en la captura de datos, estos límites deben actualizarse antes del levantamiento censal.

En este sentido, el LUC se define como una línea imaginaria que separa el área urbana del área rural, que tiene una finalidad estadística censal y su fijación corresponde a criterios técnicos propios y particulares del INE (INE Chile, 2015). Su delimitación tiene como objetivo principal establecer la diferencia metodológica que tendrán el área urbana y rural de la comuna. Estas corresponden a áreas que tendrán un tratamiento particular, tanto cartográfico como en su enumeración en la actividad censal. Para efectos muestrales, el LUC debe considerar las áreas de expansión urbana, de manera de no perder su funcionalidad en el corto plazo y no afectar las estimaciones de crecimiento de las áreas urbanas y rurales.

Las ciudades cambian en el tiempo y proyectar su crecimiento es una tarea que se torna bastante compleja debido a múltiples causas. En primer lugar, la evolución de las ciudades gira en torno a factores como el mercado inmobiliario que, a su vez responde al comportamiento de las condiciones económicas y también de otro tipo de componentes, tales como desastres naturales, políticas sociales, entre otros factores. Por estas razones, al momento de incluir las áreas de expansión urbana en el LUC, se consultan los “Instrumentos de Planificación Territorial” (IPT) vigentes, además de la información de crecimiento proyectado en el mediano plazo que entregan las direcciones de obras municipales.

12.2.4. Características de las áreas geográficas

Las áreas urbanas permiten constatar una cierta morfología particular, con una estructura que se diferencia de la rural. Este concepto se relaciona con tres características geográficas fundamentales: amanzanamiento, continuidad y concentración. En las Figura 11, 12 y 13 se representan estas características.

Figura 11. Amanzanamiento



Figura 12. Continuidad



Figura 13. Concentración

Debido a la diversidad de tipos de asentamiento y a la dificultad de diferenciar de manera dicotómica el espacio geográfico, muchas veces estas condiciones no se cumplen a cabalidad, sino más bien, se aprecian gradientes entre espacios que presentan estas características y otros que no.

En la Figura 14, se observa la inclusión de áreas con alta concentración de viviendas (dispuestas en manzanas) y áreas con menor concentración de viviendas, dispuestas en vías estructurantes, sin presencia de manzanas, dentro del LUC (en verde). Esto da cuenta de la complejidad en determinar el área a incluir dentro del trazado.

Figura 14. Inclusión de áreas con alta concentración de viviendas

12.2.5. Cambios en las definiciones urbano y rural

El cambio de LUC obedece a variados factores como, por ejemplo: expansión urbana, cambio de definiciones de área y cambios de criterios. En línea con esto, es posible evidenciar que los Censos 2002 y 2017 presentan ciertas diferencias en lo que respecta a definición de área urbana y rural. Si bien, los criterios de cantidad de población en cada área geográfica en general se mantienen, se diferencian en otros aspectos. Por ejemplo, el Censo 2002 comprende el criterio de que, si un centro cumple funciones de turismo⁵⁷ y recreación con más de 250 viviendas concentradas y que no alcanzan el requisito de población, se considera urbano de todas maneras, en contraposición con el Censo 2017 donde para definir una entidad como urbana se deben cumplir a cabalidad los criterios de continuidad, concentración y amanzanamiento.

Cuando se habla de área, se hace referencia a si es urbano o rural, pero también existe el concepto de entidades, las cuales hacen referencia al tipo de asentamiento que componen estas áreas y se diferencian por su categoría⁵⁸. En el MMV 2020 la subdivisión muestral en el área urbana no hace distinción, ya que se trabaja ambas entidades urbanas (ciudad y pueblo) de la misma forma. Sin embargo, para el área rural donde todas las entidades pertenecen al mismo estrato rural, se construye el marco separando las aldeas del resto; es decir, se conformaron UPM rurales con el mismo tamaño, pero no se mezclaron las aldeas con el resto de las entidades.

El MMV 2020 contiene variables que permiten la identificación de la división político-administrativa del país y la identificación de las áreas

⁵⁷ En Censo 2017 se elimina este criterio, ya que no existe un catastro oficial del Estado de centros turísticos.

⁵⁸ Las entidades urbanas se componen de ciudades y pueblos, mientras que las entidades rurales corresponden a aldea, caserío, asentamiento minero, asentamiento pesquero, fundo – estancia – hacienda, parcela, comunidad indígena, campamento, veranada – majada – aguada y otros.

urbanas y rurales en cada comuna. A diferencia del marco utilizado en la ENUT 2015 (MMM 2015) - que definía, en el área urbana los estratos CD en el marco de manzanas; y RAU, en el marco de manzanas y en el de secciones - en el MMV 2020 solo se diferencia la parte urbana de la rural. Esta diferencia se debe principalmente a que, con el crecimiento del área urbana, se vuelve cada vez más complejo organizar el muestreo desde el punto de vista operacional al tener que extraer información desde dos marcos (marco muestral de manzanas y marco muestral de secciones) que están contruidos de manera distinta y que incurren en distintos procesos de actualización. Dado que el nuevo marco está constituido por UPM, correspondientes a áreas geográficas homogéneas, en términos del número de viviendas definidas como viviendas particulares ⁵⁹ observadas en el Censo de Población y Vivienda de 2017 (y que en el MMV se denominan “viviendas marco”) (INE, 2020), no es necesario contar con todas las desagregaciones del marco antiguo, lo cual permite tener una división más simple, estructurada y organizada.

De las 319 comunas del MMV 2020 que contienen área urbana, existen 24 que son exclusivamente urbanas. Por su parte, el estrato rural se encuentra en 321 comunas, de las cuales 26 son exclusivamente rurales.

En la Tabla 30 se presenta una comparación de las definiciones de las entidades urbano (en el MMV 2020 y CD-RAU en el MM 2015) y rural.

⁵⁹ Se excluyen las viviendas de temporada y colectivas.

**Tabla 30. Comparación de definiciones para entidades urbano y rural
MMV 2020 y MM 2015**

Entidad	MMV 2020	MM 2015
Urbano	Asentamiento humano con continuidad y concentración de construcciones en un amanzanamiento regular, con población mayor a 2.000 habitantes, o entre 1.001 y 2.000 habitantes, donde menos de 50% de la población que declara haber trabajado, se dedica a actividades primarias. Por ejemplo, en la comuna de Coquimbo se identifican cuatro entidades urbanas: Coquimbo, Tongoy, El Peñón y Guanaqueros; entidades que en su conjunto conforman el área urbana de la comuna de Coquimbo.	Centro urbano con menos de 40.000 y más de 2.000 habitantes. Esta clasificación se da cuando en la comuna no existe una ciudad (CD), por lo que cada uno de sus centros urbanos se denominan simplemente como urbanos (U).
Rural	Asentamiento humano concentrado o disperso que posee 1.000 o menos habitantes, o entre 1.001 y 2.000 habitantes, donde más de 50% de la población que declara haber trabajado se dedica a actividades primarias. Por ejemplo, la comuna de Talca posee tres entidades rurales: Villa Illinois, El Porvenir y El Oriente.	Conformado por el conjunto de entidades clasificadas como rurales. Tienen un tamaño poblacional menor a 1.000 habitantes o entre 1.001 y 2.000 habitantes con predominio de población económicamente activa (PEA) dedicada a actividades primarias.
Ciudad (CD)	No aplica	Es un gran centro urbano conformado por uno o un conjunto de centros urbanos

		adyacentes con 40.000 o más habitantes.
Resto de Área Urbana (RAU)	No aplica	Conformado por un conjunto de centros urbanos que totalizan menos de 40.000 y más de 2.000 habitantes. Esta clasificación se da cuando en una comuna existe una ciudad (CD) y entonces todos los centros urbanos restantes, si es que existen, se denominan resto de área urbana (RAU).

Las Figura 15 y Figura 16 presentan la estratificación de ambos marcos en el área urbana y rural.

Figura 15. Estratificación geográfica del MM 2015

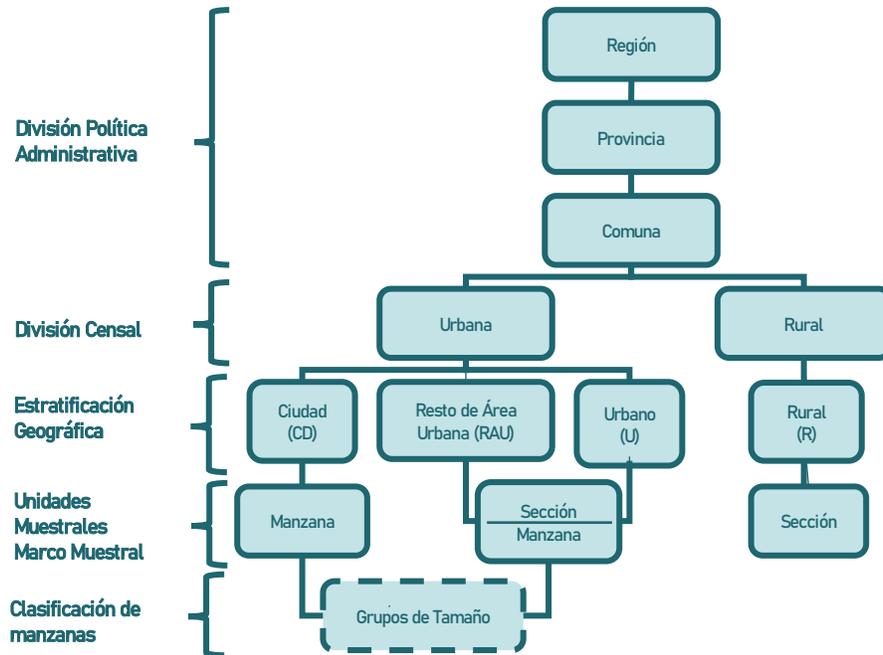


Figura 16. Estratificación geográfica del MMV 2020



En el MMV 2020, para la conformación de UPM urbanas, se consideraron como punto de partida 559 entidades urbanas definidas por el Censo 2017, según la división político-administrativa de las comunas, y la identificación de entidades urbanas tales como ciudades y pueblos. Por otro lado, para la conformación de UPM rurales, se consideraron todas las entidades rurales levantadas en el Censo 2017, correspondientes a 319 comunas a nivel nacional.

12.2.6. Comparación MMV 2020 Y MM 2015

El MMV 2020 se diferencia del Marco Muestral 2015 (MM 2015) en diversos aspectos entre los que se encuentran los siguientes: el MM 2015 está conformado por dos listados de unidades, el Marco Muestral de Manzanas (MMM 2015), que comprende las ciudades y grandes centros urbanos; y el Marco Muestral de Secciones (MMS 2015), para los estratos RAU (Resto de Área Urbana) y rural.

El MMM 2015 es el producto de los procesos de actualización permanente entre 2007 y 2009, generados sobre el marco muestral conformado a partir del Censo 2002, así como de una mejora puntual con foco en la Encuesta CASEN en 2014 para algunos centros urbanos de interés de la muestra. Por su parte, el marco muestral de secciones⁶⁰ (MMS 2002) fue construido a partir del Censo 2002 con una posterior subdivisión en estratos geográficos y secciones, para áreas urbanas y rurales, contando con una sola actualización según Precenso 2016 (MMS 2016). Mientras que los Marcos Muestrales de Manzanas y Secciones, en conjunto, se subdividen en área urbana, RAU y rural, el MMV 2020 es un único listado que comprende la división de área urbana y rural, sin considerar la división RAU.

⁶⁰ Cabe destacar, que este marco no se emplea para la ENUT, ya que este estudio solo tiene cobertura en el área urbana del país.

Por otro lado, el MMV 2020 contiene un total de 35.859 UPM, homogéneas en términos de su tamaño, lo que hace posible la inclusión de unidades que, en el MMM 2015 no tenían posibilidad de ser incluidas en la muestra (manzanas de 7 viviendas o menos, pertenecientes al grupo de tamaño 0⁶¹).

Los Marcos Muestrales de Manzanas y Secciones poseían problemas de cobertura (unidades no contenidas en el marco) y especificidad de la información (unidades que están en el marco con información imprecisa) producto de la ausencia de actualizaciones periódicas en las características de sus unidades primarias de muestreo. A su vez, tener marcos desactualizados implica, en casos extremos, zonas geográficas (con viviendas particulares ocupadas en su interior) sistemáticamente excluidas de los estudios sociales, debido a que los parámetros contenidos en el marco de selección podían diferir respecto de la realidad, por lo que la probabilidad de selección contenida en el marco podía estar desactualizada.

El MMV 2020 es un marco único para todo el territorio nacional y cuenta con una nueva conformación de UPM, actualización de los límites urbanos, estratificación socioeconómica⁶² y una nueva clasificación de áreas especiales, en reemplazo de lo que tradicionalmente se conocía como ADA y ACO.

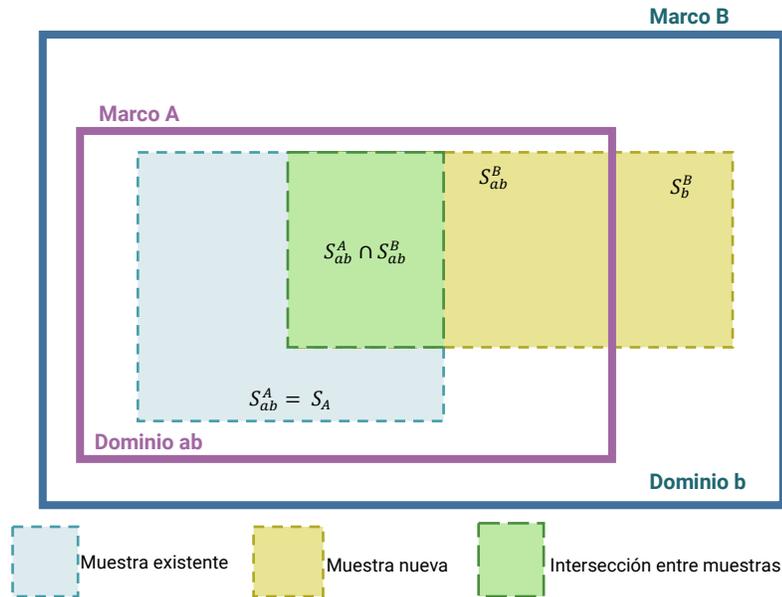
La entrada en operación del MMV 2020 fue un hito relevante en torno a la generación de procesos estadísticos de muestreo para hogares, no obstante, se mantiene el desafío de mantener los procesos de

⁶¹ Para fines de selección muestral, las manzanas del MMM se clasifican según 5 grupos de tamaño, lo cual permite generar grupos homogéneos de manzanas respecto al número promedio de viviendas para encuestar.

⁶² Para más detalles ver documento "Estratificación socioeconómica del marco muestral de viviendas 2017 (MMV 2017)" (Guerrero, Guerrero, Marín, & Pizarro, 2019).

actualización continuos. Al respecto, la Figura 17 presenta el traslape del MM 2015 (marco A en la figura) con el MMV 2020 (marco B).

Figura 17. Esquema de configuración de marcos y muestras



El marco A está completamente contenido en el marco B, es decir, no tiene unidades que no estén cubiertas por el marco B, mientras que el marco B, sí posee unidades que no están cubiertas por el marco A, por ejemplo, las correspondientes a las áreas de exclusión ADA. Esto muestra gráficamente que el nuevo marco de viviendas presenta una mayor cobertura que el anterior, lo que ayuda a mejorar la precisión de las estimaciones.

12.2.7. Conformación de las Unidades Primarias de Muestreo

A diferencia del Marco Muestral de Manzanas 2015 (MMM 2015) utilizado en ENUT 2015, el MMV 2020 está constituido por unidades primarias de muestreo, en adelante, UPM, que corresponden a áreas geográficas homogéneas en términos del número de viviendas particulares ocupadas observadas en el Censo 2017 o en su

actualización más reciente; haciendo distinción del área urbana y rural. En el área urbana, el tamaño medio de las UPM es de 200 viviendas, con un intervalo de 20%, es decir, la cantidad de viviendas por UPM se encuentra en el rango [160-240]. Por otra parte, en el área rural el tamaño medio es de 90 viviendas con un intervalo de 20%, quedando la UPM conformadas por una cantidad de viviendas que se encuentra en el intervalo [70-110] con algunas excepciones⁶³. Para la construcción de las UPM, el criterio fundamental es mantener unidades compactas y homogéneas, con el objetivo de que el muestreo bietápico sea eficiente y que la muestra pueda entrar en operación en forma óptima⁶⁴.

⁶³ Estas excepciones se definen con base a límites que fueron determinados en el contexto de la definición de umbrales de crecimiento y decrecimiento para la actualización del MMV. Sus valores pueden ser utilizados de forma referencial.

⁶⁴ El uso de UPM homogéneas en tamaño, trae bondades desde el punto de vista del diseño muestral, ya que, al seleccionar un número similar de viviendas en cada uno de ellos, se reduce la variabilidad de los factores de expansión, disminuyen los posibles valores atípicos y, por tanto, mejora la precisión de las estimaciones.

12.3. Anexo 1 – Afiche II ENUT



12.4. Anexo 2 – Díptico II ENUT

¿Cómo se puede corroborar que la persona encuestadora pertenece al INE?

Siempre portan una credencial con foto y número de RUN que las identifica como parte del INE.

Revise los datos en: www.ine.cl, sección Verificación de Encuestadores, ingresando el RUN de la persona.

Puede que su hogar sea visitado nuevamente o **le contacten por teléfono para supervisar** y corroborar los datos de la entrevista realizada.

Su información es confidencial y se encuentra protegida por ley: Está protegida con sanciones legales que buscan resguardar sus datos, amparados en el secreto estadístico (Ley 17.374).

www.ine.gob.cl/enut
¡Muchas gracias por participar en la encuesta!

INSTITUTO NACIONAL DE ESTADÍSTICAS
Morandé 801, piso 22, Santiago, Chile
Teléfono: (56) 23246 1010
www.ine.gob.cl

**Encuesta Nacional sobre
Uso del Tiempo**

¡TU TIEMPO CUENTA!

Periodo de recolección:
Agosto a diciembre 2023

La ENUT permite conocer las actividades que realizan las personas en Chile

Corroboración de identidad de la encuestadora o encuestador:
<https://verificador.ine.cl>

¡Participa con confianza!

Tu hogar fue seleccionado para contestar la II Encuesta Nacional de Uso del Tiempo (II ENUT).

La II ENUT es una encuesta que permite conocer las distintas actividades que realizan las personas en Chile: cuánto tiempo trabajan, cuánto dedican a quehaceres domésticos y a cuidar a otras personas y también qué hacen con su tiempo libre.

¿A quiénes va dirigida la encuesta?

La II ENUT se realiza a todas las personas de **12 años y más que viven en el hogar.**

La II ENUT tiene **dos cuestionarios:**

- 1. Cuestionario del Hogar:** el primero pide información de caracterización sobre todas las personas que viven en el hogar, y lo contesta una persona de 18 años o más que sepa esta información.
- 2. Cuestionario sobre Uso del Tiempo:** El segundo pregunta por distintas actividades que realizan durante un día de semana y otro de fin de semana. Esto lo debe contestar cada persona del hogar que tengan 12 años o más.

Para conocer cómo los hogares de Chile distribuyen su tiempo, necesitamos que respondan todas las personas de 12 años y más.

¡El tiempo de todas las personas cuenta!
Por esto, si no están todas las personas del hogar, tendremos que volver a visitarlas para conocer también sus respuestas.

¿Por qué es importante que responda la II ENUT?

La II ENUT nos informa sobre el tiempo que dedicamos a cuidar de otras personas y a los quehaceres domésticos, relevando la importancia que tienen estas tareas en el desarrollo y bienestar de la sociedad.

Esto permite mejorar las políticas públicas para la vida de las personas.

¿Por qué escogieron mi vivienda?

Su vivienda fue escogida bajo criterios estadísticos para representar a otros hogares similares al suyo.

¡Su participación es importante!

**Encuesta Nacional sobre
Uso del Tiempo**

12.5. Anexo 3 – Carta de apertura hogar II ENUT



Su respuesta representa a Chile

**Carta de solicitud de información
II Encuesta Nacional sobre Uso del Tiempo 2023**

Señor(a)
Jefe(a) de Hogar
Presente

Estimado(a) Señor(a):

Tengo el agrado de comunicar a usted que el Instituto Nacional de Estadísticas (INE) está desarrollando la segunda versión de la Encuesta Nacional sobre Uso del Tiempo 2023, en el marco de la generación de estadísticas oficiales del país.

Esta encuesta tiene como finalidad caracterizar el tiempo destinado por las personas a las distintas actividades de la vida diaria, como el trabajo remunerado, el trabajo no remunerado y las actividades personales. Con esta información se pueden hacer visibles los aportes que realizan las mujeres y hombres a la mantención de sus hogares y de la sociedad en general. Además, esta información permitirá que las instituciones del Estado puedan diseñar políticas públicas que apunten a mejorar la calidad de vida de las personas, por medio de la conciliación de la vida laboral y familiar, relevando la importancia de los cuidados en el desarrollo de la sociedad.

El periodo efectivo de levantamiento de la encuesta será entre los meses de **agosto a diciembre del presente año**.

En este contexto, informo a Ud. que su vivienda ha sido seleccionada de forma aleatoria para participar en esta encuesta, representando a su vez a otros hogares de su comuna. Todas las respuestas que su hogar suministre serán utilizadas en forma estrictamente confidencial y están protegidas por el Secreto Estadístico, según lo dispuesto en la Ley N.º 17.374. Aquella ley faculta al INE para solicitar información a cualquier persona natural o jurídica que habita en el territorio nacional, a la vez que establece penas de cárcel a quien divulgue sus datos.

Usted puede comprobar la identidad de la persona encuestadora a través de la credencial institucional que ésta portará o ingresando su número de RUN (sin puntos, con guión y dígito verificador) en el sitio web <https://verificador.ine.cl/>. En caso de dudas, también puede consultar, vía correo electrónico, a la coordinación del proyecto, a la dirección de correo electrónico enut2023@ine.gob.cl.

Agradeciendo su valiosa cooperación para el cumplimiento de este proceso de gran relevancia para el país, saluda atentamente,

ine.gob.cl



DANIELA MORAGA FARIAS
Subdirectora de Operaciones
Instituto Nacional de Estadísticas



II ENUT

Encuesta Nacional sobre
Uso del Tiempo

ine.gob.cl