

Instituto Nacional de Estadísticas

Documento Metodológico Encuesta Nacional de Empleo (ENE)

2021

INSTITUTO NACIONAL DE ESTADÍSTICAS

Subdirección Técnica

Departamento de Estadísticas del Trabajo

Departamento de Metodologías e Innovación Estadística

Departamento de Operaciones de Estadísticas Sociales

DOCUMENTO METODOLÓGICO ENCUESTA NACIONAL DE EMPLEO (ENE)

Abril / 2022

Versión 2.0

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

Código postal: 8340148

Teléfono: 562 3246 1010

Sitio web: www.ine.cl

Correo electrónico: ine@ine.cl

Facebook: [@ChileINE](#)

Twitter: [@INE_Chile](#)

Índice

Índice	2
1 Introducción	8
2 Contexto metodológico	10
2.1 Historia de la Encuesta Nacional de Empleo	10
2.2 Antecedentes de la actual versión de la Encuesta Nacional de Empleo	11
2.3 Objetivos de la Encuesta Nacional de Empleo	12
2.3.1 Objetivo general	12
2.3.2 Objetivos específicos	12
3 Marco de referencia	13
3.1 Marco conceptual.....	13
3.1.1 Trabajo en la ocupación	14
3.1.2 Personas desocupadas	18
3.1.3 Población fuera de la fuerza de trabajo	19
3.2 Caracterización de la población en edad de trabajar.....	21
3.2.1 Caracterización sociodemográfica.....	21
3.2.2 Caracterización de la población ocupada.....	22
3.2.3 Personas desocupadas	24
3.2.4 Fuerza de la fuerza de trabajo.....	24
3.3 Referencias nacionales e internacionales	25
3.3.1 Referencias internacionales	25
3.3.2 Referencias Nacionales.....	27
4 Diseño estadístico.....	28
4.1 Universo	28
4.2 Población objetivo	28
4.3 Marco estadístico	28
4.3.1 Marco muestral maestro	28
4.3.2 Marco muestral de la Encuesta Nacional de Empleo	38
4.4 Cobertura geográfica.....	43
4.5 Unidades estadísticas	43
4.5.1 Unidad de muestreo	43
4.5.2 Unidad de información y análisis.....	43
4.6 Definición de variables	44
4.7 Nomenclatura utilizada	45
4.8 Diseño según tipo de fuente: Muestra.....	47
4.8.1 Diseño muestral.....	47
4.8.2 Selección de unidades muestrales	48
4.8.3 Cálculo del factor de expansión.....	69
4.9 Medidas de precisión	77
4.9.1 Varianza de los estimadores.....	77
4.9.2 Errores absolutos y relativos	78
4.9.3 Representatividad de las estimaciones	79
4.10 Estimación de parámetros	83

4.10.1	Estimación de Parámetros.....	83
5	Metodología de Recolección de datos	89
5.1	Técnica y métodos de recolección de los datos.....	89
5.1.1	La muestra y el proceso de distribución operativa	89
5.1.2	Selección de la muestra.....	89
5.1.3	Periodo de recolección y distribución operativa de la muestra.....	92
5.1.4	Organización del Trabajo de Campo.....	92
5.1.5	Modificaciones en la metodología de recolección de datos	93
5.1.6	Planes especiales implementados durante la pandemia	94
5.1.7	Estructura organizativa del equipo de recolección	96
5.2	Diseño de instrumento de captura de información	100
5.2.1	Diseño de la encuesta.....	100
5.2.2	Estructura del cuestionario	101
5.3	Prueba y evaluación instrumento de captura de información.....	104
5.3.1	Prueba Piloto ENE 2019.....	104
5.3.2	Pruebas de implementaciones	105
5.4	Método de capacitación para la recolección de datos	106
5.4.1	Tipos de Capacitaciones	106
5.4.2	Métodos de capacitación	107
5.5	Métodos de Supervisión.....	108
5.5.1	Diseño metodológico de la Supervisión	108
5.5.2	Fases operativas de la supervisión	108
5.5.3	Impacto de la pandemia en el proceso de supervisión	110
5.5.4	Supervisión de gabinete y control de calidad de los datos	111
6	Metodología de procesamiento de la información	115
6.1	Procesamiento de la información	115
6.1.1	Integración de datos.....	115
6.1.2	Clasificación y codificación	116
6.1.3	Análisis de consistencia de los datos.....	117
6.1.4	Edición, imputación y resultados del procesamiento.	117
7	Metodología de análisis.....	119
7.1	Preparación de los borradores de resultados	119
7.1.1	Análisis de datos preliminar	119
7.1.2	Estándar para la evaluación de la calidad de las estimaciones	119
7.1.3	Construcción de principales variables analíticas	120
7.1.4	Registros Administrativos	124
7.2	Procedimiento para la anonimización de los datos	125
7.3	Interpretación de resultados.....	125
7.3.1	Generación de contenido para el boletín y otros productos de difusión.	125
7.3.2	Generación de series estadísticas	126
7.4	Validación de resultados	126
7.5	Finalización de resultados	127
8	Alcance de la presentación de resultados.....	128
8.1	Productos estadísticos a publicar	128

8.2	Difusión de los resultados	130
8.2.1	Difusión.....	130
8.2.2	Consideraciones sobre el uso de la información.....	132
8.2.3	Indicaciones para el uso de los datos	133
8.3	Accesibilidad a los datos y Marco legal	133
9	Glosario.....	135
10	Bibliografía.....	138
11	Anexos	144
11.1	Anexo N° 1. Número de UPM, viviendas y PET según estrato geográfico de la ENE	144
11.2	Anexo N° 2. Estratificación muestral del marco muestral de la ENE	150
11.3	Anexo N° 3. Códigos de visita de la ENE.....	156
11.4	Anexo N° 4. Tasas de no logro por estrato geográfico, según información de la ENE 2018....	158
11.5	Anexo N° 5. Estimaciones de población para la Encuesta Nacional de Empleo.....	162
11.5.1	Informe estimaciones de población para la Encuesta Nacional de Empleo.....	162
11.5.2	Interpolación de población comunal a los 15 de cada mes.	162
11.6	Anexo N° 6. Listado de marginales provinciales.....	165
11.7	Anexo N° 7. Cálculo de los factores de expansión en el período de transición	166
11.8	Anexo N° 8. Construcción de insumo para la calibración por población extranjera en factores de expansión de la ENE	171
11.8.1	Metodología para la obtención de cifras mensuales de población extranjera	172
11.8.2	Resultados del modelo	174
11.8.3	Alcances técnicos, usos actuales y prospectivos de la estimación.....	176
11.8.4	Especificaciones técnicas del modelo.....	177
11.9	Anexo N°9 Flujo operativo Supervisión de Gabinete	181
11.10	Anexo N°10 Acceso a los datos	184

Índice de Tablas

Tabla 1. Ejemplos de trabajo en la ocupación (medido en la ENE) y otras formas de trabajo (actividades productivas no consideradas)	16
Tabla 2. Criterios para clasificar a la población en edad de trabajar como personas ocupadas	17
Tabla 3. Excepciones al cumplimiento de criterios para clasificar como persona ocupada	18
Tabla 4. Criterios para clasificar como persona desocupada	18
Tabla 5. UPM con tamaño fuera de rango por región	36
Tabla 6. Estratificación socioeconómica del MMV 2017 por región	37
Tabla 7. UPM especiales por región	38
Tabla 8. Cantidad de estratos geográficos por región y área	39
Tabla 9. Estratos que no alcanzan mínimo de UPM	40
Tabla 10. UPM especiales y excluidas por región, en el marco muestral de la ENE	41
Tabla 11. Cantidad de estratos de muestreo por región y área	42
Tabla 12. Características del parámetro de interés	48
Tabla 13. Efectos del diseño en combinaciones seleccionadas de tamaño de la muestra conglomerada y correlación intraclase	55
Tabla 14. Efectos de diseño por dominios de estudio, para distintos tamaños de viviendas promedio por UPM	56
Tabla 15. Errores muestrales de la tasa de desocupación para los dominios urbanos según número promedio de vivienda por UPM	57
Tabla 16. Errores muestrales de la tasa de desocupación para los dominios rurales según número promedio de vivienda por UPM	58
Tabla 17. Cálculo de tamaños definitivos	59
Tabla 18. Distribución de la muestra y resultados de simulación	60
Tabla 19. Distribución de la Muestra total mensual año 2020	91
Tabla 20. Técnicas de recolección	95
Tabla 21. Distribución equipos de recolección y supervisión a nivel nacional	97
Tabla 22. Prueba piloto	105
Tabla 23. Series o cuadros estadísticos	129
Tabla 24. Series o cuadros estadísticos de Informalidad Laboral	130
Tabla 25. Número de UPM, viviendas y PET según estrato geográfico de la ENE	144
Tabla 26. Estratificación muestral del marco muestral de la ENE	150
Tabla 27. Códigos de registro de visita de la Tarjeta de Registro del Hogar empleada en el diseño muestral anterior de la ENE	156
Tabla 28. Códigos de visita o disposición de la ENE 2021	156
Tabla 29. Tasas de no logro por estrato geográfico, según información de la ENE 2018	158

Tabla 30. Listado de marginales provinciales.....	165
Tabla 31. Estimación y stock anual de población extranjera (al 30 de junio de cada año) a partir del saldo migratorio de las Proyecciones de las Proyecciones de población base 2017	173
Tabla 32. Matriz de actividades flujo de recuperación de datos ENE	182

Índice de Figuras

Figura 1. Clasificación de la población dentro y fuera de la fuerza de trabajo en la ENE	14
Figura 2. Clasificación de la población de acuerdo con su vínculo con la fuerza de trabajo (versión extendida)	21
Figura 3. Estratificación geográfica del MMV 2017	30
Figura 4. Estratificación de entidades urbanas	31
Figura 5. Numeración por método de serpenteo	32
Figura 6. Estratificación de entidades rurales	32
Figura 7. Esquema de rotación de la muestra de viviendas.....	65
Figura 8. Incorporación de la muestra MMV 2017 a partir de las rotaciones	66
Figura 9. Incremento gradual de la muestra teórica durante el período de transición de la ENE.....	67
Figura 10. Periodos de recolección y semana de referencia.....	92
Figura 11. Organización del Trabajo de Campo	92
Figura 12. Organigrama DOES	96
Figura 13. Simbología utilizada en el cuestionario de la ENE.....	104
Figura 14. Fases operativas de la Supervisión	108
Figura 15. Pestaña Excel de Consultas a regiones.....	113
Figura 16. Pestaña Excel Reprocesamiento.....	114
Figura 17. Variación de Población Mensual. Total Nacional	163
Figura 18. Variación de Población Mensual Ajustada. Total Nacional.....	164
Figura 19. Esquema de configuración de marcos y muestras	166
Figura 20: Esquema actualizado de configuración de marcos y muestras.....	167
Figura 21. Estimación anual de población extranjera a partir del saldo migratorio de las Proyecciones de población base 2017	174
Figura 22. Estimación anual de población extranjera a partir del saldo migratorio de las Proyecciones de población base 2017 y la estimación de los modelos estudiados.....	175
Figura 23. Estimación mensual de población extranjera según modelo logístico, y estimación anual de población extranjera a partir del saldo migratorio.	175
Figura 24. Estimación de población extranjera según modelo logístico, Censo 2017 y estimación de la población extranjera residente al 31 de diciembre de 2018.	176
Figura 25. Resultados del modelo no lineal logístico de 4 parámetros	178

Figura 26. Supuestos del modelo de regresión no lineal: función de media correcta.....	178
Figura 27. Supuestos del modelo de regresión no lineal: homocedasticidad.....	179
Figura 28. Supuestos del modelo de regresión no lineal: Normalidad	179
Figura 29. Supuestos del modelo de regresión no lineal: Independencia	180
Figura 30. Flujo proceso recuperaciones	181
Figura 31. Selección de opción de descarga de base de datos en página Web institucional del INE	184
Figura 32. Formatos de bases de datos de la ENE disponibles	185
Figura 33. Selección de base trimestral a descargar	185

1 Introducción

La Encuesta Nacional del Empleo (ENE), realizada por el Instituto Nacional de Estadísticas (INE) de Chile, tiene como objetivo clasificar y caracterizar a la población en edad de trabajar según su condición de actividad. En sus resultados entrega información estadística fundamental sobre la oferta en el mercado del trabajo chileno, lo que permite orientar la toma de decisiones y la evaluación de las políticas económicas y sobre ocupación a nivel nacional.

Desde el año 2010, la ENE ha tenido una serie de modificaciones tendientes a transformarla en una encuesta que capture de mejor manera los cambios en la ocupación de las personas, asociados a las profundas transformaciones sociales y económicas del mundo en las últimas décadas. Para orientar la actualización necesaria para este propósito, la ENE ha seguido una serie de recomendaciones internacionales, a fin de asegurar comparabilidad con otros países y contar con datos confiables. En esta armonización, para la medición de estadísticas laborales se han seguido los estándares establecidos por la Organización Internacional del Trabajo (OIT) y también la Organización para la Cooperación y el Desarrollo Económico (OCDE).

Es importante señalar que, precisamente en el año 2010, se produjo el cambio más relevante en la historia de la ENE. En dicha instancia se realizaron importantes redefiniciones conceptuales y metodológicas, que tomaron como referencia la 19° Conferencia Internacional de Estadísticos del Trabajo de la OIT (OIT, 2013). Dicha actualización le permitió al INE liderar en la región el desarrollo de las encuestas de fuerza de trabajo en las distintas oficinas nacionales de estadísticas, al establecer para la ENE las más recientes recomendaciones internacionales y, de paso, facilitar futuras actualizaciones sin comprometer la continuidad de la serie de datos.

Por otro lado, a partir del año 2020, la ENE tuvo una nueva modificación, incorporando nuevas dimensiones de análisis en el cuestionario y actualizando el marco muestral de la encuesta, el diseño metodológico y el método de calibración de los factores de expansión, tomando como referencia el Censo de 2017.

El presente documento es resultado de un intenso trabajo de fortalecimiento de la ENE, iniciado en 2018, el cual contó con la asesoría de la Organización Internacional del Trabajo (OIT) y de Grupo de Especialistas en Encuestas y Empleo, conformado durante 2019 por el INE, el cual estuvo compuesto por expertos y expertas que se desempeñan en la academia y en organismos internacionales, los cuales ven materias de estadísticas, diseño y análisis de encuestas sociales, entre otras (INE, 2020b; OIT & CEPAL, 2019). La revisión periódica y el fortalecimiento de las operaciones estadísticas, con la participación de expertos y expertas externos a la institucionalidad pública, constituyen buenas prácticas internacionales, especialmente en relación a los procedimientos o supuestos que utilizan las encuestas, los que deben ser revisados con cierta periodicidad considerando el dinamismo propio de los fenómenos sociales, demográficos y económicos de los diferentes países (OCDE, 2015).

El documento se encuentra organizado en capítulos, en los que se aborda la actualización de la metodología de la Encuesta Nacional de Empleo, con énfasis en las modificaciones incorporadas a partir del año 2020.

El segundo capítulo aborda los antecedentes y el contexto metodológico en el que se enmarca el desarrollo de la ENE.

El tercer capítulo desarrolla el marco conceptual de la ENE, incluyendo la información relevante para el análisis conceptual de este producto.

El capítulo cuatro desarrolla el diseño estadístico de la ENE, las características del marco muestral y la estimación de los parámetros utilizados.

El quinto capítulo aborda la metodología de recolección de los datos, el instrumento de captura de información y los métodos de supervisión que se realizan en la recolección de los datos.

El sexto capítulo presenta una síntesis de los procedimientos que se emplean en el procesamiento de los datos.

El séptimo capítulo incluye los métodos estadísticos que se utilizan para interpretar y explicar los resultados de la ENE.

Finalmente, en el octavo capítulo, se presentan los productos estadísticos que se ponen a disposición, y las formas en que se divulga la información de la ENE.

2 Contexto metodológico

2.1 Historia de la Encuesta Nacional de Empleo

Las estadísticas sobre fuerza de trabajo tienen una larga tradición en el INE, siendo –junto al Censo de Población y Vivienda y el Índice de Precios al Consumidor (IPC) – las mediciones más antiguas que se realizan en la institución. En particular, la ENE tiene origen en el año 1966, en el marco de la discusión internacional acerca de la pertinencia de contar con un sistema de encuestas de hogares por muestreo para obtener información sobre fuerza de trabajo y condiciones de vida, siendo una de las encuestas continuas más antiguas de América Latina (Cecchini, 2005).

Desde ese entonces la ENE ha experimentado diversas modificaciones, fundamentalmente asociadas con la actualización de los marcos muestrales (Banco Central de Chile, 2005), pero también a las implicancias conceptuales y metodológicas relacionadas con la adopción de estándares y buenas prácticas internacionales. Considerando los alcances de estas modificaciones, se identifican tres grandes períodos en la historia de la ENE.

El primer período, comprendido entre 1966 y 1985, se puede definir como de instalación de la encuesta como fuente estadística para mediciones oficiales sobre la fuerza de trabajo chilena. La principal característica del período es la búsqueda de una metodología estable en el tiempo, observándose continuos ajustes en las definiciones conceptuales, el tamaño muestral, en el diseño del instrumento, en los niveles de estimación y en la periodicidad de la publicación. Sin ir más lejos, hasta 1972 el estudio era conocido como Encuesta Continua de Mano de Obra (ECMO), usando como marco muestral el Censo de Población de 1960. Entre 1973 y 1974 no se llevó a cabo y en 1975 pasó a denominarse –por primera vez– Encuesta Nacional de Empleo, esta vez utilizando como marco el Censo de 1970 y cambiando los conceptos para identificar personas ocupadas y desocupadas (Jadresic, 1986).

El segundo período abarca entre 1986 y 2009, en el que la ENE se consolida a nivel nacional como una de las principales encuestas de hogares del país, al estandarizar tendencias de mejora que venían del período anterior (como la publicación mensual a partir de trimestres móviles). El cuestionario se estabilizó en un set mínimo de preguntas, lo mismo que el diseño muestral al establecer permanentemente dominios de estudio a nivel regional, pero también incorporando otras desagregaciones menores (como algunas provincias y ciudades). Sin embargo, las rápidas transformaciones sociales y económicas observadas en el período, principalmente asociadas a la masiva incorporación femenina a la fuerza de trabajo y la aparición de nuevas modalidades de trabajo, llevó a la necesidad de actualizar la encuesta.

Finalmente, el período comprendido desde 2010 a la fecha se ha caracterizado por la incorporación continua de las discusiones y estándares internacionales en la medición de empleo, con mayores niveles de detalle en la caracterización de las diferentes categorías de la condición de actividad de la población en edad de trabajar y con marcos muestrales más eficientes. Otras características relevantes son la flexibilidad del cuestionario para incorporar nuevas dimensiones de análisis y una activa política de transparencia que pone a disposición de las personas usuarias de información tabulaciones y bases de datos.

2.2 Antecedentes de la actual versión de la Encuesta Nacional de Empleo

El año 2010, la ENE tuvo la mayor modificación conceptual y metodológica de toda su historia, motivada por la incorporación de los estándares más actualizados para encuestas de fuerza laboral, lo suficientemente exhaustivos como para dar cuenta de fenómenos que han cambiado por completo la fisonomía del mercado laboral, pero a la vez en línea con toda la discusión internacional sobre temas de ocupación y desocupación, lo que permitirá realizar modificaciones en el futuro sin comprometer la continuidad de la serie ni la comparabilidad con otros países.

En detalle, se integraron nuevas baterías de preguntas que permiten clasificar de mejor modo a la población en edad de trabajar (PET), es decir de quince años o más, en relación a su condición de actividad, al establecer definiciones y secuencias más estrictas, así como también ampliando la cobertura de indicadores para caracterizar desde diversos focos la ocupación, la desocupación y las personas fuera de la fuerza de trabajo. De igual modo, se incorporaron nuevos temas relacionados con la calidad del empleo, el trabajo de tiempo parcial involuntario y la subutilización de la propia fuerza de trabajo, así como también la identificación de la fuerza de trabajo potencial, conocida como personas inactivas potencialmente activas, entregando una interesante herramienta para analizar el comportamiento del mercado laboral.

Otra importante precisión a destacar es la inclusión de preguntas sobre informalidad laboral a partir del trimestre julio-septiembre de 2017, tanto para personas ocupadas dependientes como independientes, incorporando en la producción continua de la ENE nuevos indicadores como la tasa de ocupación informal, la tasa de ocupación en el sector informal y el total de personas ocupadas informales, con diversas desagregaciones como grupo de ocupación, rama de actividad económica y categoría laboral (INE, 2020a).

Finalmente, a partir del año 2020 se incorporó en la ENE una nueva serie de actualizaciones tendientes a capturar de mejor manera formas de ocupación emergentes (como por ejemplo, la realizada a través de aplicaciones digitales), así como también nuevas relaciones de trabajo, incluyendo dentro de éstas las relaciones de autoridad entre los trabajadores y las unidades económicas en las que trabajan o para las que realizan el trabajo y también como los riesgos económicos derivados de la situación contractual u otras condiciones en las que se realiza el trabajo (INE, 2021; OIT, 2018).

2.3 Objetivos de la Encuesta Nacional de Empleo

2.3.1 Objetivo general

El objetivo de la Encuesta Nacional de Empleo es clasificar y caracterizar a la población en edad de trabajar (definida como aquellas personas de quince años o más), con residencia habitual en viviendas particulares ocupadas en Chile, según su condición de actividad.

Esto, cumpliendo los estándares internacionales de medición de las estadísticas del trabajo, a fin de proporcionar indicadores que permitan analizar y comparar las características ocupacionales del país, de forma coyuntural y estructural. La clasificación y caracterización de la condición de actividad permite identificar a personas ocupadas, desocupadas y fuera de la fuerza de trabajo, que son las desagregaciones básicas en encuestas de fuerza laboral. Por consiguiente, es posible estimar los indicadores tradicionales, que son las tasas de desocupación, ocupación y participación (INE, 2020b).

2.3.2 Objetivos específicos

A partir de la definición del objetivo general, se desprenden tres objetivos específicos:

1. Cuantificar a la población de quince años o más, de acuerdo con su clasificación en tres grupos mutuamente excluyentes: personas ocupadas y personas desocupadas (ambas constitutivas de la denominada fuerza de trabajo), así como también, las personas fuera de la fuerza de trabajo (denominado popularmente como personas inactivas).
2. Caracterizar a la población de personas ocupadas, respecto a las siguientes modalidades de inserción en el mercado laboral: rama de actividad económica, grupo ocupacional, situación en el empleo, informalidad laboral y horas trabajadas.
3. Calcular mensualmente los indicadores tradicionales de encuestas de fuerza laboral (tasas de desocupación, de ocupación y de participación) a nivel nacional y en los correspondientes niveles de estimación definidos en el diseño muestral.

3 Marco de referencia

El presente capítulo expone una revisión de los principales conceptos que utiliza la ENE para clasificar y caracterizar a la población en edad de trabajar, según su situación en la fuerza de trabajo.

Como sustento conceptual, la ENE recoge los más actualizados criterios y las buenas prácticas de clasificación emanados de las resoluciones de las Conferencias Internacionales de Estadísticos del Trabajo (CIET) de la Organización Internacional del Trabajo (OIT).¹

3.1 Marco conceptual

La vigente ENE utiliza como marco conceptual la resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo, de la 19º CIET del año 2013, principalmente, del trabajo en la ocupación. En dicha resolución se exponen una serie de criterios y definiciones que guían y norman a la encuesta, y que se revisarán en la presente sección.

La población en edad de trabajar (PET) es un concepto central de las estadísticas laborales y se define como aquellas personas de quince años o más. De acuerdo a la OIT,² el tamaño y la composición de esta población puede afectar considerablemente al mercado de trabajo y a la economía. Una población en edad de trabajar cada vez mayor ofrece oportunidades de crecimiento económico y, al mismo tiempo, crea inconvenientes para la creación de empleo y la integración de los nuevos participantes en el mercado laboral. En cambio, una población en edad de trabajar cada vez más reducida puede crear inconvenientes para el crecimiento económico, la competitividad, la dependencia de la población, etc.

Sin embargo, no todas las personas que forman parte de la población en edad de trabajar participan activamente en el mercado laboral. Algunos tienen trabajo, otros buscan trabajo, otros se desalientan, se dedican únicamente a otras actividades o no están interesados en el mercado laboral, por lo que se necesitan estadísticas que nos permitan comprender cómo se relacionan las personas con éste y cómo cambia con el tiempo. Estas estadísticas requieren definiciones claras para garantizar la coherencia y la claridad de la medición, la información y la interpretación.

Tal y como se define en las normas internacionales, existen tres categorías principales, personas ocupadas, personas desocupadas y personas fuera de la fuerza de trabajo, por lo que la población activa está formada por las personas en edad de trabajar que participan activamente en el mercado de trabajo, es decir la suma de las personas ocupadas y las desocupadas (personas cesantes y personas que buscan trabajo por primera vez). En conjunto, estos dos grupos de población en edad de trabajar representan la oferta de mano de obra para la producción de bienes y servicios a cambio de la remuneración o beneficio que existe en un país en un momento dado. Por otro lado, las personas fuera de la fuerza de trabajo son aquellas personas en edad de trabajar que durante el período de referencia no estaban ni en la ocupación ni en la desocupación

Cobra relevancia en este esquema inicial el concepto de *periodo de referencia*. Para efectos de la ENE, éste se establece como la semana inmediatamente anterior a la entrevista, entre lunes y domingo. Las

¹ Las CIET se realizan cada cinco años, con la participación de expertos de los gobiernos, principalmente de los ministerios responsables del trabajo y de las oficinas nacionales de estadística, así como de las organizaciones de empleadores y de trabajadores. Para mayor información, véase página Web de OIT con detalles de la instancia: <https://ilo.stat.ilo.org/es/about/standards/cls/>

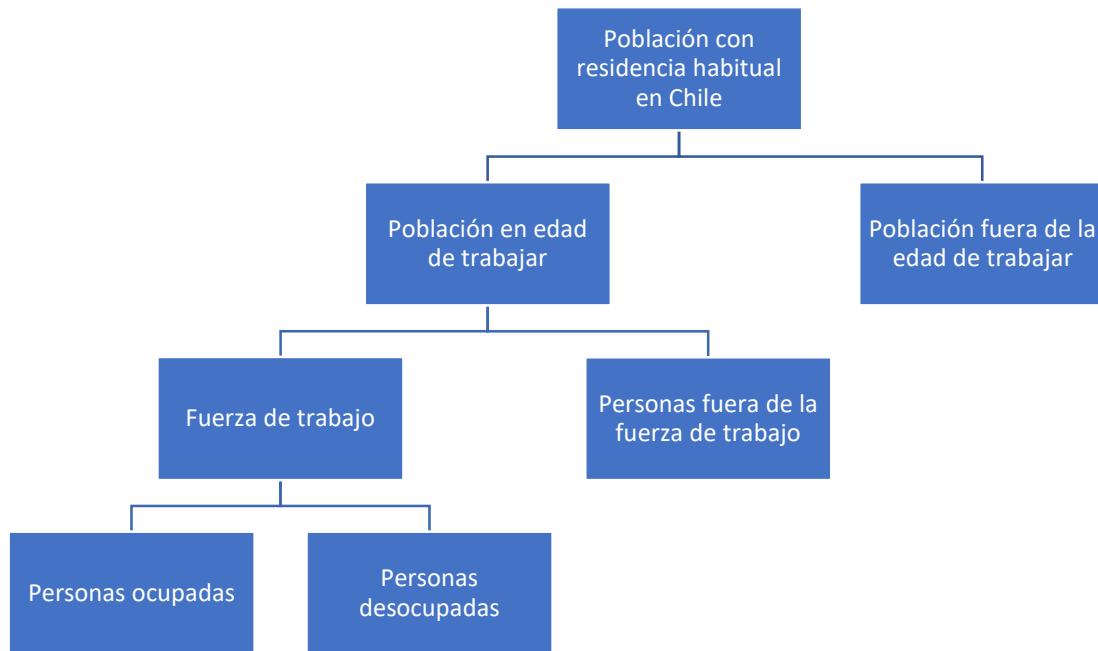
² Véase “Estadísticas sobre la población en edad de trabajar y la fuerza de trabajo”, disponible en <https://ilo.stat.ilo.org/es/topics/population-and-labour-force/>

preguntas utilizadas para identificar si una persona clasifica en la ocupación o no, se refieren específicamente a este periodo de tiempo.

Respecto a la situación en la fuerza de trabajo, para la clasificación de la población en alguna de las tres categorías mencionadas anteriormente, el cuestionario de la ENE sigue una lógica de acuerdo a un conjunto de reglas de prioridad: i) la ocupación tiene prioridad sobre las otras dos categorías; y ii) la categoría de la desocupación la tiene sobre fuera de la fuerza de trabajo. Estas tres categorías de la situación en la fuerza de trabajo son mutuamente excluyentes y exhaustivas. (OIT, 2013a).

La Figura 1 resume la clasificación de la población de acuerdo a su situación respecto a la fuerza laboral, diferenciando claramente entre quienes son parte y quienes están fuera de ella.

Figura 1. Clasificación de la población dentro y fuera de la fuerza de trabajo en la ENE



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

3.1.1 Trabajo en la ocupación

El marco conceptual de la OIT identifica cinco formas de trabajo, mutuamente excluyentes, las que se distinguen en función del destino previsto de la producción (para uso final propio o para el consumo de terceros) y en la naturaleza de la transacción (transacciones monetarias o no monetarias y transferencias). Las cinco formas de trabajo son:

1. Trabajo de producción para el autoconsumo;
2. Trabajo en la ocupación;
3. Trabajo en formación no remunerado;
4. Trabajo voluntario; y
5. Otras actividades productivas.

Las actividades realizadas para terceros a cambio de remuneración o beneficios están comprendidas en el trabajo en la ocupación. Esta forma de trabajo establece el alcance de las actividades de referencia para las estadísticas de la fuerza de trabajo. El concepto de fuerza de trabajo se refiere a la oferta de trabajo disponible en un momento dado para la producción de bienes y servicios a cambio de remuneración o beneficios. Es decir, corresponde a todas aquellas personas de quince años o más que, durante la semana de referencia, cumplen los requisitos para ser incluidas en la categoría de personas ocupadas o desocupadas o bien, planteado de otra forma, personas que se encuentran vinculadas con el mercado laboral (OIT, 2013a). **El trabajo en la ocupación es el foco principal de análisis en la ENE.**

Para diferenciar entre personas en la fuerza de trabajo y personas fuera de la fuerza de trabajo, se recurre a dos conceptos: i) recibir una remuneración o beneficio por el bien producido o el servicio prestado; y ii) haber dedicado al menos una hora a la actividad.

3.1.1.1 Criterio de remuneración o beneficio

El criterio de remuneración o beneficio se refiere al trabajo realizado en el contexto de una transacción a cambio de ingresos a ser pagados en forma de sueldos o salarios por el tiempo trabajado o el trabajo efectuado, o en forma de beneficios derivados de los bienes producidos y servicios prestados. Aquel intercambio supone transacciones de mercado y puede presentarse en dinero o en especies.

Las actividades consideradas dentro del trabajo en la ocupación son el conjunto de acciones que contribuyen a producir bienes y servicios, sean o no de carácter legal, y que se dan en un marco de transacciones de mercado, que suponen un consentimiento entre las partes. Asimismo, en ese grupo no se consideran otras formas de trabajo como el trabajo de producción para el autoconsumo, como son las tareas domésticas y de cuidados no remunerados. Tampoco incluyen el trabajo voluntario no remunerado ni el trabajo en formación no remunerado. Estas actividades productivas se recogen en las otras cuatro formas de trabajo distintas a la ocupación.

3.1.1.2 Criterio de una hora

El criterio de una hora, establecido por la OIT en la 13° CIET de octubre 1982 y ratificado en la 14° CIET de octubre de 1987 (OIT, 1982, 1987), señala que "estar trabajando" significa haber contribuido a la producción de bienes y servicios tal como se define en las cuentas nacionales, mediando un sueldo o un salario, en duración o en especie, por una duración de al menos una hora durante el período de referencia.

La relevancia de este criterio es que permite contar con cierta complementariedad con el Sistema de cuentas nacionales de las Naciones Unidas y las normas relativas a otros sistemas estadísticos. Es importante señalar que el criterio de una hora permite no excluir las ocupaciones precarias, asegurando la cobertura de todas las actividades productivas llevadas a cabo (incluidas las actividades a tiempo parcial o con carácter temporal, ocasional o esporádico) y asegurando una medición exhaustiva de todos los insumos de trabajo relacionados con la producción.

Considerando el criterio de la remuneración o beneficio, así como también el criterio de una hora, la ENE cuenta con un marco conceptual que permite identificar a las personas ocupadas, identificadas como aquellas que realizaron trabajo en la ocupación al menos una hora, realizado para terceros cambio de remuneración o beneficios.

En la Tabla 1 se presentan ejemplos del trabajo en la ocupación, capturado en la ENE, y de otras formas de trabajo o actividades productivas que la encuesta no contempla.

Tabla 1. Ejemplos de trabajo en la ocupación (medido en la ENE) y otras formas de trabajo (actividades productivas no consideradas)

Trabajo en la ocupación	Otras formas de trabajo
Trabajo doméstico remunerado para terceros	Realizar quehaceres domésticos del hogar, propio o ajeno, sin recibir pago
Elaboración de productos en el hogar para la venta por encargo	Elaboración de productos en el hogar para autoconsumo
Aprendiz o personal en formación en conocimientos relacionados con la actividad productiva de una empresa que percibe una remuneración	Formación no relacionada con las actividades productivas de una empresa, aunque se reciba un subsidio financiero
Trabajo en fundaciones o en organizaciones sin fines de lucro por remuneración (administrativos, investigadores, coordinadores)	Voluntariado no remunerado en fundaciones o en organizaciones sin fines de lucro
Cuidado de niños, ancianos, enfermos o población con discapacidad, sean familiares o no, por una remuneración	Cuidado de niños, ancianos, enfermos o población con discapacidad, sean familiares o no, sin recibir una remuneración por ello
Venta de productos en la calle o en su casa, ya sea comprados o elaborados en el hogar	--
Preparación o promoción de productos en la vía pública para su posterior venta	--
Actividades de caza y pesca con fines de venta en el mercado	Actividades de caza y pesca con fines de recreación y esparcimiento
Recolección de material de desecho (latas, cartones, vidrio, etc.) para su venta	--
Trabajo en las tierras propias o de otra persona para obtener un producto con destino previsto de venta en el mercado	--
Trabajo en permuta, donde se intercambia trabajo entre familiares con fines productivos, como la explotación agrícola en el predio de otro agricultor (minga).	--
Trabajador independiente que destina tiempo a planificar, mejorar negocio o buscar clientes.	---
Gerencia de la propia empresa o hacienda, aunque no se trabaje en la producción	Inversión en negocio sin contribuir a su gestión o funcionamiento, como tener acciones de una sociedad
Trabajo sin remuneración en una empresa familiar, pero con una participación en las ganancias o beneficio	Trabajo sin remuneración para un integrante del hogar que no es dueño de empresa o hacienda, sino que es, a su vez, perceptor de sueldo
Trabajo sin remuneración para un miembro del hogar que es familiar	--
Reparación de útiles de trabajo para su utilización futura. Por ejemplo, cuando un pescador repara su red para futuras expediciones de pesca con fines de intercambio en el mercado	Reparación de la vivienda donde se reside

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

3.1.1.3 Identificación de personas ocupadas

Las personas ocupadas son todas las personas en edad de trabajar que, durante el período de referencia, cumplen con los siguientes criterios (Tabla 2):

Tabla 2. Criterios para clasificar a la población en edad de trabajar como personas ocupadas

Criterio	Periodo de referencia
Producir un bien o realizar un servicio	Durante la semana de referencia
Duración de al menos una hora en la actividad	Durante la semana de referencia
Recibir una remuneración o beneficio	El pago por la actividad puede ser a futuro

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Entre la población ocupada, se puede hacer una primera distinción entre personas ocupadas “tradicionales” y “no tradicionales”, lo que se relaciona con la concepción de que cada persona tiene de la labor que desempeña. Así, quien identifique en la primera pregunta de la encuesta que desarrolló un trabajo, y que recibirá un pago por ello, en dinero o en especies, queda clasificado, de acuerdo con los criterios especificados, como un “ocupado/a tradicional”. Mientras que quienes no identifiquen desde un principio su actividad como un trabajo se clasifican como “ocupados/as no tradicionales”, mediante una mayor indagación respecto de la actividad realizada.

La mayor precisión en la clasificación de las personas ocupadas permite incluir personas que desarrollan trabajos que, por sus características, ya sea en horas trabajadas, frecuencia con que lo realizan o lugar de trabajo, no habrían sido identificadas, en principio, como trabajo en la ocupación. Un ejemplo de ello suelen ser los trabajos que desarrollan quienes venden productos por catálogo, personas que, además de esta actividad, algunas veces se dedican labores relacionadas con tareas de cuidado o quehaceres en el hogar. Otro ejemplo son los estudiantes que trabajan haciendo ayudantías o clases particulares y que por ello también reciben un pago.

3.1.1.4 Excepciones en la identificación de personas ocupadas

Ahora bien, entre la población ocupada, se deben considerar dos excepciones a los criterios establecidos previamente: los/as trabajadores/as familiares no remunerados y las personas ocupadas ausentes.

Los/as trabajadores/as familiares no remunerados son quienes realizaron algún trabajo en la ocupación para un familiar, por al menos una hora durante la semana de referencia, sin recibir remuneración o beneficios, siempre y cuando el familiar sea propietario de la empresa o negocio y no sea un trabajador dependiente. Un ejemplo de estos son las familias que trabajan en un puesto en la feria y que no tienen establecido un pago para sus integrantes.

Las personas ocupadas ausentes son personas que, teniendo actualmente una ocupación, estuvieron ausentes de la misma durante la semana de referencia. Para ser considerados como tal, deben mantener un vínculo con el trabajo, formal o informal, determinado de acuerdo con la certeza de la persona entrevistada en reincorporarse o no al trabajo en un tiempo determinado, o en la percepción o no de recibir una remuneración o beneficio. En el caso del trabajo dependiente, se recurre al criterio de vínculo fuerte, donde se considera que mantiene un vínculo con el empleo producto de una relación contractual o acuerdo de palabra. Es decir, estuvo ausente por razones de vacaciones, licencias, o capacitaciones. Dentro de este grupo es posible distinguir:

- Con vínculo por manutención de pago: personas ocupadas ausentes de su ocupación, pero que mantienen un sueldo o ganancia.
- Con vínculo por pronto retorno: personas ocupadas ausentes de su ocupación, pero que retornaran en un plazo de cuatro semanas.³

El resumen de lo explicado previamente se muestra en la Tabla 3.

Tabla 3. Excepciones al cumplimiento de criterios para clasificar como persona ocupada

Tipo de excepción	Criterio que no cumple	Justificación
Trabajo para un familiar	Criterio de remuneración o beneficio	La persona trabaja para la empresa o negocio de algún familiar al menos una hora en la semana de referencia
Persona ocupada ausente	Criterio de una hora	La persona no realizó su trabajo la semana de referencia, por ejemplo, por licencia, vacaciones, etc.

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

3.1.2 Personas desocupadas

Son clasificadas como personas desocupadas aquellas de quince años o más que durante el período de referencia estuvieron:

- Sin trabajo en la ocupación, es decir, que no tenían un empleo asalariado o una ocupación independiente en la semana de referencia, como se definen en el punto anterior.
- Disponibles para trabajar, es decir, disponible para trabajar en un empleo asalariado o una ocupación independiente durante el período de referencia (las dos semanas siguientes a la semana de referencia).
- En busca de trabajo en la ocupación, es decir, que habían tomado medidas concretas para buscar un empleo asalariado o iniciar una ocupación independiente en las últimas cuatro semanas.

Esto se resume en la Tabla 4:

Tabla 4. Criterios para clasificar como persona desocupada

Criterio	Periodo de referencia
No haber trabajado en una ocupación al menos una hora	Durante la semana de referencia
Haber buscado trabajo activamente	Durante las últimas cuatro semanas
Estar disponible para trabajar	Desde el último día de la semana de referencia y hasta quince días después

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

³ Durante el año 2020 y de acuerdo con la implementación de la Ley de Protección al Empleo, las personas ocupadas acogidas a ella, se consideran como las personas ocupadas ausentes, ya que se encuentran ausentes de su trabajo, pero mantienen un vínculo con su empleador y reciben un pago, con cargo a su seguridad social, pese a que su retorno al trabajo pueda superar el plazo inicialmente establecido de cuatro semanas.

El criterio de búsqueda considera haber realizado acciones específicas, en un período reciente (cuatro semanas). La búsqueda abarca al trabajo dependiente e independiente, a tiempo completo, parcial, eventual, de temporada u ocasional, y en general, cualquier tipo de trabajo en la ocupación, de acuerdo con lo previamente definido.

Las medidas concretas para buscar trabajo en la ocupación aluden a búsqueda activa, que incluyen consultar directamente con empleadores; revisión o respuesta a anuncios (en diarios, Internet); poner anuncios (en diarios, Internet); enviar currículos a empresas o instituciones; participar en pruebas, concursos o entrevistas para contratación; consultar en agencias de empleo; pedir a conocidos o familiares que lo recomendaran o le avisaran de algún trabajo; inscripción en la Oficina Municipalidad de Información Laboral (OMIL); realizar gestiones para establecerse por su cuenta; buscar clientes o pedidos, entre otras.

Asimismo, las medidas concretas para encontrar trabajo en la ocupación deben haber tenido lugar dentro de un período reciente especificado, establecido como las últimas cuatro semanas. De esta manera se asegura capturar si existe una real presión por entrar en la fuerza de trabajo.

Por su parte, se considera disponibilidad para trabajar el hecho de que, si se produce una oportunidad de trabajo en la ocupación, la persona debe estar dispuesta a trabajar en un período que no exceda los 15 días o dos semanas siguientes a la semana de referencia de la encuesta.

Se establece que si una persona satisface las condiciones de la definición de desocupación se considerará como tal, aunque compatibilice esta situación con otras como, por ejemplo, la de estudiante o labores del hogar.

Las personas desocupadas se subdividen en:

- **Cesantes:** son todos las personas desocupadas que han tenido anteriormente un trabajo en la ocupación que durara por lo menos un mes.
- **Busca trabajo por primera vez:** población que, habiendo cumplido con los criterios para ser clasificado como persona desocupada, no ha tenido un trabajo en la ocupación que durara por lo menos un mes anteriormente.

Existen ciertas excepciones a la regla en casos muy específicos. Por ejemplo, no se considerarán desocupadas a las personas que no trabajaron durante la semana de referencia, que estaban disponibles para trabajar y que no buscaron empleo, porque ya encontraron uno con el que tienen un vínculo formal (contrato o gestiones realizadas), al cual se incorporarán en más de cuatro semanas respecto a la semana de referencia. Este grupo de personas se conoce como iniciadores y se contabilizan dentro del grupo de las personas fuera de la fuerza de trabajo.

3.1.3 Población fuera de la fuerza de trabajo

La población fuera de la fuerza de trabajo abarca a todas las personas de quince años o más que no tuvieron una ocupación y tampoco buscaron y/o estuvieron disponibles para trabajar en la ocupación. En otras palabras, que no quedan clasificadas como ocupadas ni desocupadas durante la semana de referencia. Estas personas, por tanto, están fuera de la fuerza de trabajo.

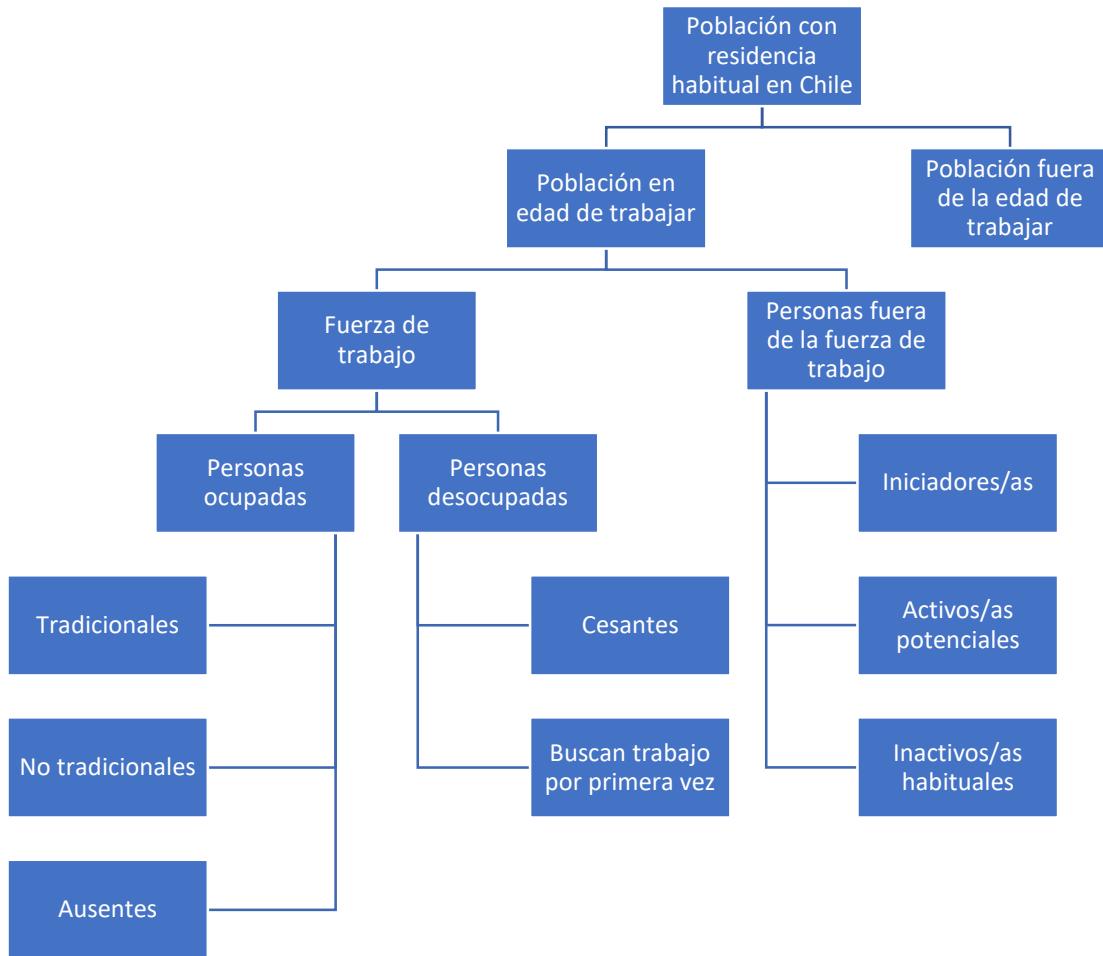
En este grupo es posible identificar tres subgrupos, de acuerdo con su disponibilidad o la potencialidad con la que puedan transitar dentro o fuera de la fuerza de trabajo (ya sea como personas ocupadas o desocupadas). Estos son:

- **Personas inactivas potenciales:** los inactivos potencialmente activos o fuerza de trabajo potencial son aquellos que buscaron trabajo en la ocupación, pero no estarían disponibles para trabajar; o no buscaron, pero estarían disponibles. Estos casos se evalúan independientemente de la razón por la cual se dan esas situaciones. La potencialidad en quienes están fuera de la fuerza de trabajo hace referencia a una presunción de la factibilidad de que una persona que no está presionando a la fuerza de trabajo, efectivamente lo haga en algún momento.
- **Personas inactivas habituales:** son aquellas personas que no buscaron trabajo en la ocupación y no están disponibles para comenzar uno.
- **Personas iniciadores/as:** Son personas que no han buscado trabajo en la ocupación en las últimas cuatro semanas debido a que iniciarán pronto un trabajo y que pueden o no estar disponibles para trabajar en las dos semanas siguientes a la semana de referencia.⁴

Todo lo presentado hasta este momento en materia de clasificación de la población en edad de trabajar se resume en la Figura 2.

⁴ Estos últimos, si señalan estar disponibles, tanto la OIT como los países miembros de la Organización para la Cooperación y el Desarrollo Económico (OCDE), recomiendan incorporarlos dentro de las personas desocupadas, y, por lo tanto, debiesen ser incluidos dentro del cálculo de la tasa de desocupación. En caso contrario, como es la situación de Chile, se les debe considerar fuera de la fuerza de trabajo, pues no buscaron ni estarán disponibles para trabajar. En general, la disponibilidad supone una mayor participación -en el corto plazo- en la fuerza de trabajo, pues, en esencia, ser iniciador (disponible o no) es una etapa intermedia entre estar fuera de la fuerza de trabajo y formar parte de ella. De acuerdo con la gradación de las personas fuera de la fuerza de trabajo, según su potencialidad de hacer parte de la misma, los iniciadores son el subconjunto con mayor potencialidad. Para fines analíticos y comparativos a nivel internacional, la ENE publica todos los meses la tasa de desocupación incluyendo a los iniciadores disponibles, como una medida complementaria a la tasa de desocupación tradicional y como parte de los indicadores de subutilización de la fuerza de trabajo.

Figura 2. Clasificación de la población de acuerdo con su vínculo con la fuerza de trabajo (versión extendida)



Fuente: Elaborado por el Instituto Nacional de Estadísticas

3.2 Caracterización de la población en edad de trabajar

Una vez identificada y clasificada la población en edad de trabajar según su situación con la fuerza de trabajo, la ENE la caracteriza de acuerdo con una serie de variables sociodemográficas y económicas.

3.2.1 Caracterización sociodemográfica

La captura de preguntas sociodemográficas permite caracterizar a la población a partir de ciertas variables básicas, dentro de las cuales se indaga sobre el sexo, la edad, fecha de nacimiento, el estado civil, el nivel educacional, el parentesco, el proveedor principal, la nacionalidad, migración, pueblos indígenas, entre otras.

Para la caracterización de la población según su nivel educativo, se utiliza la Clasificación Internacional Normalizada de la Educación (CINE), concebida por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) en 1997 (UNESCO, 2006). Este instrumento sirve para la compilación y presentación de estadísticas de educación en los distintos países y, por tanto, permite la comparabilidad a nivel internacional.

Respecto a la nacionalidad y a la migración internacional, se utilizan los Códigos Uniformes de País o de Zona para Uso Estadísticos (“*Standard Country or Area Codes for Statistical Use*”), desarrollado por la División de Estadísticas de Naciones Unidas,⁵ que permite identificar el nombre del país de la persona incluida en la población objetivo de la encuesta.

3.2.2 Caracterización de la población ocupada

3.2.2.1 Por situación en el empleo

Una primera caracterización general de la población ocupada se hace de acuerdo con la situación en el empleo, cuyo marco conceptual proviene de la Clasificación Internacional de la Situación en el Empleo (CISE) del año 1993, de la Organización Internacional del Trabajo (OIT, 1993).

La CISE clasifica la ocupación de una persona en un momento dado, con arreglo al tipo de contrato explícito o implícito de trabajo del titular con otras personas u organizaciones. Los criterios básicos utilizados para definir los grupos de la clasificación son el tipo de riesgo económico y el tipo de autoridad:

- **Tipo de riesgo económico:** que determina la solidez del vínculo entre la persona y el empleo.
- **Tipo de autoridad:** que tienen o tendrán los titulares sobre los establecimientos y sobre otros trabajadores/as.

En una primera instancia, la situación en el empleo puede ser dependiente, es decir, que existe un vínculo contractual o acuerdo de trabajo que acredita subordinación a un empleador, donde puede tener o no personas trabajadoras bajo su responsabilidad; o independiente, trabajador/a sin vínculo contractual o acuerdo de trabajo, que puede tener o no trabajadores/as bajo su responsabilidad, asumiendo el riesgo económico inherente del ejercicio de su actividad.

3.2.2.2 Por tipo de ocupación

Además de la situación en el empleo, la población ocupada se caracteriza a través de la aplicación de la Clasificación Internacional Uniforme de Ocupaciones (CIUO), que es una herramienta para organizar las ocupaciones en una serie de grupos definidos en función de las tareas que se desarrolla. La CIUO-08 define el empleo como “un conjunto de tareas y cometidos desempeñados por una persona, o que se prevé que ésta desempeñe, incluido para un empleador o por cuenta propia”. A su vez, se define la ocupación como “un conjunto de empleos cuyas principales tareas y cometidos se caracterizan por un alto grado de similitud” (INE, 2018).

Los empleos se clasifican por ocupación con respecto al tipo de trabajo realizado o por realizar. Los criterios básicos utilizados para definir el sistema de grandes grupos, subgrupos principales, subgrupos y grupos primarios son el “nivel de competencias” y la “especialización de las competencias” requeridos para efectuar eficazmente las tareas y cometidos de las ocupaciones.

En sus comienzos la ENE utilizó la versión CIUO-88, la cual se encontraba estructurada en 10 grandes grupos. Sin embargo, a partir del año 2018 este clasificador fue adaptado a la realidad chilena en su versión CIUO-08.CL, manteniendo los principios básicos y la estructura de la clasificación, pero adaptando la versión.

⁵ Véase <https://unstats.un.org/unsd/methodology/m49/>

3.2.2.3 Por rama de actividad económica

El marco conceptual de la Clasificación Industrial Internacional Uniforme (CIIU) permite caracterizar el sector económico en que se desarrolla el empleo o la ocupación, considerando la rama de actividad económica. En este marco, es preciso comprender el concepto de unidad económica, referido aquellas entidades (institución, empresa, negocio o persona) que se dedican a la producción de bienes, compra o venta de mercancía y prestación de servicios públicos y/o privados.

Actualmente, y de manera oficial desde el año 2016, la ENE utiliza la Clasificación de Actividades Económicas Nacional de Encuestas Sociodemográficas (CAENES), es una adaptación del CIIU4.CL 2012, desarrollado por el INE siguiendo recomendaciones internacionales (INE, 2016). Esta surge de la necesidad de recopilar información actualizada sobre los sectores productivos en encuestas de hogares, como también para cumplir el compromiso que la Institución asumió con la Organización para la Cooperación y Desarrollo Económico (OCDE).

La tercerización se caracteriza por volver difusos los canales de comunicación entre la persona trabajadora y quienes ejercen posiciones de autoridad sobre ella. Además, puede darse una situación de “doble jefatura”, donde persona trabajadora deba responder por sus acciones tanto a las jefaturas de la unidad donde trabaja como a los de la unidad donde está contenido el acuerdo de trabajo.

En la ENE se identifican distintas situaciones de tercerización en el mercado laboral nacional:

- Con un contratista o subcontratista de bienes o servicios
- Con una empresa de servicio temporales o suministradora de trabajadores
- Con un enganchador (contratista agrícola)

3.2.2.4 Según formalidad del trabajo en la ocupación

La Organización Internacional del Trabajo (OIT) señala que para realizar un completo seguimiento del mercado laboral no es suficiente observar únicamente los indicadores tradicionales como la tasa de desocupación, ocupación y participación. Se requiere complementar el análisis con indicadores que permitan visibilizar otros grupos de la población que se encuentran en una situación desfavorable en términos de su vínculo y participación en el mercado laboral. Por lo que se propone indagar, entre otros temas, la informalidad laboral (OIT, 2003, 2013b).

En Chile se adopta la siguiente definición de sector y ocupación informal (INE, 2020):

- **Sector informal:** Son todas aquellas unidades económicas de mercado que no cuenten con registro en el Servicio de Impuestos Internos (SII) y tampoco puedan ser clasificadas como quasi-sociedades, ya que no poseen una contabilidad completa o simplificada que les permita realizar una efectiva separación de sus gastos, por el lado de las personas trabajadoras por cuenta propia o empleadores, o a su vez, aquellas empresas que no tengan una oficina contable o no cuenten con los servicios de un contador, desde el punto de vista de los/as trabajadores/as asalariados.
- **Ocupación informal:** Son todas aquellas personas asalariada o trabajadoras del servicio doméstico que no cuentan con cotizaciones de salud (Isapre o Fonasa) y previsión social (AFP) por concepto de su vínculo laboral con un empleador. Asimismo, se consideran como ocupados/as informales por definición a todos los familiares no remunerados del hogar, además de los/as trabajadores/as por cuenta propia y empleadores propietarios/as de una unidad económica del sector informal.

La elaboración de indicadores oficiales sobre la ocupación informal y la ocupación en el sector informal fue posible a través de la implementación de un conjunto de nuevas preguntas al interior del flujo habitual del cuestionario de la ENE, enfocadas en la identificación y caracterización de la situación de formalidad desde distintas dimensiones. A contar de julio 2017, se dio inicio al levantamiento oficial de las nuevas preguntas de informalidad laboral y, por tanto, el primer dato publicado de la serie fue el trimestre móvil julio-septiembre 2017.

Adicionalmente, otras variables de interés para categorizar a la población ocupada refieren a las prestaciones laborales, seguridad social, horas trabajadas, distinguiendo entre horas acordadas, efectivas y habituales, entre otras.

3.2.3 Personas desocupadas

En el caso de la población desocupada, además de las variables de caracterización sociodemográfica, podemos utilizar otras variables de orden económico para clasificar a uno de los subgrupos que lo componen, específicamente, a los/as cesantes, es decir, aquellas personas que han tenido anteriormente un empleo que duró por lo menos un mes.

La Organización Internacional del Trabajo (OIT) plantea que el estudio y seguimiento del mercado laboral no debe ser descrito únicamente con la tasa de desocupación, sino que además debe ser complementado con indicadores sobre la subutilización de la fuerza de trabajo, que permitan describir con mayor profundidad el mercado laboral, de tal forma de visibilizar otros grupos de la población que se encuentran en una situación desfavorable en términos de su vínculo y participación en la fuerza de trabajo. A partir de este lineamiento, la ENE incorpora indicadores complementarios a los de medición de la desocupación (Herrera, Huaracán, & Inostroza, 2019).

3.2.4 Fuera de la fuerza de trabajo

Como se mencionó anteriormente, en este grupo es posible identificar tres subgrupos, de acuerdo con su disponibilidad o la potencialidad con la que puedan transitar dentro o fuera de la fuerza de trabajo (ya sea como personas ocupadas o desocupadas), determinando quienes son personas inactivas potencialmente activas, inactivas habituales o personas iniciadoras.

Además de caracterizarles sociodemográficamente, es posible clasificar la población fuera de la fuerza de trabajo según razones para esa situación. Estas consideran motivaciones económicas, estacionales y personales para no entrar a la fuerza de trabajo, que permite entender la inactividad no como una condición permanente sino como una condición que puede cambiar por determinadas razones, en períodos cortos de tiempo.

El modo de clasificar las razones declaradas por las personas para estar fuera de la fuerza de trabajo son:

- **Razones estacionales:** la población fuera de la fuerza de trabajo por razones estacionales está constituida por personas que declaran “esperar la estación de mayor actividad” y que, al mismo tiempo, señalan estar disponibles. Conforma al grupo que tiene mayor posibilidad de ingresar a la fuerza de trabajo en un breve plazo de tiempo o en cuanto cambien las condiciones en el mercado laboral. Aunque no se trata de una población muy numerosa, pueden llegar a tener un impacto en las medidas de presión de oferta laboral.

- **Razones de desaliento:** quienes declaran esta razón son personas que no buscaron trabajo en la ocupación en las últimas cuatro semanas, se cansaron de buscar un empleo porque creen que no lo encontraran, debido a su edad o porque deben hacer demasiados trámites para iniciar una actividad por su cuenta, entre otras razones, pero estarían disponibles para iniciar un trabajo en las dos semanas siguientes a la semana de referencia. Al igual que quienes declaran razones estacionales, generalmente esta población se suma a los marginalmente vinculados siendo, de hecho, el grupo más numeroso entre éstos.
- **Razones personales temporales:** Dicen relación con motivos de no búsqueda de empleo por razones de salud, de embarazo, o de responsabilidades familiares de corta duración. En caso de declarar disponibilidad se clasifican como potenciales.
- **Razones familiares permanentes:** considera a todas aquellas personas que declaran razones de cuidado y/o quehaceres domésticos para no pertenecer a la fuerza de trabajo.
- **Razones de estudio:** Como la anterior, también es una categoría importante entre las personas inactivas, asociada a la condición de estudiantes.
- **Razones de pensión o montepiado:** Son las personas que gozan de un ingreso permanente como consecuencia del trabajo realizado en su vida laboral activa, o bien por viudez. La potencialidad de este grupo es muy reducida, predominando la habitualidad en su condición.
- **Razones de jubilación:** Se trata de una población con características muy similares a la anterior. En esencia, son personas que perciben un flujo de ingresos asociado a su desempeño laboral previo.
- **Razones de salud permanentes:** Se trata de personas que señalan que o no buscaron empleo, o no estuvieron disponibles para trabajar, debido a que su salud no se los permite.
- **Sin deseo o necesidad de trabajar:** Son personas que declaran no querer o no necesitar trabajar por lo que su potencial de participación es baja.
- **Otras razones:** es una categoría de personas fuera de la fuerza de trabajo que considera a todos aquellas razones que no califican en las otras categorías de inactividad.

3.3 Referencias nacionales e internacionales

3.3.1 Referencias internacionales

Como sustento conceptual de lo descrito en el capítulo anterior, y siendo los principales trabajos de las organizaciones expertas en materia de empleo, la ENE recoge los más actualizados criterios y las buenas prácticas de clasificación emanados de las **resoluciones de las Conferencias Internacionales de Estadísticos del Trabajo (CIET) de la Organización Internacional del Trabajo (OIT)**⁶.

En las CIET se formulan recomendaciones sobre temas seleccionados de estadísticas del trabajo en forma de resoluciones y directrices, que son parte del conjunto de normas internacionales, referidas a conceptos,

⁶ Las CIET se realizan cada cinco años, con la participación de expertos de los gobiernos, principalmente de los ministerios responsables del trabajo y de las oficinas nacionales de estadística, así como de las organizaciones de empleadores y de trabajadores

definiciones, clasificaciones y otros procedimientos metodológicos que se acuerdan como representativos, sobre estadísticas del trabajo.

En particular, el marco conceptual de la ENE considera como principales referencias para su desarrollo conceptual la *“Resolución sobre estadísticas de la población económicamente activa, del empleo, del desempleo y del subempleo”*, adoptada por la 13° Conferencia Internacional de Estadísticos del Trabajo (1982); la *“Resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo”*, de la 19° CIET (2013)⁷; además de la *“Resolución sobre las estadísticas de las relaciones de trabajo”* concebida en la 20° CIET del año 2018. Estos criterios son, al mismo tiempo, los mismos que adoptan OCDE y EUROSTAT para todos sus países miembros.

La 19° CIET (2013) adoptó cinco resoluciones relativas a las estadísticas del trabajo en la ocupación y subutilización de la fuerza de trabajo, el trabajo forzoso, las cooperativas y la migración laboral, además del funcionamiento de la Conferencia. En detalle, la resolución sobre las estadísticas del trabajo, ocupación y subutilización de la fuerza de trabajo identifica cinco formas de trabajo, mutuamente excluyentes, las que se distinguen en función del destino previsto de la producción (para uso final propio o para el consumo de terceros) y de la naturaleza de la transacción (transacciones monetarias o no monetarias y transferencias).

Para la medición del **trabajo en la ocupación**, la 19° CIET introdujo, además, el criterio de haber trabajado al menos una hora en el período de referencia, permitiendo no excluir las ocupaciones precarias, asegurando la cobertura de todas las actividades productivas llevadas a cabo (incluidas las actividades a tiempo parcial o con carácter temporal, ocasional o esporádico) y asegurando una medición exhaustiva de todos los insumos de trabajo relacionados con la producción. A su vez, aquel criterio permite mantener la consistencia con los conceptos de frontera de producción y actividad económica del Sistema de Cuentas Nacionales, de acuerdo al cual se calcula el valor agregado de la economía y, por ende, el producto interno bruto (PIB).

La 20° CIET (2018), adoptó cuatro resoluciones, relativas a las estadísticas sobre las relaciones de trabajo, trabajo infantil e indicadores de los Objetivos de Desarrollo Sostenible (ODS) sobre los derechos laborales y la juventud. En cuanto a las estadísticas de las relaciones de trabajo, se abordan: a) Las relaciones de autoridad entre los trabajadores y las unidades económicas en las que trabajan o para las que realizan el trabajo y b) los riesgos económicos derivados de la situación contractual u otras condiciones en las que se realiza el trabajo.

Para la medición de la informalidad laboral en la ENE utiliza como marco conceptual las *“Directrices sobre una definición estadística de empleo informal”* de la 17° CIET de 2003, y el documento *“La medición de la informalidad: Manual estadístico sobre el sector informal y el empleo informal”* elaborado por la OIT en el año 2013.

La ENE adopta los conceptos que son parte de las directrices provenientes de las resoluciones y documentos antes mencionados para su elaboración metodológica y conceptual, es así que, en resumen la incorporación de los estándares internacionales en la ENE se traduce en concreto en:

⁷ La reunión de la decimonovena CIET se celebró en 2013, pero las reuniones preparatorias comenzaron con anticipación al año 2010, permitiendo que la nueva versión de la Encuesta Nacional de Empleo pudiera incorporar las más recientes recomendaciones de la OIT.

- Utilizar un concepto actualizado de la población ocupada, incorporando los criterios de haber trabajado al menos una hora en la semana de referencia y recibir una remuneración o beneficio por la actividad realizada.
- Distinguir, entre las personas ocupadas, a los/as ocupados/as “tradicionales” de los/as “no tradicionales”, en consideración de si la persona entrevistada reconoce o identifica en primera instancia, su actividad como trabajo.
- Actualizar la conceptualización de población desocupada, incorporando los criterios de haber buscado empleo en las últimas cuatro semanas y de estar disponible para trabajar en, al menos, las dos semanas siguientes a la de referencia.
- Clasificar las personas fuera de la fuerza de trabajo, según razones de inactividad, así como también según su potencialidad de participar en la fuerza de trabajo.
- Establecer las relaciones de autoridad y riesgo económico entre los/as trabajadores/as y las unidades económicas para las que trabajan.
- Incluir y actualizar el marco conceptual de informalidad laboral, introduciendo la perspectiva laboral basada en las características y condiciones del puesto de trabajo que ocupa una persona.

3.3.2 Referencias Nacionales

La ENE no cuenta con referencias conceptuales nacionales, ya que, como se ha descrito a lo largo de este documento, ha sido pionera en lo referido a la elaboración de estadísticas de la fuerza de trabajo, en la inclusión de elementos internacionales referidos a empleo/ocupación y como insumo para la generación de políticas públicas alusivas al trabajo en la ocupación.

Por otro lado, debido a lo descrito, este producto es referente para el desarrollo de otros productos estadísticos que aborden como tema el empleo y el trabajo en la ocupación. Esto se ha observado para productos estadísticos tanto internos del INE como externos.

4 Diseño estadístico⁸

4.1 Universo

El universo de la Encuesta Nacional de Empleo (ENE) está conformado por todas las personas que habitan en las viviendas particulares ocupadas presentes en el territorio nacional y que son cubiertas por el marco de muestreo definido para la encuesta (MMV 2017).

4.2 Población objetivo

La población objetivo de la ENE es la población en edad de trabajar, es decir, todas las personas de quince años y más, que residen habitualmente en viviendas particulares ocupadas en el país.

4.3 Marco estadístico

4.3.1 Marco muestral maestro

Un marco de muestreo se define como la lista o los procedimientos que permiten identificar a todos los elementos de una población objetivo (Groves, 2004, pág. 68). Los marcos constituyen el insumo fundamental para el diseño de encuestas por muestreo y, por consiguiente, resulta de suma importancia que su metodología de conformación sea debidamente documentada. Entre las propiedades deseables para un marco muestral, se encuentran la exhaustividad, la exactitud y su actualización (Naciones Unidas, 2009).

La propiedad de exhaustividad corresponde a que el marco muestral cubra efectivamente la totalidad de la población de interés, mientras que la propiedad de exactitud hace referencia a contar con información fidedigna, ausencia de duplicidad de registros, etc. La propiedad de actualización se refiere a la capacidad de un marco muestral de mantener información reciente y oportuna. Si bien un nivel de actualización perfecto es imposible, es importante que un marco muestral cuente con procedimientos que permitan minimizar los errores no muestrales que surjan de la desactualización.

El marco muestral que se usa con mayor frecuencia en las encuestas sociales está constituido por una lista de unidades geográficas⁹, caracterizadas y actualizadas con base en información censal, desde donde se realiza la selección, para que posteriormente las viviendas seleccionadas al interior de estas áreas pasen a integrar la muestra de cada encuesta. En este sentido, es deseable que un marco cubra en forma exhaustiva a la población de interés, que se encuentre debidamente actualizado y que disponga de información relevante para la construcción de estratos de muestreo adecuados.

⁸ Parte importante de la información de este capítulo se elaboró de acuerdo al plan de actualización del marco muestral de la Encuesta Nacional de Empleo (ENE), que se concretaría durante el año 2020, sin embargo, es relevante señalar que a partir de la irrupción de la pandemia Covid-19 en Chile (marzo de 2020), estos planes se vieron mermados, por lo que el Instituto Nacional de Estadísticas (INE) tuvo que recurrir a otras estrategias para evitar una reducción drástica de la muestra y eventuales problemas de sesgos. Toda esta información y los pasos que el INE siguió para dar continuidad a la ENE fueron documentados en las separatas técnicas que se fueron publicando mensualmente junto a los resultados de la coyuntura de empleo desde marzo de 2020 en adelante. Para más información, se recomienda revisar la página web institucional en www.ine.cl, sección ocupación y desocupación.

⁹ De acuerdo con Naciones Unidas (2009) este tipo de marcos se denomina marco de áreas, siendo aquel que comprende las unidades geográficas de un país en un orden jerárquico.

La muestra de la ENE, vigente desde enero de 2020, fue seleccionada desde el Marco Muestral de Viviendas¹⁰, elaborado a partir de los resultados del Censo de Población y Vivienda 2017, y los listados de viviendas asociados al Precio 2016 (en lo sucesivo, MMV 2017). El MMV 2017, cuenta con actualización permanente en los lapsos intercensales, en base a registros administrativos y de ser necesario a partir de un proceso de enumeración de unidades geográficas priorizadas.

El MMV 2017 está constituido por unidades primarias de muestreo (UPM), correspondientes a áreas geográficas, homogéneas en términos del número de “viviendas marco”, definidas como aquellas particulares con moradores presentes o ausentes¹¹ observadas en el Censo de Población y Vivienda de 2017, haciendo distinción del área urbana y rural. Para el área urbana, el tamaño medio es de 200 viviendas, con un intervalo del 20%, es decir, la cantidad de viviendas por UPM se encuentra en el rango [160-240]. Por otra parte, el área rural tiene un tamaño medio de 90 viviendas y un intervalo del 20%, quedando su cantidad de viviendas por UPM en el intervalo [70-110].

Dentro de los atributos del nuevo marco, destaca la generación de UPM homogéneas en tamaño, una estratificación socioeconómica del marco y la identificación de áreas de tratamiento especial, permitiendo disponer de estimaciones más precisas derivadas de encuestas que utilicen este marco.

4.3.1.1 Cobertura geográfica del marco muestral

El nuevo marco muestral 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.

4.3.1.2 Estratificación geográfica del marco muestral

El MMV 2017 contiene variables que permiten la identificación de la división política administrativa (DPA) de país¹², así como también la identificación de las áreas rurales y urbanas en cada comuna del país.

La definición de urbano y rural viene dada de las especificaciones del Censo de población y Vivienda 2017. Así las definiciones de entidades urbanas y rurales determinadas en el Censo 2017 (INE, 2019a):

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

¹⁰ Hasta diciembre de 2019, la muestra de la ENE fue seleccionada desde un marco muestral conformado por dos listados de unidades, el Marco Muestral de Manzanas actualizado con base en los listados de Precio 2016 (MMM 2016) para el estrato CD que comprendía las ciudades y grandes centros urbanos; y el Marco Muestral de Secciones también actualizado en parte, con base en los listados de Precio 2016 (MMS 2016) para los estratos RAU (Resto de Área Urbana) y rural.

¹¹ Se excluyen las viviendas de temporada y colectivas

¹² Región, provincia y comuna.

dedica a actividades primarias. Por ejemplo, la comuna de Talca posee tres entidades rurales: Villa Illinois, El Porvenir y El Oriente.

De las 345 comunas, 319 comunas tienen área urbana, de estas existen 27 comunas que contienen el estrato urbano en forma exclusiva. Por su parte, el estrato rural se encuentra en 319 comunas, de las cuales 27 lo contienen en forma exclusiva.

Figura 3. Estratificación geográfica del MMV 2017



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

4.3.1.3 Conformación de Unidades Primarias de Muestreo

Una de las principales características que tiene el MMV 2017 es la definición de UPM de tamaño homogéneo, en términos de viviendas particulares, para las áreas rurales y urbanas. La forma de construir las UPM, con base en información del Censo 2017 y listados del Precenso 2016, permite contar con unidades más estables en el tiempo, disminuir la pérdida de unidades en el levantamiento y reducir el riesgo de errores no muestrales. Asimismo, el uso de UPM homogéneas trae también bondades desde el punto de vista del diseño muestral, ya que, al seleccionar un número similar de viviendas en cada uno de ellos, reduce la variabilidad del valor de los factores de expansión, disminuye posibles *outliers* y, por tanto, mejora la precisión de las estimaciones.

a) Criterios para UPM urbanas

Para la conformación de UPM urbanas, se tomó como punto de partida las 559 entidades urbanas definidas por el Censo 2017, las cuales responden a la división político-administrativa en el nivel de comuna, y a la identificación de entidades urbanas tales como ciudades y pueblos.

Figura 4. Estratificación de entidades urbanas



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

El tamaño medio de las UPM se fija en 200 viviendas, estableciendo un rango de variabilidad entre 160 y 240, con ciertas excepciones. En el conteo, se consideran sólo “viviendas marco”, definidas como aquellas viviendas particulares ocupadas, con moradores presentes o ausentes, y las viviendas desocupadas, según lo reportado en el Censo 2017. Se excluyen las viviendas de temporada, colectivas y de otro uso.

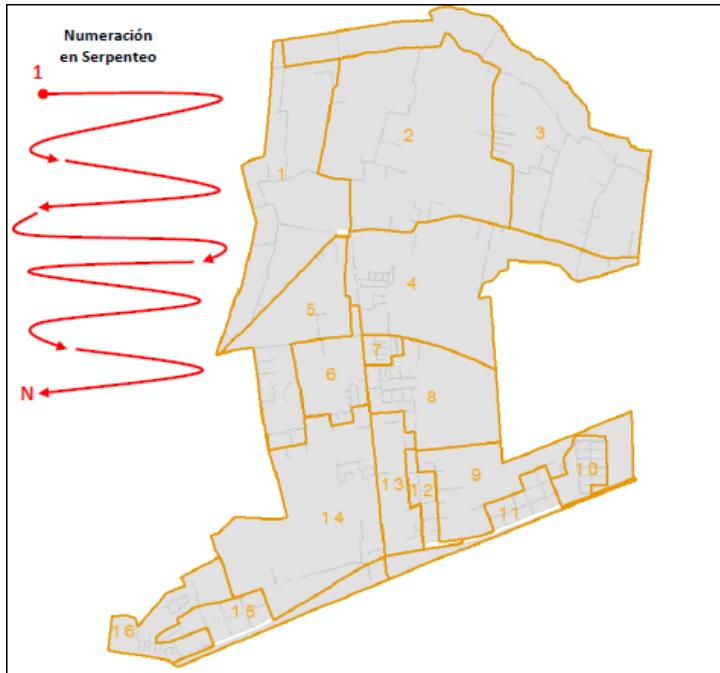
Para la construcción de las UPM, el criterio fundamental es mantener unidades compactas y homogéneas, con el objetivo de que el muestreo bietápico sea eficiente y la muestra pueda entrar en operación en forma óptima. Por esta razón, existen excepciones para el rango de tamaño de las UPM con el propósito de conservar la compacidad. Para el caso de las excepciones que exceden la cota superior, se aceptan UPM de manzanas que tengan entre 240 y 300 viviendas, ya que dividir esas manzanas produciría unidades por debajo del rango e iría en contra del principio de homogeneidad. Para las manzanas de conformación compleja, como son los casos de manzanas con formas irregulares, en las que unas rodean a otra, el criterio anterior se flexibiliza hasta 360 viviendas antes de dividirlas.

Por otra parte, para las excepciones por debajo de la cota inferior del rango, se debe considerar que, se ha establecido que deben existir al menos dos UPM en el interior de cada división urbana (ciudades, pueblos, etc.). Por esta razón, en los casos en que es necesario dividir un territorio en dos, la primera opción de división será que ambas tengan un peso de viviendas lo más cercano posible, aunque éstos queden fuera de rango. No aplicarán para estos casos dejar una óptima y otra fuera de rango, aunque esto siempre dependerá de la geografía de las manzanas y la relación con su entorno¹³.

¹³ Para mayores antecedentes respecto a los criterios de conformación de UPM urbanas, véase documento metodológico “Unidades primarias de muestreo. Nuevo marco muestral de viviendas”, (INE, 2019b, págs. 36-48).

El método numeración de las UPM es el de serpenteo, que corresponde a un ordenamiento en sentido Norte-Sur, partiendo desde el extremo superior izquierdo, siguiendo la forma de “Serpentín”. Para estos casos el nivel es comunal y por estratos (Figura 5).

Figura 5. Numeración por método de serpenteo



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

b) Criterios para UPM rurales

Para la conformación de UPM rurales, se consideraron todas las entidades rurales levantadas en el Censo 2017, correspondientes a 319 comunas a nivel nacional.

Figura 6. Estratificación de entidades rurales



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

El tamaño medio de las UPM se fija en 90 viviendas, considerando un rango entre 70 y 110, con ciertas excepciones. Con el objetivo de mantener unidades homogéneas.

Para el caso de las áreas rurales, aquellas zonas de denominación rural disperso y las entidades aldea, se trabajan en forma independiente. Por esta razón, el conteo y agrupación de viviendas para la formación de UPM se realiza a nivel de sector censal, teniendo en cuenta también la conectividad entre las UPM, con el objetivo de hacer más eficiente la operación de levantamiento.

Siempre que no contravenga los principios de homogeneidad y conectividad, es deseable que las UPM rurales tengan sus límites sustentados en elementos existentes de la cartografía, tales como ejes viales, cursos hídricos, límites DPA y límites con fines censales. Al igual que en el caso urbano, el método de numeración de las UPM es el serpenteo, de norte a sur, partiendo desde el extremo superior izquierdo.

En términos generales, los casos en que los tamaños de las UPM rurales se encuentran fuera del rango establecido de viviendas, obedecen a situaciones particulares de la conformación geográfica del país. Por ejemplo, se permiten UPM con tamaño por debajo del rango cuando se trata de áreas atrapadas entre el límite comunal y el área urbana. De la misma forma, se permiten comunas con una sola UPM rural que excede el rango, lo cual respeta más el principio de homogeneidad que dividir esos casos en dos UPM con tamaño por debajo del mínimo. Entre este tipo de casos especiales se puede mencionar las UPM con entidades insulares, las UPM en donde la conectividad se da por fuera de la comuna o incluso de la región, los casos en que el recorrido de la UPM atraviesa un área urbana, entre otros¹⁴.

4.3.1.4 Estratificación socioeconómica del marco muestral

Entre las novedades incorporadas en el MMV 2017 se cuenta con una actualización de la estratificación socioeconómica, la cual se realiza a nivel de UPM y se basa en un análisis de diversas variables recopiladas en el Censo 2017. Se obtiene como resultado una variable de estratificación con tres categorías, denominada nivel socioeconómico (NSE). En lo sucesivo se detalla el proceso de obtención de dicha variable¹⁵.

Para la determinación de los estratos socioeconómicos, se consideraron diversos métodos y niveles de agregación de la información. Los métodos considerados fueron diversos algoritmos de clasificación multivariados, tales como *Principal Component Analysis* (PCA), *K-medias*, y *Principal Component Analysis by means of Alternating Least Squares* (PRINCALS). Se consideró también la posibilidad de aplicar los algoritmos con la base a nivel de personas para luego ser agrupadas en UPM, así como también trabajar directamente a nivel de UPM desde un principio. Otros puntos de atención fueron la cantidad de estratos a construir y las variables a utilizar.

A partir de las distintas combinaciones de técnicas, agrupaciones, cantidades de estratos y variables, se obtuvieron 180 vectores de estratificación, que consideraban entre tres y cinco estratos. Para evaluar el desempeño de éstos, se calculó el efecto diseño de 24 indicadores, con lo cual fue posible escoger la estratificación que redujera la varianza de éstos. Los indicadores seleccionados fueron:

¹⁴ Para mayores antecedentes respecto a los criterios de conformación de UPM urbanas, véase documento metodológico “Unidades primarias de muestreo. Nuevo marco muestral de viviendas”, junio 2019. pp 99-119.

¹⁵ Mayores antecedentes disponibles en el documento “Estratificación socioeconómica del marco muestral de viviendas 2017 (MMV 2017)”, (INE, 2019c).

- Tasa de ocupación por sexo
- Tasa de desocupación por sexo
- Tasa de inactividad por sexo¹⁶
- Porcentaje de extranjeros
- Porcentaje de personas dentro de la fuerza de trabajo primaria respecto de la fuerza de trabajo por sexo
- Porcentaje de hogares unipersonales
- Porcentaje de población femenina ocupada por rama de actividad económica.
- Porcentaje de población masculina ocupada por rama de actividad económica.
- Porcentaje de personas según nivel educacional por sexo.

Posterior a la evaluación, el método utilizado para la estratificación del MMV 2017 fue el de estratificación óptima de la primera componente principal del PCA, que consideró los indicadores de porcentaje de personas en la educación superior, tasa de ocupación, porcentaje de viviendas con índice de materialidad alto y el indicador de total de hijos nacidos vivos. La elección del número de estratos (tres o cuatro), se basó 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 tres estratos.

Cabe mencionar que estos resultados se presentan para 35.090 UPM mientras que el MMV 2017 cuenta con 35.149 UPM, esto se debe principalmente a que, al cruzar el marco con la información de la base censal, se evidencian UPM cuyas viviendas no cuentan con la información. Se atribuye esta ocurrencia al transcurso del tiempo desde el levantamiento del Censo, por ejemplo, que el marco contenga viviendas que en el momento del levantamiento censal no eran particulares ocupadas.

En consecuencias quedaron 59 UPM sin clasificación socioeconómica, ya que no se disponía de información censal requerida para aplicar la metodología de estratificación socioeconómica. A estos casos se les asignó el nivel socioeconómico (NSE) de las UPM colindantes, y en caso de que colindara con UPM de distintos NSE, se le asignó el nivel preponderante entre las UPM colindantes. Este proceso se realizó mediante la observación cartográfica de las UPM.

4.3.1.5 Áreas especiales

En la construcción de un marco muestral, es necesario identificar áreas especiales susceptibles de exclusión. Estas áreas suelen estar afectadas por factores que dificultan su levantamiento, que pueden ser climáticos, de transporte, costo de acceso y de acceso restringido, entre otros. Para el caso del MMV 2017, se definieron Áreas Especiales (AE).

Para estas áreas especiales, el MMV 2017 aporta información mediante ocho tipologías que dan cuenta sobre las complejidades presentes en la realización de trabajos operativos, información que resulta de gran utilidad para la coordinación de las actividades en terreno, a saber:

¹⁶ Este indicador se calculó especialmente para este ejercicio a partir de la condición de actividad en la fuerza de trabajo, para definir a las personas que se encuentran fuera de ella (conocidos como inactivos) y eso sobre el total de hombres o mujeres de la UPM.

1. Tiempo de trayecto
2. Clima
3. Altitud
4. Transporte especial
5. Insularidad
6. Acceso Pedestre
7. Estado de los caminos
8. Ingreso restringido

Lo anterior permite calcular a nivel de UPM el porcentaje de viviendas que presentan dificultades de acceso, así como también considerar las razones específicas por las que cada una de estas ha sido identificada como Área Especial. En última instancia, lo anterior permite que la exclusión se lleve a cabo a nivel de UPM, en lugar de comuna, con lo cual el MMV 2017 consigue un mayor nivel de cobertura que su predecesor.

El trabajo a nivel de UPM permite también evaluar la exclusión en conjunto con los equipos regionales de modo que se posibilita la coordinación y valorización del esfuerzo operativo con información más detallada para el trabajo de campo.

4.3.1.6 Conformación definitiva del marco muestral maestro

Se obtuvo un total de 35.149 UPM separadas en dos estratos urbano y rural para un total de 6.145.493 viviendas marco, obteniendo para el estrato urbano un total de 26.195 UPM en 5.336.492 viviendas marco, por otra parte, para el estrato rural se obtuvieron un total de 8.954 UPM en 809.001 viviendas marco.

Si bien se establecieron límites para la cantidad de viviendas por UPM, sucedieron casos que obligaron a realizar excepciones. Para el área urbana, se obtuvieron 1.417 UPM que se encontraban fuera del rango entre 160 y 240 viviendas. Para el área rural, se obtuvieron 198 UPM por fuera del rango entre 70 y 110 viviendas.

En la estratificación socioeconómica, se clasificaron 35.090 UPM en estratos bajo (1), medio (2) y alto (3), obteniéndose 12.728 UPM en el estrato bajo, 15.518 UPM en el estrato medio y 6.844 UPM en el estrato alto a nivel nacional.

Tabla 5. UPM con tamaño fuera de rango por región

Región	Urbano			Rural		
	Total UPM	UPM < 160 viviendas	UPM > 240 viviendas	Total UPM	UPM < 70 viviendas	UPM > 110 viviendas
Total	26.195	363	1.054	8.954	65	133
Arica y Parinacota	331	2	5	74	-	-
Tarapacá	504	10	15	66	1	1
Antofagasta	895	7	19	70	3	1
Atacama	462	1	8	145	1	1
Coquimbo	1.092	20	31	646	2	13
Valparaíso	3.004	46	115	683	13	20
Metropolitana	11.135	188	580	919	9	19
O'Higgins	1.196	25	56	954	13	20
Maule	1.374	15	38	1.171	7	27
Ñuble	591	4	16	634	-	5
Biobío	2.372	22	99	754	8	10
Araucanía	1.209	9	27	1.189	-	5
Los Ríos	495	2	10	482	1	2
Los Lagos	1.100	10	33	996	2	5
Aysén	153	-	-	119	2	-
Magallanes	282	2	2	52	3	4

Fuente: Elaborado por el Instituto Nacional de Estadísticas

Tabla 6. Estratificación socioeconómica del MMV 2017 por región

Estratificación socioeconómica													
	No clasificado			Baja			Media			Alta			
	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Total	
Nacional	48	11	59	5.846	6.882	12.728	13.611	1.907	15.518	6.690	154	6.844	
Arica y Parinacota	-	-	-	108	70	178	178	4	182	45	-	45	
Tarapacá	2	1	3	133	60	193	248	4	252	121	1	122	
Antofagasta	1	-	1	203	51	254	481	17	498	210	2	212	
Atacama	-	-	-	226	124	350	161	19	180	75	2	77	
Coquimbo	-	-	-	356	590	946	568	54	622	168	2	170	
Valparaíso	4	3	7	671	404	1.075	1.780	268	2.048	549	8	557	
RM	35	4	39	1.159	338	1.497	5.726	524	6.250	4.215	53	4.268	
O'Higgins	-	-	-	356	746	1.102	696	205	901	144	3	147	
Maule	1	-	1	459	1.023	1.482	735	145	880	179	3	182	
Ñuble	-	-	-	225	579	804	295	50	345	71	5	76	
Biobío	3	2	5	887	681	1.568	1.053	66	1.119	429	5	434	
La Araucanía	-	1	1	476	1.066	1.542	558	118	676	175	4	179	
Los Ríos	-	-	-	191	416	607	236	60	296	68	6	74	
Los Lagos	1	-	1	368	683	1.051	563	286	849	168	27	195	
Aysén	-	-	-	17	44	61	121	58	179	15	17	32	
Magallanes	1	-	1	11	7	18	212	29	241	58	16	74	

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Con base en la información sobre características especiales señaladas para cada UPM del MMV 2017, se determina el número de UPM que presenta al menos una de dichas características y se señalan como UPM especiales. Finalmente, se determinan 973 UPM especiales lo que representa un 2,77% del total de UPM del MMV 2017 y contemplan 87.831 viviendas, alrededor de 1,43% del total de viviendas del MMV 2017.

Tabla 7. UPM especiales por región

Región	Total UPM en el marco	Total UPM especiales	% UPM Especiales	Total Viviendas marco	Total Viviendas marco especiales	% Viv. Marco especiales
Nacional	35.149	973	2,77%	6.145.493	87.831	1,43%
Arica y Parinacota	405	9	2,22%	72.602	828	1,14%
Tarapacá	570	22	3,86%	107.473	2.046	1,90%
Antofagasta	965	10	1,04%	186.917	878	0,47%
Atacama	607	11	1,81%	105.554	924	0,88%
Coquimbo	1.738	45	2,59%	278.060	4.030	1,45%
Valparaíso	3.687	38	1,03%	673.541	3.413	0,51%
Metropolitana	12.054	17	0,14%	2.352.958	1.543	0,07%
O'Higgins	2.150	71	3,30%	331.870	6.331	1,91%
Maule	2.545	171	6,72%	386.081	15.778	4,09%
Ñuble	1.225	17	1,39%	179.827	1.526	0,85%
Biobío	3.126	75	2,40%	553.538	6.709	1,21%
La Araucanía	2.398	162	6,76%	355.092	14.679	4,13%
Los Ríos	977	77	7,88%	143.005	6.937	4,85%
Los Lagos	2.096	187	8,92%	316.294	16.847	5,33%
Aysén	272	44	16,18%	41.682	3.996	9,59%
Magallanes	334	17	5,09%	60.999	1.366	2,24%

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

4.3.2 Marco muestral de la Encuesta Nacional de Empleo

El marco muestral de la Encuesta Nacional del Empleo se conforma desde el marco muestral maestro según los objetivos de la encuesta, los niveles de estimación determinados y las áreas especiales que se haya determinado excluir.

4.3.2.1 Estratificación geográfica

La estratificación geográfica comienza desde la división político-administrativa y la identificación de áreas en sectores urbanos y rurales presentes en el marco muestral maestro. En este caso, el objetivo de la estratificación es crear una división geográfica que cumpla con los requisitos mínimos de UPM para disponer de suficientes unidades al realizar las rotaciones de la encuesta, y que refleje en alguna medida la aglomeración espacial del mercado laboral. Se establece como requisito para los estratos un mínimo de 36 UPM en cada uno de ellos, de modo que existan suficientes para cubrir los procesos de actualización muestral durante los períodos intercensales. Adicionalmente, la estratificación se realiza por separado para el área urbana y para el área rural, de modo que el segmento rural de una comuna podrá compartir estrato con otras, mientras el segmento urbano constituye un estrato por sí mismo.

Con el objetivo de que la estratificación tenga en cuenta la aglomeración de la población, el proceso de construcción de estratos comenzó a partir de grupos de comunas que comparten conurbaciones y que cumplen con el criterio de contener al menos 36 UPM. En las regiones en que estas conurbaciones no

coincidieron con las capitales regionales, éstas fueron consideradas como estrato también, siempre y cuando cumplieran con la cantidad mínima de UPM.

Luego, se procedió a agrupar el resto de las comunas que pertenecen a una misma provincia en estratos. En estos casos, podía suceder que una provincia completa constituya un estrato, o bien, que contenga a varios si es que se estima que su población es excesiva. Posteriormente, para aquellos casos identificados como de población excesiva, se aplicó un tratamiento de *Principal Component Analysis* (PCA) que permitió crear estratos en casos como las conurbaciones de Gran Concepción y Gran Santiago, o de la provincia de Linares, entre otras.

Finalmente, se obtuvieron 163 estratos geográficos distribuidos en todo el territorio nacional, salvo las zonas de exclusión que se detallan en el apartado siguiente. Cabe mencionar que, entre los 163 estratos geográficos, existen once que no cuentan con el mínimo de UPM requerido, ya que no tenían otras comunas con las cuales agruparse, por lo que se genera una excepción a la regla. La distribución regional de los estratos se muestra en la Tabla 8, mientras que los casos excepcionales de estratos con menos de 36 UPM se detallan en la Tabla 9. El listado completo de los estratos geográficos queda disponible en Anexo N° 1.

Tabla 8. Cantidad de estratos geográficos por región y área

Región	Área		
	Urbano	Rural	Total
Nacional	88	75	163
Arica y Parinacota	1	1	2
Tarapacá	2	2	4
Antofagasta	3	2	5
Atacama	3	3	6
Coquimbo	5	5	10
Valparaíso	12	9	21
Metropolitana	15	9	24
O'Higgins	6	6	12
Maule	9	10	19
Ñuble	4	4	8
Biobío	7	5	12
Araucanía	6	4	10
Los Ríos	5	5	10
Los Lagos	6	6	12
Aysén	2	2	4
Magallanes	2	2	4

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Tabla 9. Estratos que no alcanzan mínimo de UPM

Región	Área	Estrato geográfico	UPM	Razón
Tarapacá	Urbano	Prov. Del Tamarugal	27	No existen otras áreas urbanas en la región, fuera de la capital regional.
Tarapacá	Rural	Conurbación Iquique - Alto Hospicio	13	Se busca preservar la capital regional para asegurar presencia de muestra en el estrato.
Antofagasta	Rural	Prov. Antofagasta (Antofagasta, Mejillones, Sierra Gorda, Tal Tal)	20	No existen otras áreas rurales en la provincia.
Atacama	Rural	Conurbación Copiapó (Copiapó, Tierra Amarilla)	33	Se busca preservar la capital regional para asegurar presencia de muestra en el estrato.
Atacama	Rural	R.R Atacama (Prov. Chañaral y Comuna Caldera)	30	Se decide agregar Caldera a un estrato fuera de su provincia con el objetivo de preservar la conurbación de la capital regional y obtener dos estratos de tamaño aproximadamente similar.
Valparaíso	Urbano	R.P. Valparaíso (Casablanca)	30	No existen otras áreas urbanas colindantes, fuera de la capital regional.
Valparaíso	Rural	Conurbación Valparaíso (Valparaíso, Concón, Quilpué, Villa Alemana)	33	Se busca preservar la capital regional para asegurar presencia de muestra en el estrato.
Valparaíso	Rural	R.P. Valparaíso (Puchuncaví, Quintero)	34	No existen otras áreas rurales en la provincia.
Metropolitana	Urbano	R.P. Cordillera	35	Todas las áreas urbanas colindantes componen conurbaciones.
Magallanes	Rural	Punta Arenas + Laguna Blanca + Río Verde + San Gregorio	28	Se busca preservar la provincia de la capital regional para asegurar presencia en el estrato.
Magallanes	Rural	R.R. Magallanes	23	Resto de las comunas de la región de Magallanes.

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

4.3.2.2 Conformación del marco muestral

Con la inclusión de los estratos geográficos, el marco muestral MMV 2017 ya cuenta con toda la información necesaria para ser utilizado en la ENE, según el diseño muestral definido. Sin embargo, es necesario definir las UPM que se van a excluir del estudio. Para ello, en primer lugar, se identifican a las UPM susceptibles de exclusión, es decir, las áreas especiales del MMV 2017, como también, aquellas que pertenecen a zonas que se encontraban excluidas en el estudio previo. Para decidir cuáles de las UPM identificadas serán finalmente excluidas, el equipo técnico del INE trabajó, en coordinación con los equipos regionales de recolección, de modo que la decisión incluya también las capacidades instaladas de cobertura¹⁷.

El resultado fue la exclusión de 355 UPM, que corresponde a 1% de las UPM a nivel nacional, donde diez de las 345 comunas que cubre el marco muestral maestro se ven excluidas en su totalidad, observándose una concentración relevante en la región de Los Lagos, donde la provincia de Palena se ve totalmente excluida. Asimismo, las viviendas del marco excluidas del estudio alcanzan 0,57% a nivel nacional. La Tabla 10 muestra la distribución regional de UPM y viviendas marco, especiales y excluidas.

¹⁷ A cada equipo regional le fue remitido el correspondiente listado de UPM susceptibles de exclusión, para que en función de sus conocimientos del terreno determinaran si era necesario excluir la UPM o no del estudio.

Tabla 10. UPM especiales y excluidas por región, en el marco muestral de la ENE

Región	Total UPM en el marco	Total UPM especiales	% UPM Especiales	Total UPM excluidas	% UPM excluidas	Total Viviendas marco	Total Viviendas marco especiales	% Viv. Marco especiales	Total Viviendas marco excluidas	% Viv. Marco excluidas
Nacional	35.149	973	2,77%	355	1,01%	6.145.493	87.831	1,43%	35.287	0,57%
Arica y Parinacota	405	9	2,22%	8	1,98%	72.602	828	1,14%	736	1,01%
Tarapacá	570	22	3,86%	13	2,28%	107.473	2.046	1,90%	1.195	1,11%
Antofagasta	965	10	1,04%	4	0,41%	186.917	878	0,47%	359	0,19%
Atacama	607	11	1,81%	6	0,99%	105.554	924	0,88%	511	0,48%
Coquimbo	1.738	45	2,59%	0	0,00%	278.060	4.030	1,45%	0	0,00%
Valparaíso	3.687	38	1,03%	28	0,76%	673.541	3.413	0,51%	3.921	0,58%
Metropolitana	12.054	17	0,14%	0	0,00%	2.352.958	1.543	0,07%	0	0,00%
O'Higgins	2.150	71	3,30%	6	0,28%	331.870	6.331	1,91%	524	0,16%
Maule	2.545	171	6,72%	5	0,20%	386.081	15.778	4,09%	457	0,12%
Ñuble	1.225	17	1,39%	8	0,65%	179.827	1.526	0,85%	726	0,40%
Biobío	3.126	75	2,40%	24	0,77%	553.538	6.709	1,21%	2.064	0,37%
La Araucanía	2.398	162	6,76%	49	2,04%	355.092	14.679	4,13%	4.494	1,27%
Los Ríos	977	77	7,88%	23	2,35%	143.005	6.937	4,85%	2.023	1,41%
Los Lagos	2.096	187	8,92%	167	7,97%	316.294	16.847	5,33%	16.895	5,34%
Aysén	272	44	16,18%	11	4,04%	41.682	3.996	9,59%	1.177	2,82%
Magallanes	334	17	5,09%	3	0,90%	60.999	1.366	2,24%	205	0,34%

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

4.3.2.3 Estratificación del marco muestral ENE

El diseño muestral de la ENE considera información tanto de los estratos geográficos descritos como de la estratificación socioeconómica presentes en el MMV 2017. Para ello, se conforman estratos de muestreo como el cruce entre la estratificación geográfica y socioeconómica, siempre que los estratos geográficos cumplan ciertas condiciones. Se establece que, para desagregar un estrato geográfico según su nivel socioeconómico (NSE) debe:

- Contener más de 150 UPM.
- Poseer suficiente heterogeneidad de condición socioeconómica, esto es,
 - El estrato geográfico no debe agrupar más de 80% de las UPM en un solo NSE o;
 - debe presentar una proporción de UPM en torno a 40% en alguna categoría de nivel socioeconómico (NSE) y en torno a 60% en otro.
- Asegurar que todos los estratos resultantes cumplan con el mínimo de 36 UPM.

Bajo estos criterios se obtiene un total de 232 estratos de muestreo, de los cuales 122 provienen directamente desde la estratificación geográfica, y 110 incorporan la información de la estratificación socioeconómica. La distribución regional de los estratos se muestra en la Tabla 11.

Tabla 11. Cantidad de estratos de muestreo por región y área

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

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

El listado completo de los estratos de muestreo queda disponible en Anexo N° 2.

4.4 Cobertura geográfica

La cobertura geográfica de la ENE considera las 16 regiones del país, buscando presencia en todo el territorio nacional y abarcando 97% de las comunas del país. No obstante, se excluyen algunas áreas geográficas que por presentar ciertas características especiales (Insularidad, clima, altitud, tiempo de traslados, estado de los caminos, etc.), son excluidas del marco de muestreo de la encuesta.

4.5 Unidades estadísticas

Al ser la ENE una encuesta por muestreo, las unidades estadísticas las conforman unidades de muestreo, unidades de información y unidades de análisis. Las unidades de muestreo son aquellas que son susceptibles de ser seleccionadas en las etapas de muestreo definidas en el diseño, pudiendo estar conformadas por una unidad elemental o por un conjunto de ellas (conglomerado). Por otra parte, las unidades de información, son aquellas desde donde se obtiene la información de interés, pudiendo ésta coincidir con la unidad la unidad de análisis que, es aquella unidad sobre la cual se puede realizar la medición del fenómeno objeto de estudio.

4.5.1 Unidad de muestreo

En la ENE existen dos unidades de muestreo.

- a) **Unidad primaria de muestreo (UPM):** definida como un conglomerado homogéneo, en términos de la cantidad de viviendas particulares¹⁸ que los componen.
- b) **Unidad secundaria de muestreo (USM):** que corresponde a viviendas particulares ocupadas dentro de las UPM seleccionadas.

Como se mencionó anteriormente en el apartado [“4.3.1 Marco muestral maestro”](#) el tamaño promedio de las UPM en el área urbana es de 200 viviendas particulares, con un rango de 160 a 240 viviendas particulares, excluyendo viviendas de temporada. Mientras que, el tamaño promedio de las UPM rurales es de 90 viviendas particulares, con un rango de 70 a 110 viviendas particulares, excluyendo las de temporada.

4.5.2 Unidad de información y análisis

Para el cumplimiento del objetivo de la encuesta, se utiliza como unidades de información y análisis a las personas que cumplen con el requisito de ser parte de la población objetivo, esto es, que tengan quince años o más y tengan residencia habitual en hogares ubicados en viviendas particulares ocupadas dentro del territorio nacional. Sin embargo, se registran las características sociodemográficas de todas las personas residentes en la vivienda seleccionada mediante el cuestionario construido para estos fines.

¹⁸ Excluyendo las viviendas de temporada.

4.6 Definición de variables

En esta sección se describen las principales variables que son capturadas y construidas para la publicación de datos, por lo que se describen las principales variables analíticas de la ENE.

Una variable analítica es aquella que se genera en función de otras variables y que se incluye en el análisis estadístico. Las principales variables analíticas son:

- a) Variable **cae general**, es la Condición de Actividad Económica General. Corresponde a la clasificación de todas las personas de quince años y más, según su situación laboral, basado en un código sumario empleo a diez categorías. Las categorías observadas se componen de la información obtenida de las secciones y preguntas específicas que estructuran el cuestionario.
- b) Variable **cae_especifico**, Condición de Actividad Económica Específica. Corresponde a la clasificación de todas las personas de quince años y más, según su situación laboral, basado en un código sumario empleo a 29 categorías.
- c) Variable **activ**, resumen la variable anterior en tres categorías jerárquicas y excluyentes.
- d) La variable **categoria_ocupacion**, corresponde a las personas ocupadas según adaptación chilena de la Clasificación Internacional de la Situación en el Empleo del año 1993 (CISE-93).
- e) Variable **habituales**, corresponde a las horas usualmente trabajadas en la semana de referencia.
- f) Variable **efectivas**, corresponde a las horas efectivamente trabajadas en la semana de referencia. Este dato proporciona una “fotografía” de las horas trabajadas en el periodo específico.
- g) La variable **CINE**, es la Clasificación Internacional Normalizada de la Educación del año 1997 (CINE 97).
- h) Variable **sector**, corresponde a la población ocupada por sector de acuerdo a la informalidad laboral.
- i) Variable **ocup_form**, corresponde a la población ocupada según informalidad laboral.
- j) Variable **fact_cal**, es Factor de expansión trimestral calibrado, que permite expandir datos muestrales a poblacionales.

Estas variables se presentan en detalle en el capítulo [“6. Metodología de procesamiento de la información”](#).

4.7 Nomenclatura utilizada

Las nomenclaturas utilizadas en la Encuesta Nacional de Empleo provienen de los principales clasificadores utilizados para caracterizar a la población en edad de trabajar y, en la mayoría de los casos, a las personas ocupadas.

a) Clasificación Internacional Normalizada de la Educación del año 1997 (CINE 97)

Forma parte de la familia internacional de Clasificaciones Económicas y Sociales de las Naciones Unidas, las cuales son empleadas a nivel mundial en la elaboración de estadísticas con el objetivo de acopiar y analizar datos comparables a nivel internacional de manera consistente. Dentro de estas, la CINE representa una clasificación de referencia que permite ordenar los programas educativos y sus respectivas certificaciones por niveles de educación y campos de estudio (UNESCO).

Actualmente, a nivel nacional la categoría contempla el siguiente desglose que se aplica a todas las personas en edad en trabajar que conforman la población objetivo de la encuesta:

1. Nunca estudió
2. Educación preescolar
3. Educación primaria (nivel 1)
4. Educación primaria (nivel 2)
5. Educación secundaria
6. Educación Técnica (Educación Superior no Universitaria)
7. Educación universitaria
8. Post títulos y maestría
9. Doctorado
10. Nivel ignorado

En el caso de la categoría de personas ocupadas, existen varios clasificadores que contribuyen a caracterizar a las personas ocupadas, de acuerdo a sus actividades realizadas, a la empresa o lugar donde trabajan, entre otras dimensiones.

b) Clasificación Internacional Uniforme de Ocupaciones del año 2008, adaptación nacional (CIUO 08.CL)

El Clasificador Chileno de Ocupaciones, CIUO 08.CL, corresponde a la adaptación nacional de la Clasificación Internacional Uniforme de Ocupaciones, CIUO 08. El Clasificador Chileno de Ocupaciones, CIUO 08.CL es el estándar oficial para el procesamiento y análisis de las ocupaciones, que permite organizarlas de acuerdo al tipo de tareas realizadas en un puesto de trabajo y las competencias requeridas para ello, dependiendo del nivel y la especialización de estas competencias.

1. Directores, gerentes y administradores
2. Profesionales, científicos e intelectuales
3. Técnicos y profesionales de nivel medio
4. Personal de apoyo administrativo
5. Trabajadores de los servicios y vendedores de comercios y mercados

6. Agricultores y trabajadores calificados agropecuarios, forestales y pesqueros
7. Artesanos y operarios de oficios
8. Operadores de instalaciones, máquinas y ensambladores
9. Ocupaciones elementales
10. No identificado

c) Clasificación Internacional Industrial Uniforme revisión 4.CL según la Clasificación de Actividades Económicas Nacional de Encuestas Sociodemográficas (CAENES)

El CAENES es una clasificación basada en el Clasificador Chileno de Actividades Económicas CIIU4. CL 2012. Por este motivo, se rige por los mismos conceptos y criterios del Clasificador Internacional CIIU Rev.4. En términos operativo corresponde al clasificador de la rama de actividad económica de la empresa o institución donde la persona ocupada trabajar o de la que es dueña. Posee 21 ramas de actividad.

- A. Agricultura, ganadería, silvicultura y pesca
- B. Explotación de minas y canteras
- C. Industrias Manufactureras
- D. Suministro de electricidad, gas, vapor y aire acondicionado
- E. Suministro de agua; evacuación de aguas residuales, gestión de desechos y descontaminación
- F. Construcción
- G. Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas
- H. Transporte y almacenamiento
- I. Actividades de alojamiento y servicio de comidas
- J. Información y comunicaciones
- K. Actividades financieras y de seguros
- L. Actividades inmobiliarias
- M. Actividades profesionales, científicas y técnicas
- N. Actividades de servicios administrativos y de apoyo
- O. Administración pública y defensa; planes de seguridad social de afiliación obligatoria
- P. Enseñanza
- Q. Actividades de atención de la salud humana y de asistencia social
- R. Actividades artísticas, de entretenimiento y recreativas
- S. Otras actividades de servicios
- T. Actividades de los hogares como empleadores; actividades no diferenciadas de los hogares como productores de bienes y servicios para uso propio
- U. Actividades de organizaciones y órganos extraterritoriales.

d) Clasificación Internacional de la Situación en el Empleo del año 1993 (CISE 93)

La CISE-93 clasifica los empleos según el tipo de riesgo económico, solidez del vínculo entre la persona y el empleo, y el tipo de autoridad, que tienen o tendrán los titulares sobre los establecimientos y sobre otros trabajadores. La actual CISE 93 utilizada en la ENE corresponde a la adaptación chilena y posee las siguientes categorías:

1. Empleador
2. Cuenta propia
3. Asalariado sector privado
4. Asalariado sector público
5. Personal de servicio doméstico puertas afuera
6. Personal de servicio doméstico puertas adentro
7. Familiar o personal no remunerado

4.8 Diseño según tipo de fuente: Muestra

El diseño muestral de la ENE, tiene como objetivo, generar estimaciones de los principales indicadores¹⁹ de la población en edad de trabajar, respecto a su vínculo con la fuerza de trabajo, basado en una mejor cobertura y distribución de las unidades muestrales, a fin de obtener estimaciones con mayores niveles de precisión para los dominios de estudio definidos, implementando además, mecanismos (enumeración y rotación de unidades muestrales) que permitan actualizar el marco muestral con regularidad, y de ser necesario la muestra, reduciendo así los errores no muestrales asociados a la desactualización del marco muestral.

4.8.1 Diseño muestral

4.8.1.1 Nivel de estimación

En correspondencia con uno de los objetivos planteados en el nuevo diseño muestral de la ENE, de generar estimaciones con mejores niveles de precisión estadística, se calcularon los tamaños muestrales para los dominios de estudio, en función a la nueva estratificación del marco muestral.

Los dominios de estudio o niveles de estimación, para los que el tamaño muestral de la ENE garantiza estimaciones representativas del parámetro de interés son: nacional, nacional urbano, nacional rural, regional, el área urbana de las 16 regiones y el área rural de las regiones de Coquimbo, Valparaíso, O'Higgins, El Maule, Biobío, La Araucanía, Los Lagos, Metropolitana, Los Ríos y Ñuble.

4.8.1.2 Estrategia muestral

El diseño muestral de la ENE 2020 se define como probabilístico, estratificado y bietápico.

Esta estrategia muestral se determina, en virtud de los objetivos definidos en el diseño muestral; probabilístico, ya que dada la amplia cobertura es necesario generar inferencias a partir de una muestra, estratificado porque permite controlar mejor la distribución espacial de la muestra y

¹⁹ Tasa de desocupación, tasa de ocupación y tasa de participación.

bietápico que permite reducir costos dado que, en una misma área se puede obtener información de distintas unidades muestrales.

Respecto a la estratificación, los estratos de muestreo están conformados por áreas geográficas según la división político administrativa, y, en donde sea posible, a partir de la combinación de éstas con el nivel socioeconómico, según lo descrito en el apartado [“4.3.2.3 Estratificación del marco muestral ENE”](#) de este documento.

4.8.2 Selección de unidades muestrales

4.8.2.1 Consideraciones para el tamaño muestral

El cálculo del tamaño muestral utiliza información del Censo de Población y Vivienda 2017 y del marco muestral del producto, a partir de los cuales se elaboran simulaciones que permiten:

- Estimar efectos de diseño (*deff*): que surgen de considerar un muestreo complejo en lugar de un muestreo aleatorio simple
- Estudiar el comportamiento de la muestra
- Evaluar la pertinencia de distintos niveles de estimación, y
- Tomar diversas decisiones relativas al tamaño muestral.

El principal parámetro de interés, para la obtención del tamaño muestral, es la tasa de desocupación, calculada como el número de personas desocupadas sobre el número de personas que participan en la fuerza de trabajo, y se estima mediante un estimador de razón.

Tabla 12. Características del parámetro de interés

Parámetro	Descripción
Parámetro a estimar	Tasa de desocupación
Variable de diseño	$X_k = \begin{cases} 0, & \text{si la persona } k \text{ pertenece a la fuerza de trabajo y no está desocupada} \\ 1, & \text{si la persona } k \text{ pertenece a la fuerza de trabajo y está desocupada} \end{cases}$
Estimador asociado	Estimador de razón $p = \frac{\text{Número de personas desocupadas}}{\text{Número de personas en la fuerza de trabajo}}$

Las simulaciones de tamaño muestral consisten en una serie de escenarios, en los cuales se seleccionan UPM desde el marco muestral del producto, para luego seleccionar viviendas pertenecientes a dichas UPM utilizando como marco muestral el Censo 2017. A partir de estas simulaciones se determina encuestar diez viviendas por cada UPM, salvo en las áreas urbanas de las regiones de Atacama, Aysén y Magallanes, donde se fijan ocho viviendas por UPM²⁰. Asimismo, es estimado el efecto de diseño a nivel de región-área que se utiliza en el cálculo del tamaño definitivo.

²⁰ En el apartado 4.8.1.7 se desarrolla la metodología implementada para determinar el número óptimo de viviendas por UPM.

4.8.2.2 Cálculo del tamaño muestral

Como se ha mencionado, el diseño muestral de la ENE se basa en un muestreo complejo, estratificado con dos etapas de selección, por lo que el cálculo del tamaño muestral es obtenido a partir de las siguientes etapas:

1. **Obtención de parámetros censales:** A partir del Censo de Población y Vivienda 2017²¹, se calculó la varianza de la variable de interés a nivel del cruce región-área (S_{ra}^2).

$$S_{ra}^2 = P_{ra} \cdot (1 - P_{ra}) \quad (1)$$

Donde, P_{ra} corresponde a la proporción de personas que se encuentran en situación de desempleo²² al nivel región-área.

2. **Cálculo de tamaño muestral inicial.** Se obtiene un primer tamaño muestral de personas a través de la fórmula del muestreo aleatorio simple, considerando un error absoluto esperado (e_{ra}). Junto con ello, se establece un nivel de confianza de 95% asociado al percentil $t_{1-\alpha/2, v-1}$ de la distribución *t-student* con v grados de libertad. No obstante, debido a que al momento de realizar inferencias a partir de la encuesta los grados de libertad varían para cada dominio de estudio, se opta utilizar un valor fijo de dos (2) como medida de simplicidad y robustez.

$$\tilde{m}_{0,ra} = \frac{2^2 \cdot S_{ra}^2}{e_{ra}^2} \quad (2)$$

Donde,

$\tilde{m}_{0,ra}$: Tamaño de muestra de personas para el dominio región-área, bajo muestreo aleatorio simple.

3. **Incorporación de efecto diseño.** Debido a que se utiliza un diseño complejo multietápico, es necesario considerar la incorporación del efecto diseño que da cuenta de la pérdida de eficiencia en el muestreo complejo respecto al muestreo aleatorio simple. Por esta razón, el tamaño de muestra inicial se incrementa a partir del efecto de diseño según ecuación:

²¹ Es importante considerar que existen diferencias en la forma de determinar la desocupación en el Censo de Población y Vivienda 2017 y en la ENE, en lo referente al periodo de referencia y a la cantidad de preguntas que se realizan para determinar la condición de desocupación. Sin embargo, dicha estimación se puede tomar como una aproximación.

²² $P_{ra} = \frac{\text{Total de personas desocupadas}}{\text{Total de personas en la fuerza de trabajo}}$

$$\tilde{m}_{1,ra} = \tilde{m}_{0,ra} \cdot deff_{ra} \quad (3)$$

Donde,

$\tilde{m}_{1,ra}$: Tamaño de muestra de personas para el dominio región-área corregido por efecto de diseño.

$deff_{ra}$: Efecto de diseño a nivel de viviendas para el dominio región-área. Estimado a partir de simulaciones²³ muestrales sobre el marco de selección y la base de personas del Censo 2017.

4. **Tamaño a nivel de viviendas:** Obtenido el tamaño muestral a partir de la ecuación (4) con información a nivel de personas en la fuerza de trabajo²⁴, se requiere obtener el tamaño en términos de número de viviendas y de UPM. Para ello, se divide el tamaño muestral a nivel de personas por el número de personas promedio en las viviendas que participan de la fuerza de trabajo (\overline{FT}_{ra}) obtenida con base en el Censo 2017.

$$m_{1,ra} = \frac{\tilde{m}_{1,ra}}{\overline{FT}_{ra}} \quad (4)$$

Donde,

$m_{1,ra}$: Tamaño de muestra de viviendas para el dominio región-área.

5. **Ajuste por población finita:** Con el tamaño a nivel de viviendas, es posible utilizar la información del marco muestral para introducir el ajuste de muestreo para poblaciones finitas. En esta instancia, se utiliza la cantidad de viviendas en el marco de la ENE (M_{ra}).

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

Donde,

$m_{2,ra}$: Tamaño de muestra de viviendas para el dominio región-área, ajustado por poblaciones finitas.

²³ Para más información consultar el Informe de Simulaciones de tamaños muestrales de la ENE 2020.

²⁴ Población Económicamente Activa (PEA) o Fuerza de Trabajo: Personas en edad de trabajar que, durante la semana de referencia, cumplen los requisitos para ser incluidas en la categoría de personas ocupadas o desocupadas.

6. **Ajustes operacionales ($m'_{2,ra}$):** Los tamaños obtenidos se redondean hacia arriba al múltiplo de 10 superior más cercano, con el fin de asegurar posteriormente, que en todas las UPM se levanten, en teoría, diez viviendas. Luego, se contrastan los tamaños obtenidos, con los tamaños de la versión anterior de la encuesta y se confirma su factibilidad operativa. Posteriormente, los tamaños se distribuyen en los estratos de muestreo, según los siguientes criterios:

- a. Todos los estratos deben contar con muestra de al menos seis²⁵ UPM.
- b. Ningún estrato puede presentar una fracción de muestreo superior a 50%.

La aplicación de estos criterios puede modificar los tamaños muestrales obtenidos hasta el punto anterior, sin embargo, se busca asegurar que no haya un impacto significativo sobre los niveles de error esperados.

4.8.2.3 Distribución del tamaño muestral

La afijación del tamaño muestral se realiza con el objetivo de identificar a nivel de estrato la cantidad de unidades que corresponde seleccionar en cada uno de ellos. El tamaño resultante de este proceso es el que se informa en la última columna de la Tabla 17. El método utilizado es la afijación óptima de Neyman, que asigna mayor muestra a los estratos de mayor tamaño, así como a los que presentan mayor desviación estándar del parámetro de interés, siguiendo la siguiente fórmula:

$$m_h = m'_{2,ra} \cdot \frac{s_h \cdot M_h}{\sum_h s_h \cdot M_h} \quad (6)$$

Donde,

m_h : Es la cantidad de muestra afijada al estrato h .

$m'_{2,ra}$: Es el tamaño muestral determinado para la región r y el área a . (Notar que este valor no necesariamente coincide con los informados en la columna $m_{2,ra}$ de la Tabla 17, puesto que ya cuenta con algunos ajustes provenientes de motivos operacionales.

s_h : Desviación estándar del parámetro de interés calculado con el Censo 2017, en el estrato h .

M_h : Cantidad de viviendas particulares ocupadas, en el estrato h , informadas en el marco del producto.

El tamaño de viviendas obtenido a nivel de estrato se divide por ocho o diez (ver apartado [“4.8.1.7 Tamaño muestral de viviendas por UPM”](#)) según sea el caso y se redondea hacia arriba para obtenerlo expresado en términos de UPM. Este resultado es el que se modifica para asegurar el cumplimiento de los criterios mencionados en el apartado anterior, en relación con que todos los

²⁵ Con la finalidad de asegurar un mínimo de 2 UPM mensuales y el aporte del estrato al cálculo de la varianza.

estratos tengan una muestra de al menos seis UPM y que ninguno tenga una fracción de muestreo superior al 50%.

La incorporación de estos criterios implica una distorsión a la afijación de muestra, e incluso pérdida de la muestra necesaria en ciertos niveles de estimación, ya que el criterio de la fracción de muestreo se implementa mediante un truncamiento del tamaño muestral en el estrato. Esto puede suceder, por ejemplo, en estratos que presentan alta variabilidad del fenómeno y que por lo tanto reciben un tamaño de muestra elevado en relación con otros estratos del mismo nivel de estimación, pero que terminan excediendo el 50% en la fracción de muestreo²⁶, con lo cual el nivel de estimación completo puede perder una cantidad importante de muestra. Para evitar lo anterior, se redistribuye la cantidad de muestra perdida entre otros estratos, así como también se resta en niveles de estimación que puedan haber quedado con más muestra de la necesaria debido al criterio del mínimo de UPM.

Finalmente, se obtiene el tamaño muestral por estrato a nivel de UPM, como resultado del proceso de afijación.

4.8.2.4 Cálculo de la sobremuestra

Entre los cambios planteados en el diseño muestral de la ENE 2020, se encuentra el cálculo del tamaño muestral sobredimensionado como alternativa al mecanismo de reemplazos que se utilizaba anteriormente. Esto, con el propósito de contrarrestar la pérdida de muestra en las sucesivas rondas de la encuesta y la no respuesta por parte de los encuestados.

Los estudios de carácter panel, como es el caso de la ENE, se enfrentan a considerables problemas de no respuesta, lo que se refleja en reducciones en el tamaño muestral alcanzado período a período debido a factores como la movilidad, el fallecimiento o la pérdida de cooperación de los informantes, entre otros factores.

En este sentido, para estimar el tamaño de muestra sobredimensionado, se realizaron cálculos de las tasas de no logro²⁷ de la ENE. Para dar cuenta de la pérdida muestral en cada turno de rotación. La condición de logro o no logro es determinada mediante los códigos de registro de visita²⁸ señalados en la Tarjeta de Registro del Hogar (TRH)²⁹ vigente en la versión anterior de la encuesta. Esta clasificación permite categorizar la situación de las viviendas en trece códigos diferentes y cuatro grupos.

²⁶ Estos casos se presentaron en dos estratos urbanos de la región de Aysén.

²⁷ La tasa de no logro es una aproximación a la tasa de no respuesta, la diferencia se encuentra en el denominador, ya que la tasa de no logro considera el total de unidades gestionadas, mientras que la tasa de respuesta excluye en el denominador aquellas unidades no elegibles (viviendas de temporada de otro uso, etc.).

²⁸ Para conocer más detalle sobre los códigos de registro de visitas empleados, ver Anexo N° 3. Códigos de visita de la EN.

²⁹ A partir de enero 2020, se implementaron mejoras en el instrumento de recolección que anteriormente se encontraba incluido en la Tarjeta de Registro del Hogar (TRH), lo cual forma parte de la estrategia de fortalecimiento y actualización del diseño muestral de la ENE. Las mejoras se enfocaron en la clasificación de las viviendas en función de códigos de disposición final y obedeciendo a estándares internacionales, así como, la captura de información complementaria durante el proceso de recolección.

La tasa de logro sobre viviendas gestionadas³⁰ (TL_{gest}) se calcula mediante la siguiente expresión:

$$TL_{gest} = \frac{\text{Viviendas logradas}}{\text{Viviendas logradas} + (\text{Viviendas tipo A}) + (\text{Viviendas tipo B}) + (\text{Viviendas tipo C})} \quad (7)$$

Y la tasa de no logro (TNL_{gest}) mediante la fórmula:

$$TNL_{gest} = 1 - TL_{gest} \quad (8)$$

Esta forma de cálculo se emplea con la finalidad de obtener la sobredimensión que se requiere realizar al tamaño muestral que permite compensar la no respuesta, independiente de la causa que la origine.

Para el cálculo de las tasas de no logro, se utilizaron los listados de las viviendas gestionadas de la muestra, entre enero 2018 y diciembre 2018. Contemplando que, la muestra ENE está distribuida en tres meses, se reconstruyeron los diferentes trimestres móviles dentro del año 2018.

Con el objetivo de calcular los indicadores de logro al mayor nivel de desagregación posible³¹, acorde al nuevo diseño muestral, fue necesario ubicar las viviendas de las muestras correspondientes al año 2018, dentro de los nuevos estratos geográficos, debido a que la muestra ENE 2020 presenta una estratificación diferente a la del diseño muestral anterior.

Se procedió a calcular las tasas de no logro para los 163 estratos geográficos del nuevo diseño, por ronda³² y trimestre móvil. Luego se promediaron los valores trimestrales, para obtener un solo indicador por estrato geográfico y ronda. Finalmente, se calcula un único valor para la tasa de no logro a nivel de estrato geográfico, promediando las tasas de no logro por ronda³³ dentro de cada estrato. Se estableció una cota superior para las tasas de no logro de 30% debido a las restricciones presupuestarias.

En el Anexo N° 4, se pueden observar los valores obtenidos para las tasas de no logro a nivel de estratos geográficos.

³⁰ Se define como vivienda gestionada, toda vivienda enviada a levantar en terreno por lo que, incluye viviendas originales y de reemplazo, indistintamente de su resultado final, es decir, si constituye una vivienda lograda o no lograda.

³¹ Para el caso de la ENE el nivel para el cual fue posible realizar el cálculo de las tasas de sobremuestreo, fue a nivel de estrato geográfico, tomando en cuenta que los datos utilizados para la estimación provenían del diseño muestral anterior.

³² Entiéndase por ronda, el número de entrevistas que es realizada a las mismas unidades de análisis de manera consecutiva.

³³ Aunque el diseño muestral anterior consideraba seis rondas para el estrato CD, nueve para el RAU y doce para el rural, tomando en cuenta que bajo el nuevo diseño se realizaron seis rondas en los estratos urbano y rural (ver sección 4.8.2.3), solo se analizaron las tasas de no logro para las seis primeras seis rondas.

La fórmula para calcular el tamaño de muestra sobredimensionado, viene dada por:

$$(n * \bar{m})_{h sm} = \frac{(n * \bar{m})_h}{1 - TNL_{h gest}} \quad (9)$$

Donde,

$(n * \bar{m})_{h sm}$: Tamaño muestral con sobremuestreo en el estrato h .

$(n * \bar{m})_h$: Tamaño de muestra objetivo en el estrato h .

$TNL_{h gest}$: Tasa de no logro sobre viviendas gestionadas en el estrato h .

4.8.2.5 Tamaño muestral de viviendas por UPM

Dado que el diseño muestral de la ENE 2020 es probabilístico, estratificado y bietápico, además de calcular el tamaño muestral de las unidades primarias de muestreo (UPM), también se debe calcular el tamaño promedio de las unidades secundarias de muestreo (Viviendas) por UPM.

Generalmente, las viviendas de un mismo conglomerado o UPM suelen ser similares en relación a los parámetros de interés para la encuesta (por ejemplo, ingreso, educación, ocupación, etc.), lo que puede deberse al hecho de que las personas que allí habitan, presentan niveles de ingreso similares, pueden compartir las mismas actitudes hacia los problemas del momento y con frecuencia se ven expuestos a las mismas condiciones ambientales (clima, enfermedades infecciosas, catástrofes naturales, etc.). Como las unidades pertenecientes al mismo conglomerado suelen ser semejantes entre sí, la correlación intraclasa es casi siempre positiva (Naciones Unidas, 2007).

Como bondad de la conglomeración destaca que reduce de manera notable los costos asociados al levantamiento de las encuestas, pero, por otro lado, a mayor correlación entre los elementos que pertenecen al conglomerado se origina un mayor incremento de la varianza, incluso una pequeña correlación positiva puede tener un gran efecto, debido al factor multiplicador $(\tilde{n} - 1)$. En consecuencia, se hace fundamental lograr el máximo equilibrio entre minimizar los costos y maximizar la precisión, a la hora de fijar el número de unidades de muestra objetivo dentro de cada UPM.

Para determinar cómo puede minimizarse o controlarse el componente de la conglomeración en el efecto del diseño conviene examinar la siguiente definición matemática (Naciones Unidas, 2009)³⁴:

³⁴ Esta expresión no es, estrictamente hablando, la fórmula del *Deff*, ya que no tiene en cuenta la estratificación ni el otro factor que interviene cuando los conglomerados no tienen un tamaño uniforme. Aun así, dado que el componente de la conglomeración es el factor predominante en *Deff*, puede usarse como una forma aproximada para mostrar cómo afecta la conglomeración en el diseño de la muestra y qué puede hacerse para controlarlo.

$$deff = 1 + (\tilde{n} - 1)\rho \quad (10)$$

Donde,

$deff$: Efecto del diseño.

\tilde{n} : Número de unidades promedio de la población objetivo³⁵ en la muestra de USM dentro del conglomerado.

ρ : Valor del coeficiente de homogeneidad o correlación intraclase.

La expresión anterior corresponde a una función multiplicativa de la correlación intraclase (ρ) y el tamaño del conglomerado (\tilde{n}). Dado que no se puede ejercer control sobre la correlación intraclásica de cualquiera de las variables sometidas al estudio, sí se puede ajustar el tamaño de la muestra para controlar así el efecto de diseño.

En la Tabla 13, se puede observar cómo influyen en el efecto del diseño el tamaño medio de la muestra objetivo por UPM y la correlación intraclase.

Tabla 13. Efectos del diseño en combinaciones seleccionadas de tamaño de la muestra conglomerada y correlación intraclase

\tilde{n}	ρ													
	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,1	0,15	0,2	0,35	0,5
5	1,0	1,1	1,1	1,2	1,2	1,2	1,3	1,3	1,4	1,4	1,6	1,8	2,4	3,0
10	1,1	1,2	1,3	1,4	1,5	1,5	1,6	1,7	1,8	1,9	2,4	2,8	4,2	5,5
15	1,1	1,3	1,4	1,6	1,7	1,8	2,0	2,1	2,3	2,4	3,1	3,8	5,9	8,0
20	1,2	1,4	1,6	1,8	2,0	2,1	2,3	2,5	2,7	2,9	3,9	4,8	7,7	10,5
30	1,3	1,6	1,9	2,2	2,5	2,7	3,0	3,3	3,6	3,9	5,4	6,8	11,2	15,5
50	1,5	2,0	2,5	3,0	3,5	3,9	4,4	4,9	5,4	5,9	8,4	10,8	18,2	25,5
75	1,7	2,5	3,2	4,0	4,7	5,4	6,2	6,9	7,7	8,4	12,1	15,8	26,9	38,0
100	2,0	3,0	4,0	5,0	6,0	6,9	7,9	8,9	9,9	10,9	15,9	20,8	35,7	50,5
125	2,2	3,5	4,7	6,0	7,2	8,4	9,7	10,9	12,2	13,4	19,6	25,8	44,4	63,0
150	2,5	4,0	5,5	7,0	8,5	9,9	11,4	12,9	14,4	15,9	23,4	30,8	53,2	75,5
175	2,7	4,5	6,2	8,0	9,7	11,4	13,2	14,9	16,7	18,4	27,1	35,8	61,9	88,0
200	3,0	5,0	7,0	9,0	11,0	12,9	14,9	16,9	18,9	20,9	30,9	40,8	70,7	100,5
225	3,2	5,5	7,7	10,0	12,2	14,4	16,7	18,9	21,2	23,4	34,6	45,8	79,4	113,0
250	3,5	6,0	8,5	11,0	13,5	15,9	18,4	20,9	23,4	25,9	38,4	50,8	88,2	125,5
275	3,7	6,5	9,2	12,0	14,7	17,4	20,2	22,9	25,7	28,4	42,1	55,8	96,9	138,0

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

³⁵ En el caso de la ENE, el valor de \tilde{n} que se utiliza es igual al número de viviendas del conglomerado, multiplicado por el promedio de personas en la fuerza de trabajo por vivienda.

Acogiendo las directrices prácticas de Naciones Unidas (2009), estimaciones del efecto de diseño ($deff$) por encima de 3,0 no son aceptables, por lo que se resalta en la tabla 13 previa aquellos escenarios que resultan factibles de aplicar.

Para definir el promedio de viviendas por UPM en el rediseño ENE 2020, se calcularon los coeficientes de correlación intraclase (ρ) para cada dominio de estudio definido, a partir de los datos del Censo 2017. En la Tabla 14, pueden observarse los efectos de diseño obtenidos para cada dominio de estudio, en función a los coeficientes de correlación intraclase y distintos valores de unidades promedio de la población objetivo en la muestra de USM dentro de los conglomerados³⁶. Como puede apreciarse todos los efectos de diseño obtenidos son menores a tres, por lo que una muestra de seis, ocho, diez o doce viviendas por UPM, garantiza estimaciones con un efecto de diseño aceptable.

Tabla 14. Efectos de diseño por dominios de estudio, para distintos tamaños de viviendas promedio por UPM

Región	Urbano								Rural							
	ρ	Pers. FFTT x Viv.	deff aproximado según \tilde{n}				ρ	Pers. FFTT x Viv.	deff aproximado según \tilde{n}				6	8	10	12
			6	8	10	12			6	8	10	12				
Nacional	0,010	1,5	1,08	1,11	1,15	1,18	0,012	1,3	1,08	1,10	1,13	1,16				
Arica y Parinacota	0,008	1,6	1,07	1,09	1,12	1,14	0,012	1,5	1,09	1,13	1,16	1,19				
Tarapacá	0,010	1,7	1,09	1,12	1,15	1,19	0,010	1,3	1,07	1,10	1,13	1,16				
Antofagasta	0,014	1,7	1,13	1,17	1,22	1,27	0,025	1,5	1,19	1,27	1,34	1,41				
Atacama	0,015	1,5	1,11	1,16	1,20	1,24	0,017	1,3	1,11	1,16	1,20	1,24				
Coquimbo	0,007	1,4	1,05	1,08	1,10	1,12	0,011	1,2	1,07	1,09	1,12	1,14				
Valparaíso	0,006	1,4	1,05	1,06	1,08	1,10	0,007	1,4	1,05	1,07	1,09	1,11				
Metropolitana	0,010	1,7	1,09	1,13	1,16	1,20	0,005	1,6	1,04	1,06	1,07	1,09				
O'Higgins	0,008	1,4	1,06	1,08	1,11	1,13	0,007	1,4	1,05	1,07	1,09	1,11				
Maule	0,007	1,4	1,05	1,07	1,09	1,11	0,011	1,2	1,07	1,10	1,12	1,15				
Ñuble	0,007	1,3	1,05	1,07	1,09	1,11	0,009	1,1	1,05	1,07	1,09	1,11				
Biobío	0,008	1,4	1,06	1,08	1,10	1,12	0,016	1,2	1,09	1,13	1,17	1,20				
La Araucanía	0,007	1,4	1,05	1,07	1,09	1,11	0,013	1,1	1,07	1,10	1,13	1,15				
Los Ríos	0,006	1,4	1,05	1,06	1,08	1,10	0,008	1,2	1,05	1,07	1,09	1,11				
Los Lagos	0,007	1,5	1,06	1,08	1,10	1,12	0,008	1,2	1,05	1,07	1,09	1,11				
Aysén	0,007	1,5	1,06	1,08	1,10	1,12	0,007	1,3	1,05	1,07	1,09	1,11				
Magallanes	0,003	1,5	1,03	1,04	1,05	1,06	0,003	1,5	1,02	1,03	1,04	1,05				

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

En este punto, para definir el número promedio de viviendas por UPM del nuevo diseño ENE, se realizaron simulaciones para evaluar los errores de muestreo de los principales indicadores de la encuesta, verificando que estuvieran dentro de los límites de tolerancia³⁷, que los tamaños

³⁶ Número de viviendas del conglomerado (m) multiplicado por el promedio de personas en la fuerza de trabajo por vivienda (Pers. FFTT).

³⁷ En cada dominio de estudio, para la tasa de desocupación, se fija como error máximo admisible 2%, en términos absolutos, y para el nivel de personas desocupadas, se fija como error máximo admisible 30%, en términos relativos.

muestrales resultantes estuvieran acordes a las capacidades operativas instaladas y en función a ello, determinar el número promedio de viviendas por UPM que cumpla con los objetivos del nuevo diseño muestral.

La Tabla 15 muestra los errores absolutos obtenidos para la tasa de desocupación y los errores relativos obtenidos para el nivel de población desocupada en los dominios urbanos mediante las simulaciones, y como se puede observar en algunos casos el error absoluto resulta mayor al 2%, por ejemplo en el dominio Atacama urbano en todos los casos la estimación del error absoluto es mayor al 2%, y en las áreas urbanas de Tarapacá y Ñuble con doce viviendas por UPM, también se obtuvieron errores absolutos por encima de 2%. En cuanto al error relativo, éste resulta mayor a 30% en el dominio urbano de la región de Aysén en todos los casos (ocho, diez y doce viviendas) y en la región de Magallanes con diez y doce viviendas.

Tabla 15. Errores muestrales de la tasa de desocupación para los dominios urbanos según número promedio de vivienda por UPM

Región	Tasa de Desocupación			Nivel de población desocupada		
	Error absoluto ³⁸			Error relativo		
	8 Viviendas	10 Viviendas	12 Viviendas	8 Viviendas	10 Viviendas	12 Viviendas
Nacional Urbano	0,30%	0,30%	0,40%	4,60%	5,10%	5,50%
Arica y Parinacota	1,50%	1,70%	1,90%	20,50%	22,80%	25,00%
Tarapacá	1,80%	2,00%	2,10%	24,90%	28,00%	30,00%
Antofagasta	1,50%	1,60%	1,80%	18,40%	20,50%	22,40%
Atacama	2,20%	2,40%	2,60%	23,10%	25,90%	28,50%
Coquimbo	1,50%	1,70%	1,80%	18,70%	20,40%	21,90%
Valparaíso	0,80%	0,90%	1,00%	12,50%	13,80%	15,20%
Metropolitana	0,50%	0,60%	0,60%	8,40%	9,30%	10,00%
O'Higgins	1,30%	1,40%	1,50%	18,80%	21,10%	23,10%
Maule	1,20%	1,30%	1,40%	17,30%	19,10%	20,90%
Ñuble	1,70%	1,90%	2,10%	21,40%	23,40%	25,60%
Biobío	0,90%	1,00%	1,10%	10,50%	11,70%	12,80%
La Araucanía	1,30%	1,40%	1,60%	18,40%	20,60%	22,40%
Los Ríos	1,50%	1,70%	1,80%	22,00%	24,60%	27,10%
Los Lagos	1,10%	1,20%	1,30%	17,20%	19,40%	21,20%
Aysén	1,70%	1,70%	1,80%	34,30%	34,60%	36,70%
Magallanes	1,30%	1,40%	1,50%	29,50%	32,60%	34,70%

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Para la toma de decisiones en los dominios rurales se definió simular como mínimo diez (10) viviendas por UPM y no ocho como se estableció para las áreas urbanas, esto considerando que las UPM rurales son áreas geográficas extensas con mayor dispersión geográfica de sus viviendas, en consecuencia, con el fin de hacer uso eficiente de las capacidades operativas instaladas.

³⁸ Para más información consultar el Informe de Simulaciones de tamaños muestrales de la ENE 2020.

En el caso de los dominios rurales, la Tabla 16 muestra los errores absolutos de la tasa de desocupación obtenidos mediante las simulaciones, como se puede observar en la mayoría de los dominios, para el caso de diez viviendas por UPM, la mayoría de los errores absolutos resultaron mayores a 2%, por ende en los casos de doce y catorce viviendas por UPM, los errores absolutos resultaron aún mayores.

Tabla 16. Errores muestrales de la tasa de desocupación para los dominios rurales según número promedio de vivienda por UPM

Región	Tasa de Desocupación			Nivel de población desocupada		
	Error absoluto ³⁹			Error relativo		
	10 Viviendas	12 Viviendas	14 Viviendas	10 Viviendas	12 Viviendas	14 Viviendas
Nacional Rural	0,70%	0,80%	0,80%	11,80%	12,90%	13,70%
O'Higgins	2,10%	2,20%	2,30%	37,40%	41,10%	42,40%
Maule	2,10%	2,30%	2,40%	35,10%	38,10%	40,10%
Ñuble	2,40%	2,80%	2,90%	33,90%	38,30%	40,20%
La Araucanía	2,60%	2,80%	3,00%	35,60%	39,70%	41,60%
Los Ríos	2,30%	2,50%	2,70%	39,70%	42,70%	46,10%
Los Lagos	1,60%	1,60%	1,80%	35,40%	37,70%	38,80%
Metropolitana	1,70%	1,90%	2,00%	38,50%	40,70%	45,20%

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Considerando los resultados de las simulaciones, se determina que el número de viviendas a levantar por UPM será de diez viviendas en estratos urbanos y rurales, con excepción de los estratos urbanos de las regiones de Atacama, Aysén y Magallanes, donde se determina levantar ocho viviendas por UPM, dado que los errores de muestreo con diez viviendas o más resultaban superiores a los límites de tolerancia.

Por último, la tabla 17 presenta los pasos de cálculo realizados hasta obtener el tamaño de muestra objetivo definitivo, los tamaños a nivel de viviendas y UPM, y los errores de simulación asociados, para todos los niveles de estimación.

³⁹ Para más información consultar el Informe de Simulaciones de tamaños muestrales de la ENE 2020.

Tabla 17. Cálculo de tamaños definitivos

Nivel de estimación	Parámetros de cálculo						Tamaños				
	Error esperado	Tasa desocupación	Varianza desocupación	Efecto diseño	FT*	Viviendas marco	$\tilde{m}_{0,ra}$	$\tilde{m}_{1,ra}$	$m_{1,ra}$	$m_{2,ra}$	$m'_{2,ra}$
Región área urbana											
Arica y Parinacota	1,72%	0,08	0,07	2,03	1,94	65.926	980	1.993	1.025	1.010	1.010
Tarapacá	2,00%	0,08	0,07	2,03	2,07	101.541	710	1.438	694	690	690
Antofagasta	1,65%	0,08	0,08	2,08	2,06	180.888	1.144	2.384	1.157	1.150	1.150
Atacama	1,92%	0,1	0,09	2,2	1,76	92.934	967	2.131	1.211	1.196	1.200
Coquimbo	1,69%	0,09	0,08	2,02	1,71	219.741	1.124	2.267	1.327	1.319	1.320
Valparaíso	0,93%	0,07	0,06	2,07	1,67	610.005	3.003	6.230	3.722	3.699	3.700
Metropolitana	0,59%	0,07	0,06	1,97	2,07	2.269.501	7.185	14.122	6.830	6.810	6.810
O'Higgins	1,39%	0,07	0,07	1,89	1,68	246.081	1.345	2.538	1.509	1.499	1.500
Maule	1,28%	0,07	0,06	1,83	1,6	279.511	1.582	2.900	1.811	1.799	1.800
Ñuble	1,87%	0,08	0,08	1,91	1,53	122.129	879	1.679	1.099	1.089	1.090
Biobío	0,97%	0,09	0,08	1,81	1,59	485.661	3.373	6.092	3.829	3.799	3.800
La Araucanía	1,44%	0,07	0,07	1,82	1,57	248.038	1.322	2.411	1.539	1.529	1.530
Los Ríos	1,67%	0,07	0,07	1,95	1,63	99.265	945	1.843	1.132	1.119	1.120
Los Lagos	1,22%	0,07	0,06	1,8	1,68	221.508	1.643	2.959	1.763	1.749	1.750
Aysén	1,44%	0,05	0,05	1,87	1,67	30.542	916	1.716	1.024	991	808
Magallanes	1,31%	0,04	0,04	1,76	1,78	56.428	974	1.716	966	949	952
Región área rural											
Arica y Parinacota	3,57%	0,03	0,03	2,55	1,69	5.940	80	205	121	119	120
Tarapacá	5,43%	0,04	0,04	4,37	1,75	4.737	49	215	123	119	120
Antofagasta	6,35%	0,04	0,03	6,48	1,78	5.670	34	218	122	119	120
Atacama	8,51%	0,08	0,07	4,88	1,57	12.109	39	190	121	120	180
Coquimbo	2,60%	0,07	0,06	3,67	1,42	58.319	375	1.374	965	949	950
Valparaíso	1,94%	0,05	0,05	2,61	1,64	59.615	536	1.399	851	839	840
Metropolitana	1,75%	0,05	0,05	2,25	1,87	83.457	630	1.415	756	749	750
O'Higgins	2,05%	0,06	0,05	2,92	1,62	85.265	506	1.477	909	899	900
Maule	2,09%	0,06	0,06	2,56	1,45	106.113	535	1.370	948	939	940
Ñuble	2,42%	0,08	0,07	2,11	1,26	56.972	484	1.020	811	799	800
Biobío	2,71%	0,08	0,07	2,43	1,36	66.345	388	944	697	689	690
La Araucanía	2,60%	0,07	0,07	2,42	1,28	102.560	404	980	765	759	760
Los Ríos	2,33%	0,06	0,06	2,27	1,33	41.717	430	975	732	719	720
Los Lagos	1,56%	0,05	0,04	2,19	1,39	77.891	718	1.574	1.136	1.119	1.120
Aysén	4,34%	0,03	0,03	3,07	1,46	9.963	59	183	125	123	120
Magallanes	3,38%	0,02	0,02	3,41	1,79	4.366	65	221	123	120	120

* Cantidad de personas integrantes de la fuerza de trabajo en promedio por vivienda.

Tabla 18. Distribución de la muestra y resultados de simulación

Región	Nacional				Urbano				Rural			
	Viviendas objetivo	Viviendas sobre-muestreo	UPM	Error de simulación	Viviendas objetivo	Viviendas sobre-muestreo	UPM	Error de simulación	Viviendas objetivo	Viviendas sobre-muestreo	UPM	Error de simulación
Nacional	39.480	54.126	5.511	0,30%	30.230	41.396	4.238	0,34%	9.250	12.730	1.273	0,69%
Arica y Parinacota	1.130	1.480	148	1,62%	1.010	1.300	130	1,70%	120	180	18	3,75%
Tarapacá	810	1.030	103	1,96%	690	850	85	1,99%	120	180	18	5,65%
Antofagasta	1.270	1.660	166	1,62%	1.150	1.500	150	1,64%	120	160	16	6,25%
Atacama	1.380	1.950	237	1,61%	1.200	1.680	210	1,65%	180	270	27	6,55%
Coquimbo	2.270	3.250	325	1,51%	1.320	1.910	191	1,69%	950	1.340	134	2,61%
Valparaíso	4.540	6.120	612	0,88%	3.700	4.910	491	0,93%	840	1.210	121	1,93%
Metropolitana	7.560	10.660	1.066	0,58%	6.810	9.560	956	0,59%	750	1.100	110	1,76%
O'Higgins	2.400	3.170	317	1,20%	1.500	1.970	197	1,39%	900	1.200	120	2,05%
Maule	2.740	3.810	381	1,12%	1.800	2.600	260	1,29%	940	1.210	121	2,11%
Ñuble	1.890	2.520	252	1,54%	1.090	1.480	148	1,87%	800	1.040	104	2,46%
Biobío	4.490	6.120	612	0,91%	3.800	5.110	511	0,97%	690	1.010	101	2,74%
La Araucanía	2.290	3.060	306	1,24%	1.530	2.110	211	1,43%	760	950	95	2,44%
Los Ríos	1.840	2.630	263	1,37%	1.120	1.640	164	1,63%	720	990	99	2,34%
Los Lagos	2.870	4.090	409	1,02%	1.750	2.520	252	1,22%	1.120	1.570	157	1,52%
Aysén	928	1.316	160	1,35%	808	1.136	142	1,42%	120	180	18	3,47%
Magallanes	1.072	1.260	154	1,19%	952	1.120	140	1,23%	120	140	14	3,41%

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

4.8.2.6 Selección de unidades de primera etapa

En el diseño muestral de la ENE, para la selección de las UPM o conglomerados, se emplea una selección aleatoria sistemática con probabilidad proporcional al tamaño. En lo referido a la información asociada al tamaño muestral y sus desagregaciones, revisar el apartado [“4.8.1.1 Tamaño muestral”](#) presente en este mismo capítulo.

En las encuestas de hogares por muestreo, el método de selección con probabilidad proporcional al tamaño (PPT) de las UPM es el más empleado por los países en desarrollo (Naciones Unidas, 2007). Se resalta que el método tiene algunas ventajas prácticas cuando las UPM varían considerablemente de tamaño; en primer lugar, este método podría dar lugar a muestras autoponderadas; en segundo lugar, genera muestras de tamaño aproximadamente igual dentro de las UPM, lo que implica cargas de trabajo similares para los encuestadores, constituyendo así una situación deseable desde la perspectiva del trabajo en el terreno. Además, es una técnica que utiliza información auxiliar para aumentar significativamente la precisión de las estimaciones de las encuestas, y que, 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 ENE la información auxiliar está definida por el tamaño de la UPM, medido en función a la cantidad de viviendas que contiene, por lo que las UPM con mayor cantidad de viviendas, tendrán mayor probabilidad de selección.

Cabe resaltar que, aunque las probabilidades de selección en primera etapa sean desiguales, al seleccionar las unidades secundarias de muestreo (USM) con igual probabilidad, se obtiene una muestra autoponderada, es decir, todas las unidades de la muestra tendrán la misma probabilidad de selección. No obstante, dicho supuesto no se aplica fácilmente en la práctica, puesto que, en la mayoría de los casos, la información del tamaño de la UPM no suele estar actualizada, debido a los largos lapsos transcurridos entre los Censos y actualizaciones de los marcos.

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).

Al respecto, en las encuestas de hogares por muestreo del INE, y en particular en la ENE, dada la desactualización natural del marco muestral con el transcurso del tiempo y como estrategia para mitigar posibles sesgos, se realiza una actualización de los listados de viviendas de las UPM seleccionadas. Por consiguiente, la muestra de USM se realizará a partir de la nueva lista, por lo que la probabilidad de selección de las viviendas queda definida por:

$$\pi_{hij} = \pi_{hi} \cdot \pi_{hi/j} \quad (11)$$

Donde,

- π_{hij} : Probabilidad de seleccionar la vivienda j en el conglomerado i dentro del estrato h .
- π_{hi} : Probabilidad de seleccionar el conglomerado i en el estrato h .
- $\pi_{hi/j}$: Probabilidad de seleccionar la vivienda j dado que fue seleccionado el conglomerado i dentro del estrato h .

Donde la probabilidad de selección de la UPM o conglomerado con probabilidad proporcional al tamaño, según su número de viviendas dentro del estrato, queda determinada como:

$$\pi_{hi} = \frac{n_h \cdot M_{hi}}{M_h} = \frac{n_h \cdot M_{hi}}{\sum_{\forall i \in \mathcal{M}_h} M_{hi}} \quad (12)$$

Donde,

- π_{hi} : Probabilidad de seleccionar el conglomerado i en el estrato h .
- n_h : Número de conglomerados en el estrato h .
- M_{hi} : Número de viviendas en el conglomerado i del estrato h .
- M_h : Número de viviendas en el estrato h .
- \mathcal{M}_h : Conjunto de todos los conglomerados del MMV 2017 pertenecientes al estrato h .

4.8.2.7 Selección de unidades de segunda etapa

Luego de seleccionar las UPM, al interior de cada una de estas se procede a seleccionar las viviendas.

La selección de viviendas dentro de cada UPM seleccionada, tanto urbana como rural, se realiza de manera independiente, en forma sistemática y, con igual probabilidad. Al igual que en otras encuestas, al algoritmo de selección sistemático de viviendas, se introduce una semilla de aleatorización para que, independientemente de quien aplique el mismo algoritmo, bajo las mismas condiciones de ordenamiento previo⁴⁰, seleccione las mismas viviendas.

El procedimiento es implementado en el *software SAS* bajo la función *Survey Select*, lo cual puede ser resumido mediante el siguiente algoritmo:

⁴⁰ Las viviendas al interior de cada UPM son ordenadas según la variable “Número de orden vivienda”, mediante el cual se asegura el recorrido del conglomerado.

Se calcula el período K .

$$K = \frac{M'_{hi}}{m_{hi}} \quad (13)$$

Donde,

m_{hi} : Número de viviendas seleccionadas en el conglomerado i del estrato h .

M'_{hi} : Número de viviendas del conglomerado i , en el estrato h , según actualización de los listados de viviendas.

Luego se genera un número aleatorio entero A entre 1 y el período K .

Enseguida se va sumando sucesivamente el período K al arranque A para obtener distintos valores que al redondearlos van generando las sucesivas selecciones:

$$A, A + K, A + 2K, A + 3K, A, \dots, A + (m_i - 1)K \quad (14)$$

La primera selección A ya es un número redondeado a entero, la segunda es el redondeo de $A + K$, la tercera es el redondeo de $A + 2K$, y así sucesivamente hasta la m_i selección dada por el redondeo de $A + (m_i - 1)K$.

Tomando en cuenta que se actualiza el número de viviendas de las UPM seleccionadas y que dentro del estrato h la selección se realiza con igual probabilidad, se tiene que la probabilidad de seleccionar la vivienda j de la UPM i seleccionada, queda definida por la siguiente expresión:

$$\pi_{hi/j} = \frac{m_{hi}}{M'_{hi}} \quad (15)$$

Donde,

$\pi_{hi/j}$: Probabilidad de seleccionar la vivienda j dado que fue seleccionado el conglomerado i en el estrato h .

En consecuencia, la probabilidad de seleccionar, dentro del estrato h , el conglomerado i y la vivienda j , queda denotada como:

$$\pi_{hij} = \frac{n_h \cdot M_{hi}}{\sum_{v_i \in \mathcal{M}_h} M_{hi}} \cdot \frac{m_{hi}}{M'_{hi}} \quad (16)$$

En la medida en que M_{hi} difiera de M'_{hi} determina que, al interior de un estrato, las viviendas entre conglomerados tengan probabilidad de selección distinta, por lo que, el diseño muestral pierde la propiedad de ser autoponderado.

No obstante, al considerar que las nuevas UPM para el diseño muestral de la ENE, sean concebidas con tamaños aproximadamente iguales, parece no exacerbar el uso del muestreo con probabilidad proporcional al tamaño. Cabe señalar que, la diferencia entre M_{hi} y M'_{hi} puede dar cuenta de la desactualización del marco, lo que naturalmente ocurrirá con el transcurso del tiempo.

Por otro lado, la aplicación de un muestreo sistemático para la selección de UPM, a diferencia del muestreo aleatorio simple, introduce al diseño una estratificación implícita, lo que hace que la muestra esté mejor distribuida dentro de cada estrato aportando mayor cobertura y permitiendo generar estimaciones más fiables.

En resumen, la selección sistemática de UPM con probabilidad proporcional al tamaño en términos de viviendas, permite controlar el tamaño de la muestra total y por ende los costos, lo que posibilita asignar cargas de trabajo más equitativas, obtener mejor cobertura dentro de cada estrato y, además, aporta información sobre la desactualización del marco.

4.8.2.8 Rotación de las unidades muestrales

La ENE es una encuesta tipo panel, por ello las unidades de análisis son entrevistadas más de una vez en varios momentos, con la finalidad de medir el cambio en los indicadores de interés a través del tiempo. Es por ello que, con el objeto de mejorar la precisión de dichas estimaciones de cambio, la ENE emplea un diseño muestral rotativo⁴¹ para garantizar un porcentaje de traslape o superposición entre las muestras de un período a otro que, a su vez, minimice la atrición de la muestra, debido a la no respuesta por cansancio de los informantes ocasionado por las repetidas entrevistas y adicionalmente, permita refrescar la muestra.

La muestra total de la ENE se distribuye en tres submuestras⁴². De esta manera, durante los meses de enero, abril, julio y octubre de cada año, son entrevistadas las viviendas de la submuestra 1. En los meses de febrero, mayo, agosto y noviembre las viviendas de la submuestra 2 y, por último, durante los meses de marzo, junio, septiembre y diciembre las viviendas de la submuestra 3.

El diseño panel contempla en total seis rondas de aplicación para cada vivienda, por consiguiente, la muestra de UPM de cada submuestra es dividida en seis turnos de rotación. En cada trimestre móvil una proporción fija de las viviendas muestreadas (1/18) es rotada para incorporar un nuevo conjunto de viviendas (muestra de refresco), permitiendo rotar la totalidad de la muestra de viviendas en un período de 18 meses.

⁴¹ Un diseño muestral rotativo se refiere a aquel donde una fracción de la muestra, después de cada cierto número de entrevistas, es refrescada con un nuevo grupo de unidades muestrales, esto como mecanismo de actualización muestral.

⁴² Este procedimiento se lleva a cabo distribuyendo equitativamente la muestra de UPM de cada estrato de muestreo ENE, entre las tres submuestras.

Figura 7. Esquema de rotación de la muestra de viviendas

Panel de rotación	Año t												Año t+1													
	E	F	M	A	M	J	J	A	S	O	N	D	E	F	M	A	M	J	J	A	S	O	N	D		
6	R6	R6	R6																							
5	R5	R5	R5	R6	R6	R6																				
4	R4	R4	R4	R5	R5	R5	R6	R6	R6	R6	R6	R6														
3	R3	R3	R3	R4	R4	R4	R5	R5	R5	R6	R6	R6														
2	R2	R2	R2	R3	R3	R3	R4	R4	R4	R5	R5	R5	R6	R6	R6	R6	R6	R6	R6	R6	R6	R6	R6	R6		
1	R1	R1	R1	R2	R2	R2	R3	R3	R3	R4	R4	R4	R5	R5	R5	R6										
6				R1	R1	R1	R2	R2	R2	R3	R3	R3	R4	R4	R4	R5	R5	R5	R6	R6	R6	R6	R6	R6		
5						R1	R1	R1	R2	R2	R2	R3	R3	R3	R4	R4	R4	R4	R5	R5	R5	R6	R6	R6	R6	
4									R1	R1	R1	R2	R2	R2	R3	R3	R3	R4	R4	R4	R5	R5	R5	R5	R5	
3												R1	R1	R1	R2	R2	R2	R2	R3	R3	R3	R4	R4	R4	R4	
2															R1	R1	R1	R2	R2	R2	R3	R3	R3	R3	R3	R3
1																		R1	R1	R1	R2	R2	R2	R2	R2	
6																					R1	R1	R1	R1	R1	

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

La Figura 7 presenta el esquema de rotación de la muestra de viviendas. Los colores señalan el panel de rotación al que pertenece cada vivienda (el color naranja corresponde al panel 1 de cada submuestra, el verde al panel 2, el rosado al panel 3, el azul al panel 4, el gris al panel 5 y el morado al panel 6). Los números que acompañan a la letra “R” en cada recuadro, indican el número de períodos que las viviendas asociadas a cada uno de los seis paneles de rotación han sido parte de la encuesta. Como se evidencia en el esquema, una parte de la muestra es renovada mensualmente⁴³. Por ejemplo, en abril (submuestra 1), las viviendas del panel de rotación morado están en su primera ronda, mientras que las viviendas del panel de rotación gris están en su sexta y última ronda. En Julio (submuestra 1), las viviendas del panel de rotación gris son refrescadas, es decir, el grupo de viviendas de dicho grupo que cumplió seis rondas en abril salen de la muestra y es incorporado un nuevo grupo de viviendas, como se ve en el diagrama. Lo anterior es aplicado de manera análoga en cada submuestra.

La muestra de refresco generalmente es seleccionada en las mismas unidades primarias de muestreo, sin embargo, también se contempla una actualización de las UPM, permitiendo el refresco de áreas geográficas.

⁴³ Se debe recordar que la muestra total ENE está conformada por las tres submuestras, por lo que, de un mes a otro, la fracción que se renueva representa una dieciochoava parte de la muestra total de la ENE.

4.8.2.9 Incorporación de la muestra del MMV 2017

La incorporación de la nueva muestra proveniente del MMV 2017, se inicia en el mes de enero 2020, de modo que, para conformar la muestra del trimestre móvil noviembre, diciembre 2019 y enero 2020 (NDE 2019) es incorporada una fracción de la muestra proveniente del nuevo marco, a razón de 1/6 de la muestra trimestral⁴⁴ (aproximadamente 9.000 viviendas).

Paralelamente 1/6 de la muestra proveniente del marco muestral anterior⁴⁵ (aproximadamente 6.300 viviendas) es restada de la muestra total, generándose de esta manera un incremento de la muestra. Este procedimiento se realiza de forma análoga en las submuestras de febrero, marzo, abril, mayo y junio 2020, por lo que el tamaño muestral irá aumentando gradualmente. Es así como, en el trimestre móvil de abril, mayo y junio 2020 (AMJ 2020) la muestra ENE tendrá un nuevo tamaño muestral⁴⁶ aproximado de 54.000⁴⁷ viviendas, todas provenientes del nuevo MMV 2017.

Figura 8. Incorporación de la muestra MMV 2017 a partir de las rotaciones

2019		2020											
N	D	E	F	M	A	M	J	J	A	S	O	N	D
R6	R6												
R5	R5												
R4	R4												
R3	R3	R4	R4	R4									
R2	R2	R3	R3	R3									
R1	R1	R2	R2	R2									
		R1	R1	R1	R2	R2	R2						
		R1	R1	R1	R2	R2	R2						
		R1	R1	R1	R2	R2	R2						
					R1	R1	R1						
					R1	R1	R1						
					R1	R1	R1						

Submuestra 1 muestra anterior
Submuestra 2 muestra anterior
Submuestra 3 muestra anterior

Submuestra 1 muestra nueva
Submuestra 2 muestra nueva
Submuestra 3 muestra nueva

→ UPM del MMV 2017

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Esta metodología representa una incorporación paulatina de la nueva muestra, controlando de esta forma, los efectos que pudieran producir en las estimaciones de los parámetros de interés, los cambios en el diseño muestral. Sin embargo, es importante resaltar que durante los primeros cinco meses del año 2020, no es posible garantizar en su totalidad los niveles de precisión planteados para el nuevo diseño, ya que los tamaños muestrales se irán incrementando gradualmente, hasta llegar al tamaño

⁴⁴ Considerar que este 1/6 es calculado sobre el nuevo tamaño muestral de la ENE, el cual contempla una sobremuestra.

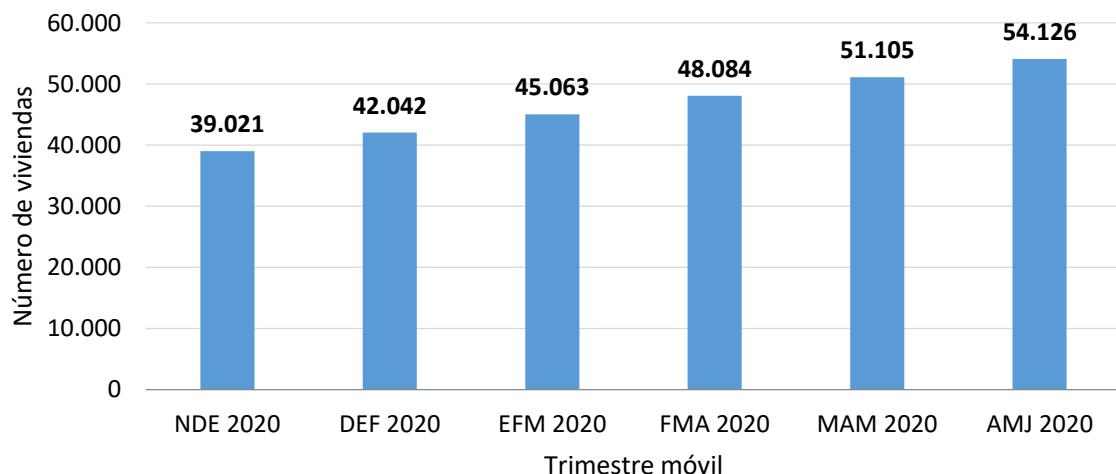
⁴⁵ Este 1/6 es calculado sobre el tamaño de muestra del diseño muestral de la ENE vigente hasta diciembre 2019.

⁴⁶ El tamaño muestral del diseño anterior era de 38.000 viviendas y no contemplaba sobremuestra.

⁴⁷ Incluyendo la sobremuestra.

de muestra definido para alcanzar los niveles de precisión óptimos, en el mes de junio 2020⁴⁸ (ver Figura 9).

Figura 9. Incremento gradual de la muestra teórica durante el período de transición de la ENE



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

a) Incorporación de la nueva muestra ENE por turnos

Para garantizar que, durante los seis meses de transición, la muestra proveniente del marco MMV 2017 sea incorporada a razón 1/6, luego de seleccionar las UPM, dentro de cada estrato de muestreo del nuevo diseño se asignó de forma sistemática un número entre 1 y 6, haciendo referencia al mes durante el cual sería incorporada la UPM, garantizando de esta forma, incorporar durante cada uno de los seis meses 1/6 de la muestra del MMV 2017 aproximadamente, en cada estrato de muestreo. En este caso, la distribución para los turnos de incorporación se pudo realizar a nivel de UPM dado que en el nuevo diseño la cantidad de viviendas a levantar por UPM es igual (diez viviendas), con excepción de los estratos urbanos de las regiones de Atacama, Aysén y Magallanes que es de ocho viviendas.

b) Selección de las viviendas a restar en cada mes

La selección de la muestra proveniente de los marcos MMS2016 y MMM2016⁴⁹ que se debían restar en cada mes, se realizó dentro de cada estrato de muestreo del diseño muestral anterior, de forma independiente para cada submuestra y a nivel de vivienda, debido a la gran diferencia en la construcción de las unidades primarias de muestreo entre los marcos MMS 2016, MMM 2016 y MMV 2017.

⁴⁸ Durante los meses de incorporación progresiva (etapa de transición) de la muestra serán evaluados los errores de estimación de los niveles definidos en el estudio, siendo posible que algunas estimaciones no alcancen los errores óptimos para ser oficiales.

⁴⁹ Hasta diciembre de 2019, la muestra de la ENE fue seleccionada desde un marco muestral conformado por dos listados de unidades, el Marco Muestral de Manzanas actualizado con base en los listados de Precio 2016 (MMM 2016) para el estrato CD que comprendía las ciudades y grandes centros urbanos; y el Marco Muestral de Secciones también actualizado en parte, con base en los listados de Precio 2016 (MMS 2016) para los estratos RAU (Resto de Área Urbana) y rural.

Para seleccionar las viviendas de la muestra proveniente del marco muestral anterior que serían restadas de la muestra en cada uno de los meses durante el período de transición, se priorizaron las viviendas, bajo los siguientes criterios:

- 1. Viviendas que estén ubicadas dentro de una UPM seleccionada en la muestra proveniente del MMV 2017:** Para esto fue necesario ubicar cada vivienda dentro de las UPM del nuevo marco. Sin embargo, en las secciones, al no existir automatización cartográfica, solo fue posible determinar intersecciones entre las secciones y UPM.

Las viviendas identificadas en este primer paso fueron priorizadas para ser restadas en primera instancia, reduciendo así la probabilidad de seleccionar en la nueva muestra, una vivienda que ya pertenecía a la muestra ENE seleccionada desde el marco muestral anterior, evitando un doble levantamiento.

- 2. Viviendas con mayor número de rondas realizadas:** En segundo orden de prioridad, se restaron aquellas viviendas con mayor número de rondas de entrevista cumplidas, hasta alcanzar la fracción de 1/6 de la muestra. Así, aquellas viviendas que tenían cinco, cuatro y hasta tres visitas en los períodos anteriores fueron priorizadas para ser restadas.

En algunos estratos de muestreo, se identificó que la cantidad de viviendas de la muestra anterior que están localizadas dentro de una UPM seleccionada desde el marco MMV 2017, es mayor a 1/6 de la muestra anterior. En estos casos se flexibilizó la fracción definida permitiendo restar en el primer grupo hasta 1/5 de las viviendas dentro del estrato de muestreo anterior.

4.8.3 Cálculo del factor de expansión

Los factores de expansión pueden definirse como el número de unidades en la población representadas por un elemento de la muestra, y la suma de ellos es igual al tamaño de la población, de modo que la muestra completa representa a toda la población (Lohr, 2000).

Para generar las estimaciones de las variables de interés para la ENE se emplea el factor de expansión a nivel de persona, sin embargo, para su cálculo es necesario calcular también el factor de expansión a nivel de vivienda, siendo este último aplicado a todos los hogares de la vivienda y todas las personas del hogar, calibrado a los stocks poblacionales según la metodología de calibración definida. El diseño muestral de la ENE es bietápico, por tanto, la fórmula de cálculo del factor de expansión debe considerar las probabilidades de selección de las unidades de primera etapa (UPM) y segunda etapa (viviendas).

Los componentes del factor de expansión se obtienen como el producto de los inversos de las probabilidades de selección de las unidades de primera etapa (UPM) y las probabilidades de selección de las unidades de segunda etapa (viviendas) seleccionadas desde los conglomerados ya muestreados, por lo que depende de:

- a) El número de conglomerados seleccionados dentro de cada estrato h , el número de viviendas en el estrato h y en el conglomerado i seleccionado, según información reportada por el marco muestral.
- b) El número de viviendas seleccionadas y existentes en el conglomerado i , según actualización de los listados de viviendas del conglomerado seleccionado.

Adicionalmente, debido a que los marcos muestrales no son perfectos y que pueden existir desactualizaciones observadas en el proceso de levantamiento, así como también no respuesta de las unidades⁵⁰, los factores de expansión de selección consideran cuatro ajustes, que son aplicados secuencialmente:

1. Ajuste por elegibilidad desconocida
2. Ajuste por no elegibilidad
3. Ajuste por no respuesta
4. Ajuste a stocks poblacionales

Dichos ajustes permiten, por una parte, distribuir los pesos de las unidades de elegibilidad desconocida entre las unidades elegibles y no elegibles; y, por otra parte, desechar los pesos de las unidades no elegibles, ya que por definición estas unidades no contienen a la población objetivo, por lo que hacer inferencias sobre ellas sería incorrecto. Finalmente, por diversos motivos, en algunas unidades, aun siendo elegibles, no se obtiene una respuesta de los informantes y por tanto su peso debe ser distribuido entre otras unidades que debería representarlos en la muestra.

En los apartados siguientes se presenta la metodología utilizada para llevar a cabo cada uno de los ajustes antes señalados.

⁵⁰ Las mejoras al instrumento de recolección permiten recopilar, con mayor nivel de detalle, información sobre el resultado o estado de disposición de cada unidad muestral seleccionada. Los códigos de visita dan cuenta de la elegibilidad de cada unidad muestral y permiten clasificar las viviendas en tres grupos específicos: elegibles, de elegibilidad desconocida y no elegibles. Información que es utilizada posteriormente para realizar los correspondientes ajustes al factor de expansión (ver Anexo N° 3. Códigos de visita de la ENE).

4.8.3.1 Factor de expansión de selección de conglomerados

Dado el carácter bietápico del diseño muestral de la ENE, en una primera etapa se seleccionan las unidades primarias de muestreo con probabilidad proporcional al tamaño, en función al número de viviendas particulares ocupadas que la UPM registraba en el Censo de Población y Vivienda 2017 o conforme a la actualización del marco más reciente, por lo que el ponderador de selección de cada UPM expresado como el inverso de la probabilidad de selección, viene dado por:

$$w_{hi} = (\pi_{hi})^{-1} = \frac{M_h}{n_h \cdot M_{hi}} \quad (17)$$

Donde,

w_{hi} : Factor de expansión de selección del conglomerado i en el estrato h .

π_{hi} : Probabilidad de selección del conglomerado i en el estrato h .

M_h : Número de viviendas particulares en el estrato h al Censo 2017 o al año de actualización del marco muestral.

M_{hi} : Número de viviendas particulares del conglomerado i en el estrato h al Censo 2017 o al año de actualización del marco muestral.

n_h : Número de conglomerados seleccionados en la muestra del estrato h .

Este ponderador se interpreta como “el total de unidades de su tipo que representa en la población”.

4.8.3.2 Factor de expansión de selección de viviendas

El ponderador de selección de viviendas w_{hij} , indica el número de viviendas en la población que representa la vivienda seleccionada en la muestra y es calculado como el inverso de la probabilidad de selección de cada vivienda.

La probabilidad total de seleccionar una vivienda j en el estrato h , está dada por el producto de seleccionar un conglomerado i en el estrato h y la probabilidad de seleccionar una vivienda j , condicional a haber seleccionado el conglomerado i (selección de la vivienda j dentro del conglomerado i seleccionado previamente). Lo anterior, se puede expresar como:

$$\pi_{hij} = \pi_{hi} \cdot \pi_{hj/i}, \text{ con } \pi_{hj/i} = \frac{m_{hi}}{M'_{hi}} \quad (18)$$

Finalmente, el ponderador de selección de viviendas, está dado por:

$$w_{hij} = (\pi_{hij})^{-1} = \frac{1}{\pi_{hi}} \cdot \frac{1}{\pi_{hj/i}} = \frac{M_h}{n_h \cdot M_{hi}} \cdot \frac{M'_{hi}}{m_{hi}} \quad (19)$$

Donde,

w_{hij} : Factor de expansión de selección de la vivienda j dentro del conglomerado i en el estrato h .

π_{hij} : Probabilidad de selección de la vivienda j dentro del conglomerado i en el estrato h .

- π_{hi} : Probabilidad de selección del conglomerado i en el estrato h .
- $\pi_{hi/j}$: Probabilidad de selección de la vivienda j dado que fue seleccionado el conglomerado i en el estrato h .
- M_h : Número de viviendas particulares en el estrato h al Censo 2017 o al año de actualización del marco muestral.
- n_h : Número de conglomerados en la muestra en el estrato h .
- M_{hi} : Número de viviendas particulares del conglomerado i en el estrato h al Censo 2017 o al año de actualización del marco muestral.
- m_{hi} : Número de viviendas seleccionadas en el conglomerado i en el estrato h .
- M'_{hi} : Número de viviendas particulares ocupadas, según actualización de los listados de viviendas, en el conglomerado i en el estrato h .

4.8.3.3 Ajuste por elegibilidad

Los ajustes por elegibilidad tienen como objetivo corregir las posibles fallas de haber incluido en la muestra viviendas no elegibles y aquellas clasificadas como de elegibilidad desconocida.

a) Ajuste por elegibilidad desconocida

En principio, todas las viviendas tienen un estatus de elegibilidad conocida, ya sea elegible o no elegible. Más aun las viviendas pueden ser catalogadas como particulares ocupadas o no, esto ya que los listados de direcciones de las viviendas se encuentran definidos en base a la información censal o de los procesos de actualización de los listados de viviendas. Sin embargo, es posible que durante el trabajo de campo de la encuesta no se pueda comprobar el estatus de algunas viviendas, debido al impedimento de acceso (por clima, seguridad, entre otros) a las viviendas o simplemente la muestra no fuera gestionada oportunamente, quedando clasificadas en el estado de “Elegibilidad desconocida”. Como medida de mitigación de esta situación, se realiza un ajuste a los factores de expansión de vivienda con la finalidad de redistribuir los pesos de dichas unidades de forma proporcional entre las unidades elegibles y las no elegibles.

Así, el ponderador de ajuste es definido mediante la siguiente expansión:

$$\hat{R}_{h,know} = \frac{\sum_{i \in \Omega_h} \sum_{j \in \Theta_i} w_{hij}}{\sum_{i \in \Omega_h} \sum_{j \in \Theta_{i,know}} w_{hij}} \quad (20)$$

Donde,

- Θ_i : Conjunto de viviendas j seleccionadas en el conglomerado i .
- $\Theta_{i,know}$: Conjunto de viviendas j seleccionadas en el conglomerado i y clasificadas como elegibles o no elegibles (elegibilidad conocida).
- Ω_h : Conjunto de conglomerados i seleccionados en el estrato h .
- w_{hij} : Factor de expansión de selección de la vivienda j dentro del conglomerado i en el estrato h .

Así, $\hat{R}_{h,now}$ es calculado al interior de cada estrato de muestreo a partir del cociente entre la estimación de viviendas totales (elegibles, no elegibles y elegibilidad desconocida) y la estimación del total de viviendas de elegibilidad conocida (elegibles y no elegibles).

Luego, se calcula el factor de expansión de selección de vivienda ajustado por elegibilidad desconocida (w'_{hij}), como:

$$w'_{hij} = \hat{R}_{h,now} \cdot w_{hij} \quad (21)$$

b) Ajuste por no elegibilidad

La población objetivo de la ENE está conformada por aquellas personas de quince años o más que residen habitualmente en las viviendas particulares ocupadas. Sin embargo, pese a los esfuerzos por mantener el marco de muestreo lo más actualizado posible, siempre existe la posibilidad de incluir en la muestra viviendas que no contienen a la población objetivo, debido a errores en el proceso de actualización o por cambios en el uso de la vivienda.

Este ajuste se implementa asignando un valor nulo (*missing value*) en el ponderador de selección de viviendas corregido por elegibilidad desconocida (w'_{hij}) a las viviendas clasificadas como no elegibles.

Este ajuste lo podemos describir de la siguiente forma:

$$R_{hij,eleg} = \begin{cases} 1 & , \quad \text{Si la vivienda } j \text{ del conglomerado } i \text{ en el estrato } h \text{ es elegible} \\ . & , \quad \text{Si la vivienda } j \text{ del conglomerado } i \text{ en el estrato } h \text{ no es elegible} \end{cases} \quad (22)$$

El factor de expansión de viviendas, ajustado por elegibilidad desconocida y por no elegibilidad, queda:

$$w''_{hij} = w'_{hij} \cdot R_{hij,eleg} \quad (23)$$

4.8.3.4 Ajuste por no respuesta

Este ajuste considera solo las viviendas elegibles, es decir, aquellas viviendas particulares ocupadas; y contempla la probabilidad de no respuesta asociada a aquellas viviendas, donde los informantes rechazan responder la encuesta, donde no fue posible establecer contacto o no se pudo llevar a cabo la entrevista por otros motivos⁵¹.

El ajuste por no respuesta busca corregir los pesos para que la encuesta, a partir de las unidades que responden, sea representativa de toda la población. Como resultado, los pesos de las unidades que respondieron a la encuesta se incrementan para compensar el peso de las unidades que no respondieron.

Para ello, se obtiene el ponderador $\hat{R}_{h,R}$ definido como la razón entre la estimación del total de viviendas elegibles en la población, calculada a partir del factor de expansión definido en el apartado 4.8.3.3 - b (corregido por no elegibilidad); y la estimación del total de viviendas de las que se obtiene

⁵¹ En correspondencia con los códigos de disposición final definidos en el instrumento de recolección a partir de enero 2020.

respuesta de sus informantes. Así, el ajuste por no respuesta queda determinado por la siguiente expresión:

$$\hat{R}_{h,R} = \begin{cases} \frac{\sum_{i \in \Omega_h} \sum_{j \in \Theta_{i,eleg}} W''_{hij}}{\sum_{i \in \Omega_h} \sum_{j \in \Theta_{i,eleg,R}} W''_{hij}} & , \quad \text{Si la vivienda } j \text{ es elegible y responde} \\ . & , \quad \text{Si la vivienda } j \text{ es elegible y no responde} \end{cases} \quad (24)$$

Donde,

- $\Theta_{i,eleg}$: Conjunto de viviendas j seleccionadas en el conglomerado i , clasificadas como elegibles.
- $\Theta_{i,eleg,R}$: Conjunto de viviendas j logradas en el conglomerado i , clasificadas como elegibles y que responden la encuesta.
- Ω_h : Conjunto de conglomerados i seleccionados en el estrato h .
- w''_{hij} : Factor de expansión de selección de la vivienda j dentro del conglomerado i en el estrato h , ajustado por no elegibilidad.

Luego se calcula el ponderador de selección de vivienda ajustado por no respuesta (w_{hij}^{NR}), como:

$$w_{hij}^{NR} = \hat{R}_{h,R} \cdot w''_{hij} \quad (25)$$

Este factor de expansión es aplicado a las k personas dentro de la vivienda j en el conglomerado i del estrato h , escribiéndose como w_{hijk}^{NR} .

4.8.3.5 Ajuste a stocks poblacionales (Calibración)

Los ponderadores a nivel de vivienda calculados hasta este punto sirven para llevar dichas unidades a los respectivos stocks poblacionales. Sin embargo, para la ENE, las unidades de análisis son las personas de quince años o más que residen en viviendas particulares ocupadas, por lo que, conocidas las proyecciones de población basadas en los Censos de población y vivienda, se utilizan dichos stocks para corregir posibles desbalances de los totales poblacionales en variables como edad y sexo.

Si el factor de expansión de selección de viviendas ajustado por no respuesta es asignado a cada persona dentro de la vivienda, se obtiene una estimación del total de personas. Sin embargo, este valor difiere de las estimaciones provenientes de otras fuentes, como las proyecciones de población generadas a partir de modelos demográficos. Es por ello que, con el objetivo de ajustar la población de la encuesta a los stocks poblacionales calculados mediante los modelos, se realiza un ajuste de calibración a estos stocks mediante la incorporación de información auxiliar, originando estimaciones aproximadamente insesgadas y con varianza más pequeña que la del estimador de Horvitz y Thompson.

Los métodos de calibración se basan en algoritmos iterativos que ajustan sucesivamente los factores de expansión para que alcancen totales poblacionales de las variables auxiliares, mientras se minimiza la distancia entre los factores originales y los calibrados.

La idea de mantener cerca los ponderadores calibrados es poder asociarles por aproximación cualquier propiedad buena de estimación de los ponderadores iniciales. Por ejemplo, si los ponderadores iniciales producen estimadores insesgados, los nuevos ponderadores calibrados producirán estimadores aproximadamente insesgados.

Los estimadores calibrados presentan propiedades importantes. Por ejemplo, si los totales poblacionales de las variables auxiliares (marginales) se han publicado, la calibración garantiza que las estimaciones de la encuesta sean coherentes con las estimaciones que ya han sido publicadas.

La segunda propiedad es su simplicidad, es decir, los estimadores de calibración son lineales. Esto quiere decir que a cada elemento de la muestra se le asigna un solo peso para generar la estimación de cualquiera de las variables de interés de la encuesta. El cálculo de las estimaciones para totales, medias, razones, entre otros parámetros es sencillo utilizando un *software* estadístico estándar.

Otra propiedad es su flexibilidad para incorporar información auxiliar incluyendo variables continuas, discretas o de ambos tipos al mismo tiempo. Adicionalmente, los estimadores de calibración también pueden, de cierta forma, contrarrestar el sesgo generado por la no respuesta y por la cobertura insuficiente en subpoblaciones debido a la imperfección del marco muestral (IBGE, 2004).

Por su parte, el método multiplicativo o *raking*, asegura factores de expansión positivos, debido a que la función de distancia no se encuentra definida para valores negativos (Vanderhoef, 2001). Este último es el método escogido para la calibración de la ENE⁵².

La función de distancia $G(w^{cal}, w^{NR})$ se define como:

$$G(w^{cal}, w^{NR}) = w^{cal} \log(w^{cal}/w^{NR}) - w^{cal} + w^{NR} \quad (26)$$

Donde:

w^{cal} : Factor de expansión calibrado a nivel de la marginal correspondiente.

w^{NR} : Factor de expansión corregido por no respuesta.

Lo anterior permite generar un factor de expansión calibrado a nivel de la marginal correspondiente, el cual se obtiene mediante la siguiente expresión:

$$w^{cal} = w^{NR} \cdot F(w^{cal}, w^{NR}) \quad (27)$$

Donde:

$F(w^{cal}, w^{NR})$: Función inversa de $\partial G(w^{cal}, w^{NR})/\partial w^{cal}$ también denotada por $g^{-1}(w^{cal}, w^{NR})$.

⁵² Para más información consultar ver Anexo N° 8. Construcción de insumo para la calibración por población extranjera en factores de expansión de la ENE.

Las marginales definidas⁵³ son las siguientes:

- A nivel nacional: según sexo, trece tramos de edad que se encuentran distribuidos de [0-14], [15-19], [20-24], [25-29], [30-34], [35-39], [40-44], [45-49], [50-54], [55-59], [60-64], [65-69] y [70+].
- A nivel nacional: estimación de la población extranjera⁵⁴.
- A nivel regional: Población total según sexo.
- A nivel regional: Población en Edad de Trabajar (PET) según sexo.
- A nivel regional: Población total Área Urbana según sexo.
- Región Metropolitana: Población total para seis grupos de edad por sexo, los cuales se encuentran distribuidos de [15-24], [25-34], [35-44], [45-54], [55-64] y [65+].
- Regiones (sin RM): Población total para 3 grupos de edad que se encuentran distribuidos de [15-34], [35-54] y [55+].

Para la calibración de los ponderadores de la ENE se utilizan las proyecciones de población con base en el Censo 2017, mensualizadas según la metodología descrita en el Anexo N° 5.

Adicionalmente, a las marginales definidas previamente, se agregan marginales para la PET a nivel de provincia y grandes conurbaciones⁵⁵.

Para definir las provincias a calibrar se definieron los siguientes criterios:

- El coeficiente de variación de la Población en Edad de Trabajar (PET) de la provincia, calculado con el ponderador de selección base, debe ser menor al 10% en todas las 117 muestras entre los trimestres de EFM 2010 y SON 2019.
- Se debe considerar a la provincia que incluya la comuna que es capital regional.
- Considerar aquellas provincias con una PET mayor a 100.000 habitantes en octubre de 2019.
- Luego de aplicar esos criterios, se consideraron algunos aspectos adicionales:
- La provincia de Diguillín ingresará a la marginal desde el trimestre ASO de 2018. Se incluye desde esa fecha por tres razones:
 - Provincia que incluye la capital regional.
 - La región de Ñuble entró en vigor en septiembre del 2018.

⁵³ La justificación de incorporar información poblacional de varios tramos de edad e información sobre el total (*stock*) de la población extranjera se detalla en (INE, 2019d).

⁵⁴ Revisar el documento de trabajo “Fundamentos de la nueva metodología de calibración de los factores de expansión de la Encuesta Nacional de Empleo”, (INE, 2019d)

⁵⁵ Si bien se incorporaron de forma retrospectiva estimaciones para los nuevos dominios de estudios a nivel provincial y grandes conurbaciones en la ENE, a raíz de los efectos de la pandemia del COVID-19 en nuestro país y específicamente en la recolección de la encuesta, desde el trimestre móvil enero-marzo de 2020 hasta mayo-julio de 2021 se interrumpió la difusión de estas estimaciones ante posibles sesgos provocados, por lo que no se cuenta con esa información, retomando la publicación a modo de marcha blanca en el trimestre móvil junio-agosto de 2021. Para mayor información ver notas técnicas publicadas en la página web www.ine.cl.

- Esta provincia no cumplía el requisito de CV (<10%) previo a 2018.
- La provincia de Arica, aunque cumple los tres requisitos, no ingresa a la marginal, debido a que su complemento en la región, la provincia de Parinacota posee una población muy pequeña y desde el año 2016 no forma parte de la muestra.
- Para observar el listado de las marginales provinciales ver el Anexo N° 6.
- Por otro lado, se consideran cuatro grandes conurbaciones:
- La Serena-Coquimbo, ingresa desde el trimestre EFM de 2010.
- Gran Valparaíso, ingresa desde el trimestre EFM de 2010.
- Gran Santiago, ingresa desde el trimestre EFM de 2010.
- Gran Concepción, ingresa desde el trimestre EFM de 2010.

Inicialmente se había incorporado la gran conurbación de Antofagasta, pero posteriormente se decidió excluir esta calibración debido a que esta conurbación corresponde sólo a la comuna de Antofagasta y en esta comuna se concentra más del 95% de la población de la provincia de Antofagasta, generando así el riesgo de que las restantes comunas de la provincia tengan baja presencia en la muestra y por lo tanto las muestras a nivel de provincia y de comuna tienden a coincidir.

4.9 Medidas de precisión

Actualmente los resultados de las encuestas son empleados, entre otras bondades, en la toma de decisiones de políticas públicas, por lo que, evidentemente los objetivos analíticos van mucho más allá de solo presentar resultados concernientes a la estimación de un parámetro, sino que cobra relevancia que las estimaciones incluyan medidas de precisión o exactitud de las estimaciones.

4.9.1 Varianza de los estimadores

Una de las medidas claves de precisión en las encuestas por muestreo es la varianza de muestreo, un indicador de la variabilidad que se produce al elegir una muestra, en lugar de confeccionar una lista de toda la población. Aunque existen otras medidas de precisión, cabe mencionar que guardan relación algebraica entre sí, por lo que es posible deducir la expresión de cualquiera de ellas a partir de las otras mediante la aplicación simple de operaciones matemáticas.

Los métodos exactos de estimación de la varianza son la mejor forma de estimar la varianza, sin embargo, dado que el nuevo diseño muestral de la ENE al igual que el predecesor corresponde a un diseño complejo, se imposibilita la aplicación del cálculo de la varianza por medio de los métodos exactos.

En consecuencia, la estimación de la varianza de las principales estimaciones de la ENE se obtiene a partir del Método del Conglomerado Último. Esto considerando que la mayor contribución a la varianza de un estimador, en un diseño multietápico, se origina a partir de las Unidades Primarias de Muestreo (UPM). El Método del Conglomerado Último permite simplificar un diseño complejo multietápico en un caso especial del muestreo de conglomerados sin submuestreo, es decir en una etapa. La denotación de conglomerado último refiere al conjunto de unidades muestrales de última etapa que pertenecen a una unidad primaria de muestreo, independiente del número de etapas de selección que se efectúan al interior de la primera etapa de selección.

Este método permite aproximar la varianza de los estimadores considerando también la estratificación asociada al diseño muestral.

Otro elemento importante a considerar para la estimación de la varianza en los estimadores de la ENE, es que los factores de expansión han sido calibrados mediante el método *Raking* (ver apartado [“4.8.3.5 Ajuste a stocks poblacionales \(Calibración\)”](#)) por lo que, de acuerdo con (Deville & Särndal, 1992) la estimación de la varianza se obtiene empleando aproximaciones basadas en el estimador de regresión generalizado (GREG)⁵⁶, que emplea la función de distancia chi-cuadrado.

Cabe mencionar que el GREG forma parte de una familia de estimadores de calibración, en donde se encuentra el método multiplicativo o *Raking* y como lo demuestran (Deville & Särndal, 1992), bajo condiciones de regularidad, constituyen estimaciones asintóticamente equivalentes, por lo que se puede emplear la estimación de la varianza del GREG. La estimación de la varianza a través del GREG tiene la bondad que constituye la expresión menos compleja de la familia de distancias mencionadas, y aunque se han realizado esfuerzos por derivar expresiones para los demás miembros de la familia, estas han resultado muy complicadas, por lo que en la práctica se aproxima mediante la expresión asociada al GREG (Deville & Särndal, 1992).

⁵⁶ Para más información ver (Särndal, Swensson, & Wretman, 1991, pág. 235).

Las fórmulas para la estimación de la varianza mediante el GREG se encuentran implementadas en el paquete *Survey* del *software R*, por lo que se ha operacionalizado el cálculo de los estimadores de la ENE en este *software*, siendo necesario especificar las características de diseño muestral complejo y la generación de los pesos calibrados por el método *Raking*⁵⁷.

4.9.2 Errores absolutos y relativos

Como se ha mencionado, una encuesta por muestreo está sujeta a los errores de muestreo, que se presentan debido a que la investigación estadística se hace solo sobre una parte representativa de la población y no sobre su totalidad. Sin embargo, existen distintas formas de medir estos errores para las variables de interés como, por ejemplo, la varianza, la desviación estándar o algunas transformaciones de ellas, como el error absoluto, el coeficiente de variación, el error relativo, entre otros. En la ENE se utilizan principalmente, las tres últimas medidas de error, que ayudan a determinar la precisión estadística de las estimaciones.

4.9.2.1 Error absoluto e intervalo de confianza

Para el cálculo del intervalo de confianza de 95% para la estimación de la variable Y , se utiliza el error absoluto (error de estimación que está expresado en las mismas unidades de la variable en estudio⁵⁸).

El error absoluto permite observar cuánto es la diferencia máxima esperada entre el valor estimado y el verdadero parámetro poblacional, donde dicha estimación se puede realizar a distintos niveles de confianza, siendo la más común la del 95%. A partir de este error se construye el intervalo de confianza de la variable de interés Y .

El error absoluto se obtiene del producto entre la desviación estándar ($S(\hat{Y})$) de la estimación y el valor crítico asociado al percentil de la distribución *t-student* con v grados de libertad ($t_{1-\frac{\alpha}{2},v}$), donde el primer argumento hace referencia al nivel de confiabilidad con el cual se desea hacer inferencia (95%), mientras que el segundo argumento corresponde a los v grados de libertad con los que se calcula el valor crítico⁵⁹. Su expresión algebraica está dada por:

$$E_{abs}(\hat{Y}) = t_{1-\frac{\alpha}{2},v} \cdot S(\hat{Y}) \quad (28)$$

Alternativamente, en términos probabilísticos, se puede expresar como:

$$\Pr(|\hat{Y} - Y| \leq E_{abs}(\hat{Y})) = 1 - \alpha \quad (29)$$

Donde \hat{Y} e Y son los valores estimado y verdadero en la población, respectivamente.

⁵⁷ El desarrollo de la aproximación de la varianza para el método *Raking* en base a la varianza del estimador GREG se encuentra desarrollada en (Särndal, Swensson, & Wretman, 1991).

⁵⁸ Cuando la variable Y es un porcentaje o tasa, el error absoluto se expresa en puntos porcentuales.

⁵⁹ Los grados de libertad corresponden a la diferencia entre el número de conglomerados no vacíos (con presencia de la variable de interés) y el número de estratos no vacíos.

Por su parte, el intervalo de confianza es el intervalo estimado que incluye el promedio de las estimaciones de todas las muestras posibles con una probabilidad del 95%. Este intervalo tiene un límite inferior y un límite superior dentro de los cuales se espera que se encuentre el valor real o poblacional con el nivel de confianza del 95%.

El intervalo de confianza es construido a través del error absoluto, considerando como el valor mínimo del parámetro poblacional el parámetro estimado menos el error absoluto, y como valor máximo el parámetro estimado más el error absoluto. Matemáticamente, el intervalo de confianza es igual a:

$$IC(\hat{y})_{95\%} = (\hat{y} - E_{abs}(\hat{y}); \hat{y} + E_{abs}(\hat{y})) = \left(\hat{y} - t_{1-\frac{\alpha}{2}, v} \cdot S(\hat{y}); \hat{y} + t_{1-\frac{\alpha}{2}, v} \cdot S(\hat{y}) \right) \quad (30)$$

4.9.2.2 Error relativo y coeficiente de variación

Adicionalmente existen otros indicadores, que dan cuenta de los errores de muestreo, en términos relativos, el error relativo y el coeficiente de variación.

El error relativo, expresa la magnitud del error de muestreo en términos porcentuales, reflejando la máxima variación porcentual que se puede esperar entre el valor estimado y el valor poblacional con una probabilidad de $(1 - \alpha) \cdot 100\%$, y se obtiene de la siguiente forma:

$$E_{rel}(\hat{y}) = t_{1-\frac{\alpha}{2}, v} \cdot \frac{\hat{\sigma}_{\hat{y}}}{\hat{y}} = t_{1-\frac{\alpha}{2}, v} \cdot CV_{\hat{y}} \quad (31)$$

Donde, $\frac{\hat{\sigma}_{\hat{y}}}{\hat{y}}$ representa el coeficiente de variación $CV_{\hat{y}}$ asociado a la estimación de la variable \hat{y} .

La gran utilidad tanto del error relativo como del coeficiente de variación es que pueden ser expresados en porcentajes, lo que los distingue de otras mediciones del error, al permitir comparar entre diferentes estimaciones, soslayando las diferencias entre unidades de medida.

4.9.3 Representatividad de las estimaciones

Basados en los Lineamientos para el Uso de Medidas de Precisión en Encuestas de Hogares, el que una estimación sea o no aceptable depende del enfoque planteado en el fluograma para publicar o suprimir estimaciones (INE, 2020), el cual contempla los siguientes elementos:

4.9.3.1 Elementos relacionados con el tamaño muestral

Tamaño muestral, sin pérdida de generalidad, al señalar el término “tamaño muestral” se hará referencia al total de unidades de análisis consideradas para obtener las estimaciones. El umbral de aceptación considera un tamaño muestral mayor o igual a 60 unidades.

Grados de libertad, los grados de libertad (gl) se definen como el número de observaciones independientes que son libres de variar (número de individuos en los datos) menos el número de parámetros estimados (número de relaciones impuestas a los datos), lo que en un diseño muestral complejo se obtiene como la diferencia entre el número de conglomerados no vacíos (con presencia de la variable de interés) y el número de estratos no vacíos. Además, según (Heeringa, West, &

Berglund, 2010), éstos también, dan cuenta de la precisión con la que se están estimando los parámetros a través de la muestra, por lo que, a mayores grados de libertad, las estimaciones serán más precisa, y viceversa. El umbral de aceptación contempla que la estimación evaluada tenga nueve o más grados de libertad.

Si el nivel de desagregación evaluado no cumple con el criterio del tamaño muestral o con los grados de libertad, entonces se debe suprimir la estimación para dicho nivel.

4.9.3.2 Elementos relacionados con la precisión de la estimación

Si el estimador es proporción será evaluado por medio del error estándar (ee), los umbrales quedaran definidos por la siguiente función:

$$\text{Máximo ee tolerable} = \begin{cases} \sqrt[3]{p^2}/9, & \text{si } p \leq 0,5 \\ \sqrt[3]{(1-p)^2}/9, & \text{si } p > 0,5 \end{cases}$$

Si el nivel de desagregación evaluado cumple con los criterios del tamaño muestral y de los grados de libertad, pero no cumple los criterios anteriores, se publica la estimación advirtiendo que se trata de una estimación poco fiable.

Si el estimador no es proporción será evaluado a través del coeficiente de variación, y se considera:

$$\text{Máximo CV tolerable} = \begin{cases} CV \leq 0,15 \rightarrow \text{Aceptable} \\ 0,15 < CV \leq 0,30 \rightarrow \text{Estimación poco fiable} \end{cases}$$

Si el nivel de desagregación evaluado cumple con los criterios del tamaño muestral y de los grados de libertad, pero presenta un coeficiente de variación mayor a 0,30 se debe suprimir la estimación para dicho nivel.

Posteriormente, la publicación del tabulado depende del número de estimaciones que se han suprimido o generado con alerta de poca fiabilidad, en este caso, se recomienda que, si estas estimaciones suman más de 50% de las que forman el tabulado, este último no debe ser publicado.

4.9.3.3 Tasa de respuesta y tasa de logro objetivo de la recolección

La tasa de respuesta 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.

$$\text{Tasa de respuesta} = \frac{E}{E + R + NC + ED}$$

Donde:

E: número de unidades entrevistadas

R: número de unidades rechazadas

NC: número de unidades no contactadas

O: número de unidades no entrevistada por otros motivos

ED: número de unidades de elegibilidad desconocida

La tasa de logro objetivo, es un complemento de la tasa de respuesta, ya que permite evaluar la calidad del proceso de recolección en relación a los niveles de precisión establecidos en el diseño muestral de cada estudio.

Describe el número de unidades muestrales a encuestar en un período determinado y donde se obtiene una entrevista (completa o parcial), en relación a 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).

Por lo general, este indicador se emplea para estudios donde el diseño muestral considera un número de unidades como sobremuestra con el propósito de mitigar la falta de respuesta.

$$\text{Tasa de logro objetivo} = \left[\frac{E}{Mo} \right] * 100$$

Donde,

E: número de unidades con entrevista completa

Mo: Total de unidades muestra objetivo en el período t

Las viviendas de la muestra pueden estar en diferentes estados según su nivel de logro. A continuación se describen las principales categorías:

- **Entrevistadas (E):** viviendas u hogares contenidos en la muestra en las que fue posible hacer contacto y aplicar la encuesta, obteniendo información suficiente para el estudio, de acuerdo a la evaluación de suficiencia de datos.
- **Rechazo (R):** viviendas u hogares contenidos en la muestra en los que fue posible hacer contacto, sin embargo, un informante idóneo no acepta participar en el estudio desde un inicio, no existiendo recolección de información de estas unidades.
- **Otras no entrevistadas (O):** viviendas u hogares contenidos en la muestra en los que fue posible interactuar presencialmente con un informante idóneo, pero no fue posible aplicar la encuesta por razones distintas al rechazo, por ejemplo, por problemas de comunicación con el informante debido al idioma; porque el informante estaba impedido física y/o mentalmente para contestar o por muerte del informante o de algún familiar, entre otros motivos.
- **Break-off:** «Se refiere a unidades en las que fue posible hacer contacto y realizar la entrevista, sin embargo, no reportan la información suficiente requerida para ser consideradas en la categoría entrevistadas, por lo que no son incluidas en los resultados finales de la encuesta.

Esta categoría constituye una clasificación generada post-recolección de datos, por lo que no es labor del encuestador distinguir estos casos al momento de realizar el trabajo de recolección de datos. La categoría de clasificación “break-off”, se asigna en la fase de procesamiento, al momento de determinar la cantidad de información recogida en cada unidad y decidir si la encuesta cumple con la información suficiente para ser incluida en la base de datos final» INE (2021). «Guía para la implementación de indicadores de rendimiento y calidad en el monitoreo del proceso de recolección de datos (documento interno)».

4.9.3.4 Tasas de supervisión

La etapa de supervisión en la ENE cumple el rol de garantizar la correcta ejecución del trabajo de recolección de la encuesta, con el objetivo de confirmar que la información recolectada es fidedigna y además identificar problemáticas, así como errores u omisiones que afecten el cumplimiento de los objetivos del estudio.

Respecto a las medidas de precisión relacionadas con la supervisión de la recolección, se establecieron las siguientes:

- **Tasa de supervisión:** Es el porcentaje de unidades supervisadas respecto a la muestra teórica del periodo.

$$\text{Tasa de supervisión} = \frac{\text{Total de unidades supervisadas}}{\text{Muestra teórica}} * 100$$

- **Tasa de supervisión indirecta re entrevista:** Es el porcentaje de unidades con entrevista completa supervisada respecto al total de unidades entrevistadas para el periodo.

$$\text{Tasa de supervisión re entrevista} = \frac{\text{Unidades con entrevista completa supervisadas}}{\text{Total de unidades entrevistadas}} * 100$$

- **Tasa de corrección consultas recuperaciones:** Es el porcentaje entre el número de consultas corregidas más el número de consultas por región respecto al número de consultas recibidas a nivel nacional durante el período.

$$\text{Tasa corrección} = \left(\frac{N \text{ consulta corregida} + n \text{ consulta consultada por región}}{N^o \text{ de consultas recibidas en el periodo a nivel nacional}} \right) * 100$$

- **Tasa de supervisión indirecta unidades no entrevistadas:** Es el porcentaje de unidades no entrevistadas supervisadas respecto del total de unidades no entrevistadas en el periodo.

$$\text{Tasa de supervisión unidades no entrevistadas} = \left(\frac{\text{Unidades no entrevistadas supervisadas}}{\text{Total de unidades no entrevistadas}} \right) * 100$$

Se abordará más en detalle el trabajo de supervisión de la ENE, en el próximo capítulo.

4.10 Estimación de parámetros

4.10.1 Estimación de Parámetros

La ENE incluye en sus diferentes productos de difusión la estimación de diferentes parámetros, que son principalmente niveles y razón. Al generar la estimación para cada parámetro, se estiman también sus correspondientes medidas de precisión, tales como, el error estándar, el coeficiente de variación, los grados de libertad y los intervalos de confianza, detallados en el apartado anterior de este documento “[4.9 Medidas de precisión](#)”. Cabe señalar que todas las estimaciones de los parámetros que se publican incluyen los criterios de calidad estadística establecidos por el INE para evaluar estimaciones en encuestas de hogares⁶⁰.

La nomenclatura utilizada para los estimadores es:

w_{hijk}^{cal} : Ponderador para la persona k ajustado por proyecciones de población y asociado a la UPM i en el estrato h .

$P_{hijk} = 1$, para toda persona k de la vivienda j de la UPM i en el estrato h .

P_{hijk_c} : Persona k de la vivienda j de la UPM i en el estrato h , que cumplen con la característica C .

Es decir: $P_{hijk_c} = \begin{cases} 1 & , si la persona cumple con la característica C \\ 0 & , en otro caso \end{cases}$

Y_{hijk} = Valor de la variable Y medida en la persona k de la vivienda j de la UPM i en el estrato h .

Y_{hijk_c} = Valor de la variable Y medida en la persona k de la vivienda j de la UPM i en el estrato h que cumple con la característica C .

Es decir: $Y_{hijk_c} = \begin{cases} Y_{hijk} & , si la persona cumple con la característica C \\ 0 & , en otro caso \end{cases}$

También puede ser expresado como:

$$Y_{hijk_c} = Y_{hijk} \cdot P_{hijk_c} = \begin{cases} Y_{hijk} \cdot 1 = Y_{hijk} & , si la persona cumple \\ & con la característica C \\ Y_{hijk} \cdot 0 = 0 & , en otro caso \end{cases}$$

Nota: recordar que $\hat{P}_h = P_h$, es decir, por construcción, la estimación de personas dentro de cada estrato y nivel, coincide con la proyección de población.

⁶⁰ Departamento de Metodología e Innovación Estadística . (Febrero de 2020). Estándar para la evaluación de la calidad de las estimaciones en encuestas . Obtenido de Instituto Nacional de Estadísticas : <https://www.ine.cl/docs/default-source/institucionalidad/buenas-pr%C3%A1cticas/clasificaciones-y-estándares/est%C3%A1ndar-evaluaci%C3%B3n-de-calidad-de-estimaciones-publicaci%C3%B3n-27022020.pdf>

4.10.1.1 Estimación de niveles

Dentro de los parámetros que se calculan está la estimación de niveles de diferentes categorías, las cuales se utilizan para la estimación de razón.

Para la estimación de niveles totales se emplea el estimador de Horvitz-Thompson, el cual tiene las características de ser insesgado y consistente. Estas estimaciones se realizan con los factores de expansión respectivos y a través de *software* estadísticos como *R* o *STATA*. Estos indicadores son:

1. **Población en edad de trabajar (PET):** es población actualmente residente en el país de quince años y más.
2. **Fuerza de trabajo (FT):** son personas en edad de trabajar que durante la semana de referencia cumplen los requisitos para ser incluidas en la categoría de personas ocupadas o desocupadas.
3. **Personas ocupadas (O):** son todas las personas en edad de trabajar, que durante la semana de referencia dedicaron al menos una hora a alguna actividad para producir bienes o servicios a cambio de una remuneración o beneficios.
4. **Personas desocupadas (DO):** corresponde todas las personas en edad de trabajar que no estaban ocupadas durante la semana de referencia, que habían llevado a cabo actividades de búsqueda de un puesto de trabajo durante las últimas cuatro semanas (incluyendo la de referencia) y que estaban disponibles para trabajar en las próximas dos semanas (posteriores a la de referencia).
5. **Fuera de la fuerza de trabajo (FFT):** son las personas en edad de trabajar que durante la semana de referencia no son incluidas en la categoría de personas ocupadas o desocupadas.
6. **Personas Iniciadoras disponibles (ID):** son todas las personas fuera de la fuerza de trabajo que esgrimen no haber buscado trabajo en las últimas cuatro semanas debido a que iniciarán pronto una actividad laboral y que, al mismo tiempo, declaran disponibilidad.
7. **Población ocupada que trabaja a tiempo parcial involuntario (TPI):** corresponden a aquellas personas ocupadas que trabajan habitualmente dos tercios de la jornada completa establecida y además señalan estar disponibles para trabajar más horas ya sea de manera inmediata o dentro de los próximos quince días.
8. **Personas ocupadas que buscaron empleo (OBE):** corresponden a personas ocupadas que declararon haber buscado empleo en la semana de referencia.
9. **Fuerza de trabajo potencial – Personas inactivas potencialmente activas (FTP):** corresponde a aquellas personas que se encuentran fuera de la FT, sin embargo declaran alguna intención de entrar al mercado laboral, a través de la búsqueda de empleo o la disponibilidad para trabajar.
10. **Fuerza de trabajo ampliada (FTA):** se define como la suma de la fuerza de trabajo más las personas iniciadoras disponibles y la fuerza de trabajo potencial (personas inactivas potencialmente activas).

Las siguientes expresiones representan la fórmula generalizada empleada para obtener estimaciones del total de una variable según el nivel de estimación requerido.

- Total estimado de la variable \mathbf{Y} en el estrato \mathbf{h} .

$$\hat{Y}_h = \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk} \quad (328)$$

- Total estimado de la variable \mathbf{Y} en el nivel de estimación requerido.

$$\hat{Y} = \sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk} = \sum_h \hat{Y}_h \quad (33)$$

La estimación del total \hat{Y} en el nivel de estimación requerido se obtiene a partir de la suma de los totales de cada uno de los estratos que conforman dicho nivel de estimación.

- Total estimado de la variable \mathbf{Y} en personas con la característica \mathbf{C} en el estrato \mathbf{h} .

$$\hat{Y}_{h_c} = \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk_c} \quad (34)$$

- Total estimado de la variable \mathbf{Y} en personas con la característica \mathbf{C} en el nivel de estimación requerido.

$$\hat{Y}_c = \sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk_c} = \sum_h \hat{Y}_{h_c} \quad (35)$$

4.10.1.2 Estimación de indicadores de razón

Una razón matemática, es un vínculo entre dos magnitudes que son comparables entre sí. Es el resultado obtenido cuando una de las magnitudes o cantidades se divide por otra, y, por tanto, pueden expresarse como fracciones o como números decimales. Tanto el numerador como el denominador pueden ser variables cuantitativas o cualitativas.

El estimador de razón es un estimador aproximadamente insesgado ya que es el cociente de dos variables aleatorias y en general, la esperanza del cociente de dos variables aleatorias no es igual que el cociente de las esperanzas, pero el sesgo es despreciable en muestras grandes, véase (Cochran, 1998).

Los indicadores de razón que se calculan en cada periodo son los siguientes:

1. **Tasa de desocupación:** expresa el porcentaje de personas desocupadas dentro de la fuerza de trabajo.

$$Tasa\ de\ desocupación = \frac{desocupados/as}{fuerza\ de\ trabajo} * 100$$

2. **Tasa de ocupación:** expresa el porcentaje de personas ocupadas dentro de la población en edad para trabajar.

$$Tasa\ de\ ocupación = \frac{ocupados/as}{población\ en\ edad\ de\ trabajar} * 100$$

3. **Tasa de participación:** es el porcentaje de personas que participa en el mercado laboral o componen la fuerza de trabajo (población ocupada y desocupada) dentro del total de personas que tienen edad para trabajar.

$$Tasa\ de\ participación = \frac{ocupados/as + desocupados/as}{población\ en\ edad\ de\ trabajar} * 100$$

4. **Tasa de desocupación con iniciadores disponibles (SU1):** es la tasa de desocupación incluyendo a las personas iniciadoras disponibles tanto en el numerador como en el denominador, de modo que pueda expresar una tasa ampliada agregando a la población que si bien no está aún en la fuerza de trabajo o se encuentra inactiva ha manifestado disponibilidad para trabajar y declarado que empezará una actividad laboral en el futuro próximo.

$$SU1 = \frac{desocupados/as + iniciadores/as disponibles}{fuerza\ de\ trabajo + iniciadores/as disponibles} * 100$$

5. **Tasa combinada de desocupación y tiempo parcial involuntario (SU2):** Es la tasa de desocupación con las personas iniciadoras disponibles que además de incluirlas, en ambas partes de la fracción, también lo hace con las personas ocupadas a tiempo parcial involuntario en el numerador, tal que se evidencie la desocupación de tiempo completo.

$$SU2 = \frac{desocupados/as + iniciadores/as disponibles + ocupados/as a tiempo parcial involuntario}{fuerza\ de\ trabajo + iniciadores/disponibles} * 100$$

6. **Tasa combinada de desocupación y fuerza de trabajo potencial (SU3):** es la tasa de desocupación con la fuerza de trabajo ampliada, compuesta por la fuerza de trabajo más las personas iniciadoras disponibles y la fuerza de trabajo potencial (personas inactivas potencialmente activas), dando cuenta de la desocupación de quienes participan del mercado del trabajo y quienes potencialmente podrían entrar a participar en el corto plazo.

$$SU3 = \frac{desocupados/as + iniciadores/a disponibles + inactivos/as potencialmente activos/as}{fuerza\ de\ trabajo + iniciadores\ disponibles/as + inactivos/as potencialmente activos/as} * 100$$

7. **Tasa de presión laboral:** es la razón entre las personas desocupadas más las ocupadas que buscan empleo y las personas iniciadoras disponibles, sobre la fuerza de trabajo más las personas iniciadoras disponibles. Evidencia el porcentaje de personas que presionan el mercado laboral dentro de la fuerza de trabajo, pues incluye a quienes declararon disponibilidad y la posibilidad de empezar pronto un trabajo.

$$Tasa\ presión\ laboral = \frac{desocupados/as + iniciadores/as disponibles + ocupados/as que buscan\ empleo}{fuerza\ de\ trabajo + iniciadores/as disponibles} * 100$$

8. **Tasa de desocupación ajustada estacionalmente:** es la tasa de desocupación que elimina los efectos de los factores exógenos estacionales de naturaleza no económica que influyen en su comportamiento coyuntural.
9. **Tasa de ocupación informal:** expresa el porcentaje de personas que están ocupadas en actividades informales dentro del total de población ocupada.

$$\text{Tasa de ocupación informal} = \frac{\text{personas ocupadas informales}}{\text{población ocupada total}} * 100$$

10. **Tasa de ocupación en el sector informal:** Es el porcentaje de personas que trabajan en unidades económicas informales dentro del total de personas ocupadas.

$$\text{Tasa de ocupación en el sector informal} = \frac{\text{personas ocupadas en unidades económicas informales}}{\text{población ocupada total}} * 100$$

La siguiente expresión representa la fórmula generalizada empleada para obtener estimaciones de razón entre dos variables según el nivel de estimación requerido.

- **Estimador de razón entre las variables X e Y**

La razón entre las variables X e Y viene dada por:

$$\hat{R}_{h-y/x} = \frac{\hat{Y}_h}{\hat{X}_h} = \frac{\sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk}}{\sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot X_{hijk}} \quad (296)$$

El estimador de la razón entre las variables X e Y equivale al cociente obtenido al dividir los estimadores del total de dichas variables.

De forma similar obtenemos el estimador de razón en el nivel de estimación requerido, como una razón combinada.

$$\hat{R}_{y/x} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk}}{\sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot X_{hijk}} \quad (37)$$

4.10.1.3 Estimación de indicadores complementarios

De forma complementaria a los principales indicadores de ocupación y desocupación mencionados anteriormente, se estiman otro tipo de indicadores, cuya evolución es relevante analizar para dar cuenta de la relación de la oferta de trabajo con la economía. Estos indicadores permiten, además, realizar comparaciones internacionales.

1. **Rama de actividad económica:** Distribución de las incidencias de las distintas ramas de actividad económica, de acuerdo con CAENES, en la población ocupada.
2. **Categoría ocupacional:** Distribución de las incidencias de las categorías de la ocupación en la población ocupada. Las categorías incluidas son las definidas por la CISE 93.

3. **Grupo ocupacional:** Distribución de las incidencias de las categorías de grupo ocupacional, de acuerdo con la CIUO 08.CL, en la población ocupada.
4. **Horas habituales y efectivas:** Distribución del volumen de trabajo en la ocupación, medido en horas efectivas trabajadas por la población ocupada. Mientras que a partir de las horas habitualmente trabajadas, se realiza una separación entre trabajo de tiempo parcial voluntario e involuntario, con objeto de evidenciar la subutilización de la fuerza de trabajo.

5 Metodología de Recolección de datos

5.1 Técnica y métodos de recolección de los datos

En este capítulo se presenta la descripción de procedimientos y estrategias para llevar a cabo la recolección de datos de la Encuesta Nacional de Empleo (ENE). En primer lugar, se da a conocer la conformación y selección de la muestra de la ENE, incluyendo el proceso de distribución operativa; en segundo término, se expone los períodos de recolección; además, se da a conocer la organización del trabajo de campo, indicadores de promedio de carga de trabajo asignada y número de entrevistas promedio por encuestador/a – recolector/a.

Lo anterior, se describe considerando las modificaciones en la metodología de recolección de datos dados por el cambio en el marco muestral, implementación de la nueva hoja de ruta, la incorporación de la respuesta parcial en las preguntas con las opciones no sabe y no responde como alternativas en el cuestionario central e instrumento de captura, y también considerando los efectos de la contingencia sanitaria por la pandemia Covid-19.

5.1.1 La muestra y el proceso de distribución operativa

Como se comenta en el capítulo anterior, desde enero de 2020 la muestra de la ENE es seleccionada desde el Marco Muestral de Viviendas, elaborado a partir de los resultados del Censo de Población y Vivienda 2017, y los listados de viviendas asociados al Precio 2016 (en lo sucesivo, MMV 2017).

El MMV 2017 está constituido por unidades primarias de muestreo (UPM), correspondientes a áreas geográficas, homogéneas en términos del número de “viviendas marco”, definidas como viviendas particulares con moradores presentes o ausentes observadas en el Censo de Población y Vivienda de 2017, haciendo distinción del área urbana y rural. Para el área urbana, el tamaño medio es de 200 viviendas, con un intervalo del 20%, es decir, la cantidad de viviendas por UPM se encuentra en el rango [160-240]. Por otra parte, el área rural tiene un tamaño medio de 90 viviendas y un intervalo del 20%, quedando su cantidad de viviendas por UPM en el intervalo [70-110].

El número de unidades a entrevistar por UPM son diez en estratos urbanos y rurales, con excepción de los estratos urbanos de las regiones de Atacama, Aysén y Magallanes, donde se determinó levantar ocho viviendas por UPM.

Como se señala, las UPM pueden tener en promedio 200 viviendas, seleccionando diez posibles viviendas a entrevistar por UPM, por lo que, a diferencia de la antigua selección por manzana (en zonas urbanas) y sección (en zonas rurales), la distribución geográfica entre una vivienda y otra, puede ser muy extensa, las UPM también cuentan con manzanas, pero la selección de viviendas no estaba anclada a esta unidad. Este incremento en la distancia entre una vivienda y otra, implica un impacto en la cantidad de intentos de contactos para cada vivienda, sobre todo en el ámbito rural.

En la sección [“5.1.3 Período de recolección y distribución operativa de la muestra”](#) se profundiza en la distribución operativa de la muestra.

5.1.2 Selección de la muestra

En cuanto a la selección de la muestra, esta obedece a un proceso probabilístico, estratificado y bietápico. Los estratos de muestreo corresponden a la combinación “Estrato geográfico – Estrato Socioeconómico”. Luego de seleccionar las UPM (con una selección aleatoria sistemática con

probabilidad proporcional al tamaño), al interior de cada una de estas se procede a seleccionar las viviendas, que corresponden a la Unidad Secundaria de Muestreo (USM). La selección de viviendas dentro de cada UPM seleccionada, tanto urbana como rural, se realiza de manera independiente, en forma sistemática y, con igual probabilidad de selección.

Antes de la contingencia sanitaria, el sistema de rotación de la muestra seleccionada consistía en sustituir cada trimestre una sexta parte del total de viviendas; por lo tanto, cada trimestre, “salen” de la muestra aquellas viviendas que ya han sido consideradas en seis rondas (siendo entrevistadas o no) y “entran” otras sustituyendo a estas últimas. Las viviendas seleccionadas se entrevistan en seis ocasiones o rondas, a la primera visita se le denomina “apertura”, a lo largo de 18 meses, con intervalos de tres meses entre cada entrevista, con el fin de actualizar la información recabada.

En cada ronda, la vivienda y/u hogar puede ser visitada más de una vez para obtener la información de todas las personas pertenecientes a la misma, estas visitas se transforman en intentos de contactos telefónicos para aquellas viviendas que cuentan con un número de contacto, mientras que siguen siendo visitadas presencialmente aquellas en las que no existe tal. A raíz de la pandemia se produjo una desactualización en el marco muestral, lo que provocó que el sistema de rotación sufriera cambios debido a la falta de información de contacto, por lo que cierta proporción de la muestra no pudo ser sustituida y por ende permaneció por más de seis rondas en la muestra. Para mitigar este efecto adverso, el INE generó operativos de recuperación de contacto, en la medida que las restricciones sanitarias y de movilidad lo permitían, sin embargo, la renovación de la muestra ha demorado mucho más de lo que se tenía considerado en un principio.

De esta manera, podemos observar en la Tabla 19 que la submuestra 1 del año 2020 se repite en los meses de enero, abril, julio y octubre, mientras que la submuestra de febrero en mayo, agosto y noviembre y por último la submuestra de marzo se repiten en junio, septiembre y diciembre.

Tabla 19. Distribución de la Muestra total mensual año 2020

Región	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
Tarapacá	345	331	340	345	331	340	345	331	340	345	331	340
Antofagasta	435	459	454	435	459	454	435	459	454	435	459	454
Atacama	486	497	477	486	497	477	486	497	477	486	497	477
Coquimbo	873	901	901	873	901	901	873	901	901	873	901	901
Valparaíso	1849	1846	1855	1849	1846	1855	1849	1846	1855	1849	1846	1855
O'Higgins	891	909	900	891	909	900	891	909	900	891	909	900
Maule	1036	1030	1041	1036	1030	1041	1036	1030	1041	1036	1030	1041
Biobío	1694	1711	1729	1694	1711	1729	1694	1711	1729	1694	1711	1729
La Araucanía	839	851	848	839	851	848	839	851	848	839	851	848
Los Lagos	1144	1183	1178	1144	1183	1178	1144	1183	1178	1144	1183	1178
Aysén	393	361	399	393	361	399	393	361	399	393	361	399
Magallanes	352	342	352	352	342	352	352	342	352	352	342	352
Metropolitana	3068	3093	3084	3068	3093	3084	3068	3093	3084	3068	3093	3084
Los Ríos	692	701	668	692	701	668	692	701	668	692	701	668
Arica y Parinacota	443	430	442	443	430	442	443	430	442	443	430	442
Ñuble	587	582	595	587	582	595	587	582	595	587	582	595
Total Nacional	15127	15227	15263									

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

5.1.3 Periodo de recolección y distribución operativa de la muestra

La muestra está subdivida en submuestras mensuales-trimestrales, es decir, la muestra teórica mensual se repite cada tres meses. Además, las submuestras mensuales se subdividen por semanas de recolección, lo que permite cumplir con el criterio de **semana de referencia**. Este criterio corresponde al periodo en el que están referidos los datos lo que quiere decir que la entrevista se lleva a cabo consultando por la semana inmediatamente anterior a la semana que se realiza la entrevista.

De esta manera, el levantamiento comienza la segunda semana del mes en curso, para obtener información desde la primera semana del mes, terminando la primera semana del mes siguiente, recolectando así los datos la última semana del mes en estudio.

La Figura 10 ejemplifica los periodos de recolección (semanas de levantamiento) y su relación con las semanas de referencia.

Figura 10. Periodos de recolección y semana de referencia

Mes 1				Mes 2				Mes 3			
Semana 1	Semana 2	Semana 3	Semana 4	Semana 1	Semana 2	Semana 3	Semana 4	Semana 1	Semana 2	Semana 3	Semana 4
Semana de referencia	Levantamiento			Semana de referencia	Levantamiento						
	Semana de referencia	Levantamiento			Semana de referencia	Levantamiento					
		Semana de referencia	Levantamiento			Semana de referencia	Levantamiento				
			Semana de referencia	Levantamiento			Semana de referencia	Levantamiento			

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

5.1.4 Organización del Trabajo de Campo

La recolección cuenta con tres actividades principales: planificación, distribución y trabajo de campo, los cuales se desarrollan en un determinado orden y tiempo, para dar cumplimiento tanto a las semanas de levantamiento como a los plazos de apertura y cierre de periodo.

Figura 11. Organización del Trabajo de Campo



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

5.1.4.1 Planificación levantamiento

Para dar inicio a las labores de recolección de datos, la persona encargada de la coordinación realiza la planificación del levantamiento mensual. Antes del comienzo de la submuestra, se utiliza como

referencia las características propias de cada lugar y de la muestra en sí, como la distancia, clima, cantidad de UPM, personal de recolección, presupuesto, vehículos, entre muchos otros. Además, se consideran en la planificación del levantamiento los distintos procesos a ejecutar, como entrevistar telefónica o presencialmente, recuperar contactos o trabajo de gabinete con los mapas móviles.

5.1.4.2 Distribución

Para realizar la distribución de las entrevistas se considera la cantidad de personas efectivamente disponibles para la recolección, buscando la equidad en la carga asignada y cercanía al domicilio del recolector/a lo que es importante, ya que ante un eventual proceso de recuperación de datos, esto reducía los posibles desplazamientos. Esta medida se toma por dos razones, la primera es para evitar los riesgos de contagio, entendiendo el contexto de la pandemia Covid-19, y por otro lado, facilita el acceso al lugar permitiendo flexibilidad en el horario para contactar al informante. Por último, también se toma en consideración para la distribución de las encuestas a realizar la cantidad de unidades (viviendas) sin datos de contacto, con números malos, con/sin datos precargados en el cuestionario sociodemográfico, entre otros elementos.

La distribución del promedio de la carga de trabajo diaria asignada corresponde a 3,6 encuestas, de estas el número promedio de entrevistas por recolector/a es 2,8 encuestas.

5.1.4.3 Trabajo de Campo

La actividad del trabajo de campo corresponde al levantamiento o recolección de información propiamente tal, que inicial y tradicionalmente era presencial, sin embargo, a raíz de la pandemia se rotó a una metodología de recolección telefónica y en algunos casos presencial. A partir, de esta experiencia, el INE determinó que para futuras recolecciones de la ENE esta se realizará de manera mixta, es decir, contemplará la manera presencial y telefónica.

Si bien la unidad de selección en la muestra es la vivienda, la recolección de la información se realiza a nivel de las personas residentes habituales de las viviendas seleccionadas. La información que se recolecta es la que está presente en los instrumentos que se revisarán más adelante. Esta es la tarea base y más importante del proceso de recolección, pues de la precisión y veracidad con la que se capturen los datos, dependerá la calidad de la información que presente la encuesta.

5.1.5 Modificaciones en la metodología de recolección de datos

La metodología del trabajo de campo tuvo una importante variación durante el año 2020, por una diversidad de factores, i) el cambio en el marco muestral, pasando de manzanas y secciones a UPM, ii) la modificación en el modo de recolección de información, del cuestionario en papel al uso de dispositivo móvil de captura (DMC), iii) pasar de la modalidad presencial de recolección, a realizar la entrevista en modalidad mixta (telefónica y presencialmente), este último cambio se inserta en el contexto de la contingencia sanitaria por la pandemia Covid-19.

5.1.5.1 Ajuste en los métodos de Recolección

Durante enero, febrero y hasta la primera quincena de marzo de 2020, se llevó a cabo el mismo proceso que se venía realizando desde hace años para la ENE, el cual era completamente presencial. Este proceso consistía en que el recolector o recolectora visitaba la vivienda y se procedía a obtener la entrevista y, junto con lo anterior, se le indicaba al informante que volvería a ser entrevistado/a en tres meses.

Durante el periodo antes mencionado, la muestra estaba dividida en dos fuentes: el Marco Muestral de Manzanas (MMM) y el Marco Muestral de Secciones (MMS) o “muestra antigua”. Los datos de esta muestra se ingresaban, en un cuestionario en formato papel, y se pasaban al módulo de digitación del sistema de Coordinación Operativa de la Encuesta Nacional de Empleo (COENE). A partir del cambio al Marco Muestral de Viviendas (MMV 2017) o “muestra nueva”, se registra la información en el DMC y se realiza la gestión operativa por parte de la coordinación en el Sistema Integrado de la Encuesta Nacional de Empleo (SIENE). Por lo anterior es que el sistema COENE se terminó de utilizar en junio del 2020.

Las viviendas gestionadas bajo el sistema COENE se recolectaban utilizando un cuestionario físico, de acuerdo a la modalidad papel y lápiz (conocida como PAPI⁶¹). Una vez gestionada la muestra, se revisaba la completitud del cuestionario y secuencia, para luego ingresar los resultados en COENE. Si en el proceso de la digitación se detectaban inconsistencias, se enviaban esos cuestionarios a la sección de supervisión y verificación para que fueran revisados y consultados con los equipos recolectores/as; todo el flujo de material y la gestión del mismo quedaban registrado en el sistema COENE.

En cambio, las unidades que se gestionan en SIENE solo se cargan al DMC (modo CAPI⁶²), específicamente a una aplicación diseñada especialmente para la ENE por el equipo de informáticos del INE, el que por medio de validadores y restricciones programadas, detecta en el momento faltas de completitud o incoherencia, quedando en manos de la persona responsable de la coordinación de recolección el trabajo de cómo y cuándo gestionar cada vivienda. Por lo que, la revisión de consistencia deriva en la aplicación de los criterios de suficiencia definidos por la ENE, estas son validaciones que como mínimo se deben cumplir para que una encuesta entre en el procesamiento para obtener los datos de un período. El resultado de esta etapa es un reporte con la información que se debe recuperar y que es gestionado por la sección de supervisión y verificación.

5.1.6 Planes especiales implementados durante la pandemia

Ante la crisis sanitaria producto de la pandemia Covid-19 y ya que, gran parte del país se encontró en cuarentena desde la segunda quincena de marzo 2020, se tomó la decisión de realizar la recolección de información de la ENE de manera telefónica, esta fue una medida que buscó resguardar la seguridad y salud de los/as recolectores/as e informantes. Esto implicó un ajuste en la muestra, debiendo usar solo aquellas viviendas en las que existe al menos un contacto telefónico,

⁶¹ PAPI, Paper-and-Pencil Interviewing, entrevistas con papel y lápiz.

⁶² CAPI, computer-assisted personal interview, entrevista personal asistida por computador.

para ambos sistemas. En la siguiente Tabla 20 se pueden observar las diferentes técnicas de recolección según el periodo de recolección.

Tabla 20. Técnicas de recolección

Periodo	Entrevista presencial	Entrevista telefónica	Encuesta web
Enero 2020 – marzo (primera quincena 2020)	▪		
Marzo (segunda quincena 2020) – en adelante	▪	▪	▪

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Además de este cambio en la modalidad, se implementó el modo web de captura (CAWI⁶³), para los casos en los que la persona informante se niega a responder la encuesta por teléfono, dándole la alternativa de hacerlo en una plataforma web⁶⁴ diseñada en *Survey Solutions*⁶⁵, entregando para ello un usuario y contraseña, únicos por hogar. Este cuestionario de tipo autoaplicado es una réplica del cuestionario PAPI, manteniendo las mismas secciones, módulos, validaciones y restricciones, a excepción de la Hoja de Ruta (instrumento de la ENE que se detallará más adelante), ya que para fines operativos se considera un rechazo esta modalidad, mientras que si esta encuesta cumple con los criterios de suficiencia entra en el procesamiento de la ENE y se considera una encuesta lograda y completa según el criterio técnico, el modo web dejó de implementarse como técnica de recolección en el mes de octubre 2021.

Por todo lo anterior, se debió elaborar y ajustar los protocolos de recolección y realizar una serie de instructivos para los/as recolectores/as en cuanto a la forma de realizar entrevistas telefónicas, el uso de la mensajería, formas de contacto, entre otras.

A partir de agosto de 2020 se homologaron las variables del marco muestral proveniente de los sistemas de gestión COENE y SIENE, dando paso a que la totalidad de los resultados de la muestra con contacto se pudieran registrar en el DMC. A raíz de este cambio y de la inevitable desactualización de la información de contacto (muestra con y sin contacto telefónico), en septiembre de 2020 comenzó un nuevo proceso de trabajo en terreno, denominado “Operativos de recuperación de contactos”, en el que los recolectores/as, de manera voluntaria y de acuerdo con las nuevas restricciones sanitarias y de movilidad, volvieron al trabajo presencial para poder registrar el contacto según el caso. Esta medida permitió poder ir actualizando paulatinamente el marco muestral de la encuesta, y volver a un esquema de rotación en la ENE.

⁶³ CAWI, Computer Assisted Web Interviewing. Entrevista web asistida por computador

⁶⁴ Disponible en: www.ine.cl haciendo click en el banner.

⁶⁵ La plataforma *Survey Solutions* dejó de implementarse en la recolección de la ENE a partir de octubre de 2021, pues se trabajará en una nueva versión web para la encuesta.

5.1.7 Estructura organizativa del equipo de recolección

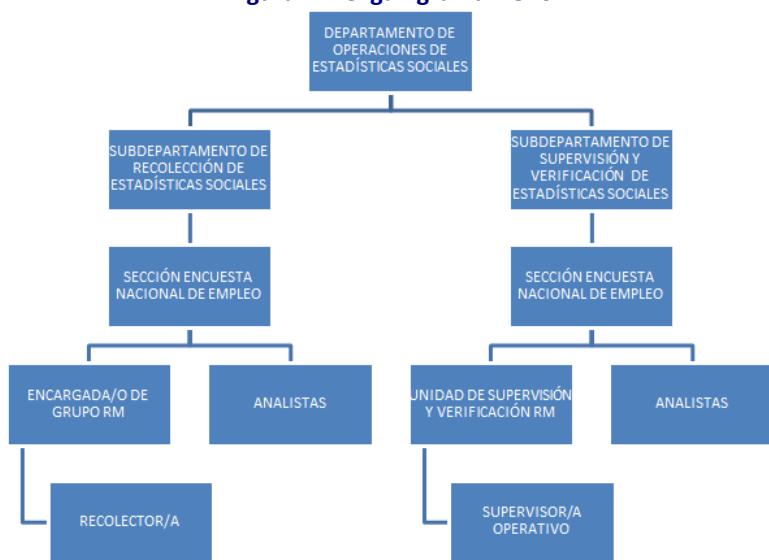
De acuerdo con la estructura organizativa del INE, en la Subdirección de Operaciones (SDO) se encuentra el Departamento de Operaciones Estadísticas Sociales (DOES), el que se compone por el Subdepartamento de Recolección de Estadísticas Sociales (SRES) y el Subdepartamento de Supervisión y Verificación de Estadísticas Sociales (SSVES). El DOES es el responsable de liderar los procesos de la recolección de datos de las operaciones estadísticas oficiales de hogares, de acuerdo con los diseños, estándares metodológicos y buenas prácticas entregados por los equipos de la Subdirección Técnica (SDT) y del Departamento de Metodologías e Innovación Estadística de la institución (DMIE). En relación a la Encuesta Nacional de Empleo, cada subdepartamento lleva a cabo su cometido a través de una sección que se encarga de coordinar al equipo operativo de la Región Metropolitana y a los analistas a su cargo, los que se comunican permanentemente con los equipos regionales, dando a conocer los lineamientos de los procedimientos a realizar y que monitorean el cumplimiento de los mismos.

5.1.7.1 Organigrama del Equipo de trabajo

El organigrama del equipo de trabajo que a continuación se presenta solo hace referencia al equipo vinculado a la ENE, dejando fuera a las demás operaciones categorizadas como proyectos sociales, tampoco incluye a las unidades de digitación y de administración de recursos, ya que producto del uso del DMC, como se explicó anteriormente, estas unidades debieron ser reconvertidas. A diciembre de 2020 la dotación del equipo operativo a nivel central y regional estaba compuesta por 295 personas.

En regiones los equipos del trabajo de campo, supervisores/as y recolectores/as están bajo una única coordinación del encargado/a regional, quien a su vez está bajo la dependencia del jefe/a de operaciones.

Figura 12. Organigrama DOES



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

5.1.7.2 Distribución de equipos Recolección y Supervisión a nivel nacional

Como se muestra en la Tabla 21 la distribución de los recolectores/as y supervisores/as de la Encuesta Nacional de Empleo no es homogénea por región, ya que responde a la cantidad de unidades muestrales asignadas para ser entrevistadas. Es por ello que la Región Metropolitana, Biobío y Valparaíso concentran cerca del 50% de la dotación total de recolectores y recolectoras.

La dotación de recolectores y recolectoras en las regiones están a cargo del coordinador/a regional, solo en la Región Metropolitana el equipo de supervisores posee una coordinación independiente a la de recolección.

Tabla 21. Distribución equipos de recolección y supervisión a nivel nacional

Región	Coordinador Sección Recolección	Coordinador Sección Supervisión	Analista Recolección	Analista Supervisión	Coordinador Recolección	Coordinador Supervisión	Recolector/a de datos	Supervisor operativo	Total general
Tarapacá	-	-	-	-	1	-	4	1	6
Antofagasta	-	-	-	-	1	-	8	2	11
Atacama	-	-	-	-	1	-	8	2	11
Coquimbo	-	-	-	-	1	-	10	4	15
Valparaíso	-	-	-	-	1	-	27	8	36
O'Higgins	-	-	-	-	1	-	12	2	15
Maule	-	-	-	-	1	-	13	3	17
Biobío	-	-	-	-	1	-	22	6	29
La Araucanía	-	-	-	-	1	-	10	2	13
Los Lagos	-	-	-	-	1	-	14	3	18
Aysén	-	-	-	-	1	-	7	1	9
Magallanes	-	-	-	-	1	-	6	2	9
Metropolitana	1	1	2	9	6	1	48	7	75
Los Ríos	-	-	-	-	1	-	8	2	11
Arica y Parinacota	-	-	-	-	1	-	5	2	8
Ñuble	-	-	-	-	1	-	10	1	12
Nacional	1	1	2	9	21	1	212	48	295

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

5.1.7.3 Perfiles de cargo equipos Recolección y Supervisión

El equipo operativo ENE está compuesto por una familia de cargos de coordinadores/as, supervisores/as, analistas, operativos/as y supervisores/as operativos, que se describen a continuación.

a) Coordinación Sección Recolección

La coordinación de la sección de recolección de la ENE, es un cargo único y centralizado, que debe coordinar con los analistas y los equipos regionales todos los elementos asociados a la recolección, de tal manera de garantizar la tasa de respuesta requerida para la encuesta.

b) Coordinador/a Sección Supervisión

La coordinación de la sección de supervisión y verificación de la ENE, es un cargo único y centralizado, que debe coordinar con los analistas y equipos regionales todos los elementos asociados a la supervisión, de tal manera de garantizar la calidad de los datos recolectados.

c) Coordinador/a Recolección

En quince de las dieciséis regiones, corresponde al cargo de la coordinación regional el que cumple con el rol de jefatura directa de los recolectores/as y supervisores/as., La persona asignada al cargo es quien realiza la distribución de la muestra, asigna carga, realiza trámites administrativos y organiza los operativos de recuperación de contacto para actualizar el marco. El coordinador/a cuenta con la capacidad de combinar gestión y planificación, implementando y controlando las actividades de la recolección como de la supervisión y sus necesidades; junto con el manejo de los conocimientos conceptuales y metodológicos de la ENE, con el objetivo de brindar el soporte necesario a su equipo para cumplir las metas y/o indicadores establecidos. Por su parte, en la Región Metropolitana, existen seis encargados de grupo y sus labores abordan el proceso de recolección.

d) Coordinador/a Supervisión

Este cargo solo se encuentra presente en la Región Metropolitana. La persona coordinadora de supervisión es la responsable de coordinar la ejecución del proceso de supervisión, monitorear el cumplimiento de las actividades de supervisión, monitorea rendimientos de productividad requerida, es responsable de que el equipo de supervisores cumpla los procedimientos y protocolos definidos para la encuesta.

Sus principales actividades son:

1. Planificar el trabajo de supervisión de su región, cumpliendo con la calendarización proporcionada desde la Sección de Supervisión y Verificación.
2. Asignar la cargar a cada supervisor/a a cargo, haciendo seguimiento operativo de las actividades del proceso de supervisión de la encuesta.
3. Retroalimentar al equipo de supervisor y supervisoras respecto a su desempeño, orientando en las mejoras operativas en el proceso de supervisión.

e) Analistas recolección

La sección de recolección que desempeña sus funciones en la Región Metropolitana cuenta con dos analistas, quienes realizan las labores de monitoreo de tasas de respuesta, apoyo en el uso de DMC, capacitaciones, implementaciones, entre otras labores que apoyan tanto de manera nacional como regional.

f) Analistas supervisión

La sección de supervisión que desempeña sus funciones en la Región Metropolitana cuenta con nueve analistas, quienes dan seguimiento operativo a los procesos de recolección en terreno, estados de avance y contingencias detectadas. Una parte de los analistas se ocupan del monitoreo operativo de la recolección y otra del análisis de los datos recolectados. Los encargados del monitoreo operativo, son los encargados de coordinar con los equipos regionales todos los elementos asociados a la supervisión y realizan los análisis de consistencia de contenido y forma del proceso de recolección de datos en terreno, con el fin de detectar hallazgos los cuales son comunicados mediante retroalimentación, entre sus actividades principales se encuentran:

1. Monitorear la ejecución de la supervisión de la recolección de datos de la ENE en los tiempos y condiciones definidas para el proceso.
2. Generar informes con indicadores del proceso de supervisión para cada una de las regiones.
3. Retroalimentar a los equipos de campo regionales en los temas que se identifican deficientes durante el proceso de supervisión.
4. Recuperar datos una vez cerrado el periodo de la recolección.

Por su parte, los analistas que desempeñan monitoreo y análisis, entre sus actividades principales se encuentran:

1. Disponer oportunamente de base de datos recolectados a las coordinaciones regionales de equipo de supervisoras y supervisores operativos.
2. Gestionar recuperaciones de datos con las coordinaciones operativas regionales de recolección.
3. Realizar jornadas de capacitación focalizadas en inducción o fortalecimiento de instrucciones y procedimientos de la recolección de datos.
4. Elaborar y emitir informes de hallazgos respecto a resultados de revisión de datos y procedimientos establecidos.

g) Recolector/a de datos

Los/as recolectores/as de datos tienen la función principal de realizar la ejecución de la recolección mediante la aplicación de cuestionarios, siguiendo los lineamientos técnicos y operativos definidos para el estudio. Para ello deben:

1. Recepcionar cargas de unidades a entrevistar.
2. Preparar los materiales y antecedentes necesarios para la recolección de la información, de acuerdo con los procedimientos establecidos por la coordinación de la Encuesta Nacional del Empleo (ENE).
3. Realizar entrevistas presenciales en terreno y/o telefónicamente para recolectar información siguiendo los procedimientos y estándares definidos por la coordinación de la Encuesta Nacional del Empleo (ENE).

h) Supervisor/a operativo

Los/as supervisores/as operativos son los principales responsables de controlar el cumplimiento de la metodología y procedimientos definidos para la encuesta. Es decir, son quienes deben controlar la calidad de la recolección de datos, para lo cual deben aplicar instrumentos que permitan verificar el cumplimiento de protocolos de identificación de la unidad a encuestar y el registro fidedigno de datos en función de los requerimientos técnicos del cuestionario de entrevista, esto se denomina supervisión indirecta. Además, ocasionalmente acompañan al recolector/a para observar su desempeño en la parte conductual y de manejo de aplicación del instrumento, lo cual está definido como tipo de supervisión directa.

Entre sus actividades se encuentran:

1. Recepcionar cargas de unidades a supervisar y ejecutar la aplicación y digitación⁶⁶ de cuestionario según tipo de supervisión.
2. Informar oportunamente a su coordinador o coordinadora cualquier hallazgo de inconsistencia detectado durante la supervisión.

5.2 Diseño de instrumento de captura de información

El diseño de un instrumento de captura de información se crea a partir de una serie de eventos y procesos, en lo que están involucradas todas las partes que intervienen en la recolección de datos. Es decir, hay participación de todos los equipos que componen la Encuesta Nacional de Empleo como producto final, para elaborar un cuestionario que cumpla no solo con los fines técnicos, sino que además sea una herramienta de fácil uso para los recolectores y recolectoras.

5.2.1 Diseño de la encuesta

Esta etapa inicial considera el trabajo de diseño conceptual y diseño metodológico y estadístico de la encuesta. Dentro del diseño conceptual del instrumento se consideraron los siguientes aspectos para su elaboración:

- Análisis de los requerimientos de información y diseño del marco conceptual de la encuesta, que incluye las siguientes tareas:
 - a) Revisar los antecedentes existentes en torno a los conceptos que se quieren medir con la encuesta, las recomendaciones y lineamientos internacionales, evidencia de otros países, entre otros. Parte de estos elementos se revisaron en la sección [“3.1 Marco conceptual”](#) de este documento.
 - b) Establecer temas, categorías, variables y clasificaciones a utilizar.
- Diseño y evaluación de los instrumentos de la ENE (cuestionario central, cuestionario sociodemográfico, hoja de ruta), que implica:

⁶⁶ Se digita hasta que entra en uso el DMC en octubre de 2020.

- a) Redactar y ordenar las preguntas y sus alternativas de respuestas, así como también sus aspectos de orden en el cuestionario, como secuencia, agrupamiento temático, etcétera.
- Elaborar manuales técnicos y de trabajo en terreno, y materiales de apoyo.
- Someter cuestionarios a evaluaciones y validaciones.
- Definir criterios de validación de la información obtenida en el trabajo en terreno, que implica:
 - a) Definir rangos de aceptación para valores de respuesta en cada variable, así como relaciones entre variables para verificar la congruencia en la información.
 - b) Determinar proceso de evaluación de consistencia de la información y alternativas para resolver posibles problemas, como valores fuera de rango, multirespuestas, incongruencia, respuesta a preguntas no aplicables.
- Diseño de la estrategia de presentación de resultados, que conlleva las siguientes tareas:
 - a) Definir los productos de acuerdo a las necesidades de los distintos tipos de usuarios.
 - b) Establecer el contenido específico de cada producto.
 - c) Definir los medios de presentación.

5.2.2 Estructura del cuestionario

La Encuesta Nacional de Empleo está formada por tres instrumentos, con distintos objetivos específicos, que buscan abordar los objetivos generales y específicos de la ENE. Estos instrumentos son, en el orden a abordar en el trabajo en terreno:

5.2.2.1 Hoja de ruta

La Hoja de Ruta corresponde a un cuestionario de autoreporte para el/la encuestador/a, el cual se emplea para recoger principalmente información acerca del proceso de recolección y los intentos de contacto a la unidad muestral. Busca caracterizar el trabajo de campo de la encuesta de acuerdo con un procedimiento estandarizado, respondiendo a 16 preguntas.

Este instrumento se considera para todos los intentos de contacto con la vivienda, no solo cuando existe una entrevista, por ejemplo en caso de no obtener la entrevista, se registran todos los contactos realizados a esa unidad muestral. En ese sentido, la hoja de ruta permite capturar el historial completo de los intentos de contacto realizados con las unidades muestrales, además de datos auxiliares sobre la vivienda.

Al hacer un historial de los intentos de contactos realizados, la hoja de ruta permite recoger información sobre el estado de disposición de las unidades seleccionadas en una muestra.

5.2.2.2 Cuestionario sociodemográfico

El Cuestionario Sociodemográfico recoge información brindada por el informante sobre sus características sociodemográficas y las de su hogar. Como instrumento, tiene un doble objetivo: por

un lado, busca identificar la cantidad de hogares al interior de una vivienda seleccionada, y por otro, busca caracterizar a todos los miembros de cada hogar.

Para caracterizar a los miembros del hogar, se deben responder 28 preguntas, que abordan características atingentes a la relación de los miembros del hogar con la fuerza de trabajo. Primero, se abordan características básicas como parentesco con el proveedor principal del hogar, y después se registran elementos sociodemográficos como: sexo, edad, estado conyugal, migración, nacionalidad, pertenencia a un pueblo originario y educación.

5.2.2.3 Cuestionario central

Este cuestionario aborda los elementos que permiten clasificar y caracterizar a la población en edad de trabajar, por lo que sólo es contestado por los miembros del hogar de quince años o más.

La construcción del cuestionario considera el orden de prioridad definido por la OIT para las encuestas de fuerza de trabajo. Este orden de prioridad indica que, lo primero que se debe realizar en el cuestionario es identificar a la población ocupada y no ocupada. Esto en la ENE se realiza en el módulo “A. Identificación de personas ocupadas y no ocupadas”, que tiene nueve preguntas.

Posteriormente se caracteriza a la población ocupada de acuerdo con las variables principales, como rama económica a la que pertenece la actividad principal desarrollada, categoría en la ocupación, relación contractual o acuerdo de trabajo, informalidad laboral, jornada laboral, entre otras. En la ENE, esta caracterización contempla los módulos “B. Situación en el empleo” de siete preguntas, “C. Caracterización de trabajadores dependientes” con nueve preguntas, “D. Caracterización de trabajadores independientes” que contiene cinco preguntas, “E. Prestaciones laborales y protección social” de tres preguntas, “F. Características de la actividad principal” que tiene siete preguntas, “G. Segunda actividad” con dos preguntas, “H. Horas trabajadas” módulo de dieciocho preguntas e “I. Personas ocupadas que buscaron empleo” con cuatro preguntas. De acuerdo con la descripción de la situación de las personas ocupadas, se van respondiendo los módulos y las secuencias de la encuesta que correspondan.

Por último, quienes fueron clasificados como personas no ocupadas en la sección A, son clasificados según si se encuentran dentro de la fuerza de trabajo como personas desocupadas o si se encuentran fuera de ella (población inactiva). En la ENE, esta clasificación se realiza en la sección “J. Caracterización de personas no ocupadas” que contiene diecisiete preguntas. En total, el cuestionario central de la ENE posee 81 preguntas, sin embargo, debido a los diferentes flujos que se presentan dentro de este cuestionario, ninguna persona encuestada pasa por todas las preguntas de este instrumento.

Ante el impacto de la pandemia en el mercado laboral, el Instituto Nacional de Estadísticas, decidió realizar algunas modificaciones en el cuestionario central de la ENE, con el objetivo de poder incorporar algunas preguntas, que pudieran dar cuenta de la situación actual de las personas ocupadas, que vieron disminuidos sus ingresos producto de la pandemia, y de las personas no ocupadas, que esgrimieron haber perdido o dejado su ocupación durante este período. Al mismo tiempo, se generó un instructivo para potenciar el módulo A del cuestionario, en lo relativo a la

captura de las personas ocupadas ausentes, ya que esta categoría incluye a las personas acogidas a la Ley de Protección al Empleo. Si bien el objetivo de la ENE no es medir el impacto de la pandemia en la fuerza de trabajo o fuera de ella, igualmente se trabajó para que a través del cuestionario se pudieran capturar estas implicancias, pese a que originalmente no fueron diseñadas para ello. Cabe destacar que durante este proceso se trabajó en conjunto con organismos internacionales, como la OIT que asesoró al INE en esta y en otras materias.

5.2.2.4 Tipos de pregunta incluidas en el cuestionario

Los instrumentos antes mencionados que forman parte de la ENE tienen una funcionalidad semiestructurada, es decir, se compone de preguntas pre-codificadas y preguntas abiertas. Al existir diversas formas de clasificar las preguntas del formulario de esta encuesta, se debe mencionar que las preguntas están elaboradas de tal forma que las respuestas que se ofrezcan reúnan dos condiciones imprescindibles para cumplir con el objetivo de cada una de ellas; el de ser excluyentes y ser exhaustivas. De esta forma, en las respuestas se presentan todas las posibilidades para que ningún encuestado las deje sin contestar por no encontrar la respuesta que represente de mejor manera su situación y, por otra parte, no se pueda elegir dos respuestas a la misma pregunta.

Existen diversas clasificaciones que se le pueden dar a las preguntas de un cuestionario, según el tipo de respuesta que nos permite obtener, según la función que tengan dentro del cuestionario, según el diseño o presentación y según el tipo de respuesta esperada. A continuación, se detallan los tipos de preguntas:

a) Segundo tipo de respuestas

La libertad que se le concede a las persona entrevistada al momento de realizar una pregunta es el factor que determina el tipo de preguntas en esta clasificación. Las preguntas “abiertas” no establecen ningún tipo de respuesta previa, por lo que el informante puede responder con sus propias palabras, y se emplean generalmente en situaciones cuando no se puede presumir la respuesta. En las preguntas “cerradas”, la persona entrevistada no puede salirse de las categorías consideradas. Mientras que en las preguntas “semicerradas”, además de la lista cerrada de respuestas, el informante puede añadir espontáneamente una respuesta que no esté en esa lista, generalmente en la opción “otro” o “especifique”.

b) Segundo su función dentro del cuestionario

Existen tres categorías de preguntas que cumplen distintas funciones dentro del cuestionario. Una “pregunta filtro” tiene como función el evitar que se formulen preguntas posteriores a aquellas personas que no corresponden, ya sea por su condición en la categoría ocupacional y se señalan de manera clara. En este apartado también se encuentran aquellas preguntas que cumplen una función de “control o de identificación”, que por lo general cumplen un rol en la identificación de la muestra, control del trabajo de terreno, identificación del entrevistador/a, etc. Existen también aquellas “preguntas de consistencia”, que tienen como finalidad asegurar la veracidad y fiabilidad de las respuestas. Se trata de preguntas similares, pero que han sido redactadas de forma distinta

para verificar congruencia entre sí, o también aquellos casos que, con una pregunta de distinta índole, asegura la respuesta de una pregunta anterior que tenga directa relación en el contenido.

c) **Según su diseño o presentación**

Existen distintas formas de presentar una pregunta en el cuestionario. Una de ellas son las “preguntas en batería”, las cuales permiten utilizar la misma mecánica para preguntar sobre aspectos diferentes.

d) **Según el tipo de respuesta que se espera**

En esta categoría de preguntas encontramos aquellas que son de carácter “cuantitativo”, por lo tanto, la respuesta es un número, y aquellas preguntas de carácter “cualitativo”, que miden características no numéricas.

En el cuestionario ENE se pueden encontrar una serie de símbolos que implican una instrucción específica. A continuación, el detalle de cada una de ellas:

Figura 13. Simbología utilizada en el cuestionario de la ENE

SIMBOLOGÍA		
[]	Referencia al nombre de la persona a la cual se está entregando información	 Salto a otra pregunta sujeto a condiciones
	Salto inmediato a otra pregunta	 Atención en la instrucción

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

5.3 Prueba y evaluación instrumento de captura de información

Para las pruebas y evaluación del instrumento, se han presentado diversas instancias, en las cuales se revisa tanto la consistencia del cuestionario, el funcionamiento del dispositivo o usabilidad, así como también que las implementaciones respondan a las necesidades del producto.

5.3.1 Prueba Piloto ENE 2019

En agosto de 2019 se realizó la última prueba piloto enfocada en la usabilidad del dispositivo y flujo del cuestionario, durante tres meses (agosto, septiembre y octubre). Esta prueba estuvo a cargo del Departamento de Implementación y Mejora de Operaciones Estadísticas (DIMOE), perteneciente a la Subdirección de Operaciones.

En esta prueba piloto de la ENE 2019 participaron once regiones, con una carga total de 10.902 viviendas divididas en los tres meses que duró. A diferencia de los procesos de la ENE antes descritos, la carga de esta prueba se envió a regiones desde nivel central, y cada coordinador/a realizó la distribución dependiendo los recolectores/as asignados para ello, para esta prueba no existió el periodo de referencia de las semanas de levantamiento, por lo que se distribuyó en función

de las necesidades de cada región. A continuación, se presenta la Tabla 22 con la distribución de la carga de la prueba piloto por regiones.

Tabla 22. Prueba piloto

Región	Agosto	Septiembre	Octubre	Total por Región
Antofagasta	333	329	306	968
Atacama	144	138	126	408
Coquimbo	287	243	267	797
Valparaíso	556	595	583	1734
O'Higgins	95	72	45	212
Maule	263	265	263	791
Ñuble y Biobío	352	376	329	1057
Los Ríos	330	271	290	891
Los Lagos	541	513	475	1529
Aysén	137	171	153	461
Metropolitana	694	676	684	2054
Total nacional	3732	3649	3521	10902

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Para esta prueba piloto se contrataron a recolectores/as y coordinadores/as específicos para la prueba, lo que permitió evitar la resistencia al cambio y la adaptación a la metodología existente, así como cualquier tipo de sesgo que pueda existir en recolectores/as de mayor experiencia.

Uno de los cambios más relevantes de la prueba piloto con respecto al levantamiento oficial que se realizaba de forma paralela en esa época, fue la intervención y los cambios que hubo en la hoja de ruta, ya que los códigos de visita que se utilizaron en un principio tuvieron restricciones en sus repeticiones, debido a que eran muy acotados. Este hallazgo permitió realizar modificaciones para evitar problemas durante el levantamiento.

Durante la prueba piloto se detectaron problemas de usabilidad relacionados con el MDM (*master data management*) herramienta que permitía controlar los dispositivos y distribuir las aplicaciones para la recolección. Se llegó a la conclusión que era una herramienta poco eficiente debido a que no era administrada directamente por el equipo informático de la Subdirección de Tecnologías de la Información (SDTI) de la ENE, lo que generaba una demora en la resolución de los problemas que se podían tener en el trabajo de campo.

Los inconvenientes que se produjeron durante la aplicación de la prueba piloto de la ENE 2019 sirvieron para mejorar la usabilidad del instrumento en el dispositivo móvil. Al respecto, esta prueba concluyó que el cuestionario funcionaba sin problemas, tanto en sus saltos como en el flujo del mismo. Además, en términos operativos los/as recolectores/as que realizaron la prueba piloto generaron una evaluación favorable en términos de la usabilidad de la aplicación del instrumento y la correcta ejecución de la encuesta.

5.3.2 Pruebas de implementaciones

Debido al proceso de mejora continua de la encuesta, el instrumento ha experimentado algunas modificaciones a lo largo del tiempo, con el objetivo de incorporar las recomendaciones

internacionales y nuevos estándares institucionales de medición y con ello mejorar la caracterización de las personas y su vínculo con la fuerza de trabajo. En ese sentido, se agregan o quitan preguntas de acuerdo a la coyuntura, como por ejemplo las preguntas incorporadas durante la pandemia para medir el impacto en los ingresos del trabajo⁶⁷, sin afectar a la batería de preguntas centrales de la encuesta, de modo de no alterar los indicadores y resultados que de ellas se desprenden. Las modificaciones de fraseos, saltos, validaciones, entre otros, afectan directamente al cuestionario diagramado (versión en papel) y a la aplicación en el DMC. Luego de realizar cualquiera de estas modificaciones la aplicación pasa por un riguroso testeo que permite prevenir errores de funcionamiento. Dependiendo del tipo y magnitud de la implementación, las pruebas pueden durar desde una semana a tres meses.

5.4 Método de capacitación para la recolección de datos

La ENE al ser una encuesta de continua realiza capacitaciones periódicas tanto a personas recolectoras como a supervisores y supervisoras. Debido al cambio que hubo en la ENE el 2020, se llevaron a cabo varias instancias de capacitación inicialmente. Estos cambios fueron referidos al sistema de captura de información, pasando de papel a un sistema en dispositivo móvil, y también con respecto a la estructura y preguntas del cuestionario. Estas modificaciones fueron realizadas al mismo tiempo, por lo que hubo que llevar a cabo capacitaciones que entregaron los nuevos lineamientos e instrucciones referidas al correcto uso del DMC, la aplicación de la ENE, y las definiciones técnicas para el nuevo cuestionario. Posterior a esas capacitaciones iniciales, se ha seguido continuamente capacitando, buscando dar retroalimentación de los elementos que se van encontrando y de las mejoras que se introducen en el cuestionario y en la aplicación en el DMC.

5.4.1 Tipos de Capacitaciones

5.4.1.1 Capacitación Inicial ENE 2020

Por los cambios antes mencionados en la ENE en el año 2020, existieron dos instancias iniciales de capacitación, la primera, sobre la actualización de cuestionario, que se realizó durante la primera semana de enero de 2020. Luego de esto, se realizó una segunda capacitación durante la última semana de enero y la primera de febrero de 2020, que estuvo enfocada en dos ámbitos, el uso de DMC para recolectores/as, coordinadores/as, supervisores/as y el uso del sistema SIENE para coordinadores/as y supervisores/as, ambas instancias se realizaron de manera presencial.

5.4.1.2 Capacitaciones Especiales

Estas capacitaciones están enfocadas en aquellas temáticas particulares del producto y que requieren de una preparación, tiempo y enfoque específico, las cuales involucran a varios profesionales de la ENE que se desempeñan en los distintos ámbitos de acción del producto.

⁶⁷ Nota técnica sobre preguntas incorporadas por contingencia COVID-19: https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/contingencia-covid-19/separata-t%C3%A9cnica-n-3-contingencia-covid-19.pdf?sfvrsn=da5cedde_4

Una de estas capacitaciones especiales, realizada durante 2020, tuvo relación con la inclusión de preguntas sobre los efectos de la pandemia, tanto para las personas ocupadas como no ocupadas. Por lo que se llevaron a cabo una serie de capacitaciones a cargo del equipo técnico de la ENE, para explicar los lineamientos detrás de cada set de preguntas. Estas capacitaciones tuvieron la participación SDRES para complementar las instrucciones sobre el funcionamiento del DMC. Estas capacitaciones especiales se realizaron en modalidad de videollamada.

5.4.1.3 Capacitación Continua

Tal como se ha reiterado en varias ocasiones, la ENE se recolecta todos los meses en un proceso continuo, por lo anterior, también se deben realizar capacitaciones de forma continua que den cuenta de los hallazgos y situaciones que se deban mejorar. Por lo que este proceso comprende el reforzamiento y actualización de información, enfocado en dos ámbitos distintos. Primero, los relacionados a la actualización continua del cuestionario y/o retroalimentación para abordar los errores más recurrentes. Segundo, los cambios que se producen en la aplicación móvil de la ENE o cambios operativos. Todas estas capacitaciones continuas desde el 2020 abarcan a los equipos técnicos y operativos, y se han realizado a través de video llamada debido a la pandemia Covid-19.

5.4.2 Métodos de capacitación

Debido a los tipos de capacitación antes mencionados, estas instancias se van modificando de acuerdo a su objetivo, y mejorando a través del tiempo, especialmente en el caso de la capacitación continua. Sin embargo, la capacitación básica de la recolección de datos de la ENE viene dada principalmente por dos procesos, el primero, referido a la capacitación técnica del cuestionario ENE y de sus instrumentos, y después una capacitación operativa, donde se enseña sobre el uso del DMC y la parte operativa del proceso. Estas se realizan principalmente a los/as recolectores/as de la ENE al ingresar al equipo operativo, y a lo largo del año según se requiera reforzar ciertos elementos, además en algunas ocasiones se capacita a supervisores/as y analistas según sea necesario.

La capacitación técnica se divide en la exposición de los tres instrumentos de la ENE: Hoja de Ruta, Cuestionario Sociodemográfico y Cuestionario Central, y sus lineamientos técnicos. Para la capacitación de cada instrumento se utiliza como material de apoyo, presentaciones en formato *power point* las cuales abarcan los contenidos necesarios para conocer los instrumentos de la ENE, sus elementos técnicos y los procedimientos principales a considerar. Estas capacitaciones son realizadas por analistas del Subdepartamento de Estadísticas Continuas del Trabajo (SECT), perteneciente a la Subdirección Técnica, quienes manejan los contenidos técnicos necesarios para realizarlas. Esta capacitación tiene un tiempo de duración de doce horas aproximadamente y generalmente se realiza en un periodo de tres días.

La capacitación operativa en relación al subproceso del uso del dispositivo móvil y la aplicación de la ENE es realizada por analistas del Departamento de Operaciones Estadísticas Sociales. Esta capacitación abarca los aspectos generales de la aplicación, la utilización de bases de datos, la asignación de cargas, el método de sincronización y la forma de utilizar los instrumentos dentro de la aplicación ENE. Esta capacitación utiliza como material de apoyo el dispositivo móvil y la Guía de Usuario de este. Por último, esta tiene una duración de dos horas y se realiza de forma periódica según se evalúe su necesidad o exista algún cambio en la aplicación.

5.5 Métodos de Supervisión

5.5.1 Diseño metodológico de la Supervisión

El Subdepartamento de Supervisión y Verificación de Estadísticas Sociales (SSVES) es el responsable de la supervisión y verificación de la Encuesta Nacional de Empleo (ENE) y, por lo tanto, debe garantizar la correcta ejecución del trabajo de recolección, asegurando que el/la recolector/a de datos aplique de manera correcta el instrumento en base a los procedimientos metodológicos definidos; esto con el objetivo de confirmar que la información recolectada es fidedigna. A su vez, la supervisión tiene la finalidad de identificar problemáticas periódicas o esporádicas, desviaciones, errores u omisiones que afecten el cumplimiento de los objetivos del estudio.

El proceso de supervisión permite identificar errores no muestrales, que se pueden clasificar en errores de observación y de medición. Los primeros están relacionados con errores de cobertura y la falta de respuesta, y los segundos identifican problemas de diseño del cuestionario que afecta la captura de datos y el actuar de la persona que realiza la entrevista.

Cabe señalar que la implementación de la supervisión se inició en enero de 2020, cuando el equipo técnico de la ENE, pertenecientes al Subdepartamento de Estadísticas Continuas del Trabajo (SECT), de la Subdirección Técnica (SDT), realizó una capacitación nacional, en la que se comunicó una nueva estrategia de supervisión, de acuerdo con los cambios implementados en la encuesta, la que tendría una mayor participación del equipo técnico para dar a conocer oportunamente lineamientos y/o cualquier medida correctiva, de ser necesaria, y la incorporación de un análisis histórico de los patrones de respuesta y prácticas de los recolectores y recolectoras. Esto fue posible gracias a los cambios realizados en el proceso de recolección de la ENE, que pasó del papel al dispositivo móvil de captura (DMC), por lo que ahora es posible tener acceso de inmediato a los datos brutos recolectados. Estos hitos permitieron al equipo de la sección de supervisión y verificación tener el rol de transmitir y retroalimentar los hallazgos con los equipos de regionales.

5.5.2 Fases operativas de la supervisión

La Figura 14 detalla las fases operativas de la supervisión que se implementan actualmente:

Figura 14. Fases operativas de la Supervisión



Fuente: Elaborado por el Instituto Nacional de Estadísticas

5.5.2.1 Fase Planificación

La planificación de la recolección y supervisión no se encuentran vinculados en un mismo sistema de gestión y tampoco se usa el mismo instrumento de registro: la recolección se gestiona a través del Sistema Integrado Encuesta Nacional de Empleo (SIENE), en cambio el registro de la recolección se realiza utilizando el DMC, mientras que el registro de la supervisión se realiza en papel y se digita

posteriormente en la plataforma *Survey Solutions*. Es por ello que la asignación de unidades a supervisar es realizada por un analista de la sección de supervisión y verificación, quien semanalmente extrae los resultados de la recolección desde SIENE y los carga en *Survey Solutions*, luego notifica a través de correo electrónico a los respectivos coordinadores regionales para que estos asignen carga a los supervisores.

La asignación de cargas se realiza bajo tres criterios generales: a) cumplimiento de la **tasa mínima requerida (12%)** sobre la muestra teórica, b) representatividad de resultados de recolección de la Hoja de Ruta y c) considerando las cargas gestionadas por recolector/a.

5.5.2.2 Fase Ejecución

La ejecución de la supervisión consiste en el desarrollo de la supervisión en dos modalidades: a) de la supervisión de terreno compuesta por la supervisión directa, supervisión indirecta de re-entrevista y supervisión indirecta de unidades no entrevistada; y b) la supervisión de gabinete. Sin embargo, la pandemia COVID-19 obligó a cambiar la metodología de recolección de datos de una presencial a una recolección telefónica o mixta, priorizando recursos en implementaciones que garantizaran la tasa de respuesta en la recolección, así como el cuidado de los/as recolectores/as y los/as informantes. Lo anterior, impactó también al proceso de supervisión, como se detallará más adelante.

La ejecución de la supervisión en terreno debe garantizar la tasa mínima requerida para la Encuesta Nacional de Empleo. Los tipos de supervisión en terreno que se desarrollan son:

a) Supervisión directa

El objetivo de esta supervisión es evaluar y observar in situ el cumplimiento de los protocolos operativos por parte del equipo de recolectores/as, enfocándose en la preparación de la salida a terreno, el desplazamiento a las viviendas, los intentos de contacto, el desarrollo de la entrevista y la revisión de preguntas específicas de los diferentes cuestionarios que conforman la ENE. Este tipo de supervisión se pudo realizar solo durante el periodo de febrero 2020, ya que desde marzo 2020 en adelante, producto de la pandemia COVID-19, se suspendió la recolección presencial y se activó una recolección telefónica, imposibilitando realizar la observación directa al desempeño del recolector/a.

b) Supervisión indirecta re-entrevista

Apunta a detectar inconsistencias en la entrevista, así como también eventuales falseamientos de información, implicando volver a realizar la encuesta en la vivienda entrevistada por el/la recolector/a. La información es contrastada en gabinete por la SDT a fin de evaluar eventuales medidas correctivas, la principal innovación es que se pone el foco en el informante idóneo. Producto de la pandemia, este tipo de supervisión desde la segunda quincena de marzo hasta diciembre de 2020 se realizó de manera telefónica. Posteriormente, durante el año 2021, la supervisión se ha realizado de manera mixta, es decir, de manera telefónica y presencial según el caso.

c) Supervisión indirecta de unidades no entrevistadas

El objetivo de esta supervisión es ratificar que aquellas unidades muestrales no entrevistadas hayan sido debidamente clasificadas de acuerdo a los códigos de visita de la Hoja de Ruta, considerando que cada una de éstas permanecerá seis rondas en la muestra. De igual modo, permite registrar casos en los que haya sido posible revertir un rechazo, si se logró contacto y/o entrevista con unidades donde no fue posible hacerlo antes por parte del recolector o recolectora. Producto de la pandemia, este tipo de supervisión desde la segunda quincena de marzo hasta diciembre de 2020 se realizó de manera telefónica. Posteriormente, durante el año 2021, la supervisión se ha realizado de manera mixta, es decir, de manera telefónica y presencial según el caso.

Paralelamente a la supervisión de terreno se ejecuta la supervisión en gabinete la cual debe garantizar la calidad y consistencia de los datos levantados el terreno:

d) Supervisión de gabinete

Corresponde a la revisión y monitoreo de la calidad del dato una vez realizada la entrevista. Durante 2020 y 2021 la supervisión de gabinete consistió en implementar un procedimiento de recuperaciones de datos emanados de la etapa de pre procesamiento de la encuesta. Este nuevo proceder consiste en recepcionar y gestionar semanalmente un archivo Excel desde la base una tabla con datos alojada en el servidor de la SDT con las inconsistencias y/o valores a recuperar. Para más detalles del subproceso de supervisión de gabinete dirigirse al apartado [“5.5.4 Supervisión de gabinete y control de calidad de los datos”](#).

5.5.2.3 Fase de Cierre

Para dar el cierre operativo de la supervisión, analistas de la sección de supervisión y verificación se comunican con las coordinaciones regionales solicitando verificar que todas las asignaciones hayan sido completadas. Posteriormente, se lleva a cabo la extracción de base de *Survey Solutions* con la cual se construyen los indicadores de tasas de supervisión y se elabora el reporte regional, el cual es enviado mediante correo electrónico a las coordinaciones operativas regionales.

Finalmente, se realizan jornadas mensuales de retroalimentación con los equipos operativos de supervisión regionales, donde se muestran los resultados del periodo y, en caso de existir, también se dan a conocer los resultados del análisis que haya realizado el equipo del Subdepartamento de Estadísticas Continuas del Trabajo.

5.5.3 Impacto de la pandemia en el proceso de supervisión

La pandemia obligó a realizar ajustes en el flujo de material para la supervisión de la “muestra antigua”, ya que en el caso de las encuestas en papel, al no poder dar continuidad a una digitación centralizada en la Región Metropolitana, se ajustaron los sistemas para que los recolectores/as digitaran una vez levantada la encuesta en papel, por lo que ellos mismos debieron aplicar la supervisión de gabinete y garantizar que no existiesen errores de completitud, omisión y secuencia. Esta práctica se mantuvo hasta julio de 2020 cuando se logra homologar el marco antiguo de

secciones y manzanas al marco de UPM y por ende, todas las unidades de la muestra fueron recolectadas mediante el DMC.

Por otra parte, el subproceso de supervisión mantiene en su modelo la ejecución de supervisión directa e indirecta sin que estas se encuentren vinculadas al sistema de gestión, sino que, como se mencionó, se implementan en la plataforma *Survey Solutions*

5.5.4 Supervisión de gabinete y control de calidad de los datos

Los cambios derivados del proyecto de modernización de la Encuesta Nacional de Empleo y de la pandemia Covid-19, han impactado en los procesos de aseguramiento de la calidad del dato, conllevando a la implementación del proceso operativo denominado *“Recuperación y edición de datos”* dentro de la supervisión de gabinete, que deriva de la etapa posterior a la recolección y que tiene por objetivo asegurar la calidad del dato recolectado en la encuesta mediante el procesamiento de datos, cumpliendo con el registro de datos demandados de las validaciones y criterios de suficiencia definidos por el Subdepartamento de Estadísticas Continuas del trabajo.

5.5.4.1 Recuperación de la información

En enero del 2020 con la implementación de la recolección de datos con nuevo marco muestral (UPM), nuevo cuestionario y nuevo sistema de recolección (DMC), el equipo de procesamiento del SECT en conjunto con el equipo de supervisión y verificación, han trabajado para definir procedimientos que permitan abordar la nueva tipología de errores, identificando cuáles pueden ser mediante implementaciones de funcionalidades en el DMC y en base a criterios técnicos de suficiencia, que definen la prioridad en la información a recuperar durante el mes de levantamiento, asociada a una submuestra y materializada en los reportes de consultas, manteniendo una evaluación periódica.

Por lo tanto, la supervisión de gabinete, mediante el proceso de recuperación y de edición de datos, monitorea y asegura que la calidad de los datos recolectados, tanto en el cuestionario central como el sociodemográfico por cada miembro del hogar, permita determinar si una entrevista es “completa” mediante del cumplimiento de una serie de reglas de evaluación de respuestas válidas en preguntas claves.

Para llevar a cabo la implementación del proceso de recuperación y/o edición de datos, entre los equipos SECT, SSVES y la Subdirección de Tecnologías de la Información (SDTI) se llegaron a las siguientes definiciones:

- Designación de un analista de procesamiento SECT quien está cargo de procesar la base de datos y generar semanalmente un reporte en formato Excel con las variables a recuperar o editar, el cual se envía al analista del SSVES.
- Designación de un analista SSVES quien está a cargo del proceso de la recuperación de la información asegurando la calidad del dato cumpliendo con los criterios de suficiencia.

- Designación de dos analistas SECT quienes están a cargo de revisar el reprocesamiento o información recuperada ingresada en el reporte.
- Definición de los cuatro tipos de errores que se abordarán en las recuperaciones:
 - De suficiencia: cuando falta información en alguna pregunta dentro de un módulo del cuestionario sociodemográfico o cuestionario central.
 - De inconsistencia: cuando hay inconsistencia lógica en alguna pregunta dentro de un módulo del cuestionario sociodemográfico o cuestionario central.
 - Valor atípico: cuando se detectan valores fuera de rango en alguna pregunta dentro de un módulo del cuestionario sociodemográfico o cuestionario central.
 - Omisión: corresponden a preguntas respondidas con valores 88 (No sabe) y 99 (No responde) en alguna pregunta dentro de un módulo del cuestionario sociodemográfico o cuestionario central.
- Los errores o inconsistencias a recuperar solo corresponden a las entrevistas con código de visita 11 (Entrevista completa) que no cumplen con los criterios de suficiencia y validaciones de segundo nivel.
- Los errores generados por error de levantamiento (omisión e inconsistencia) que son enviados a los equipos regionales se abordan mediante recuperación telefónica.
- Los errores generados por errores de la aplicación (validaciones, saltos y sincronización) y detectados en el reporte, se identifican y envían mediante correo electrónico a SDTI y SECT para que se implementen las soluciones correspondientes.
- El analista SSVES entrega retroalimentaciones a los equipos regionales indicando instrucciones a reforzar con el objetivo de disminuir la cantidad de casos a recuperar.

5.5.4.2 Metodología de Recuperación de información

El proceso de recuperación y/o edición de información se ha ido modificando en función de las necesidades que demanda la continuidad del producto, es por eso que una característica importante del equipo de la sección de supervisión y verificación, en su rol de monitorear la calidad de la información recolectada, es que se constituye como un equipo externo y con organización separada al equipo de recolección.

Esta característica permite una independencia del equipo, la que se traduce en un trabajo más focalizado y eficiente en la identificación de errores o aspectos a mejorar en el trabajo de los equipos de recolección a nivel nacional.

A su vez, los equipos de recolección realizan simultáneamente controles de calidad interno, a cargo principalmente por los/as coordinadores/as regionales y/o supervisores/as, por lo que el trabajo del equipo de supervisión y verificación es también complementario a dichos procesos.

Entre las tareas definidas del analista SSVES, a cargo del proceso de recuperación de la información, está asegurar que la captura de datos durante la entrevista se lleve a cabo según la metodología y criterios de suficiencia e identificar inconsistencias periódicas que provengan tanto del trabajo de

campo como inconsistencias de origen informático (DMC, aplicación, Validaciones y Saltos), buscando soluciones y homogeneizando criterios para su mejora continua.

5.5.4.3 Gestión Reporte de consultas

El proceso de recuperación y edición de información involucra la gestión de los *Reportes de consultas* la cual es abordado por los coordinadores regionales y equipo de supervisión y verificación.

Las Coordinaciones Regionales están a cargo de abordar los tres o cuatro reportes del periodo cuando el periodo tiene cuatro o cinco semanas, respectivamente. El último reporte de cada periodo lo gestiona el equipo de supervisión y verificación con el objetivo de apoyar el proceso de cierre de periodo.

El reporte de consulta es en formato Excel y la información obtenida de las recuperaciones se captura a través de dos pestañas con sus respectivos módulos que se describen a continuación:

- **Pestaña Consulta a Regiones:** contiene los casos a recuperar correspondientes a cada región. Se divide en módulos con distintas variables que entrega la información necesaria para contactar al informante como para realizar la recuperación y/o corrección de datos. Además, se agregan variables que permiten visualizar la gestión de las recuperaciones realizadas por los equipos regionales, como se muestra en la Figura 15
- **Pestaña Reprocesamiento:** en esta pestaña se ingresan todas las recuperaciones en base al formato que solicita el equipo de procesamiento SECT, como se observa en la Figura 16.

Figura 15. Pestaña Excel de Consultas a regiones

idrph	codigo	codigo	codigo	IDPERIÓ	CONSULTA	ESTADO	JUSTIFICACIÓ	OBSERVACIÓ	region	r_p_c	comun
133033	PAP116	PSS113	PSS113	170	O_B b1_2;				13	13130	SAN MIGU
133342	BLO140	PCV150	PCV150	170	O_C c5;				16	16101	CHILLAÑ
143148	FTS769	RPG108	RPG108	170	O_J j4_dia;O_J j4_sem;				13	13124	PUDAHUE
143150	FTS769	RPG108	RPG108	170	O_J j4_dia;O_J j4_sem;				13	13124	PUDAHUE
227769	FTS769	MPH948	MPH948	170	M_CS mig4;M_CS mig5;				13	13119	MAIPU
168275	FTS769	RPG108	RPG108	170	O_J j4_dia;				13	13119	MAIPU
213887	TZR153	MPH948	MPH948	170	FR_H h2_d;FR_H h4b_d;				13	13505	SAN PEDR
226832	FTS769	CVS979	CVS979	170	M_CS edu2 respecto a edu1;				13	13114	LAS COND
226833	FTS769	CVS979	CVS979	170	M_H h2_t;				13	13114	LAS COND
187270	PAP116	JTA961	JTA961	170	M_A a6 modo COVID;				13	13121	PEDRO AG
192725	JGS131	RBM178	RBM178	170	M_H h3b respecto a h4a;				12	12101	PUNTA AR
194260	TZR153	LBR121	LBR121	170	O_J j4_dia;				13	13401	SAN BERN

Fuente: Subdepartamento de Supervisión y Verificación de Estadísticas Sociales

IDPRH	Figura 16. Pestaña Excel Reprocesamiento
134877	j4_sem 3
167311	f3_anio 2001
165596	edu1 0
165596	edu2 1
132239	f6 1
132239	f6_1 1
165698	h1 2
165698	h2_d
165698	h2_de
165698	h2_h
165698	h2_t
165698	h3a_h 999
165698	h3a_d 6
165698	h3a_t 45

Fuente: Subdepartamento de Supervisión y Verificación de Estadísticas Sociales

En el caso de las recuperaciones efectuadas por los equipos regionales, son los coordinadores regionales los encargados de la distribución de los casos a los/as recolectores/as y de gestionar el reporte de consultas. En el caso de las recuperaciones realizadas por el equipo de supervisión y verificación, es el/la analista SSVES a cargo del proceso quien distribuye los casos a los/as analistas del equipo.

6 Metodología de procesamiento de la información

6.1 Procesamiento de la información

El procesamiento de la información se entiende como la operación que convierte los datos brutos, es decir la información que no ha sido organizada y/o preparada, en información adecuada para el análisis y posterior difusión.

Antes de describir el proceso, es necesario explicitar que la ENE cuenta con una recolección de información y procesamiento continuo, lo que significa que los periodos de ejecución de la recolección y el procesamiento son cortos, de cuatro a cinco semanas dependiendo el mes, por lo que ambas etapas de traslanan, es decir, que mientras se realiza el procesamiento de la información de un mes, en paralelo ya se comenzó con la recolección del mes siguiente.

Es así como, en cada periodo de recolección se obtiene información bruta de las entrevistas, provenientes de los instrumentos que se utilizan en este proceso, como la hoja de ruta, el cuestionario sociodemográfico y el cuestionario central de la ENE, que entran en el procesamiento. A continuación, se describirán las etapas de esta secuencia.

6.1.1 Integración de datos

La ENE integra datos internos y externos en sus diferentes etapas y para los distintos procesos que se efectúan.

- **Datos externos:**

- 1) **Proyecciones de población:** a contar del trimestre móvil noviembre-diciembre-enero 2019 la ENE utiliza las Estimaciones de población (mensualizadas) derivadas de las proyecciones de población base Censo 2017, considerando totales poblacionales para características como el sexo y tramos de edad, para distintas desagregaciones territoriales.
- 2) **Información del stock mensual de población extranjera:** En un contexto donde existía subcobertura de población extranjera y ante evidencia de un vínculo diferenciado que esta subpoblación tiene con la fuerza de trabajo respecto a los nacionales, fue necesario incorporar información exógena sobre la subpoblación de extranjeros. Por tanto, se hizo necesario integrar información actualizada de la población extranjera residente en el país, de manera que los resultados de la encuesta puedan entregar estimaciones insesgadas, confiables y precisas. En este sentido, se requiere contar con información mensual sobre el stock de dicha subpoblación para su uso en el ajuste o “calibración” de los factores de expansión de la ENE, y en cualquier encuesta a hogares para la cual la población extranjera le sea de particular interés.

- **Datos internos recolectados:**

Los datos que a continuación se enumeran se trabajan durante todo el periodo de procesamiento en las etapas que se desarrollarán más adelante, con estas bases de datos brutas se generan las bases trimestrales con las que se calculan los factores de expansión y que posteriormente se usa para los cálculos de las variables principales y la difusión de los indicadores.

- 1) **Base de Hoja de Ruta:** la hoja de ruta es un instrumento que contiene la información muestral de las viviendas que se recolectan en un periodo, y los registros de visitas e intentos de contacto a una vivienda.
- 2) **Base de Personas:** Es la base de datos que contiene toda la información recolectada en las entrevistas a los hogares, a partir del cuestionario sociodemográfico y el cuestionario central de la ENE.

6.1.2 Clasificación y codificación

La ENE adopta una metodología de clasificación automática de textos para el proceso de codificación de las descripciones de actividad económica y ocupación de la población ocupada entrevistada. Como es de conocimiento, el proceso estadístico de codificación en encuestas sociodemográficas consiste en el acto de identificar un texto con una categoría determinada, es decir, clasificar el texto y asignarle el código correspondiente de acuerdo con un clasificador internacional o una nomenclatura estandarizada.

Se han desarrollado algoritmos en el *software R (scripts)* los cuales se ejecutan y aplican para realizar la clasificación y codificación automática de las diferentes preguntas y categorías. Es así que el proceso comienza aplicando un código que realizará una limpieza de los textos con fines de eliminación de caracteres irrelevantes y de homologación semántica, y posteriormente se ejecuta un código que clasifica los casos requeridos. El proceso continúa con la fase de análisis *ex-post* que consiste en supervisar permanente y de forma focalizada los casos codificados, esto con fines de control de calidad para la publicación mensual de las estimaciones derivadas con la encuesta.

Los clasificadores que se utilizan dentro del proceso de la ENE se describen en el apartado [“4.7 Nomenclaturas utilizadas”](#) del capítulo [“4. Diseño estadístico”](#) de este documento. Algunos de estos son: El Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES), elaborado por el Instituto Nacional de Estadísticas (INE), para la clasificación de los textos que describen la actividad económica. En el caso de la clasificación de la ocupación, labor u oficio se utiliza la adaptación nacional del Clasificador Internacional Uniforme de Ocupación (CIUO) de la Organización Internacional del Trabajo (OIT) en su última versión de 2008.

Para mayor detalle sobre la metodología de clasificación y codificación de la ENE, se recomienda revisar en la página web institucional el documento “Sistema de Clasificación y Codificación Automática en la Encuesta Nacional de Empleo”⁶⁸ (INE, 2019)

⁶⁸ INE. (Mayo de 2019). Instituto Nacional de estadísticas. Obtenido de Sistema de Clasificación y Codificación Automática en la Encuesta Nacional de Empleo: <https://www.ine.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion>

6.1.3 Análisis de consistencia de los datos

Dentro del procesamiento de la ENE un tema importante es que exista consistencia de los datos publicados. El primer filtro que se utiliza para que el dato entregado sea válido es la existencia de validadores internos dentro del dispositivo móvil (Tablet) con el que se realiza la encuesta. Estos validadores corroboran la secuencia de preguntas que se realizan, la omisión de las preguntas, posibles inconsistencias con el valor entregado o advierte algún número extraño, y buscan evitar el registro de la información que no corresponda.

Pese a la existencia de estos validadores en línea, se realiza una revisión de consistencia durante la el procesamiento de la base de datos, verificando y controlando la consistencia de forma “manual”. Es así que, para que exista consistencia en todos los datos dentro de la base de datos se revisa que los saltos de preguntas y secuencias estén correctos, y también se verifica que los valores de respuesta estén dentro de lo que corresponde, corrigiendo errores dentro de lo que compete.

Esta revisión de consistencia se hace tanto en los datos de cuestionarios por persona como en el registro de intentos de contacto, o base de datos de hoja de ruta. En el caso de la base de datos a nivel de personas, se debe tener en cuenta que hay preguntas dentro del cuestionario que son consideradas de mayor relevancia, debido a que son fundamentales para tener la encuesta completa en función del objetivo que tiene la ENE. Por esto, se pone especial énfasis en la revisión de algunas preguntas del cuestionario, algunas de estas son: la edad, la secuencia de las preguntas “J3-J4-J5” (las que se refieren la fecha de la búsqueda de trabajo de las personas no ocupadas), la cual es relevante para poder clasificar a la población en edad de trabajar dentro o fuera de la fuerza de trabajo. En el caso de la base de la hoja de ruta se revisa y valida caso a caso los intentos de contacto al hogar para poder cerciorarse de que los códigos de visita y la disposición final de cada vivienda sea la correcta de acuerdo a la secuencia de intentos de contacto.

Finalmente, al terminar esta revisión y validación de datos en algunas ocasiones se realizan instructivos para apoyar la recolección en función de situaciones encontradas en este proceso, que tienen como objetivo la mejora continua de los datos de la ENE. Estos instructivos tienen diferentes finalidades, algunas de estas son: hallazgos importantes en la revisión y se debe retroalimentar; casos especial en cómo se debe pesquisar alguna pregunta; o debido a alguna contingencia que amerita una nueva instrucción, como por ejemplo el impacto que tuvo la pandemia del COVID-19 en la recolección y aplicación de la encuesta, entre otras razones.

6.1.4 Edición, imputación y resultados del procesamiento.

Durante el procesamiento de la ENE no se imputan respuestas, ya que el proceso de edición no se realiza a través de ningún método estadístico u otra metodología en específico. Sin embargo, en relación al análisis de consistencia que se describió anteriormente, durante el procesamiento de las bases de datos y durante la aplicación del llamado plan de análisis, que consiste en revisar las respuestas de acuerdo a los distintos perfiles dentro de la encuesta (personas ocupadas,

desocupadas e inactivas), se pueden editar respuestas basadas en los criterios de consistencia y valores fuera de rango para cerciorarse que la información sea consistente con la línea lógica de la encuesta.

Al terminar el proceso anterior, se integran las bases internas que se mencionaron en el punto [“6.1.1 Integración de datos”](#) de este apartado, y se generan las bases finales de personas, hogares y vivienda del periodo en curso. Al tener las bases de datos del periodo en curso listas, se unen a las bases de los dos periodos anteriores, produciendo la base trimestral de personas y viviendas, que son los productos finales del procesamiento de la ENE. Es con estas bases de datos que se realiza el cálculo del factor de expansión, según lo descrito en el capítulo [“4. Diseño Estadístico”](#) de este documento.

Al tener lista la base de datos con los factores de expansión, se ejecutan los cálculos agregados de la ENE y totales poblacionales a partir de microdatos, se genera la base de datos para la publicación, la misma se utiliza como insumo para la construcción de los cuadros estadísticos y *datasets* del INE.Stat, y se elabora una base especial con las variables correspondientes para ejecutar la carga del Banco de datos ENE. Además, comienza la producción estadística, generando los productos de difusión y creación de las principales variables analíticas para cumplir el objetivo de la ENE. El detalle de este proceso se revisa en los siguientes capítulos de este documento.

7 Metodología de análisis

En esta sección se describirán los procesos que se realizan en la Encuesta Nacional de Empleo (ENE) en lo referido al análisis de datos y los métodos utilizados para explicar e interpretar los resultados para su difusión y divulgación.

7.1 Preparación de los borradores de resultados

Esta etapa tiene como por objetivo preparar los resultados estadísticos, realizando una exploración al cálculo de las principales tasas e indicadores que serán publicados en el boletín mensual de empleo, examinando la tasa de desocupación, que es el indicador principal de la encuesta, y otros indicadores complementarios a ella, como las tasas de subutilización de la fuerza de trabajo, la tasa de desocupación con ajuste estacional y los resultados de las personas ocupadas según las principales variables de interés, como la rama de actividad económica, grupo ocupacional, categoría de ocupación y cualquier otro conjunto pertinente⁶⁹. Las variables y fórmulas que se calculan se presentan en el punto “[Estimación de parámetros](#)” del capítulo “[4. Diseño estadístico](#)” de este documento.

7.1.1 Análisis de datos preliminar

Como parte inicial del análisis de los datos, se comienza realizando una última validación asegurándose que la base de datos que se recibe corresponda al periodo que se debe interpretar, esté correctamente compuesta por los meses del trimestre en ejercicio y con su factor de expansión correspondiente.

Posteriormente, para verificar la consistencia de datos, se analizan cifras agregadas del período, trimestralmente y en doce meses. En esta parte se analiza la trayectoria de los datos, contrastando la información con periodos pasados o con situaciones de distinta naturaleza (económica, social o natural) que se han producido durante el periodo y que podrían haber afectado tanto el proceso de recolección como las cifras que posteriormente se obtienen. Se pone especial atención en los hallazgos de este análisis de consistencia, de modo de poder reunir toda la información y explicar de mejor manera qué está ocurriendo, además del origen de las cifras del período.

7.1.2 Estándar para la evaluación de la calidad de las estimaciones

Al terminar la revisión anterior, se construyen las principales variables analíticas (que se describen en el próximo punto), se realizan los cálculos de estimaciones de *stocks* de población según la situación en la fuerza de trabajo, la caracterización de las personas ocupadas y de sus indicadores principales, como las tasas de desocupación, ocupación, participación, informalidad laboral, subutilización de la fuerza de trabajo, entre otras. Lo anterior se incluye en los productos de difusión,

⁶⁹ Estos mismos procedimientos se aplican para la elaboración, análisis y difusión del boletín trimestral de informalidad laboral, con información que se desprende de la encuesta nacional de empleo.

como el boletín de empleo mensual y las series estadísticas (tabulados Excel) que se ponen a disposición de todos los usuarios.

Junto con todos los cálculos anteriores se aplica el estándar de calidad estadística para las estimaciones de niveles de población y tasas que se publican (INE, Febrero de 2020), este consiste en aplicar un flujo que incluye el número de observaciones o tamaño muestral, el coeficiente de variación, los grados de libertad, el error estándar, los intervalos de confianza de la estimación. A partir de esta información, tanto en el boletín como en las series o cuadros estadísticos se incorporan unas notas o advertencias que dan cuenta de la calidad estadística de la estimación al desagregarla según otras variables de interés. Además de la realización de las operaciones anteriores, en esta parte del proceso se genera el cálculo de ajuste estacional de las series de empleo y de las tasas por sexo.

7.1.3 Construcción de principales variables analíticas

Con las bases de datos, posterior al procesamiento, se construyen las principales variables analíticas de la ENE. Una variable analítica es aquella que se genera en función de otras variables y que se incluye en el análisis estadístico. Estas variables contribuyen a dar una explicación del cuestionario de la Encuesta Nacional de Empleo, y son fundamentales para el análisis estadístico que se realiza.

- Las principales variables analíticas que se construyen son:

7.1.3.1 Variable CAE General

La variable **cae_general**, indica la Condición de Actividad Económica General, la que corresponde a la clasificación de todas las personas de quince años o más según su situación en la fuerza de trabajo, basado en un código sumario empleo a diez categorías. Las categorías observadas se componen de la información obtenida de las secciones y preguntas específicas que estructuran el cuestionario.

Las categorías que componen la Condición de Actividad Económica General son las siguientes:

0. Menor de 15 años
1. Persona ocupada tradicional
2. Persona ocupada no tradicional
3. Persona ocupada ausente
4. Cesante
5. Busca trabajo por primera vez
6. Iniciador
7. Personas Inactivas que buscaron
8. Personas Inactivas que estuvieron disponibles
9. Personas Inactivas que no buscaron ni estuvieron disponibles

7.1.3.2 Variable CAE Específico

La variable **cae_especifico**, o Condición de Actividad Económica Específica, corresponde a la clasificación de todas las personas de quince años o más según su situación laboral, basado en un código sumario empleo a 29 categorías.

Las categorías que componen el **cae_especifico** son las siguientes:

0. Menor de 15 años
1. Persona ocupada tradicional
2. Persona ocupada sin remuneración tradicional
3. Persona ocupada no tradicional
4. Persona ocupada sin remuneración no tradicional
5. Persona ocupada ausente con vínculo efectivo
6. Persona ocupada ausente con pronto retorno
7. Persona ocupada ausente con sueldo o ganancia
8. Cesante
9. Busca trabajo por primera vez
10. Iniciador/a (Disponible)
11. Iniciador/a (No disponible)
12. Razones familiares permanentes (Potencial)
13. Razones familiares permanentes (Habitual)
14. Razones de estudio (Potencial)
15. Razones de estudio (Habitual)
16. Razones de jubilación (Potencial)
17. Razones de jubilación (Habitual)
18. Razones de pensión o montepiado (Potencial)
19. Razones de pensión o montepiado (Habitual)
20. Razones de salud permanentes (Potencial)
21. Razones de salud permanentes (Habitual)
22. Razones personales temporales (Potencial)
23. Razones personales temporales (Habitual)
24. Sin deseo de trabajar
25. Razones estacionales
26. Razones de desaliento
27. Otras razones (Potencial)
28. Otras razones (Habitual)

A partir de la variable anterior, se crea la variable **activ** que agrupa a las personas según su condición de actividad en la fuerza de trabajo en tres categorías, jerárquicas y excluyentes, las que corresponden a:

1. Personas ocupadas
2. Personas desocupadas
3. Fuera de la fuerza de trabajo

7.1.3.3 Variables de categoría en la ocupación

La variable **categoria_ocupacion**, corresponde a las personas ocupadas según adaptación chilena de la Clasificación Internacional de la Situación en el Empleo del año 1993 (CISE-93).

La CISE-93 clasifica los empleos según el tipo de riesgo económico, solidez del vínculo entre la persona y el empleo, y el tipo de autoridad, que tienen o tendrán los titulares sobre los establecimientos y sobre otros trabajadores. La CISE-93 utilizada en la ENE corresponde a la adaptación chilena.

Los valores que componen la variable categoría en la ocupación son las siguientes:

0. No corresponde
1. Empleador
2. Cuenta propia
3. Asalariado/a sector privado
4. Asalariado/a sector público
5. Personal de servicio doméstico puertas afuera
6. Personal de servicio doméstico puertas adentro
7. Familiar o personal no remunerado

7.1.3.4 Variables de Horas

Se construyen dos variables al respecto de las horas trabajadas.

- Variable **habituales**, que corresponde a las horas usualmente trabajadas en la semana de referencia.
- Variable **efectivas**, corresponde a las horas efectivamente trabajadas en la semana de referencia. Este dato proporciona una “fotografía” de las horas trabajadas en el periodo específico.

7.1.3.5 Variable CINE

La variable **CINE**, es la Clasificación Internacional Normalizada de la Educación del año 1997 (CINE 97). Como se desarrolló en el apartado “Nomenclatura utilizada” del capítulo “4. Diseño Estadístico”, la CINE forma parte de la familia internacional de Clasificaciones Económicas y Sociales de las Naciones Unidas, las cuales son empleadas a nivel mundial en la elaboración de estadísticas con el objetivo de acopiar y analizar datos comparables a nivel internacional de manera consistente. Dentro de estas, la CINE representa una clasificación de referencia que permite ordenar los

programas educativos y sus respectivas certificaciones por niveles de educación y campos de estudio (UNESCO).

Las categorías que son parte de la variable CINE son las siguientes:

1. Nunca estudió
2. Educación preescolar
3. Educación primaria (nivel 1)
4. Educación primaria (nivel 2)
5. Educación secundaria
6. Educación Técnica (Educación Superior no Universitaria)
7. Educación universitaria
8. Post títulos y maestría
9. Doctorado
- 999 Nivel ignorado

7.1.3.6 Variables de formalidad laboral

Se construyen dos variables para representar la informalidad en la ocupación, la primera es la variable **sector**, que clasifica a las personas ocupadas por sector ocupacional, es decir si las personas ocupadas pertenecen al sector formal, informal u hogares como empleadores. Las categorías dentro de esta variable son las siguientes:

1. Sector Formal
2. Sector Informal
3. Sector Hogares como Empleadores

La variable **ocup_form**, corresponde a una variable que clasifica a la población ocupada según informalidad en su ocupación, tal como se explicó en el capítulo [“3. Marco de Referencia”](#) del presente documento, con las personas ocupadas informales nos referimos a todos aquellos asalariados o trabajadores del servicio doméstico que no cuentan con cotizaciones de salud (Isapre o Fonasa) o previsión social (AFP) por concepto de su vínculo laboral con un empleador. Las categorías presentes en esta variable son las siguientes.

1. Persona ocupada formal
2. Persona ocupada Informal

7.1.3.7 Factor de Expansión

El factor de expansión se interpreta como la cantidad de personas en la población que representa una persona en la muestra. De acuerdo con el diseño de la muestra, corresponde aplicar este factor

a cada persona seleccionada. Para conocer sobre el diseño y cálculo de factor de expansión de la ENE se recomienda revisar el capítulo [“4. Diseño Estadístico”](#) de este documento.

La variable **fact_cal**, es el Factor de expansión trimestral calibrado, que permite expandir datos muestrales a poblacionales, de acuerdo a las proyecciones del Censo 2017 utilizando una metodología llamada *raking*, de mayor precisión.

7.1.3.8 Variables de información muestral.

Además de las variables analíticas antes descritas existen un set de variables de información muestral que se incluyen en las bases de datos. Estas variables corresponden a los datos que permiten identificar las características geográficas y del marco muestral utilizado en la encuesta, algunas de estas variables son:

- **Estrato:** Identificador único de los estratos geográficos utilizados en el diseño muestral, homologado para considerar los estratos definidos en el marco muestral de manzanas y secciones y en el marco muestral de viviendas.
- **Conglomerado:** Identificador único de unidades primarias de selección, consolidado de acuerdo a las unidades contenidas en el marco nuevo y en el marco antiguo. La variable “conglomerado” cumple la función de la variable **id_directorio**, presente en las bases hasta el trimestre octubre-diciembre de 2019.
- **id_identificacion:** Identificador único del hogar en las viviendas particulares encuestadas

7.1.3.9 Variables de identificación de la persona

Dentro de las variables requeridas para la identificación individual de miembros de un hogar, se construyen dos variables para poder identificar a las personas entrevistadas, estas son:

- **nro_linea:** Corresponde al número de la línea con el cual la persona fue identificada en el cuestionario, o al orden correlativo que la persona toma dentro de la composición de un hogar.
- **idrph:** Identificador único de personas. Esta variable permite identificar a las personas encuestadas de forma longitudinal.

7.1.4 Registros Administrativos

En paralelo a la construcción de los productos a difundir y el análisis de la ENE, comienza el trabajo con los registros administrativos de la Superintendencia de Pensiones, cuyos resultados se incluyen en una sección especial del boletín de empleo. El procedimiento que se realiza es sencillo y consiste en comparar de manera agregada el número de cotizantes por periodo con los obtenidos en la ENE, utilizando los cuadros estadísticos que la Superintendencia de Pensiones publica en su página web.

Mensualmente, se descarga el cuadro estadístico llamado “Cotizantes e ingreso imponible promedio (mensuales)”, y se obtienen los datos de “Nº de Cotizantes Dependientes”, “Nº de Cotizantes Dependientes Masculino”, y “Nº de Cotizantes Dependientes Femenino”. Los datos que se presentan por mes en el tabulado corresponden a los datos del mes anterior, es decir, por ejemplo, los datos de Julio 2021 son los datos de Junio 2021. Finalmente, se calcula de cada una de las variables mencionadas el promedio del número de cotizantes y se comparan con el número de cotizantes que se estiman trimestralmente en la ENE. Estos datos se publican en el boletín mensual, producto de difusión de la ENE que se describirá en el capítulo [“8.1 Productos estadísticos a publicar”](#) de este documento.

7.2 Procedimiento para la anonimización de los datos

Un paso dentro de la construcción final de la base de datos que se publica es la realización de un procedimiento que garantice la no determinación de la unidad estadística. En el caso de la ENE este proceso viene dado por dos elementos centrales.

El primer elemento que se realiza para la indeterminación de las viviendas y personas es dentro de la base de datos, ya que no se considera ninguna variable que permita identificar a la persona entrevistada o a la vivienda seleccionada. Es decir, no se incluyen variables como el nombre, teléfono o RUN de la persona, como tampoco se incluyen variables que permitan identificar a las viviendas desde sus características muestrales, geográficas o de manera cruzada con otros productos estadísticos. En definitiva, la estructura de la base de datos por si sola permite indeterminar y anonimizar a los informantes y a las unidades muestrales (viviendas).

Además, para minimizar aún más el riesgo de determinación de las unidades entrevistadas, se enmascaran dos identificadores, estos son: la variable **idrph**, que es el identificador único de personas, y la variable **conglomerado**, que es la identificadora del conglomerado consolidado del marco nuevo/marco antiguo. Se genera un conglomerado y un idrph ficticio que no permita realizar ninguna conexión con la vivienda seleccionada.

En el próximo capítulo se profundiza sobre la accesibilidad a los datos y el marco legal que regula la protección y la difusión de la información por parte del INE.

7.3 Interpretación de resultados

Esta etapa tiene por objetivo interpretar y explicar los resultados estadísticos obtenidos del análisis descriptivo de las cifras del período, trimestralmente y en doce meses, lo que incluye el cálculo de las principales tasas e indicadores que serán publicados.

7.3.1 Generación de contenido para el boletín y otros productos de difusión.

Posteriormente a la validación y anonimización de la base se terminan de realizar los cálculos de los contenidos que se analizan en el boletín, estos se ejecutan en una rutina automatizada en el *software* estadístico (*R Studio*) para la generación e interpretación de resultados. Además, se

producen los gráficos y tablas correspondientes a la publicación. Estos análisis descriptivos incluyen las estimaciones por niveles, variaciones interanuales, incidencias, variaciones trimestrales y/o cualquier observación que se encuentre relevante de incluir según el contexto nacional.

Una vez realizadas las estimaciones, análisis, gráficos, tablas para el boletín, se lleva a cabo un contraste de los cálculos generados, a partir de la base de datos consolidada y mediando el uso de un *software* estadístico distinto al que genera el boletín (*STATA*) y en caso de encontrar diferencias o tener comentarios que puedan complementar y mejorar el análisis, estos se aplican y se implementan las mejoras para obtener una versión final completamente revisada del boletín o cualquier otro producto a difundir en el periodo.

7.3.2 Generación de series estadísticas

A partir de la base trimestral consolidada se genera el cálculo de los diferentes series o cuadros estadísticos que se difunden, los que contienen series de datos en formato Excel y que se conocen como tabulados. En cada período se actualiza la información de cada cuadro estadístico, que posee un formato Excel previamente establecido y diseñado para ello, incorporando el último dato de cada trimestre en análisis y aplicando el estándar de calidad institucional para las estimaciones provenientes de encuestas de hogares con una nota y la respectiva advertencia.

Una vez producidos los cuadros estadísticos requeridos para su publicación se realiza un contraste de cifras. En esta validación cruzada de resultados se revisan las estimaciones y las notas de calidad estadística. Si llegasen a existir diferencias, se busca su origen y se vuelven a generar los cuadros estadísticos. Una vez que las series estadísticas estén validadas, estos se dejan listos en formato de Excel para su difusión.

7.4 Validación de resultados

Esta etapa del proceso tiene por objetivo validar la calidad del boletín mensual de acuerdo con un marco general de calidad y con las expectativas establecidas, garantizando la coherencia esperada, la completitud y suficiencia de los datos agregados para cada perfil en la Población en Edad de Trabajar. El estándar por el cual se valida el boletín y los productos de difusión de a ENE se describe brevemente en el apartado [“7.1.2 Estándar para la evaluación de la calidad de las estimaciones”](#) de este capítulo, o se puede revisar su detalle en el documento “Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares” (INE, Febrero de 2020).

Además de validar la calidad del boletín de acuerdo con el marco de calidad estadística, se realiza una validación cruzada de boletín automatizado con los cálculos directos desde la base de datos final y con las series estadísticas estimadas, asegurando así coherencia y validez entre los productos estadísticos que se difunden.

7.5 Finalización de resultados

Esta etapa tiene por objetivo realizar las revisiones pertinentes a los contenidos del boletín, en su última versión. Estas revisiones que tienen como fin garantizar la calidad de la información entregada a los usuarios, comprende la revisión, implementación de mejoras y, finalmente, el visto bueno de las jefaturas correspondientes a nivel institucional para entregar el producto final, para así iniciar el proceso difusión de los resultados.

8 Alcance de la presentación de resultados

8.1 Productos estadísticos a publicar

Los productos estadísticos que se ponen a disposición de los usuarios son los siguientes:

- **Boletín de empleo Nacional:** Documento que detalla los principales indicadores del mercado del trabajo en el país, desde la perspectiva de la oferta de trabajo, describiendo las variaciones de la situación dentro y fuera de la fuerza de trabajo, incluye la tasa de desocupación y una completa caracterización de las personas ocupadas a través de distintas categorías, como la rama de actividad económica, la categoría en la ocupación, entre otros. Incorpora, además las tasas de subutilización de la fuerza de trabajo, la tasa de desocupación con ajuste estacional, un apartado sobre informalidad laboral y la comparación con registros administrativos.
- El boletín de empleo nacional incluye una sección de anexos que contiene tablas con los niveles principales de la ENE, además de los intervalos de confianza para el trimestre en análisis e igual trimestre del año anterior.
- A nivel regional se produce un boletín de empleo, similar al nacional, que refleja la dinámica de la oferta de trabajo en cada región e incluye los principales indicadores.
- **Boletín de empleo región Metropolitana:** Documento que detalla los principales indicadores de la oferta de trabajo en la región metropolitana, describe las variaciones de la situación dentro y fuera de la fuerza de trabajo, incluye la tasa de desocupación además de distintas categorías para caracterizar a las personas ocupadas.
- **Boletín de informalidad laboral:** Documento de publicación trimestral que detalla los principales indicadores de ocupación informal en el país, describe las variaciones de las personas ocupadas formales e informales en distintas categorías, como la rama de actividad económica según la empresa que le paga el sueldo o de la que es dueña la persona ocupada y la categoría en la ocupación, entre otros. Incluye tasas de ocupación informal y la tasa de ocupación en el sector informal.
- **Boletín de empleo población extranjera:** Documento de publicación mensual que detalla los principales indicadores de ocupación de la población extranjera en el país. Dentro de sus estimaciones incluye la tasa de desocupación, ocupación y participación por sexo, así como también indicadores de subutilización, informalidad y nivel educacional, entre otros.
- **Separata Anual:** Documento que resume la dinámica de la oferta de trabajo en un año y su variación respecto al período anterior, a través de un ejercicio analítico y descriptivo que calcula datos anuales, donde se utiliza un factor de expansión trimestral ajustado por no respuesta y aplicando el ajuste de marco dual , ponderado por la proporción de personas

según trimestre calendario y región-área y, se calibra a las proyecciones de población del periodo central del año. En esta separata se incluyen los principales indicadores de la fuerza de trabajo además de la caracterización de la población ocupada, destacando las variaciones según incidencia en las desagregaciones incorporadas (sexo, rama de actividad económica, categoría ocupacional y principales tasas).

- **Separata técnica:** Documento que resume una dinámica y/o situación en particular sobre aspectos técnicos, metodológicos y/o operativos relacionados con la Encuesta Nacional de Empleo. Anteriormente su publicación era eventual y cumplía con el objetivo de apoyar la difusión de la información de manera precisa, simple y oportuna. A partir del 2020, todos los meses con la difusión de los resultados de cada período se publica una separata técnica que aborda distintos aspectos técnicos y temáticos a profundizar.
- **Base de datos:** Publicación de las bases de datos anonimizadas e indeterminadas con la información trimestral de la ENE desde año 2010 en adelante, hasta el último período publicado.
- **Series o cuadros estadísticos (Tabulados Excel):** Series de datos sobre la información principal de la encuesta a nivel nacional, según sexo, y regional, desde el año 2010 en adelante, con actualización trimestral. Presentado en cuadros estadísticos en series de tiempo, nueva calibración, proyecciones de población Censo 2017. Se complementan **18 series** para la Encuesta Nacional del Empleo y **4 series** para Informalidad.

Tabla 23. Series o cuadros estadísticos

Nº	PRODUCTO	TÍTULO*
1	ENE	Población en edad de trabajar por situación en la fuerza de trabajo
2	ENE	Población ocupada según categoría en la ocupación
3	ENE	Población ocupada según rama de actividad económica ciu-rev4.cl
4	ENE	Población ocupada por grupo de ocupación
5	ENE	Indicadores relacionados con horas trabajadas
6	ENE	Indicadores complementarios a la tasa de desocupación
7	ENE	Situación en fuerza de trabajo según provincias**
8	ENE	Situación en fuerza de trabajo según conurbaciones**
9	ENE	Comparación ENE-RR.AA
10	ENE	Tasa y niveles de desocupación con ajuste estacional
11	ENE	Población ocupada ausente
12	ENE	Población ocupada con disminución de ingresos laborales
13	ENE	Medidas ante disminución de los ingresos laborales del hogar
14	ENE	Población ocupada que emite boleta de honorarios***
15	ENE	Situación en fuerza de trabajo de personas pertenecientes a pueblos originarios***
16	ENE	Personas jóvenes que no estudian ni están ocupadas

17	ENE	Personas ocupadas según existencia de turnos de trabajo
18	ENE	Situación en la fuerza de trabajo de la población extranjera

* En 2021 el título o nombre de cada archivo disponible en la web fue actualizado.

** Series de datos descontinuadas desde el trimestre enero-marzo de 2020 en adelante. Se retoma la publicación en el trimestre junio-agosto de 2021.

*** Series de datos incluidas en la difusión en julio de 2021, pero con información desde el trimestre enero-marzo de 2020.

La Tabla 24 muestra cada serie para “informalidad laboral” (dimensión analítica que profundiza la información actualmente entregada por la Encuesta Nacional del Empleo).

Tabla 24. Series o cuadros estadísticos de Informalidad Laboral

Nº	PRODUCTO	TÍTULO
1	Informalidad laboral	Personas ocupadas formales e informales y tasa de ocupación informal, nacional y regional
2	Informalidad laboral	Personas ocupadas formales e informales por rama de actividad
3	Informalidad laboral	Personas ocupadas formales e informales por grupo de ocupación
4	Informalidad laboral	Personas ocupadas formales e informales por categoría en la ocupación

8.2 Difusión de los resultados

8.2.1 Difusión

- **Microdatos**

La difusión de los resultados y el material elaborado a partir de la Encuesta Nacional del Empleo se realiza a través de la página oficial del Instituto Nacional de Estadísticas (www.ine.cl), sección Estadísticas Sociales, subsección Mercado laboral, Ocupación y desocupación (<https://www.ine.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion>), el Banco de Datos ENE (<http://bancodatosene.ine.cl>) y la plataforma INE.STAT (<https://stat.ine.cl>)⁷⁰.

El Banco de Datos de la Encuesta Nacional de Empleo (ENE) es un sistema de información orientado a la consulta que permite a los usuarios realizar cruces de forma dinámica con la información registrada desde el trimestre móvil enero-marzo de 2010. La actualización se realiza el día de la publicación de las cifras de empleo, según el calendario oficial INE.

INE.Stat es un repositorio de datos estadísticos flexible, interactivo y de uso gratuito, administrado por el Instituto Nacional de Estadísticas (INE). INE.Stat brinda al usuario la posibilidad de buscar datos de indicadores estadísticos por fuente de información (por ejemplo, nombre de la encuesta), tema de interés, palabra clave o usando el árbol temático de la plataforma. La flexibilidad de la plataforma le permite al usuario personalizar la visualización de sus consultas y modificar la

⁷⁰ En el Anexo N°9 se detalla la forma de acceder a los datos.

disposición de las filas, columnas y filtros en las tablas. La información del INE.Stat se actualiza después de 48 horas hábiles a la publicación oficial de la ENE, el último día hábil de cada mes.

Las bases de datos de la ENE se encuentran publicadas en el portal web del INE (INE, 2021)⁷¹, desde el periodo enero-marzo de 2010 hasta la actualidad, que corresponde a la serie vigente de la encuesta. Las bases de datos de la ENE están disponibles en tres formatos: *SPSS (.sav)*, *STATA (.dta)* y *comma-separated values (.csv*, por su sigla en inglés, para trabajar con otros *software*).

Junto con la base de datos y el repositorio de INE.Stat se difunden según correspondan los siete productos que se enlistaron en la sección anterior “[Productos estadísticos a publicar](#)”, en los que se incluyen los boletines de empleo Nacional, Región Metropolitana y de informalidad laboral, las separatas, y cuadros estadísticos, entre otros.

- **Metadatos**

En relación a la publicación de metadatos, en la página web mencionada en el apartado anterior, se pone a disposición el presente documento: “Documento Metodológico de la Encuesta Nacional de Empleo”, que sintetiza los principales aspectos metodológicos del estudio, específicamente, resume el marco conceptual, el diseño estadístico de la encuesta, el proceso de recolección y supervisión de la encuesta, y la metodología de procesamiento y análisis de la ENE.

Junto a la metodología, se publica el “Libro de Códigos Base de Datos ENE” del año 2019 y 2021, este último es una actualización de la versión 2020, que tiene como propósito facilitar el manejo de los datos por parte de las diferentes personas usuarias de la información contenida en las bases. También, en la página web antes mencionada se encuentra a disposición el “Cuestionario ENE - Informalidad 2018”, el cual contiene las preguntas de la ENE que se realizan a los informantes.

Otro de los documentos adicionales que se publican es sobre la “Actualización del marco muestral” de mayo de 2019, que aborda la implementación del Marco Muestral de Manzanas (MMM) y del Marco Muestral de Secciones (MMS) actualizado con información de Precio 2016 en la ENE. También se publica el documento “Sistema de clasificación y codificación automática en la encuesta nacional de empleo” que describe los elementos del Sistema de Clasificación y Codificación Automática desarrollado en el INE para abordar las necesidades de codificación de preguntas abiertas contenidas dentro del proceso de producción estadística. En el mismo sentido, como guía práctica para codificación de actividad económica se tiene a disposición pública el documento “CAENES: Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas”.

⁷¹ INE. (2021). Mercado laboral: Ocupación y desocupación. Obtenido de Instituto Nacional de estadística: www.ine.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion

También se ponen a disposición la ficha técnica estandarizada en DDI/DCMI, glosarios, documentos de trabajo asociados a la Encuesta Nacional de Empleo e información sobre Comités y Notas técnicas, entre otros. Toda esta información se encuentra disponible en el sitio de la ENE.

- **Otros**

Junto con todos los documentos y productos de difusión, el INE pone a disposición de sus usuarios/as el portal del Sistema Integral de Información y Atención Ciudadana (SIAC), disponible en www.ine.cl/SIAC, para canalizar cualquier requerimiento adicional de información, el que será evaluado de acuerdo a dos criterios:

- Disponibilidad de la información solicitada.
- Protección de la confidencialidad de los datos, de acuerdo al secreto estadístico establecido en la Ley 17.374.

8.2.2 Consideraciones sobre el uso de la información

En relación a la información que se pone a disposición de las personas usuarias de los datos, es importante tener algunas consideraciones.

Primero, **la información de la ENE se publica el último día hábil de cada mes, según el calendario de difusión oficial que posee el INE**. Si bien los datos se publican mensualmente, las bases contienen información de un trimestre completo, inmediatamente anterior al mes de publicación. Por ejemplo, en julio de 2020 se publica la información correspondiente al trimestre móvil abril-mayo-junio. El concepto de trimestre móvil obedece a una lógica de actualización de la submuestra (mes) que conforma el trimestre que se publica en cada período, es decir, se está agregando un nuevo mes, al mismo tiempo que se debe eliminar el primero del trimestre anterior, manteniendo los otros dos meses para conformar un nuevo trimestre. Siguiendo con el mismo ejemplo, para el mes de agosto de 2020, se elimina el mes de abril y se agrega el de julio, manteniendo los meses de mayo y junio, conformando el trimestre mayo-junio-julio.

Segundo, **la actual serie de datos de la ENE comienza en el trimestre móvil enero-marzo de 2010**, y se encuentra a disposición del público en la página institucional de INE para su descarga directa. Para información anterior a este período, se deben solicitar las bases de datos a través del Sistema Integral de Información y Atención Ciudadana (SIAC), disponibles desde el trimestre diciembre-febrero de 1986 en adelante.

Tercero, **los datos de la serie vigente no son comparables con series anteriores de la ENE**, debido a la incorporación de importantes cambios conceptuales y metodológicos a partir de enero-marzo de 2010.

Finalmente, si bien se encuentra disponible información sobre la comuna de residencia habitual o de trabajo de las personas en las bases de datos, **la ENE no tiene representatividad comunal**. Su diseño garantiza errores de muestreo aceptables sólo para los dominios nacional, nacional urbano,

nacional rural, regional urbano (para todas las regiones) y regional rural (sólo para regiones Metropolitana, O'Higgins, El Maule, Ñuble, La Araucanía, Los Ríos y Los Lagos) (INE, 2020a). El INE no se puede hacer cargo de estimaciones fuera de estos dominios oficiales de estudio, siendo responsabilidad de las personas investigadoras realizar las advertencias necesarias sobre esta situación.

8.2.3 Indicaciones para el uso de los datos

La población objetivo de la ENE corresponde a la población en edad de trabajar, esto es, personas de quince años o más, por lo tanto, todas las estimaciones de la encuesta referidas a su población de interés según su situación en la fuerza de trabajo o fuera de ella se acotan a este límite inferior a nivel etario en su definición, por lo que los filtros ya se encuentran incorporados. No obstante, la información sociodemográfica está disponible para la totalidad de los integrantes de un hogar.

Dado que la unidad de análisis corresponde a personas, en un contexto de diseño muestral bietápico, se utiliza un factor de expansión. Éste se interpreta como la cantidad de personas en la población que representa una persona en la muestra. De acuerdo con el diseño de la muestra, corresponde aplicar este factor a cada persona seleccionada.

En ese sentido, se deben tener dos consideraciones:

1. Entre el comienzo de la actual serie de la ENE y el trimestre octubre-diciembre de 2019, las bases de datos de la ENE incluyen dos factores de expansión, no comparables entre sí por consideraciones metodológicas.
 - a. Factor tradicional (fact). Se basa en las proyecciones de población del Censo 2002, utilizando una metodología de cálculo que divide a la población por sexo y según dos tramos etarios (personas menores de 15 años y personas de 15 años o más) (INE, 2018).
 - b. Factor calibrado (fact_cal). Utiliza las proyecciones del Censo 2017 y una metodología de cálculo llamada *raking*, de mayor precisión.
2. A partir del trimestre noviembre-enero de 2019, las bases sólo contienen el factor calibrado.

8.3 Accesibilidad a los datos y Marco legal

De acuerdo a lineamientos institucionales, la información producida por la ENE se encuentra supeditada a los compromisos de confidencialidad y reserva de datos protegidos por secreto estadístico. Estos lineamientos pueden ser resumidos en el siguiente marco legal:

- 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 19.628 de protección de vida privada y datos.
- 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 29º 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 30º 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 y autónomas y las demás personas jurídicas estatales serán titulares del derecho de autor respecto de las obras producidas por sus funcionarios en el desempeño de sus cargos*".

9 Glosario

A continuación se presentaran los conceptos claves que permiten entender la metodología expuesta anteriormente.⁷²

- **Asalariado/a:** es la persona que trabaja para un empleador, jefe o patrón y percibe una remuneración en forma de sueldo, salario, comisión, propinas, pagos a destajo o en especie.
- **Asalariado/a privado:** esta categoría incluye a todos los asalariados que no trabajan en instituciones o empresas del Estado, independiente del sector económico donde se desempeñan.
- **Asalariado/a público:** esta categoría incluye a todos los asalariados que trabajan en instituciones o empresas del Estado, ya sea dentro de la administración pública como en otros sectores económicos (salud, enseñanza, etc.).
- **Brecha de género:** es la diferencia que se manifiesta entre la situación, condición y/o posición de las mujeres y la de los hombres, es decir, la distancia que hace falta recorrer para alcanzar la igualdad en una situación determinada.
- **Población desocupada:** todas las personas en edad de trabajar que no estaban ocupadas durante la semana de referencia, que habían llevado a cabo actividades de búsqueda de un puesto de trabajo durante las últimas cuatro semanas (incluyendo la de referencia) y que estaban disponibles para trabajar en las próximas dos semanas (posteriores a la de referencia).
- **Fuerza de trabajo:** personas en edad de trabajar que durante la semana de referencia cumplen los requisitos para ser incluidas en la categoría de las personas ocupadas o desocupadas.
- **Fuerza de trabajo potencial (Inactivos potencialmente activos):** personas en edad de trabajar que, durante la semana de referencia, no estaban ocupadas ni desocupadas y que habían llevado a cabo «actividades de búsqueda» pero no estaban «actualmente disponibles para trabajar»; o que no llevaron a cabo «actividades de búsqueda», pero que si estaban «actualmente disponibles para trabajar».
- **Fuerza de trabajo ampliada:** se define como la suma de la fuerza de trabajo más los iniciadores disponibles y la fuerza de trabajo potencial (inactivos potencialmente activos).
- **Horas efectivas:** cuantifican el tiempo dedicado por las personas al desempeño de actividades económicas que contribuyen a la producción de bienes y servicios durante un período de referencia corto (la semana de referencia de la encuesta).

⁷² Este listado de términos es extraído de: Instituto Nacional de Estadísticas. (Enero de 2019). Glosario ENE. Obtenido de Instituto Nacional de Estadísticas: <https://www.ine.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion>

- Horas habituales: corresponden a las horas comúnmente trabajadas en un período de referencia determinado (promedio de los últimos tres meses).
- Iniciadores/as disponibles: todas las personas fuera de la fuerza de trabajo que esgrimen no haber buscado trabajo en las últimas cuatro semanas debido a que iniciarán pronto una actividad laboral y que, al mismo tiempo, declaran disponibilidad. Este grupo poblacional se incluye dentro de los Inactivos, pero para ciertos efectos (de armonización estadística con la OCDE), se suma a la población desocupada para calcular una tasa de desocupación con iniciadores disponibles.
- Población Ocupada: todas las personas en edad de trabajar, que durante la semana de referencia dedicaron al menos una hora a alguna actividad para producir bienes o servicios a cambio de una remuneración o beneficios.
- Ocupación informal: comprende a todos los trabajadores dependientes que carecen de acceso al sistema de seguridad social (pensión y salud) por concepto de su vínculo laboral. Para el caso de los trabajadores independientes, se considera que poseen una ocupación informal si la empresa, negocio o actividad que desarrollan pertenece al sector informal. Familiar no remunerado son todos clasificados como informales.
- Población en edad de trabajar (PET): población actualmente residente en el país de 15 años y más.
- Población fuera de la fuerza de trabajo o inactiva: todas las personas de la población en edad de trabajar que no son clasificadas como ocupadas ni desocupadas.
- Sector informal: conjunto de unidades económicas que carecen de registro en el Servicio de Impuestos Internos (SII) y no llevan una contabilidad a través de la cual puedan separar los gastos del negocio de los del hogar.
- Tasa de desocupación: número de personas desocupadas expresado como porcentaje de la fuerza de trabajo.
- Tasa de desocupación ajustada estacionalmente: es la tasa de desocupación que elimina los efectos de los factores exógenos estacionales de naturaleza no económica que influyen en su comportamiento coyuntural. Para el ajuste estacional de las series de empleo, se utiliza el programa X13 ARIMA SEATS. Se aplica este método independientemente en la serie de personas desocupadas y de ocupadas, para luego obtener la tasa de desocupación ajustada sin estacionalidad (método indirecto). De este modo es posible comparar de forma directa los valores de la serie entre distintos momentos del tiempo.
- Tasa de desocupación con iniciadores disponibles (SU1): número de personas desocupadas más personas iniciadores disponibles expresados como porcentaje de la fuerza de trabajo más iniciadores disponibles.
- Tasa combinada de desocupación y tiempo parcial involuntario(SU2): número de personas desocupadas más personas iniciadores disponibles y personas ocupadas a tiempo parcial involuntario expresado como porcentaje de la fuerza de trabajo más iniciadores disponibles.

- Tasa combinada de desocupación y fuerza de trabajo potencial (SU3): número de personas desocupadas más personas iniciadoras disponibles e inactivas potencialmente activas expresado como porcentaje de la fuerza de trabajo ampliada.
- Tasa de ocupación: número de personas ocupadas como porcentaje de la población en edad de trabajar.
- Tasa de ocupación en el sector informal (TOSI): número de personas ocupadas que trabajan en unidades económicas informales expresado como porcentaje del total de personas ocupadas.
- Tasa de ocupación informal (TOI): número de personas con ocupaciones informales expresado como porcentaje del total de personas ocupadas.
- Tasa de participación: número de personas en la fuerza de trabajo expresado como porcentaje de la población en edad de trabajar.
- Tasa de presión laboral: número de personas desocupadas, más personas ocupadas que buscan empleo, más iniciadores disponibles, expresado como porcentaje de la fuerza de trabajo más los iniciadores disponibles.

10 Bibliografía

- 13° CIET. (1982). Obtenido de https://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/resolutions-adopted-by-international-conferences-of-labour-statisticians/WCMS_087483/lang--es/index.htm
- 19° CIET. (2013). Obtenido de https://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/resolutions-adopted-by-international-conferences-of-labour-statisticians/WCMS_234036/lang--es/index.htm
- 20°CIET. (Febrero de 2018). Resolución sobre las estadísticas de las relaciones de trabajo. Obtenido de https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms_619087.pdf
- AAPOR. (2016). Standard Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys. Obtenido de https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf
- Antía, M., & Coimbra, A. (2009). Tratamiento de la no respuesta en encuestas de panel en el caso de poblaciones finitas: "Las damas perdidas". Universidad de la República, Uruguay.
- Banco Central de Chile. (2005). Indicadores económicos y sociales de Chile: 1969-2000. Santiago. Recuperado de https://www.bcentral.cl/documents/33528/133439/bcch_archivo_098139_es.pdf
- Bankier, M. D. (1986). Estimators Based on Several Stratified Samples, with applications to Multiple Frame Surveys. Obtenido de Journal of American Statistical Association, 81, 1074-1079.
- Bates, D., & Chambers, J. (1992). Nonlinear models . Wadsworth & Brooks/Cole.
- Bates, D., & Watts, D. (1988). Nonlinear regression analysis and its applications . Wiley & Sons.
- Binder, D. A., & Theberge, A. (1988). Estimating the Variance of Raking Ratio Estimators. Obtenido de Canadian Journal of Statistics, 16 (supplement), 47-55.
- Brackstone, G. J., & Rao, J. N. (1979). An Investigation or Raking Ratio Estimators. Obtenido de Sankhya C, 41, 97-114.
- Bravo, D. (2010). Encuestas Longitudinales en Chile. Santiago, Chile.
- Cecchini, S. (2005). Indicadores sociales en América Latina y el Caribe (Vol. 34). Santiago: CEPAL. Recuperado de http://repositorio.cepal.org/bitstream/handle/11362/4735/S05707_es.pdf
- Choudhry, G. H., & Lee, H. (1987). Variance Estimation for the Canadian Labour Force Survey. Obtenido de Survey Methodology, 13, 147-161.

- Cochran, W. (1998). Técnicas de Muestreo. México D.F.: Compañía Editorial Continental, S.A. de C.V.
- Crawley, M. (2013). The R book. Chichester: Wiley & Sons.
- Deparment of Commerce Bureau of the Census. (1971). Curso Suplementario Para Un Estudio de Caso Sobre Encuestas Y Censos: Conferencias Sobre Muestreo.
- Deville J., S. C. (1993). Generalized Raking Procedures in Survey Sampling.
- Deville, J.-C., & Särndal, C.-E. (1992). Calibration Estimators in Survey Samplin. Recuperado el 2020, de Journal of the American Statistical Association, Vol. 87 (418): 376-382.
- Dobson, A. (2002). An Introduction to Generalized Linear Models. CRC Press.
- Glejberman, D. (2015). Taller Regional. Las Encuestas de Hogares en América Latina: Estado de situación y prospectiva. Usos de las Encuestas Longitudinales. Santiago de Chile.
- Groves, R. (2004). Survey Errors and Survey Cost. Nueva York: John Wiley & Sons.
- Hartley, H. O. (1962). Multiple Frame Surveys. Proceedings of the American Statistical Association, Social Statistics Sections.
- Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). Applied Survey Data Analysis. Obtenido de Boca Ratón, FL: Taylor and Francis Group, LLC.
- Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). Applied Survey Data Analysis. USA: Chapman & Hall / CRC Press.
- Herrera, D., Huaracán, V., & Inostroza, D. (2019). Indicadores de subutilización de la fuerza de trabajo en Chile. Evidencia desde la Encuesta Nacional de Empleo. Santiago: Instituto Nacional de Estadísticas (Chile). Recuperado de https://www.ine.cl/docs/default-source/documentos-de-trabajo/estadísticas-de-subutilización-de-fuerza-trabajo-ine.pdf?sfvrsn=54e122f9_2
- Hidiroglou, M. (1986). The Construction of a Self-Representing Stratum of Large Units in Survey Design. The American Statistician, 40(1), 27-31.
- IBGE. (2004). Semantic Scholar. Recuperado el 15 de 01 de 2020, de https://pdfs.semanticscholar.org/21bb/a469c8b217f276755a827006d31b4255bdc5.pdf?_ga=2.132283335.871766543.1578866611-143326825.1578866611
- IESTA. (s.f.). IESTA. Recuperado el 12 de diciembre de 2019, de http://www.iest.edu.uy/wp-content/uploads/2014/05/TJA2010_ant%C3%ADa-coimbra.pdf
- INE. (Abril de 2016a). CAENES. Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas. Obtenido de [https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/documento-clasificador-de-rama-de-actividad-econ%C3%B3mica-caenes-\(abril-2016\).pdf?sfvrsn=1b1cc016_3](https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/documento-clasificador-de-rama-de-actividad-econ%C3%B3mica-caenes-(abril-2016).pdf?sfvrsn=1b1cc016_3)

- INE. (Abril de 2016b). Separata Técnica. Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES) en la Encuesta Nacional de Empleo. Obtenido de https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/tem%C3%A1ticas/separata-tecnica-caenes-empleo-abril-2016.pdf?sfvrsn=25455d76_3
- INE. (2018a). CIUO 08.CL. Clasificador Chileno de Ocupaciones. Obtenido de http://webanterior.ine.cl/docs/default-source/publicaciones/2018/ciuo-08.cl-clasificador-chileno-de-ocupaciones.pdf?sfvrsn=50bd58d2_4
- INE. (Febrero de 2018b). Estadísticas de informalidad laboral: Marco conceptual y manual metodológico. Obtenido de https://www.ine.cl/docs/default-source/informalidad-y- condiciones-laborales/metodologia/antecedentes-metodologicos/manual-conceptual-y- metodol%C3%B3gico-informalidad-laboral.pdf?sfvrsn=afad6bfc_5
- INE. (Noviembre de 2018c). Características de la inmigración internacional en Chile, Censo 2017. Obtenido de <http://www.censo2017.cl/descargas/inmigracion/181123-documento-migracion.pdf>
- INE-DEM. (Febrero de 2019). Estimaciones de Personas Extranjeras residentes en Chile, al 31 de Diciembre 2018. Obtenido de <https://www.extranjeria.gob.cl/media/2019/07/Estimaci%C3%B3n-Poblaci%C3%B3n-Extranjera-en-Chile.pdf>
- INE. (Marzo de 2019a). Ciudades, pueblos, aldeas y caseríos 2019. Obtenido de https://geoarchivos.ine.cl/File/pub/Cd_Pb_AI_Cs_2019.pdf
- INE. (Mayo de 2019). SEPARATA Codificación Automática ENE. Obtenido de https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/tem%C3%A1ticas/separata-tecnica-codificacion-automatica-ene.pdf?sfvrsn=273dd63c_3
- Instituto Nacional de Estadísticas. (Enero de 2019). *Glosario ENE*. Obtenido de Instituto Nacional de Estadísticas: <https://www.ine.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion>
- INE. (Junio de 2019b). Unidades primarias de muestreo. Nuevo marco muestral de viviendas. Documento interno. Instituto Nacional de Estadísticas (INE).
- INE. (Octubre de 2019c). Estratificación socioeconómica del marco muestral de viviendas 2017 (MMV 2017). Documento de Trabajo. N°11. Obtenido de https://www.ine.cl/docs/default-source/documentos-de-trabajo/estratificaci%C3%B3n-socioecon%C3%B3mica-del-marco-muestral-de-viviendas-2017.pdf?sfvrsn=e4482f3a_2
- INE. (Diciembre de 2019d). Fundamentos de la nueva metodología de calibración de los factores de expansión de la Encuesta Nacional de Empleo. Documento de Trabajo. N° 12. Obtenido de <https://www.ine.cl/inicio/documentos-de-trabajo/documento/documento-de-trabajo-fundamentos-de-la-nueva-metodolog%C3%ADa-de-calibraci%C3%B3n-de-los-factores-de-expans%C3%ADn-de-la-encuesta-nacional-de-empleo>.

- INE. (2020a). Estadísticas de informalidad laboral: Marco conceptual y manual metodológico. Santiago: Instituto Nacional de Estadísticas (Chile). Recuperado de <https://www.ine.cl/docs/default-source/informalidad-y-condiciones-laborales/metodologia/antecedentes-metodologicos/manual-conceptual-y-metodologico-informalidad-laboral.pdf>
- INE. (2020b). Separata Técnica: Fortalecimiento de la Encuesta Nacional de Empleo 2018-2022. Instituto Nacional de Estadísticas (Chile). Recuperado de <https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/temáticas/separata-técnica-fortalecimiento-ene-febrero-2020.pdf>.
- INE. (Marzo de 2020). Lineamientos para el Uso de Medidas de Precisión en Encuestas de Hogares. Documento de Trabajo (próximo a publicar). Obtenido de <https://www.ine.cl/inicio/documentos-de-trabajo>.
- INE. (Febrero de 2020). *Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares*. Obtenido de INE: <https://www.ine.cl/docs/default-source/institucionalidad/buenas-pr%C3%A1cticas/clasificaciones-y-estándares/est%C3%A1ndar-evaluaci%C3%B3n-de-calidad-de-estimaciones-publicaci%C3%B3n-27022020.pdf>
- INE. (Julio de 2020). Cálculo de población extranjera en chile para la calibración de factores de expansión en la encuesta nacional de empleo. Obtenido de https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/documentos-de-trabajo/documentos-de-trabajo/c%C3%A1culo-de-poblaci%C3%B3n-extranjera-en-chile-para-la-calibraci%C3%B3n-de-factores-de-expans%C3%ADn-en-la-encuesta-nacional-de-empleo.pdf?sfvrsn=184891be_4
- INE. (2021). Separata técnica: Nuevas dimensiones de análisis Encuesta Nacional de Empleo (ENE). Santiago: Instituto Nacional de Estadísticas (Chile). Recuperado de <https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/temáticas/separata-técnica-nuevas-dimensiones-de-análisis-en-la-ene.pdf>
- Jadresic, E. (1986). Evolución del empleo y desempleo en Chile 1970-85. Series anuales y trimestrales. *Estudios CIEPLAN*, 20, 147–193.
- Kish, L. (1963). Changing strata and selection probabilities. *Proceedings of the Social Statistics Section, American Statistical Association*, 124-131.
- Konijn, H. S. (1981). Biases, Variances, and Covariances of Raking Ratio Estimators for Marginal and Cell Totals and Averages of Observed Characteristics. *Obtenido de Metrika*, 28, 109-121.
- Lohr, S. L. (2000). *Muestreo: Diseño y Análisis*. México: International Thomson Editores.
- Lynn, P. (2005). *Metodología de Encuestas Longitudinales*. Seminario Internacional de Estadística Eustat.

- Medina, F. (2001). Las Encuestas de Panel. Diseño e Implementación de un Sistema de Encuestas de Hogares, (págs. 139-145). Aguascalientes, México. Recuperado el 15 de Noviembre de 2017, de <https://www.cepal.org/deype/mecovi/docs/TALLER7/12.pdf>
- Medina, F. (2012). Esquemas de Rotación de la Muestra en Encuestas Longitudinales. Diseño de encuestas de población activa y módulos de empleo en las encuestas de hogares para medir el trabajo decente. Turín, Italia.
- Montgomery, D., Peck, E., & Vining, G. (2006). Introducción al Análisis de Regresión Lineal. México: Compañía Editorial Continental.
- 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/demographic-social/Standards-and-Methods/files/Handbooks/surveys/seriesf_96-S.pdf
- Naciones Unidas. (2009). Diseño de muestras para encuestas de hogares. Directrices prácticas (Vol. Serie F N°98). Nueva York. Obtenido de https://unstats.un.org/unsd/publication/seriesf/Seriesf_98s.pdf
- Observatorio Social Universidad Alberto Hurtado. (2008). Entrega N°5: Documento Metodológico Encuesta Panel Casen Medición 2007. Santiago, Chile.
- OCDE. (2015). Recomendación del Consejo de la OCDE sobre Buenas Prácticas Estadísticas. Recuperado de <https://www.oecd.org/statistics/good-practice-toolkit/OECD-LEGAL-0417-spa.pdf>
- OIT. (2013). 19.a Conferencia Internacional de Estadísticos del Trabajo. Informe II: Estadísticas del trabajo, el empleo y la subutilización de la fuerza de trabajo. Ginebra.
- OIT. (2018). 20.a Conferencia Internacional de Estadísticos del Trabajo. Resolución sobre las estadísticas de las relaciones de trabajo. Ginebra.
- OIT, & CEPAL. (2019). Fortalecimiento de la Encuesta Nacional de Empleo (ENE). Misión de Asistencia Técnica Conjunta al Instituto Nacional de Estadísticas. Santiago. Recuperado de https://www.ine.cl/docs/default-source/laborales/ene/antecedentes-metodologicos/metodologia/informe-ine-oit---fortalecimiento-de-la-encuesta-nacional-de-empleo.pdf?sfvrsn=7e655cd2_4
- OIT. (1982). Resolución sobre estadísticas de la población económicamente activa, del empleo, del desempleo y del subempleo, 13.a Conferencia Internacional de Estadígrafos del Trabajo. Ginebra. Recuperado de https://www.ilo.org/wcmsp5/groups/public/-/-dgreports/-/-stat/documents/normativeinstrument/wcms_087483.pdf
- OIT. (1987). Directrices sobre la incidencia de los programas de promoción del empleo sobre la medición del empleo y del desempleo, 14.a Conferencia Internacional de Estadísticos del Trabajo (octubre-noviembre de 1987). Recuperado de https://www.ilo.org/wcmsp5/groups/public/-/-dgreports/-/-stat/documents/normativeinstrument/wcms_087604.pdf

- OIT. (1993). Resolución sobre la Clasificación Internacional de la Situación en el Empleo (CISE), adoptada por la decimoquinta Conferencia Internacional de Estadísticos del Trabajo. Ginebra. Recuperado de https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087564.pdf
- OIT. (2003). 17.a Conferencia Internacional de Estadísticos del Trabajo: Directrices sobre una definición estadística de empleo informal. Ginebra. Recuperado de Directrices sobre una definición estadística de empleo informal
- OIT. (2013a). 19.a Conferencia Internacional de Estadísticas del Trabajo. Resolución I: Resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo. Ginebra. Recuperado de http://www.ilo.int/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_234036.pdf
- OIT. (2013b). La medición de la informalidad: Manual estadístico sobre el sector informal. Ginebra: Organización Internacional del Trabajo (OIT). Recuperado de https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_222986.pdf
- UNESCO. (2006). Clasificación Internacional Normalizada de Educación (CINE 1997). Ginebra: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.
- Ritz, C., & Streibig, J. C. (2008). Nonlinear regression with R. Springer.
- Särndal, C. E., Swensson, B., & Wretman, J. (1991). Model Assisted Survey Sampling. Springer.
- Statistics New Zealand. (2015). Household Labour Force Survey. Obtenido de http://archive.stats.govt.nz/browse_for_stats/income-and-work/employment_and_unemployment/hlfs-sources-methods-2015.aspx
- UNESCO. (2006). Clasificación Internacional Normalizada de la Educación CINE 1997. Obtenido de <https://unesdoc.unesco.org/ark:/48223/pf0000147002>
- Valliant, R., Dever, J. A., & Kreuter, F. (2013). Practical Tools for Designing and Weighting Survey Samples. New York: Springer.
- Vanderhoeft, C. (2001). Generalised calibration at Statistics Belgium: SPSS® module g-CALIB-S and current practices. Statistics Belgium.
- Wolter, K. (2007). Introduction to Variance Estimation (Statistics for Social and Behavioral Sciences). New York: Springer.

11 Anexos

11.1 Anexo N° 1. Número de UPM, viviendas y PET según estrato geográfico de la ENE

Tabla 25. Número de UPM, viviendas y PET según estrato geográfico de la ENE

Región	Área	Código de estrato geográfico	Estrato geográfico	UPM	Viviendas	PET	
Arica y Parinacota	Urbano	1500001	Región de Arica y Parinacota	331	65.926	161.895	
Arica y Parinacota	Rural	1500012	Región de Arica y Parinacota	74	6.676	14.963	
Tarapacá	Urbano	110011	Conurbación Iquique - Alto Hospicio	477	96.171	224.140	
Tarapacá	Urbano	120011	Prov. Del Tamarugal	27	5.370	11.128	
Tarapacá	Rural	110012	Conurbación Iquique - Alto Hospicio	13	1.027	4.430	
Tarapacá	Rural	120012	Prov. Del Tamarugal	53	4.905	13.657	
Antofagasta	Urbano	210001	Prov. Antofagasta (Conurbación Antofagasta + Mejillones + Tal Tal)	581	116.250	296.879	
Antofagasta	Urbano	220011	Prov. El Loa (San Pedro de Atacama y Conurbación Calama)	263	54.000	126.604	
Antofagasta	Urbano	230011	Prov. Tocopilla	51	10.638	22.637	
Antofagasta	Rural	210002	Prov. Antofagasta (Conurbación Antofagasta + Mejillones + Sierra Gorda + Tal Tal)	20	1.590	19.916	
Antofagasta	Rural	220012	R.R. (Prov. Tocopilla y Prov. El Loa)	50	4.439	13.636	
Atacama	Urbano	310011	Conurbación Copiapó	279	56.292	123.524	
Atacama	Urbano	310021	Prov. Chañaral + R.P. Copiapó (Caldera)	79	15.842	30.551	
Atacama	Urbano	320011	Prov. Huasco (Conurbación Vallenar, Freirina, Huasco)	104	20.800	45.821	
Atacama	Rural	310012	Conurbación Copiapó	33	2.876	5.786	
Atacama	Rural	310022	R.R. Atacama (Prov. Chañaral y R.P. Copiapó (Caldera))	30	2.485	3.030	
Atacama	Rural	330012	Prov. Huasco (Vallenar, Alto del Carmen, Freirina, Huasco)	82	7.259	11.834	
Coquimbo	Urbano	410011	Conurbación La Serena – Coquimbo	734	148.084	324.319	
Coquimbo	Urbano	410021	R.P. Elqui	53	10.134	21.697	
Coquimbo	Urbano	420011	Prov. Choapa (Illapel, Canela, Los Vilos, Salamanca)	104	21.254	44.484	
Coquimbo	Urbano	430011	Conurbación Ovalle	150	30.145	67.570	
Coquimbo	Urbano	430021	R.P. Limarí	51	10.124	20.676	
Coquimbo	Rural	410012	Conurbación La Serena - Coquimbo	131	11.955	26.519	
Coquimbo	Rural	410022	R.P. Elqui	87	7.896	15.533	
Coquimbo	Rural	420012	Prov. Choapa (Illapel, Canela, Los Vilos, Salamanca)	150	13.567	27.527	
Coquimbo	Rural	430012	Conurbación Ovalle	112	10.042	18.816	
Coquimbo	Rural	430022	R.P. Limarí	166	14.859	26.114	
Valparaíso	Urbano	510011	Conurbación Valparaíso (Valparaíso)	(1)	549	110.459	243.343

Región	Área	Código de estrato geográfico	Estrato geográfico	UPM	Viviendas	PET
Valparaíso	Urbano	510021	Conurbación Valparaíso (2) (Viña del Mar, Concón)	728	148.135	310.830
Valparaíso	Urbano	510031	R.P. Valparaíso (Casablanca)	30	6.153	14.243
Valparaíso	Urbano	510041	R.P. Valparaíso (Puchuncaví, Quintero)	79	16.383	34.061
Valparaíso	Urbano	530011	Prov. Los Andes (Conurbación Los Andes + Rinconada y San Esteban)	158	32.555	71.553
Valparaíso	Urbano	540011	Prov. Petorca	99	19.473	41.739
Valparaíso	Urbano	550011	Prov. Quillota (Conurbación Quillota - Calera - Hijuelas - La Cruz, Nogales)	304	62.748	140.150
Valparaíso	Urbano	560011	Conurbación San Antonio - Cartagena - Las Cruces + Santo Domingo	218	45.301	100.787
Valparaíso	Urbano	560021	R.P. San Antonio (Algarrobo, El Quisco)	56	11.409	21.216
Valparaíso	Urbano	570011	Prov. San Felipe	199	40.942	91.157
Valparaíso	Urbano	580011	Conurbación Valparaíso (3) (Quilpué, Villa Alemana)	483	98.237	222.296
Valparaíso	Urbano	580021	R.P. Marga Marga (Limache, Olmué)	89	18.210	41.153
Valparaíso	Rural	510012	Conurbación Valparaíso	33	2.841	5.551
Valparaíso	Rural	510022	R.P. Valparaíso (Casablanca)	38	3.492	7.226
Valparaíso	Rural	510032	R.P. Valparaíso (Puchuncaví, Quintero)	34	3.035	6.096
Valparaíso	Rural	530012	Prov. Los Andes (Los Andes, Rinconada y San Esteban)	77	6.587	15.949
Valparaíso	Rural	540012	Prov. Petorca	107	9.924	20.814
Valparaíso	Rural	550012	Prov. Quillota	104	9.260	21.152
Valparaíso	Rural	560012	Prov. San Antonio (Cartagena - Las Cruces, Algarrobo, El Quisco, Santo Domingo)	72	6.582	12.979
Valparaíso	Rural	570012	Prov. San Felipe	157	13.949	30.862
Valparaíso	Rural	580022	R.P. Marga Marga (Limache, Olmué)	53	4.701	9.771
Metropolitana	Urbano	1310011	Conurbación Gran Santiago (1) (Independencia, Quilicura, Recoleta, Huechuraba, Conchalí)	1.044	214.899	553.894
Metropolitana	Urbano	1310021	Conurbación Gran Santiago (2) (Providencia, Ñuñoa, Lo Barnechea, Las Condes, Vitacura, La Reina)	1.814	368.243	775.975
Metropolitana	Urbano	1310031	Conurbación Gran Santiago (3) (San Miguel, Peñalolén, La Florida, San Joaquín, Macul)	1.497	305.152	757.278
Metropolitana	Urbano	1310041	Conurbación Gran Santiago (4) (La Pintana, Pedro Aguirre Cerda, Lo Espejo, La Cisterna, La Granja, El Bosque, San Ramón)	1.197	242.901	662.795
Metropolitana	Urbano	1310051	Conurbación Gran Santiago (5) (Estación Central, Pudahuel, Quinta Normal, Renca, Cerrillos, Cerro Navia, Lo Prado, Maipú)	2.222	452.604	1.175.572
Metropolitana	Urbano	1310061	Conurbación Gran Santiago (6) (Santiago)	963	191.494	355.195

Región	Área	Código de estrato geográfico	Estrato geográfico	UPM	Viviendas	PET
Metropolitana	Urbano	1320011	Conurbación Gran Santiago (7) (Puente Alto)	827	170.156	445.308
Metropolitana	Urbano	1320021	R.P. Cordillera	35	7.124	18.234
Metropolitana	Urbano	1330011	R.P. Chacabuco (Colina, Tiltil)	190	39.356	97.225
Metropolitana	Urbano	1330021	Conurbación Gran Santiago (8) (Lampa)	128	26.200	58.963
Metropolitana	Urbano	1340011	Conurbación Gran Santiago (9) (San Bernardo, Calera de Tango)	445	91.635	237.921
Metropolitana	Urbano	1340021	Conurbación Buin - Paine	197	40.769	100.519
Metropolitana	Urbano	1350011	Prov. Melipilla	192	38.298	90.078
Metropolitana	Urbano	1360011	Conurbación Gran Santiago (10) (Talagante, Padre Hurtado, Peñaflor)	297	62.518	154.495
Metropolitana	Urbano	1360021	R.P. Talagante	87	18.152	44.097
Metropolitana	Rural	1310012	Conurbación Gran Santiago (1)	37	2.890	9.063
Metropolitana	Rural	1320022	Prov. Cordillera	76	6.724	17.689
Metropolitana	Rural	1330012	R.P. Chacabuco (Colina, Tiltil)	107	9.846	25.842
Metropolitana	Rural	1330022	Conurbación Gran Santiago (8) (Lampa)	74	6.780	16.334
Metropolitana	Rural	1340012	Conurbación Gran Santiago (9) (San Bernardo, Calera de Tango)	60	5.541	15.040
Metropolitana	Rural	1340022	Conurbación Buin - Paine	135	12.682	31.258
Metropolitana	Rural	1350012	Prov. Melipilla	276	24.951	57.199
Metropolitana	Rural	1360012	Conurbación Gran Santiago (10) (Talagante, Padre Hurtado, Peñaflor)	101	9.075	23.114
Metropolitana	Rural	1360022	R.P. Talagante	53	4.968	11.994
O'Higgins	Urbano	610011	Conurbación Rancagua - Machalí - Gultro - Los Lirios	597	124.471	266.228
O'Higgins	Urbano	610021	R.P. Cachapoal (Codegua, Mostazal)	47	9.394	21.864
O'Higgins	Urbano	610031	R.P. Cachapoal (Otras comunas)	255	50.988	115.789
O'Higgins	Urbano	620011	Conurbación San Fernando	113	23.774	49.901
O'Higgins	Urbano	620021	Prov. Cardenal Caro	47	9.330	18.268
O'Higgins	Urbano	630011	R.P. Colchagua	137	28.124	62.308
O'Higgins	Rural	610012	Conurbación Rancagua - Machalí - Gultro - Los Lirios	101	8.874	22.546
O'Higgins	Rural	610022	R.P. Cachapoal (Codegua, Mostazal)	41	3.575	8.195
O'Higgins	Rural	610032	R.P. Cachapoal (Otras comunas)	375	33.393	74.442
O'Higgins	Rural	620012	Conurbación San Fernando	42	3.830	8.290
O'Higgins	Rural	620022	Prov. Cardenal Caro	111	10.347	19.547
O'Higgins	Rural	630012	R.P. Colchagua	284	25.770	57.407
Maule	Urbano	710011	Conurbación Talca - Culenar	447	90.726	198.787
Maule	Urbano	710021	R.P. Talca (1) (Constitución, Empedrado, Pencahue, Curepto)	88	17.516	36.876
Maule	Urbano	710031	R.P. Talca (2) (Pelarco, Rio Claro, San Rafael, San Clemente)	54	11.772	25.756

Región	Área	Código de estrato geográfico	Estrato geográfico	UPM	Viviendas	PET
Maule	Urbano	720011	Prov. Cauquenes	83	16.414	32.609
Maule	Urbano	730011	R.P. Curicó (1) (Conurbación Curicó, Romeral, Teno)	267	55.067	121.728
Maule	Urbano	730021	R.P. Curicó (2) (Licantén, Vichuquén, Hualañe, Molina, Sagrada Familia)	105	20.967	45.425
Maule	Urbano	740011	Prov. Linares (1) (Conurbación Linares, Colbún, Yeras Buenas)	164	33.555	72.830
Maule	Urbano	740021	Prov. Linares (2) (Parral, Retiro, Longaví)	86	17.879	36.649
Maule	Urbano	740031	Prov. Linares (3) (San Javier, Villa Alegre)	80	15.615	32.896
Maule	Rural	710012	Conurbación Talca - Culenar	79	7.208	16.100
Maule	Rural	710022	R.P. Talca (1) (Conurbación Constitución, Empedrado, Pencahue, Curepto)	91	8.315	17.100
Maule	Rural	710032	R.P. Talca (3) (Pelarco, Río Claro, San Clemente, San Rafael)	166	15.483	33.695
Maule	Rural	720012	Prov. Cauquenes	76	7.033	13.077
Maule	Rural	730012	R.P. Curicó (Conurbación Curicó, Molina, Sagrada Familia)	167	15.042	33.284
Maule	Rural	730022	R.P. Curicó (Romeral, Teno)	109	10.002	22.864
Maule	Rural	730032	R.P. Curicó (Licantén, Vichuquén, Hualañe)	42	3.801	7.286
Maule	Rural	740012	Prov. Linares (1) (Conurbación Linares, Colbún, Yeras Buenas)	166	14.811	31.924
Maule	Rural	740022	Prov. Linares (2) (Parral, Retiro, Longaví)	189	17.072	36.692
Maule	Rural	740032	Prov. Linares (3) (San Javier, Villa Alegre)	86	7.803	16.087
Ñuble	Urbano	1610011	Conurbación Chillán - Chillán Viejo	336	70.344	156.551
Ñuble	Urbano	1610021	R.P. Diguillín	110	22.192	46.739
Ñuble	Urbano	1620011	Prov. Itata	53	10.996	23.139
Ñuble	Urbano	1630011	Prov. Punilla	92	18.597	39.832
Ñuble	Rural	1610012	Conurbación Chillán - Chillán Viejo	81	7.294	15.760
Ñuble	Rural	1610022	R.P. Diguillín	201	18.251	37.646
Ñuble	Rural	1620012	Prov. Itata	117	10.609	20.853
Ñuble	Rural	1630012	Prov. Punilla	235	21.544	46.203
Biobío	Urbano	810011	Conurbación Gran Concepción + Florida + Santa Juana (1) (Santa Juana, Lota, Coronel)	281	57.170	129.702
Biobío	Urbano	810021	Conurbación Gran Concepción + Florida + Santa Juana (2) (Florida, Tomé, Hualqui, Penco, Talcahuano)	447	92.101	218.980
Biobío	Urbano	810031	Conurbación Gran Concepción + Florida + Santa Juana (3) (Concepción, Hualpén, Chiguayante, San Pedro de la Paz)	931	189.042	429.134
Biobío	Urbano	820011	Prov. Arauco	210	44.059	97.483

Región	Área	Código de estrato geográfico	Estrato geográfico	UPM	Viviendas	PET
Biobío	Urbano	830011	Conurbación Los Ángeles	264	54.355	117.238
Biobío	Urbano	830021	R.P. Biobío (Cabrero, Laja, Nacimiento, Negrete, San Rosendo, Yumbel)	144	29.455	64.354
Biobío	Urbano	830031	R.P. Biobío (Antuco, Mulchén, Quilaco, Quilleco, Santa Bárbara, Tucapel)	95	19.479	41.860
Biobío	Rural	810012	Prov. Concepción (Conurbación Gran Concepción, R.P. Concepción (Florida), R.P. Concepción (Santa Juana))	130	11.678	24.546
Biobío	Rural	820012	Prov. Arauco	161	14.628	31.198
Biobío	Rural	830012	Conurbación Los Ángeles	203	18.385	40.670
Biobío	Rural	830022	R.P. Biobío (Cabrero, Laja, Nacimiento, Negrete, San Rosendo, Yumbel)	137	12.273	24.219
Biobío	Rural	830032	R.P. Biobío (Antuco, Mulchén, Quilaco, Quilleco, Santa Bárbara, Tucapel, Alto Biobío)	123	10.913	21.944
La Araucanía	Urbano	910011	Conurbación Temuco - Padre Las Casas - Cajón	570	117.806	258.871
La Araucanía	Urbano	910021	R.P. Cautín (1) (Carahue, Freire, Galvarino, Lautaro, Nueva Imperial, Perquenco, Saavedra, Teodoro Schmidt, Toltén, Cholchol)	159	32.450	70.617
La Araucanía	Urbano	910031	R.P. Cautín (2) (Cunco, Curarrehue, Gorbea, Loncoche, Melipeuco, Pitrufquén, Pucón, Villarrica)	209	42.076	87.898
La Araucanía	Urbano	920011	Conurbación Angol	100	20.314	44.819
La Araucanía	Urbano	920021	R.P. Malleco (1) (Ercilla, Los Sauces, Lumaco, Purén)	38	7.746	15.386
La Araucanía	Urbano	920031	R.P. Malleco (2) (Collipulli, Curacautín, Lonquimay, Traiguén, Victoria)	133	27.646	58.671
La Araucanía	Rural	910012	Conurbación Temuco - Padre Las Casas - Cajón	244	21.407	48.631
La Araucanía	Rural	910022	R.P. Cautín (1) (Carahue, Freire, Galvarino, Lautaro, Nueva Imperial, Perquenco, Saavedra, Teodoro Schmidt, Toltén, Cholchol)	401	36.556	76.307
La Araucanía	Rural	910032	R.P. Cautín (2) (Cunco, Curarrehue, Gorbea, Loncoche, Melipeuco, Pitrufquén, Pucón, Villarrica)	312	28.285	53.358
La Araucanía	Rural	920012	Prov. Malleco (Angol, Ercilla, Los Sauces, Lumaco, Purén, Collipulli, Curacautín, Lonquimay, Traiguén, Victoria)	232	20.806	42.794
Los Ríos	Urbano	1410011	Conurbación Valdivia	275	55.012	125.829
Los Ríos	Urbano	1410021	R.P. Valdivia (Corral, Los Lagos, Máfil, Mariquina, Paillaco)	69	13.891	31.018
Los Ríos	Urbano	1410031	R.P. Valdivia (Lanco, Panguipulli)	50	9.857	20.903
Los Ríos	Urbano	1420011	Prov. Ranco (Futrono, Lago Ranco, Río Bueno)	53	10.704	21.759
Los Ríos	Urbano	1420021	Prov. Ranco (La Unión)	48	9.801	21.073

Región	Área	Código de estrato geográfico	Estrato geográfico	UPM	Viviendas	PET
Los Ríos	Rural	1410012	Conurbación Valdivia	47	4.301	9.291
Los Ríos	Rural	1410022	R.P. Valdivia	148	13.071	26.601
Los Ríos	Rural	1410032	R.P. Valdivia (Lanco, Panguipulli)	106	9.689	19.366
Los Ríos	Rural	1420012	Prov. Ranco (Futrono, Lago Ranco, Río Bueno)	131	12.006	22.728
Los Ríos	Rural	1420022	Prov. Ranco (La Unión)	50	4.673	9.201
Los Lagos	Urbano	1010011	Conurbación Puerto Montt	442	92.056	196.779
Los Lagos	Urbano	1010021	R.P. Llanquihue (1) (Calbuco, Fresia, Los Muermos, Maullín)	71	14.234	29.665
Los Lagos	Urbano	1010031	R.P. Llanquihue (2) (Frutillar, Llanquihue)	49	9.679	21.458
Los Lagos	Urbano	1020011	Conurbación Osorno	269	54.355	117.710
Los Lagos	Urbano	1020021	Prov. Chiloé + Palena	207	42.413	85.988
Los Lagos	Urbano	1030011	R.P. Osorno	62	12.442	27.181
Los Lagos	Rural	1010012	Conurbación Puerto Montt + Cochamó	184	16.967	33.746
Los Lagos	Rural	1010022	R.P. Llanquihue (1) (Calbuco, Fresia, Los Muermos, Maullín)	179	16.646	32.336
Los Lagos	Rural	1010032	R.P. Llanquihue (2) (Frutillar, Llanquihue)	42	3.699	7.185
Los Lagos	Rural	1020012	Conurbación Osorno	55	5.175	10.987
Los Lagos	Rural	1020022	Prov. Chiloé + Palena	365	32.865	62.013
Los Lagos	Rural	1030012	R.P. Osorno	171	15.763	31.483
Aysén	Urbano	1110011	Conurbación Coyhaique	92	18.590	38.040
Aysén	Urbano	1110021	R.R. Aysén	61	12.528	24.466
Aysén	Rural	1110012	Conurbación Coyhaique	48	4.282	6.307
Aysén	Rural	1110022	R.R. Aysén	71	6.282	11.142
Magallanes	Urbano	1200011	Conurbación Punta Arenas + Cabo de Hornos + Porvenir	243	48.608	108.428
Magallanes	Urbano	1200021	Prov. Última Esperanza	39	7.820	15.427
Magallanes	Rural	1200012	Conurbación Punta Arenas + Laguna Blanca + Río Verde + San Gregorio	28	2.609	6.174
Magallanes	Rural	1200022	R.R. Magallanes	23	1.916	5.278

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

11.2 Anexo N° 2. Estratificación muestral del marco muestral de la ENE

Tabla 26. Estratificación muestral del marco muestral de la ENE

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
Arica y Parinacota	Urbano	1500001	15000011	108	21.395
Arica y Parinacota	Urbano	1500001	15000012	178	35.494
Arica y Parinacota	Urbano	1500001	15000013	45	9.037
Arica y Parinacota	Rural	1500012	15000120	66	5.94
Tarapacá	Urbano	110011	1100111	110	22.314
Tarapacá	Urbano	110011	1100112	244	49.622
Tarapacá	Urbano	110011	1100113	123	24.235
Tarapacá	Urbano	120011	1200110	27	5.37
Tarapacá	Rural	110012	1100120	13	1.027
Tarapacá	Rural	120012	1200120	40	3.71
Antofagasta	Urbano	210001	2100011	126	25.097
Antofagasta	Urbano	210001	2100012	275	54.597
Antofagasta	Urbano	210001	2100013	180	36.556
Antofagasta	Urbano	220011	2200110	263	54
Antofagasta	Urbano	230011	2300110	51	10.638
Antofagasta	Rural	210002	2100020	18	1.383
Antofagasta	Rural	220012	2200120	48	4.287
Atacama	Urbano	310011	3100111	115	23.084
Atacama	Urbano	310011	3100112	105	21.488
Atacama	Urbano	310011	3100113	59	11.72
Atacama	Urbano	310021	3100210	79	15.842
Atacama	Urbano	320011	3200110	104	20.8
Atacama	Rural	310012	3100120	33	2.876
Atacama	Rural	310022	3100220	26	2.147
Atacama	Rural	330012	3300120	80	7.086
Coquimbo	Urbano	410011	4100111	186	37.603
Coquimbo	Urbano	410011	4100112	392	78.497
Coquimbo	Urbano	410011	4100113	156	31.984
Coquimbo	Urbano	410021	4100210	53	10.134
Coquimbo	Urbano	420011	4200110	104	21.254
Coquimbo	Urbano	430011	4300111	48	9.664
Coquimbo	Urbano	430011	4300112	102	20.481
Coquimbo	Urbano	430021	4300210	51	10.124
Coquimbo	Rural	410012	4100120	131	11.955
Coquimbo	Rural	410022	4100220	87	7.896
Coquimbo	Rural	420012	4200120	150	13.567
Coquimbo	Rural	430012	4300120	112	10.042
Coquimbo	Rural	430022	4300220	166	14.859
Valparaíso	Urbano	510011	5100111	148	29.47
Valparaíso	Urbano	510011	5100112	290	58.419
Valparaíso	Urbano	510011	5100113	111	22.57

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
Valparaíso	Urbano	510021	5100211	90	18.227
Valparaíso	Urbano	510021	5100212	381	77.601
Valparaíso	Urbano	510021	5100213	257	52.307
Valparaíso	Urbano	510031	5100310	30	6.153
Valparaíso	Urbano	510041	5100410	79	16.383
Valparaíso	Urbano	530011	5300111	50	10.053
Valparaíso	Urbano	530011	5300112	108	22.502
Valparaíso	Urbano	540011	5400110	99	19.473
Valparaíso	Urbano	550011	5500111	87	17.775
Valparaíso	Urbano	550011	5500112	179	36.809
Valparaíso	Urbano	550011	5500113	38	8.164
Valparaíso	Urbano	560011	5600111	78	16.353
Valparaíso	Urbano	560011	5600112	140	28.948
Valparaíso	Urbano	560021	5600210	56	11.409
Valparaíso	Urbano	570011	5700110	199	40.942
Valparaíso	Urbano	580011	5800111	48	9.431
Valparaíso	Urbano	580011	5800112	365	74.022
Valparaíso	Urbano	580011	5800113	70	14.784
Valparaíso	Urbano	580021	5800210	89	18.21
Valparaíso	Rural	510012	5100120	33	2.841
Valparaíso	Rural	510022	5100220	38	3.492
Valparaíso	Rural	510032	5100320	33	2.939
Valparaíso	Rural	530012	5300120	76	6.493
Valparaíso	Rural	540012	5400120	106	9.825
Valparaíso	Rural	550012	5500120	104	9.26
Valparaíso	Rural	560012	5600120	72	6.582
Valparaíso	Rural	570012	5700120	155	13.762
Valparaíso	Rural	580022	5800220	50	4.421
Metropolitana	Urbano	1310011	13100111	108	22.402
Metropolitana	Urbano	1310011	13100112	642	131.764
Metropolitana	Urbano	1310011	13100113	294	60.733
Metropolitana	Urbano	1310021	13100210	1814	368.243
Metropolitana	Urbano	1310031	13100311	75	15.091
Metropolitana	Urbano	1310031	13100312	835	170.211
Metropolitana	Urbano	1310031	13100313	587	119.85
Metropolitana	Urbano	1310041	13100411	338	68.167
Metropolitana	Urbano	1310041	13100412	788	159.994
Metropolitana	Urbano	1310041	13100413	71	14.74
Metropolitana	Urbano	1310051	13100511	204	42.132
Metropolitana	Urbano	1310051	13100512	1538	313.245
Metropolitana	Urbano	1310051	13100513	480	97.227
Metropolitana	Urbano	1310061	13100610	963	191.494
Metropolitana	Urbano	1320011	13200111	62	12.733
Metropolitana	Urbano	1320011	13200112	588	119.951

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
Metropolitana	Urbano	1320011	13200113	177	37.472
Metropolitana	Urbano	1320021	13200210	35	7.124
Metropolitana	Urbano	1330011	13300111	37	7.928
Metropolitana	Urbano	1330011	13300112	99	20.614
Metropolitana	Urbano	1330011	13300113	54	10.814
Metropolitana	Urbano	1330021	13300210	128	26.2
Metropolitana	Urbano	1340011	13400111	92	18.429
Metropolitana	Urbano	1340011	13400112	292	60.013
Metropolitana	Urbano	1340011	13400113	61	13.193
Metropolitana	Urbano	1340021	13400211	48	9.847
Metropolitana	Urbano	1340021	13400212	149	30.922
Metropolitana	Urbano	1350011	13500111	54	10.681
Metropolitana	Urbano	1350011	13500112	138	27.617
Metropolitana	Urbano	1360011	13600111	63	13.157
Metropolitana	Urbano	1360011	13600112	188	39.476
Metropolitana	Urbano	1360011	13600113	46	9.885
Metropolitana	Urbano	1360021	13600210	87	18.152
Metropolitana	Rural	1310012	13100120	37	2.89
Metropolitana	Rural	1320022	13200220	76	6.724
Metropolitana	Rural	1330012	13300120	107	9.846
Metropolitana	Rural	1330022	13300220	74	6.78
Metropolitana	Rural	1340012	13400120	60	5.541
Metropolitana	Rural	1340022	13400220	135	12.682
Metropolitana	Rural	1350012	13500121	171	15.474
Metropolitana	Rural	1350012	13500122	105	9.477
Metropolitana	Rural	1360012	13600120	101	9.075
Metropolitana	Rural	1360022	13600220	53	4.968
O'Higgins	Urbano	6100111	6100111	116	23.91
O'Higgins	Urbano	6100111	6100112	364	76.037
O'Higgins	Urbano	6100111	6100113	117	24.524
O'Higgins	Urbano	6100211	6100210	47	9.394
O'Higgins	Urbano	6100311	6100310	255	50.988
O'Higgins	Urbano	6200111	6200110	113	23.774
O'Higgins	Urbano	6200211	6200210	47	9.33
O'Higgins	Urbano	6300111	6300110	137	28.124
O'Higgins	Rural	6100121	6100120	100	8.773
O'Higgins	Rural	6100221	6100220	41	3.575
O'Higgins	Rural	6100321	6100320	374	33.299
O'Higgins	Rural	6200121	6200120	40	3.664
O'Higgins	Rural	6200221	6200220	111	10.347
O'Higgins	Rural	6300121	6300121	224	20.574
O'Higgins	Rural	6300121	6300122	58	5.033
Maule	Urbano	7100111	7100111	95	19.412
Maule	Urbano	7100111	7100112	253	51.356

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
Maule	Urbano	710011	7100113	99	19.958
Maule	Urbano	710021	7100210	88	17.516
Maule	Urbano	710031	7100310	54	11.772
Maule	Urbano	720011	7200110	83	16.414
Maule	Urbano	730011	7300111	54	11.201
Maule	Urbano	730011	7300112	167	34.493
Maule	Urbano	730011	7300113	46	9.373
Maule	Urbano	730021	7300210	105	20.967
Maule	Urbano	740011	7400110	164	33.555
Maule	Urbano	740021	7400210	86	17.879
Maule	Urbano	740031	7400310	80	15.615
Maule	Rural	710012	7100120	79	7.208
Maule	Rural	710022	7100220	90	8.223
Maule	Rural	710032	7100320	166	15.483
Maule	Rural	720012	7200120	74	6.852
Maule	Rural	730012	7300120	166	14.954
Maule	Rural	730022	7300220	109	10.002
Maule	Rural	730032	7300320	42	3.801
Maule	Rural	740012	7400120	165	14.715
Maule	Rural	740022	7400220	189	17.072
Maule	Rural	740032	7400320	86	7.803
Ñuble	Urbano	1610011	16100111	78	16.165
Ñuble	Urbano	1610011	16100112	190	39.559
Ñuble	Urbano	1610011	16100113	68	14.62
Ñuble	Urbano	1610021	16100210	110	22.192
Ñuble	Urbano	1620011	16200110	53	10.996
Ñuble	Urbano	1630011	16300110	92	18.597
Ñuble	Rural	1610012	16100120	81	7.294
Ñuble	Rural	1610022	16100220	196	17.795
Ñuble	Rural	1620012	16200120	116	10.517
Ñuble	Rural	1630012	16300120	233	21.366
Biobío	Urbano	810011	8100111	171	34.372
Biobío	Urbano	810011	8100112	110	22.798
Biobío	Urbano	810021	8100211	159	32.852
Biobío	Urbano	810021	8100212	252	52.006
Biobío	Urbano	810021	8100213	36	7.243
Biobío	Urbano	810031	8100311	152	30.126
Biobío	Urbano	810031	8100312	458	93.403
Biobío	Urbano	810031	8100313	321	65.513
Biobío	Urbano	820011	8200111	146	30.444
Biobío	Urbano	820011	8200112	64	13.615
Biobío	Urbano	830011	8300111	81	16.354
Biobío	Urbano	830011	8300112	129	26.454
Biobío	Urbano	830011	8300113	54	11.547

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
Biobío	Urbano	830021	8300210	144	29.455
Biobío	Urbano	830031	8300310	95	19.479
Biobío	Rural	810012	8100120	128	11.465
Biobío	Rural	820012	8200120	156	14.204
Biobío	Rural	830012	8300120	203	18.385
Biobío	Rural	830022	8300220	131	11.717
Biobío	Rural	830032	8300320	119	10.574
La Araucanía	Urbano	910011	9100111	119	24.526
La Araucanía	Urbano	910011	9100112	288	59.39
La Araucanía	Urbano	910011	9100113	163	33.89
La Araucanía	Urbano	910021	9100210	159	32.45
La Araucanía	Urbano	910031	9100310	209	42.076
La Araucanía	Urbano	920011	9200110	100	20.314
La Araucanía	Urbano	920021	9200210	38	7.746
La Araucanía	Urbano	920031	9200310	133	27.646
La Araucanía	Rural	910012	9100120	244	21.407
La Araucanía	Rural	910022	9100220	400	36.468
La Araucanía	Rural	910032	9100320	273	24.667
La Araucanía	Rural	920012	9200120	223	20.018
Los Ríos	Urbano	1410011	14100111	55	10.799
Los Ríos	Urbano	1410011	14100112	154	31.01
Los Ríos	Urbano	1410011	14100113	66	13.203
Los Ríos	Urbano	1410021	14100210	69	13.891
Los Ríos	Urbano	1410031	14100310	50	9.857
Los Ríos	Urbano	1420011	14200110	53	10.704
Los Ríos	Urbano	1420021	14200210	48	9.801
Los Ríos	Rural	1410012	14100120	45	4.117
Los Ríos	Rural	1410022	14100220	141	12.499
Los Ríos	Rural	1410032	14100320	101	9.227
Los Ríos	Rural	1420012	14200120	124	11.383
Los Ríos	Rural	1420022	14200220	48	4.491
Los Lagos	Urbano	1010011	10100111	106	22.37
Los Lagos	Urbano	1010011	10100112	221	46.083
Los Lagos	Urbano	1010011	10100113	115	23.603
Los Lagos	Urbano	1010021	10100210	71	14.234
Los Lagos	Urbano	1010031	10100310	49	9.679
Los Lagos	Urbano	1020011	10200111	91	18.375
Los Lagos	Urbano	1020011	10200112	131	26.16
Los Lagos	Urbano	1020011	10200113	47	9.82
Los Lagos	Urbano	1020021	10200211	59	12.016
Los Lagos	Urbano	1020021	10200212	129	26.726
Los Lagos	Urbano	1030011	10300110	62	12.442
Los Lagos	Rural	1010012	10100121	57	5.263
Los Lagos	Rural	1010012	10100122	104	9.547

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
Los Lagos	Rural	1010022	10100220	152	14.208
Los Lagos	Rural	1010032	10100320	42	3.699
Los Lagos	Rural	1020012	10200120	52	4.921
Los Lagos	Rural	1020022	10200221	188	17.166
Los Lagos	Rural	1020022	10200222	97	8.667
Los Lagos	Rural	1030012	10300120	156	14.42
Aysén	Urbano	1110011	11100110	92	18.59
Aysén	Urbano	1110021	11100210	58	11.952
Aysén	Rural	1110012	11100120	48	4.282
Aysén	Rural	1110022	11100220	63	5.681
Magallanes	Urbano	1200011	12000112	185	37.054
Magallanes	Urbano	1200011	12000113	58	11.554
Magallanes	Urbano	1200021	12000210	39	7.82
Magallanes	Rural	1200012	12000120	27	2.52
Magallanes	Rural	1200022	12000220	22	1.846

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

11.3 Anexo N° 3. Códigos de visita de la ENE

Tabla 27. Códigos de registro de visita de la Tarjeta de Registro del Hogar empleada en el diseño muestral anterior de la ENE.

Código	Glosa descriptiva	
Lograda		
L 10	Entrevistada	
No lograda, Tipo "A"		
A1 11	Moradores ausentes	
A2 12	Rechazo	
A3 13	Inaccesible	
A4 14	Casos Especiales	
No lograda, Tipo "B"		
B1 21	Desocupada temporalmente	
B2 22	En reparación, ampliación o remodelación	
B3 23	Exclusión transitoria	
B4 24	De temporada	
No lograda, Tipo "C"		
C1 31	Demolida	
C2 32	Otro uso	
C3 33	En muestra	
C5 35	Mal tomada	

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Tabla 28. Códigos de visita o disposición de la ENE 2021.

Estado de elegibilidad	Tipo de contacto	Código	Glosa descriptiva
Elegibles, entrevistadas	Presencial y Telefónico	11	Entrevistada completa
		12	Entrevistada parcial
Elegibles, rechazadas	Presencial y Telefónico	21	Rechazo
Elegibles, no contactadas	Presencial	31	Vivienda ocupada, sin moradores presentes
		32	Se impide el acceso a la vivienda
	Presencial y Telefónico	33	Informante no ubicable, se contacta a otra persona
		34	Informante ausente en período de levantamiento de la encuesta
	Telefónico	71	Marca tono, pero sin respuesta
		72	Responde buzón de voz
		73	Suena ocupado
		74	Se encuentra apagado o fuera del área de cobertura
		75	Número telefónico no existe
		76	Número telefónico equivocado
		77	Sin número telefónico
		78	No se realizó la llamada
		79	Otra razón de no contacto (Especifique)

Elegible, no entrevistadas por otros motivos	Presencial y Telefónico	41	Informante impedido física y/o mentalmente para contestar
		42	Informante no habla idioma español
		43	Muerte del informante o algún familiar
		44	Concertó cita con el informante
	Telefónico	81	Calidad de audio mala o intermitente
		82	Lugar o actividad del informante no permite una entrevista
	Presencial y Telefónico	49	Otra razón de no entrevista (Especifique)
	Presencial	51	No se envió a terreno
		52	No fue posible localizar la dirección
		53	Inaccesibilidad por razones climáticas o desastres naturales
		54	Área peligrosa
	Telefónica	91	Informante/hogar ya no reside en la vivienda
	Presencial	59	Otra razón de elegibilidad desconocida (Especifique)
	Presencial y Telefónico	61	Establecimiento, oficina de gobierno u otra organización
		62	Institución (por ejemplo: hospital, cárcel, asilo de ancianos, etc.)
		63	Dormitorio colectivo (por ejemplo: militar, de trabajo, internado, hostales, residenciales, etc.)
		64	Vivienda en demolición, incendiada, destruida o erradicada
		65	Vivienda desocupada (por ejemplo en venta, arriendo, etc.)
		66	Vivienda de veraneo o de uso temporal
		67	Edificación no calificable (por ejemplo: en construcción, en reparaciones, etc.)
		69	Otra razón para ser considerada no elegible (Especifique)

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

11.4 Anexo N° 4. Tasas de no logro por estrato geográfico, según información de la ENE 2018

Tabla 29. Tasas de no logro por estrato geográfico, según información de la ENE 2018

Región	Área	Código de estrato geográfico	Tasas de no logro	Tasas de no logro (Truncada a 30%)
Arica y Parinacota	Urbano	1500001	22%	22%
Arica y Parinacota	Rural	1500012	33%	30%
Tarapacá	Urbano	120011	22%	22%
Tarapacá	Rural	110011	16%	16%
Tarapacá	Rural	120012	30%	30%
Tarapacá	Rural	110012	33%	30%
Antofagasta	Urbano	210001	23%	23%
Antofagasta	Urbano	220011	20%	20%
Antofagasta	Urbano	230011	31%	30%
Antofagasta	Rural	220012	23%	23%
Antofagasta	Rural	210002	23%	23%
Atacama	Urbano	310021	33%	30%
Atacama	Urbano	310011	27%	27%
Atacama	Urbano	320011	27%	27%
Atacama	Rural	310022	45%	30%
Atacama	Rural	310012	36%	30%
Atacama	Rural	330012	48%	30%
Coquimbo	Urbano	420011	31%	30%
Coquimbo	Urbano	430011	25%	25%
Coquimbo	Urbano	430021	26%	26%
Coquimbo	Urbano	410021	37%	30%
Coquimbo	Urbano	410011	36%	30%
Coquimbo	Rural	430012	25%	25%
Coquimbo	Rural	420012	29%	29%
Coquimbo	Rural	410012	35%	30%
Coquimbo	Rural	430022	26%	26%
Coquimbo	Rural	410022	36%	30%
Metropolitana	Urbano	1310061	35%	30%
Metropolitana	Urbano	1320011	24%	24%
Metropolitana	Urbano	1310051	25%	25%
Metropolitana	Urbano	1310041	28%	28%
Metropolitana	Urbano	1310031	25%	25%
Metropolitana	Urbano	1310021	39%	30%
Metropolitana	Urbano	1310011	31%	30%
Metropolitana	Urbano	1360021	25%	25%
Metropolitana	Urbano	1360011	28%	28%
Metropolitana	Urbano	1350011	23%	23%
Metropolitana	Urbano	1340021	23%	23%
Metropolitana	Urbano	1340011	31%	30%
Metropolitana	Urbano	1330021	25%	25%
Metropolitana	Urbano	1330011	28%	28%

Región	Área	Código de estrato geográfico	Tasas de no logro	Tasas de no logro (Truncada a 30%)
Metropolitana	Urbano	1320021	36%	30%
Metropolitana	Rural	1310012	36%	30%
Metropolitana	Rural	1360022	36%	30%
Metropolitana	Rural	1360012	36%	30%
Metropolitana	Rural	1350012	50%	30%
Metropolitana	Rural	1340022	36%	30%
Metropolitana	Rural	1340012	36%	30%
Metropolitana	Rural	1330022	23%	23%
Metropolitana	Rural	1330012	37%	30%
Metropolitana	Rural	1320022	36%	30%
Valparaíso	Urbano	510031	21%	21%
Valparaíso	Urbano	510021	27%	27%
Valparaíso	Urbano	510011	25%	25%
Valparaíso	Urbano	540011	41%	30%
Valparaíso	Urbano	580021	25%	25%
Valparaíso	Urbano	580011	17%	17%
Valparaíso	Urbano	560021	39%	30%
Valparaíso	Urbano	570011	18%	18%
Valparaíso	Urbano	550011	18%	18%
Valparaíso	Urbano	530011	19%	19%
Valparaíso	Urbano	560011	18%	18%
Valparaíso	Urbano	510041	25%	25%
Valparaíso	Rural	510032	27%	27%
Valparaíso	Rural	530012	29%	29%
Valparaíso	Rural	510022	28%	28%
Valparaíso	Rural	510012	27%	27%
Valparaíso	Rural	580022	25%	25%
Valparaíso	Rural	570012	37%	30%
Valparaíso	Rural	560012	26%	26%
Valparaíso	Rural	550012	26%	26%
Valparaíso	Rural	540012	25%	25%
O'Higgins	Urbano	630011	23%	23%
O'Higgins	Urbano	620011	20%	20%
O'Higgins	Urbano	610031	17%	17%
O'Higgins	Urbano	610021	23%	23%
O'Higgins	Urbano	620021	22%	22%
O'Higgins	Urbano	610011	24%	24%
O'Higgins	Rural	630012	37%	30%
O'Higgins	Rural	620022	20%	20%
O'Higgins	Rural	620012	17%	17%
O'Higgins	Rural	610032	17%	17%
O'Higgins	Rural	610022	20%	20%
O'Higgins	Rural	610012	39%	30%
Maule	Urbano	710031	32%	30%

Región	Área	Código de estrato geográfico	Tasas de no logro	Tasas de no logro (Truncada a 30%)
Maule	Urbano	710021	43%	30%
Maule	Urbano	710011	36%	30%
Maule	Urbano	740031	37%	30%
Maule	Urbano	740021	35%	30%
Maule	Urbano	740011	25%	25%
Maule	Urbano	730021	36%	30%
Maule	Urbano	730011	40%	30%
Maule	Urbano	720011	31%	30%
Maule	Rural	710022	26%	26%
Maule	Rural	710012	23%	23%
Maule	Rural	740032	27%	27%
Maule	Rural	740022	15%	15%
Maule	Rural	740012	14%	14%
Maule	Rural	730022	13%	13%
Maule	Rural	730012	23%	23%
Maule	Rural	730032	17%	17%
Maule	Rural	720012	18%	18%
Maule	Rural	710032	15%	15%
Ñuble	Urbano	1630011	22%	22%
Ñuble	Urbano	1620011	26%	26%
Ñuble	Urbano	1610021	32%	30%
Ñuble	Urbano	1610011	23%	23%
Ñuble	Rural	1630012	19%	19%
Ñuble	Rural	1620012	21%	21%
Ñuble	Rural	1610022	21%	21%
Ñuble	Rural	1610012	30%	30%
Biobío	Urbano	830031	30%	30%
Biobío	Urbano	830021	37%	30%
Biobío	Urbano	830011	26%	26%
Biobío	Urbano	820011	25%	25%
Biobío	Urbano	810031	26%	26%
Biobío	Urbano	810021	22%	22%
Biobío	Urbano	810011	18%	18%
Biobío	Rural	830032	35%	30%
Biobío	Rural	830022	29%	29%
Biobío	Rural	830012	39%	30%
Biobío	Rural	820012	48%	30%
Biobío	Rural	810012	57%	30%
La Araucanía	Urbano	920031	27%	27%
La Araucanía	Urbano	920021	24%	24%
La Araucanía	Urbano	920011	24%	24%
La Araucanía	Urbano	910031	23%	23%
La Araucanía	Urbano	910021	26%	26%
La Araucanía	Urbano	910011	29%	29%

Región	Área	Código de estrato geográfico	Tasas de no logro	Tasas de no logro (Truncada a 30%)
La Araucanía	Rural	920012	15%	15%
La Araucanía	Rural	910032	15%	15%
La Araucanía	Rural	910022	22%	22%
La Araucanía	Rural	910012	17%	17%
Los Ríos	Urbano	1420021	30%	30%
Los Ríos	Urbano	1420011	30%	30%
Los Ríos	Urbano	1410031	30%	30%
Los Ríos	Urbano	1410021	35%	30%
Los Ríos	Urbano	1410011	30%	30%
Los Ríos	Rural	1420022	23%	23%
Los Ríos	Rural	1420012	30%	30%
Los Ríos	Rural	1410032	32%	30%
Los Ríos	Rural	1410022	19%	19%
Los Ríos	Rural	1410012	31%	30%
Los Lagos	Urbano	1030011	39%	30%
Los Lagos	Urbano	1020021	35%	30%
Los Lagos	Urbano	1020011	29%	29%
Los Lagos	Urbano	1010031	32%	30%
Los Lagos	Urbano	1010021	26%	26%
Los Lagos	Urbano	1010011	29%	29%
Los Lagos	Rural	1030012	23%	23%
Los Lagos	Rural	1020022	29%	29%
Los Lagos	Rural	1020012	32%	30%
Los Lagos	Rural	1010032	42%	30%
Los Lagos	Rural	1010022	23%	23%
Los Lagos	Rural	1010012	27%	27%
Aysén	Urbano	1110021	34%	30%
Aysén	Urbano	1110011	27%	27%
Aysén	Rural	1110022	45%	30%
Aysén	Rural	1110012	31%	30%
Magallanes	Urbano	1200021	21%	21%
Magallanes	Urbano	1200011	13%	13%
Magallanes	Rural	1200022	12%	12%
Magallanes	Rural	1200012	12%	12%

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

11.5 Anexo N° 5. Estimaciones de población para la Encuesta Nacional de Empleo

11.5.1 Informe estimaciones de población para la Encuesta Nacional de Empleo

El presente documento describe la metodología a través de la cual se realizaron las estimaciones de población mensuales para la Encuesta Nacional de Empleo (ENE). Para el cálculo de los indicadores de la ENE se requiere contar con estimaciones de población a nivel de comunas por sexo y edad simple, y a nivel de regiones urbano/rural, también por sexo y edad simple. Para realizar estas estimaciones de población se realizó una interpolación de las Proyecciones de Población comunales anuales centradas el día 30 de junio, a nivel de sexo y edad simple, obteniendo estimaciones de población al día 15 de cada mes. Con estas estimaciones comunales se obtuvieron, por suma, las estimaciones regionales por sexo y edad simple. Por otra parte, se realizó una estimación de la población regional urbana, por sexo y edad simple, al día 15 de cada mes, interpolando las Proyecciones de Población regionales urbanas, obteniéndose la estimación de la población regional rural por diferencia entre la población urbana y el total de la población regional anteriormente obtenida por la suma de las comunas. A continuación, se especificarán los diferentes pasos seguidos en esta metodología.

11.5.2 Interpolación de población comunal a los 15 de cada mes.

Las Proyecciones de Población son calculadas para los días 30 de junio de cada año. Para obtener estimaciones de población para otras fechas, se realiza una interpolación utilizando como pivotes las proyecciones de los años que circunscriben dicha fecha. La encuesta de empleo es continua, por tanto, para obtener las estimaciones de población al día 15 de cada mes, se interpolan sucesivamente las Proyecciones de Población de dos años continuos, obteniéndose sucesivamente estimaciones de población entre el 15 de julio y el 15 de junio del año siguiente. El uso de los pivotes asegura que las estimaciones no se alejen de las Proyecciones de Población oficiales publicadas por el INE.

Dado el requerimiento para el cálculo de los indicadores de la ENE, la interpolación realizada entre los sucesivos pivotes fue una interpolación exponencial por sexo y edad simple. Con la interpolación exponencial inicialmente se obtuvieron 11 estimaciones a los días 30 de cada mes entre las dos poblaciones pivotes. Las estimaciones obtenidas se redondearon a la parte entera para obtener poblaciones sin decimales, y luego se aplicaron promedios simples entre los días 30 de cada mes consecutivo para obtener estimaciones al día 15 de cada mes, fecha que necesita la ENE para calcular los indicadores de cada trimestre móvil. Los días 15 de cada mes obtenidos a su vez fueron, alternadamente en cada grupo de edad, redondeados y truncados para evitar que el promedio se acercara a alguno de los dos meses promediados.

$$P_f = P_i \cdot e^{r \cdot \Delta t'}$$

$$r = \ln \left(\frac{P_f}{P_i} \right) \cdot \frac{1}{\Delta t}$$

Donde:

P_f : Población pivote final.

P_i : Población pivote inicial.

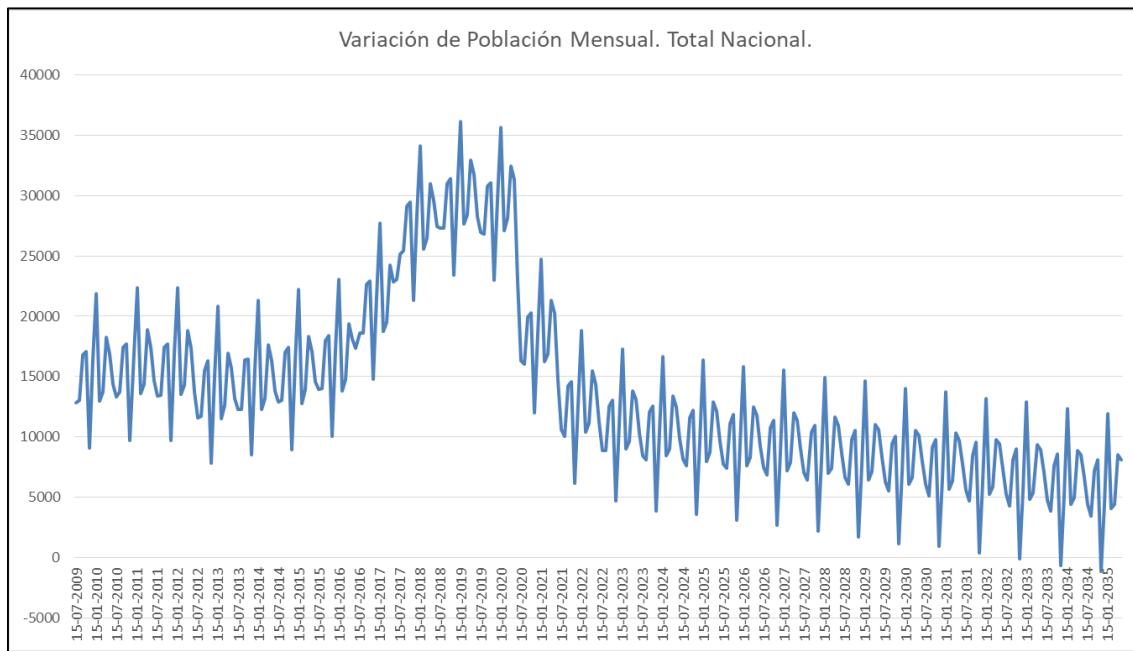
r : Tasa de crecimiento anual.

Δt : Intervalo temporal entre poblaciones pivote.

$\Delta t'$: Intervalo temporal entre población pivote y la población a estimar

Dado que la interpolación exponencial se realizó a nivel de las edades simples, en muchos casos las diferencias anuales entre las poblaciones pivotes eran muy pequeñas. Esto provocó, al ser llevada la serie a números enteros, que los incrementos (o reducciones) de población se produjeran sistemáticamente en determinados meses. Por ejemplo, si el incremento de población entre dos pivotes era de dos personas, sistemáticamente la primera adición se producía en el mes de octubre y la segunda en el mes de abril. Esto provocaba, a nivel del total nacional, regional y comunal, que el crecimiento mensual de la población fuera significativamente irregular entre los meses de un año, irregularidad que se repetía sistemáticamente de año en año.

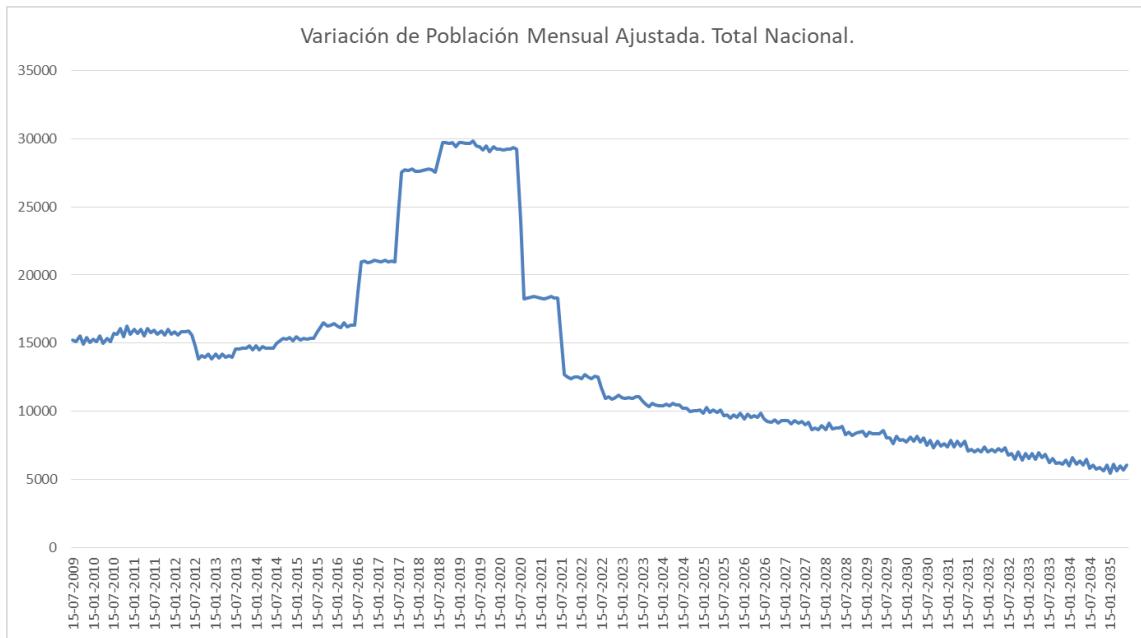
Figura 17. Variación de Población Mensual. Total Nacional



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Para eliminar estas irregularidades se optó por redistribuir los incrementos de población mensuales. Para ello se utilizó una tabla sistemática que redistribuyó los crecimientos mensuales de población en los meses que señalaba la tabla para cada edad específica. Con ello se evitaron las fluctuaciones bruscas de los crecimientos de población que se observaban entre los meses de una misma interpolación anual.

Figura 18. Variación de Población Mensual Ajustada. Total Nacional.



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

De esta manera, si bien los cambios anuales no pueden ser suavizados ya que provienen de la Proyección de Población nacional, se evitaron las fluctuaciones en los incrementos de población mensual observadas inicialmente. Con estas poblaciones por sexo y edad se conformaron las comunas, y con la suma de éstas las regiones.

En paralelo, utilizando la misma metodología, se calcularon las estimaciones de población de las regiones urbanas por sexo y edad simple. Para ello, se tomaron como pivotes las Proyecciones de Población regionales urbanas publicadas por el INE, y se siguió el mismo procedimiento anteriormente descrito para la estimación de la población comunal. La población rural a nivel regional se obtuvo por la diferencia entre la población regional total (por sexo y edad simple) y la población regional urbana.

11.6 Anexo N° 6. Listado de marginales provinciales

Tabla 30. Listado de marginales provinciales

Región	Código de provincia	Provincia	Proyección de Población PET, Octubre 2019	Trimestre	Año de ingreso
Tarapacá	11	Iquique	265.356	EFM	2010
Antofagasta	21	Antofagasta	352.067	EFM	2010
Antofagasta	22	El Loa	152.521	EFM	2010
Atacama	31	Copiapó	157.71	EFM	2010
Coquimbo	41	Elqui	435.302	EFM	2010
Coquimbo	43	Limarí	145.175	EFM	2010
Valparaíso	51	Valparaíso	665.584	EFM	2010
Valparaíso	56	San Antonio	145.215	EFM	2010
Valparaíso	57	San Felipe	131.504	EFM	2010
Valparaíso	58	Marga Marga	302.039	EFM	2010
Metropolitana	131	Santiago	4.845.166	EFM	2010
Metropolitana	132	Cordillera	539.844	EFM	2010
Metropolitana	134	Maipo	428.787	EFM	2010
Metropolitana	135	Melipilla	167.073	EFM	2010
Metropolitana	136	Talagante	261.044	EFM	2010
O'Higgins	61	Cachapoal	556.719	EFM	2010
Maulle	71	Talca	357.162	EFM	2010
Maulle	73	Curicó	249.679	EFM	2010
Maulle	74	Linares	245.335	EFM	2010
Ñuble	161	Diguillín	275.163	ASO	2018
Biobío	81	Concepción	864.48	EFM	2010
Biobío	82	Arauco	138.379	EFM	2010
Biobío	83	Biobío	335.921	EFM	2010
La Araucanía	91	Cautín	635.515	EFM	2010
Los Ríos	141	Valdivia	248.062	EFM	2010
Los Lagos	101	Llanquihue	350.096	EFM	2010
Los Lagos	103	Osorno	200.285	EFM	2010
Aysén	111	Coyhaique	48.022	EFM	2010
Magallanes	121	Magallanes	115.869	EFM	2010

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

11.7 Anexo N° 7. Cálculo de los factores de expansión en el período de transición

De acuerdo con lo expuesto en el apartado [“4.8.2.4 Incorporación de la muestra del MMV 2017”](#) de este documento, entre enero y mayo 2020 la muestra de la ENE atraviesa un proceso de transición, durante el cual están en funcionamiento dos muestras independientes, la muestra anterior (donde las secciones provienen del MMS 2016 y las manzanas del MMM 2016) y la muestra nueva proveniente el MMV 2017.

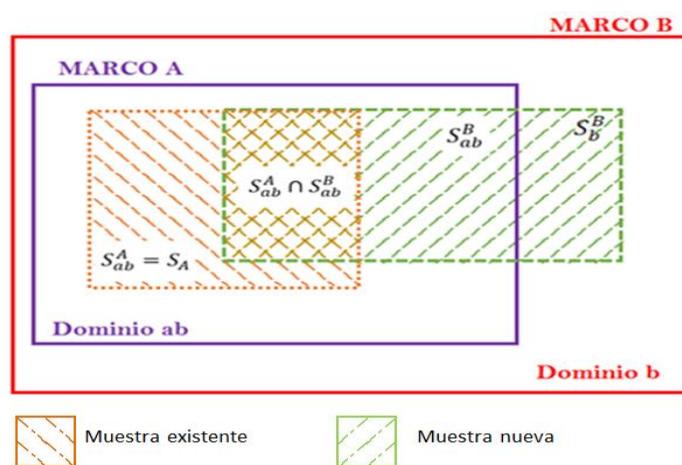
Para efectos de nomenclatura denominaremos marco “A” a la unión de los listados MMS 2016 y MMM 2016 y marco “B” al MMV 2017.

En el primer trimestre de la transición, se conserva aproximadamente 5/6 de las viviendas de la muestra anterior, y se incorpora 1/6 de la muestra nueva. Como consecuencia, los pesos de diseño calculados de la muestra nueva provenientes del marco B, serán más grandes que los pesos de muestra completa. Por lo tanto, si combináramos la muestra anterior y la muestra nueva, obtendríamos una cifra de aproximadamente el doble de la población⁷³.

La utilización simultánea de varias muestras extraídas de distintos marcos para estimar la población objetivo en un dominio o estrato, se denomina estimación de marcos múltiples, donde uno o varios marcos pueden tener cobertura incompleta, es decir, pueden existir dominios que son cubiertos por tan solo un marco (Hartley, 1962). Sin embargo, el desafío de las estimaciones de marcos múltiples se presenta en los dominios donde los marcos se solapan.

En forma práctica, como se observa en la Figura 19 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 que ahora fueron incorporadas al marco B.

Figura 19. Esquema de configuración de marcos y muestras



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

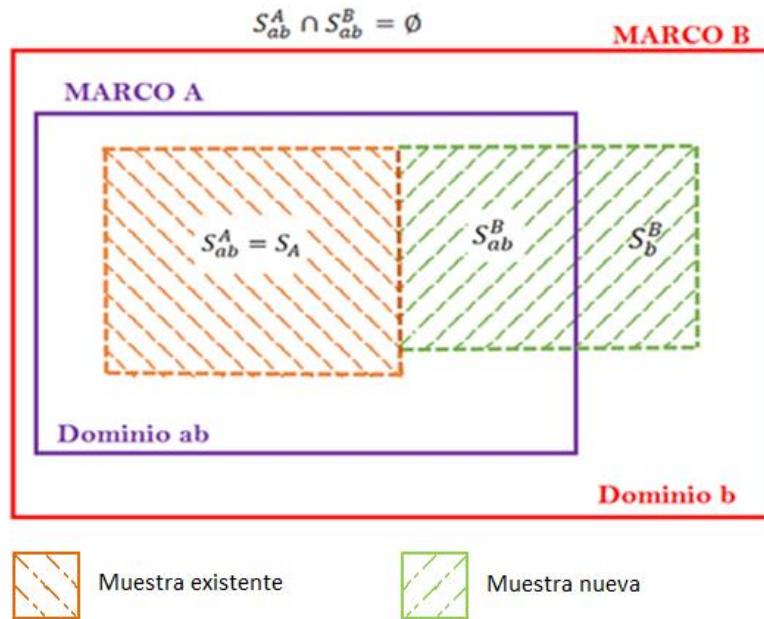
⁷³ Suponiendo que no haya falta de respuesta o cobertura insuficiente.

Si se extraen muestras independientes en cada estrato se tiene que: S_A es la muestra extraída del marco A (muestra anterior), que se puede descomponer en $S_A = S_a^A \cup S_{ab}^A = S_{ab}^A$, con $S_a^A = \emptyset$ al no tener el marco A dominios que no cubra el marco B. La muestra extraída del marco B se denomina $S_B = S_b^B \cup S_{ab}^B$, siendo S_b^B la muestra extraída del marco B que corresponde al dominio b o dominios que no son cubiertos por el marco A al ser excluidos (ADA).

En el dominio común (ab) se puede producir una intersección de las muestras $S_{ab}^A \cap S_{ab}^B$, debido a que, por un lado, la selección de viviendas se realiza de manera independiente en cada marco, y por el otro existe solapamiento, parcial o total de estas unidades. Al respecto, como estrategia para evitar el doble levantamiento de una vivienda teniendo en cuenta que la muestra del Marco A (actual) ya fue seleccionada, no se levanta la vivienda para la muestra del marco B, cuya perdida muestral se espera sea cubierta por el sobremuestreo⁷⁴.

De esta manera, el conjunto de viviendas duplicadas debido a las selecciones independientes sería vacío $S_{ab}^A \cap S_{ab}^B = \emptyset$ y el esquema de la Figura 19 se transforma al esquema de la Figura 20.

Figura 20: Esquema actualizado de configuración de marcos y muestras



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Es en el dominio ab donde se hace pertinente ajustar los factores de expansión, dado que para generar estimaciones se tienen unidades muestrales provenientes de marcos distintos S_{ab}^A y S_{ab}^B .

⁷⁴ Conforme a las previsiones tomadas para determinar la fracción de la muestra anterior que debe ser desincorporada en cada turno, se espera que los casos donde se presente un doble levantamiento sean mínimos.

Como los ponderadores de selección de viviendas apuntan al stock de viviendas de cada marco en el dominio común, esto es, a M_{ab}^A en el caso del marco A y a M_{ab}^B en el caso del marco B, y si $M_{ab}^A \approx M_{ab}^B$, la suma de los factores de viviendas de ambas muestras⁷⁵ duplican aproximadamente el stock de viviendas de uno u otro marco.

Para remediar esto se realiza un ajuste al ponderador de selección de viviendas ajustado por no respuesta, en función de la proporción de cada una de las muestras (anterior y nueva) en la muestra total en cada trimestre de la transición, lo que queda expresado de la siguiente forma⁷⁶:

$$w_{lij}^{MM} = \begin{cases} w_{h_{Aij}}^{NR} \cdot \alpha & , \text{si la muestra proviene del marco A (muestra anterior)} \\ w_{h_{Bij}}^{NR} \cdot (1 - \alpha) & , \text{si la muestra proviene del marco B (muestra nueva)} \end{cases}$$

siendo, $\alpha = \frac{m_A}{m_A + m_B}$ y $L = \{h_A \text{ y } h_B \in \text{al dominio común ab}\}$ (30)

Donde:

- w_{lij}^{MM} : Ponderador de selección de la vivienda j dentro del conglomerado i en el dominio común L , ajustado por Marcos Múltiples (MM).
- $w_{h_{Aij}}^{NR}$: Ponderador de selección de la vivienda j dentro del conglomerado i en el estrato h_A , ajustado por no respuesta de la muestra anterior.
- $w_{h_{Bij}}^{NR}$: Ponderador de selección de la vivienda j dentro del conglomerado i en el estrato h_B , ajustado por no respuesta de la muestra nueva.
- m_A : Número de viviendas logradas de la muestra anterior.
- m_B : Número de viviendas logradas de la muestra nueva.
- h_A, h_B : estratos definidos en Marco B (nuevo); h_A : estratos definidos en Marco A (actual).

Es importante mencionar que los ponderadores de selección de la muestra anterior, se calculan en función a lo definido en el diseño muestral anterior, en vista de que durante los primeros tres trimestres móviles del período de transición no se dispone de los insumos para realizar los ajustes por elegibilidad de la totalidad de la muestra proveniente del Marco A. Esto, se explica también porque el diseño muestral anterior no consideraba sobremuestreo⁷⁷ y la nueva clasificación de códigos de disposición final entrará en vigor a partir de enero 2020.

⁷⁵ Si se diera que $M_{ab}^A = M_{ab}^B$, entonces, en la primera incorporación de muestra nueva (o la primera desincorporación de muestra anterior) la suma de factores de viviendas de ambas muestras (5/6 de la muestra anterior y 1/6 de la muestra nueva) duplicaría exactamente el stock de viviendas del Marco (A o B).

⁷⁶ Esta metodología fue implementada en la Encuesta de Fuerza Laboral de Nueva Zelanda (Statistics New Zealand, 2015) con el ponderador a nivel de personas y se demostró que la expresión α , minimiza la varianza para la estimación de los totales.

⁷⁷ Efectivamente, el diseño muestral anterior consideraba reemplazo de viviendas y no sobremuestreo. Además, la clasificación de códigos de visita empleado en la Tarjeta de Registro del Hogar (TRH) no seguía lineamientos internacionales, por lo que se basaba en una clasificación *ad-hoc* de los grupos, mediante categorías de *logro* o *no logro*.

Por tanto, el cálculo del factor de expansión teórico asociado al conglomerado i en el estrato h_A , de la muestra ENE anterior, se calcula de acuerdo con la siguiente fórmula:

$$w_{h_Ai} = \frac{M_{h_A}}{n_{h_A}^{sel} \cdot M_{h_Ai}} \cdot \frac{M'_{h_Ai}}{m_{h_Ai}^{sel}} \quad (31)$$

Donde:

- M_{h_A} : Número de viviendas particulares en el estrato h_A al censo correspondiente⁷⁸, o al año de actualización del marco.
- M'_{h_Ai} : Número de viviendas particulares ocupadas, según actualización de los listados de viviendas, en la unidad de primera etapa i en el estrato h_A .
- M_{h_Ai} : Número de viviendas particulares en la unidad de primera etapa i en el estrato h_A al censo correspondiente o al año de actualización del marco.
- $n_{h_A}^{sel}$: Número de unidades de primera etapa seleccionadas en el estrato h_A .
- $m_{h_Ai}^{sel}$: Número de viviendas seleccionadas en la unidad de primera etapa i en el estrato h_A .

Luego, para dar cuenta de la calidad del levantamiento en terreno y ajustar el factor de expansión a la respuesta obtenida efectivamente en terreno, se ajusta por la no respuesta de los conglomerados y de las viviendas logradas, por lo que se calculan los siguientes dos ponderadores de ajuste:

1. **Ponderador para conglomerados (o unidades de primera etapa):** Razón entre el número de conglomerados seleccionados y el número de conglomerados efectivamente logrados.

$$\text{Ponderador conglomerado} = \frac{n_{h_A}^{sel}}{n_{h_A}} \quad (32)$$

Donde:

- $n_{h_A}^{sel}$: Número de unidades de primera etapa seleccionadas en el estrato h_A .
- n_{h_A} : Número de unidades de primera etapa logradas en el estrato h_A .

2. **Ponderador para viviendas:** razón entre el número de viviendas seleccionadas y el número de viviendas efectivamente logradas.

$$\text{Ponderador vivienda} = \frac{m_{h_Ai}^{sel}}{m_{h_Ai}} \quad (33)$$

⁷⁸ Se debe recordar que durante el periodo de incorporación de la nueva muestra y de desincorporación de la muestra anterior, la muestra total proviene de distintos marcos provenientes de censos diferentes, con distintas unidades de primera etapa y diferente estratificación.

Donde:

$m_{h_Ai}^{sel}$: Número de viviendas seleccionadas en la unidad de primera etapa i en el estrato h_A .

m_{h_Ai} : Número de viviendas logradas en la unidad de primera etapa i en el estrato h_A .

Así, el factor de expansión es ajustado por no respuesta mediante la siguiente fórmula:

$$w_{h_Aij} = w_{h_Ai} \cdot \frac{n_{h_A}^{sel} \cdot m_{h_Ai}^{sel}}{n_{h_A} \cdot m_{h_Ai}} = \frac{M_{h_A}}{n_{h_A}^{sel} \cdot M_{h_Ai}} \cdot \frac{M'_{h_Ai}}{m_{h_Ai}^{sel}} \cdot \frac{n_{h_A}^{sel} \cdot m_{h_Ai}^{sel}}{n_{h_A} \cdot m_{h_Ai}} = \frac{M_{h_A}}{n_{h_A} \cdot M_{h_Ai}} \cdot \frac{M'_{h_Ai}}{m_{h_Ai}} \quad (34)$$

Donde,

M_{h_A} : Número de viviendas particulares en el estrato h_A al censo correspondiente, o al año de actualización del marco.

M'_{h_Ai} : Número de viviendas particulares ocupadas, según actualización de los listados de viviendas, en la unidad de primera etapa i en el estrato h_A .

M_{h_Ai} : Número de viviendas particulares en la unidad de primera etapa i en el estrato h_A al censo correspondiente o al año de actualización del marco.

$n_{h_A}^{sel}$: Número de unidades de primera etapa seleccionadas en el estrato h_A .

$m_{h_Ai}^{sel}$: Número de viviendas seleccionadas en la unidad de primera etapa i en el estrato h_A .

n_{h_A} : Número de unidades de primera etapa logradas en el estrato h_A .

m_{h_Ai} : Número de viviendas logradas en la unidad de primera etapa i en el estrato h_A .

Como se puede observar, el factor de expansión base, queda en función del número de conglomerados y viviendas logradas (n_{h_A} ; m_{h_Ai}) y no de las seleccionadas inicialmente ($n_{h_A}^{sel}$; $m_{h_Ai}^{sel}$). Este factor lleva implícitamente los dos ajustes mencionados, análogo al ajuste de no-respuesta aplicado a los factores del Marco B (ver apartado 4.8.3). Entonces,

$$w_{h_Aij} = w_{h_Aij}^{NR} \quad (35)$$

Finalmente, el ponderador calibrado por proyecciones de población ($w_{h_Aij}^{cal}$), durante el período de transición se obtiene a partir del ponderador w_{lij}^{MM} (ponderador de selección ajustado por marcos múltiples).

11.8 Anexo N° 8. Construcción de insumo para la calibración por población extranjera en factores de expansión de la ENE

La llegada de población nacida en el extranjero a Chile ha aumentado sostenidamente en el último tiempo, pasando de representar el 0,8% de la población en 1992, al 1,3% en el censo de 2002 y el 4,4% en el Censo de 2017. En particular, este último censo permite observar que el 66,7% del *stock* de inmigrantes internacionales llegó al país en el período 2010-2017, concentrándose principalmente en los años 2016 y 2017 (INE, 2018c). La estimación de extranjeros residentes en Chile (INE-DEM, 2019) dio como uno de los principales resultados que el porcentaje de población extranjera con respecto a la población del país era de un 6,6⁷⁹ al 31 de diciembre de 2018, y se espera que este porcentaje aumente para el 2019.

Al ser una población más joven que la población nacida en Chile, en el censo se aprecia que alcanza a un 4,9% de la población en edad de trabajar, y un 6,5% de la fuerza de trabajo primaria (INE, 2018c), mientras que al 2018 correspondía a un 7,5% y a un 10,1% de la población en edad de trabajar y a la fuerza de trabajo primaria respectivamente⁸⁰.

Además de ser una población más joven, los nacidos en el extranjero se relacionan de manera distinta que los nacidos en Chile con el trabajo. Una evidencia de ello es que, para el Censo de 2017, el 74,1% de los primeros declaran haber trabajado la semana anterior al censo, mientras que para los segundos el porcentaje que declaró trabajar fue de un 56,9%⁸¹ (INE, 2018c).

Dada la relevancia que ha adquirido la inmigración y el impacto que ha tenido, tiene y tendrá este fenómeno sobre los diversos indicadores que miden las encuestas de hogares, en particular, sobre los del mercado laboral y la necesidad de contar con información mensual sobre la cantidad de personas extranjeras que residen en Chile, se hizo necesario contar con una estimación para el *stock* de personas extranjeras residentes en Chile, con periodicidad mensual, de modo que esta información sea un *benchmark* a incorporar como parte de la información exógena que será utilizada en el ajuste o calibración de los factores de expansión de las encuestas de hogares que lleva adelante el INE y, particularmente, como un insumo de la nueva metodología de calibración de los factores de expansión de la ENE.

Para lograr este objetivo, el Subdepartamento de Investigación Estadística, con la colaboración del Subdepartamento de Demografía, ha elaborado un modelo de cálculo de población extranjera

⁷⁹ En rigor, el porcentaje señalado no es a un porcentaje propiamente tal, ya que la relación entre ambas poblaciones no corresponde a una proporción sino a una razón. Esto es así, ya que las Estimaciones y Proyecciones de Población sólo distinguen el sexo y la edad de la población, no la nacionalidad ni cualquier otra variable.

⁸⁰ En rigor los porcentajes de 2018 corresponden a una razón cuyo numerador de la estimación al 31 de diciembre de 2018 y el denominador a la estimación y proyección al 30 de junio de 2018,

⁸¹ Se presenta una diferencia por sexo, ya que de los nacidos en Chile el porcentaje es de 69,0% en el caso de los hombres y un 45,6% para las mujeres, mientras que los nacidos en otro país el porcentaje de hombres corresponde a 82,5% y para las mujeres es de un 66%.

mensualizada, a partir del cual sea posible incorporar este componente en la calibración de los factores de expansión de la ENE.

11.8.1 Metodología para la obtención de cifras mensuales de población extranjera

La fuente de información para medir la migración internacional más usada en Chile han sido los censos de población y vivienda, los que consideran como residentes habituales a las personas que residen o tienen intenciones de residir 6 meses o más. No obstante, se reconocen como fuentes para el análisis demográfico los registros administrativos, donde la residencia teórica se relaciona con la duración de la visa, que es otorgada generalmente por períodos de 1 año, y las encuestas especializadas en temas demográficos. Cada fuente considera unidades de análisis y metodologías diferentes, por lo que se debe alertar a los lectores en tanto a las comparaciones que se realizan entre unas y otras. Además de las diferencias en la residencia habitual, los censos permiten identificar a las personas nacidas en el extranjero (nacionales o extranjeros), mientras que los registros administrativos dan cuenta de la población extranjera (nacida en el país o fuera de él).

En la actualidad, Chile no cuenta con ningún registro administrativo que, con periodicidad mensual, pueda entregar un reporte sobre el total (*stock*) de personas extranjeras que residen en Chile; de modo que esta información, como se ha mencionado, pueda ser usada para ajustar o “calibrar” los factores de expansión de la ENE y, de esta forma, incorporar el efecto que genera el aumento de la población extranjera dentro de la población residente total. Como medida complementaria, se buscó definir un modelo que permita obtener el stock de la población extranjera, con el propósito de contar con una estimación mensual para dicha subpoblación. Para este propósito, sin embargo, no se cuenta con suficiente cantidad de observaciones sobre el *stock*, ya que información directa del *stock* de personas extranjeras solo se tiene a través del Censo de 2017, sobre el cual se debe agregar el correspondiente ajuste debido a la omisión censal; y la estimación INE-DEM⁸² al 31 de diciembre de 2018 (INE-DEM, 2019). Esta cantidad de datos evidentemente es insuficiente, por lo cual se debió adaptar la información entregada por las Estimaciones y Proyecciones de Población, como medida complementaria de datos.

Las estimaciones y proyecciones de población utilizan como uno de sus insumos, una estimación del flujo migratorio internacional. Debe tenerse presente que este flujo (estimado y proyectado) de personas inmigrantes, no diferencia a la población chilena de la extranjera, es decir, tanto entre los inmigrantes como entre los emigrantes, hay chilenos y extranjeros, por lo que es necesario destacar que “extranjeros” e “inmigrantes” son dos poblaciones diferentes. No obstante, en la metodología que aquí se presenta, la totalidad de los inmigrantes son considerados como una aproximación del total de extranjeros. De este modo, en el diseño del método aquí propuesto, el flujo de extranjeros que ingresan a Chile será aproximado por el flujo estimado de la inmigración, que es utilizada por las Estimaciones y Proyecciones de Población base 2017.

⁸² Departamento de Extranjería y Migración (DEM).

Para aproximar el total (*stock*) de personas extranjeras residentes en Chile, con frecuencia mensual, para el período 2012-2025⁸³, en primer lugar, se definieron stocks anuales de personas extranjeras. Para esto se definió un stock inicial en junio 2012, a partir de la suma del flujo de inmigrantes entre los años 1992-2012 (ver Tabla 31) que utilizaron las Estimaciones y Proyecciones de Población. Luego, los siguientes stocks para los junios de cada año sucesivo, hasta el 2025, se calcularon sumando al *stock* inicial los correspondientes flujos anuales de inmigrantes de las Estimaciones y Proyecciones de Población.

Tabla 31. Estimación y stock anual de población extranjera (al 30 de junio de cada año) a partir del saldo migratorio de las Proyecciones de las Proyecciones de población base 2017

Año	1992	1993	1994	1995	1996	1997	1998
Extranjeros	1.646	2.081	2.523	4.895	5.588	6.199	6.629
Stock	-	3.727	6.250	11.145	16.733	22.932	29.561
Año	1999	2000	2001	2002	2003	2004	2005
Extranjeros	6.929	10.903	11.393	12.380	17.915	19.354	22.996
Stock	36.490	47.393	58.786	71.166	89.081	108.435	131.431
Año	2006	2007	2008	2009	2010	2011	2012
Extranjeros	25.603	29.786	28.506	28.385	43.303	48.449	32.633
Stock	157.034	186.820	215.326	243.711	287.014	335.463	368.096
Año	2013	2014	2015	2016	2017	2018	2019
Extranjeros	40.546	46.277	65.701	136.529	224.195	246.861	239.656
Stock	408.642	454.919	520.620	657.149	881.344	1.128.205	1.367.861
Año	2020	2021	2022	2023	2024	2025	
Extranjeros	107.936	39.868	25.022	21.796	20.936	20.658	
Stock	1.475.797	1.515.665	1.540.687	1.562.483	1.583.419	1.604.077	

Fuente: Elaboración propia en base al componente Inmigración de las Proyecciones de Población base 2017.

En segundo lugar, una vez construidos los totales (*stocks*) de personas extranjeras a junio, de cada uno de los años entre el 2012 y 2025, se ajustó un modelo no lineal cuya función de media corresponde a una función logística de 4 parámetros (en adelante, modelo logístico), con el cual se estiman totales mensuales que son una aproximación para el total de personas extranjeras residentes en Chile⁸⁴.

⁸³ Tener una estimación mensual posterior a 2025 no es relevante por dos motivos: 1) En 2022 se realizará un nuevo Censo y por tanto en 2025 de todos modos existirán nuevas proyecciones de población, junto con sus componentes como saldo migratorio; 2) al incorporar información alejada del momento del Censo (2017) se pierde precisión, luego incorporar cifras posteriores a 2025 redundaría en el problema de subajuste del modelo.

⁸⁴ Las estimaciones generadas por el modelo, debido a que se refieren a personas, son redondeadas al primer argumento entero, considerando cero decimales, mediante la función: *round(, digits = 0)*.

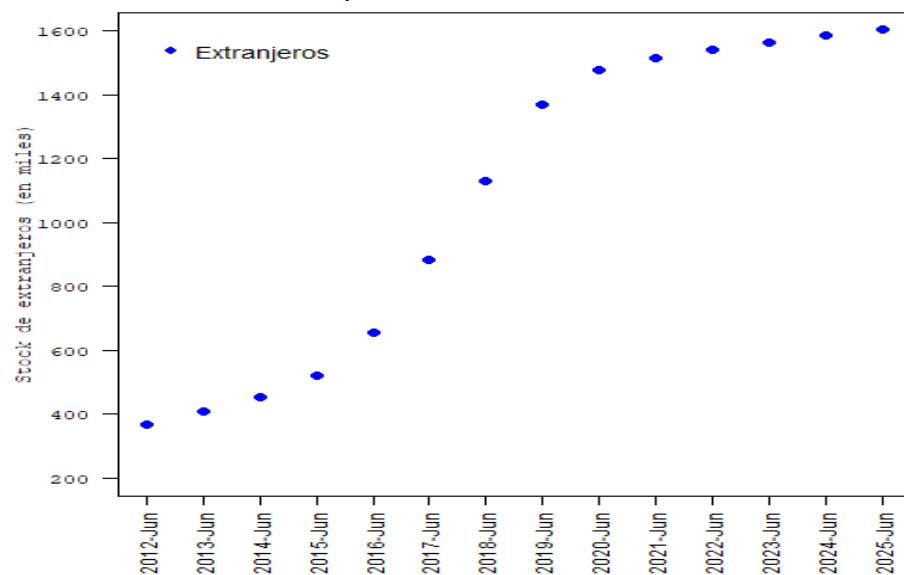
11.8.2 Resultados del modelo

En este apartado se presentan los resultados del modelo logístico. El detalle técnico se encuentra en la sección final de este anexo.

De los datos antes descritos y graficados en la Figura 21 se aprecian dos particularidades:

1. La forma funcional de Y es no lineal.
2. La forma funcional de Y tiene la forma de una función logística.

Figura 21. Estimación anual de población extranjera a partir del saldo migratorio de las Proyecciones de población base 2017.

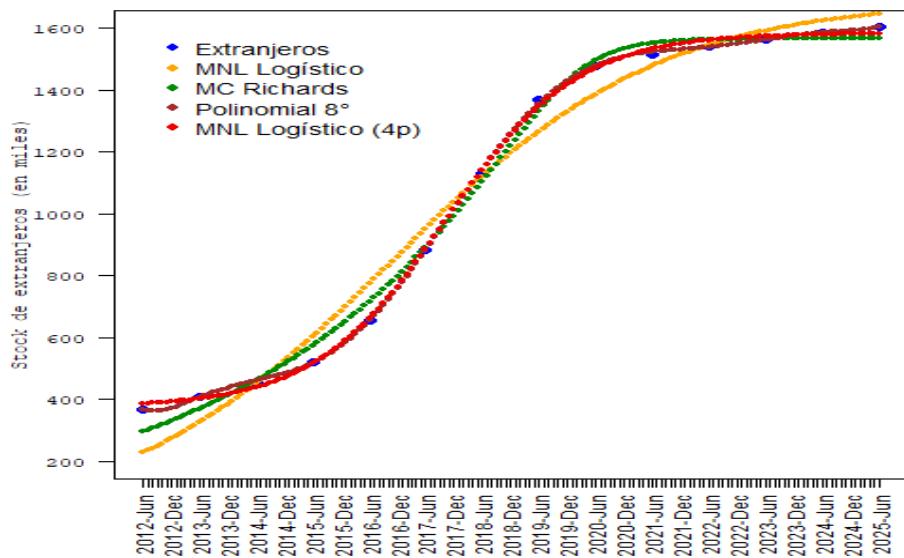


Fuente: Elaboración propia en base al componente Inmigración de las Proyecciones de Población base 2017.

En el proceso de estimación, se probaron distintas especificidades para la función de media a estimar; específicamente, se ajustaron tres funciones de media, todas correspondieron a una forma logística: Modelo de Crecimiento de Richards, la función logística de 3 parámetros y la función logística de 4 parámetros, y un modelo polinómico de 8 grados⁸⁵. La Figura 22 recoge estos resultados.

⁸⁵ La función de crecimiento exponencial se descartó debido a la presencia de asíntotas en la senda que siguen los datos, en tanto que otros modelos como el modelo de crecimiento de Gompertz, no alcanzó convergencia para su solución numérica.

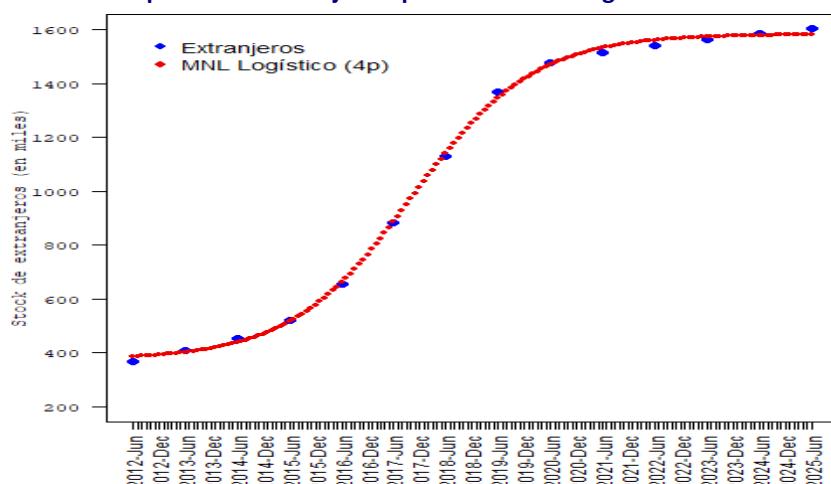
Figura 22. Estimación anual de población extranjera a partir del saldo migratorio de las Proyecciones de población base 2017 y la estimación de los modelos estudiados



Fuente: Elaboración propia en base al componente Inmigración de las Proyecciones de Población base 2017 y modelos estudiados.

Al evaluar la suma de los cuadros de la diferencia de la estimación y las estimaciones anuales, sumado al contraste con las cifras de población nacida en el extranjero para el momento del Censo 2017 (746.465 personas) y de la estimación oficial de residentes extranjeros para el 31 de diciembre de 2018 entre el INE y el Departamento de Extranjería e Inmigración (1.251.225 personas) y por la parsimonia que entrega un modelo logístico en este tipo de estimaciones, se optó por escoger como modelo final para la estimación de población extranjera, el correspondiente al del modelo logístico de 4 parámetros. La Figura 23 ilustra esta serie.

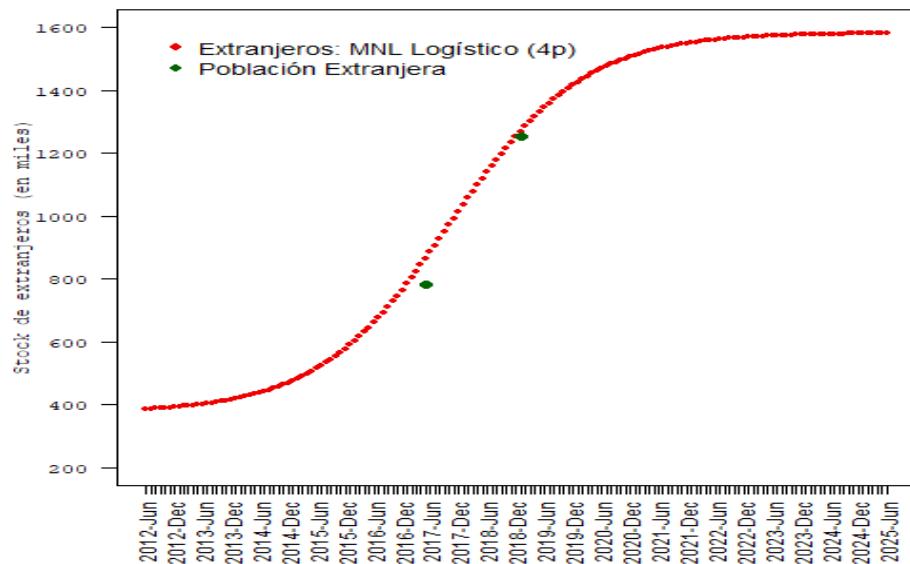
Figura 23. Estimación mensual de población extranjera según modelo logístico, y estimación anual de población extranjera a partir del saldo migratorio.



Fuente: Elaboración propia en base al componente Inmigración de las Proyecciones de Población base 2017 y la estimación del modelo no lineal logístico de 4 parámetros.

La Figura 24, por su parte, contrasta los resultados del modelo seleccionado y los stocks poblacionales para la población extranjera residente en Chile, reportados en el Censo de 2017, incorporando la omisión censal, y la Estimación de Población Extranjera al 31 de diciembre de 2018.

Figura 24. Estimación de población extranjera según modelo logístico, Censo 2017 y estimación de la población extranjera residente al 31 de diciembre de 2018.



Fuente: Elaboración propia en base al Censo 2017 (con omisión censal), la Estimación de la Población Extranjera al 31 de diciembre y la estimación del modelo no lineal logístico de 4 parámetros.

11.8.3 Alcances técnicos, usos actuales y prospectivos de la estimación

La metodología planteada en este documento busca mejorar el tratamiento de la calibración de las encuestas de hogares dado el impacto que tiene la migración internacional en los fenómenos sociales. Su uso está orientado para la encuesta de empleo y no busca reemplazar las estimaciones oficiales anuales entregadas por el INE y el DEM en conjunto con otros servicios del Estado, las cuales se construyen conceptual y metodológicamente de manera distinta a este trabajo. De hecho, las estimaciones anuales consideran información actualizada del fenómeno, producto de los cambios de las políticas migratorias. Además, se debe considerar que, desde el punto de vista demográfico, las personas que inmigran también pueden salir del país o morir, algo que no considera el modelo. Considerando lo anterior, las estimaciones derivadas del modelo serán monitoreadas de forma continua y, a partir de nueva información que pueda ser incluida dentro del modelo, sus resultados podrán ser actualizados.

En segundo lugar, incorporar información exógena sobre la población extranjera no implica que las estimaciones específicas que se puedan realizar respecto de esta subpoblación desde los datos de la ENE tengan per se la apropiada calidad estadística exigida para el resto de las estimaciones que se derivan de la encuesta. Para verificar la calidad estadística, se debe observar las medidas de precisión (errores estándar, coeficientes de variación y/o intervalos de confianza), los grados de libertad dentro de la subpoblación y el número de observaciones muestrales con las que se estiman los indicadores para esta subpoblación.

11.8.4 Especificaciones técnicas del modelo

El modelo seleccionado corresponde a un modelo no lineal, cuya función de media es una función logística de cuatro parámetros, la cual tiene el siguiente detalle⁸⁶:

$$f(x, A, B, xmid, scal) = A + \frac{B - A}{1 + e^{(xmid - x)/scal}} \quad (44)$$

Donde:

- x : Vector numérico de valores donde se evalúa el modelo.
- A : Parámetro que representa la asíntota horizontal (izquierda).
- B : Parámetro que representa la asíntota horizontal (derecha).
- $xmid$: Parámetro que representa el valor de x en el punto de inflexión de la curva.
- $scal$: Parámetro de escala numérica en el eje de la variable x .

La estimación de los cuatro parámetros de la función de media, arriba descrita, fue realizada utilizando el *software R*, a partir del package *nls()* y la siguiente rutina.

```
library(foreign)
library(NISTnls)
library(proto)
library(nls2)
library(nlrwr)
library(NRAIA)
library(nlstools)
library(nlsMicrobio)

rm(list = objects())
setwd("C://Estimacion_Extranjeros//Estimacion_OFICIAL")
inm2012 <- read.csv(file = "Inmigrantes2012.csv", header = TRUE, dec = ",")
attach(inm2012)
set.seed(5782)
NLlog4_2012 <- nls(immigration ~ SSfpl(t, A, B, xmid, scal), data = inm2012)
summary(NLlog4_2012)
```

Los resultados descritos en la Figura 25, muestran que los cuatro parámetros estimados, presentan coeficientes con un nivel de significancia al 1%, por cuanto, se rechaza la hipótesis nula de coeficientes nulos, esto es, los parámetros estimados son estadísticamente significativos.

⁸⁶ Para mayor detalle sobre estimación de modelos no lineales se puede revisar (Bates & Watts, Nonlinear regression analysis and its applications , 1988) (Bates & Chambers, Nonlinear models , 1992) y (Crawley, 2013).

Figura 25. Resultados del modelo no lineal logístico de 4 parámetros

```
> summary(NLlog4_2012)

Formula: inmigracion ~ ssfp1(t, A, B, xmid, scal)

Parameters:
Estimate Std. Error t value Pr(>|t|)    
A     3.751e+05 1.261e+04 29.75 4.31e-11 ***
B     1.583e+06 8.661e+03 182.72 < 2e-16 ***
xmid 6.435e+01 6.348e-01 101.37 < 2e-16 ***
scal  1.400e+01 5.801e-01 24.13 3.41e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

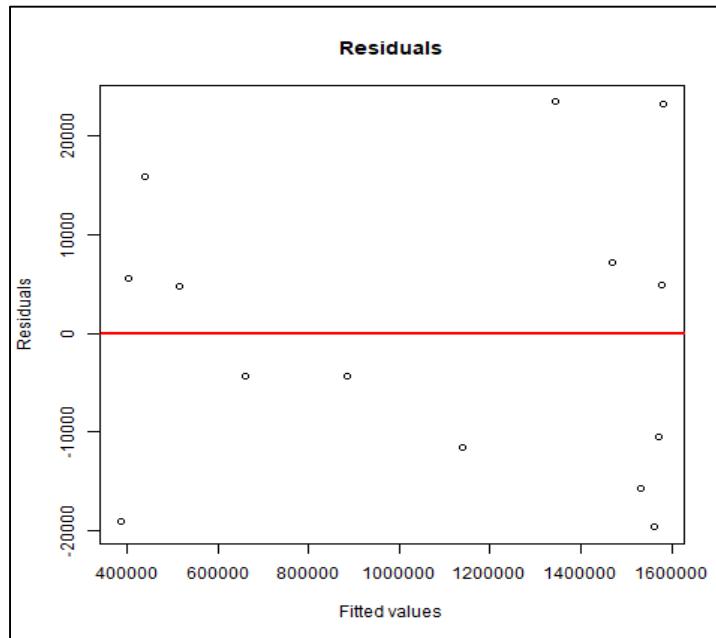
Residual standard error: 16580 on 10 degrees of freedom

Number of iterations to convergence: 0
Achieved convergence tolerance: 2.741e-06
```

Fuente: Elaboración propia en base al componente Inmigración de las Proyecciones de Población base 2017

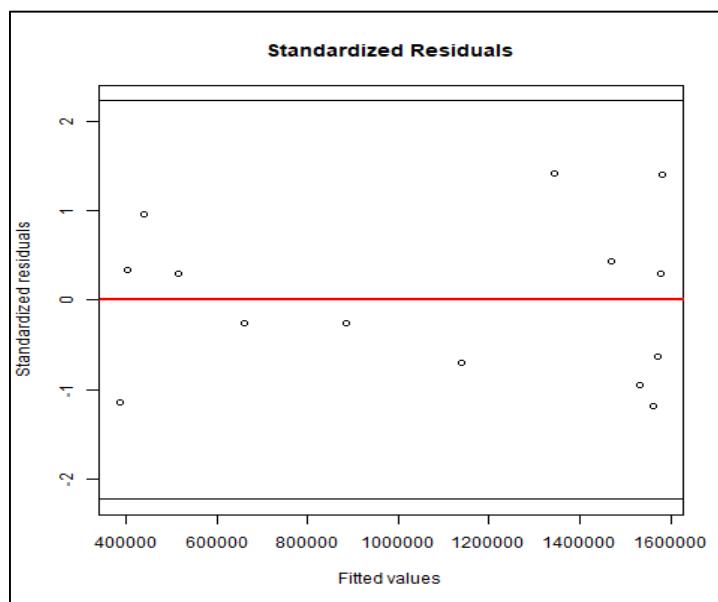
Finalmente, como señalan (Ritz & Streibig, 2008), se deben validar los supuestos detrás del modelo de regresión no lineal. Para esto, se debe verificar que: la función de media sea la correcta, homocedasticidad (con los residuos estandarizados), distribución normal de los residuos y, por último, independencia de los residuos, lo cuales se cumplen en el modelo ajustado.

Figura 26. Supuestos del modelo de regresión no lineal: función de media correcta



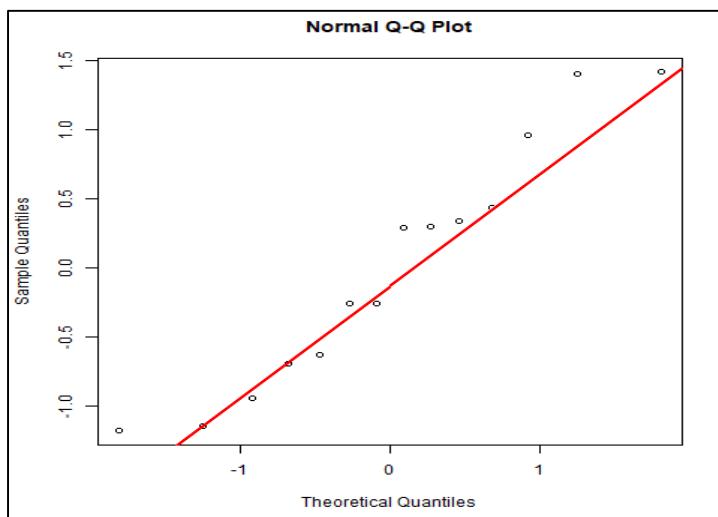
Fuente: Elaboración propia en base a los resultados del modelo no lineal ajustado.

Figura 27. Supuestos del modelo de regresión no lineal: homocedasticidad



Fuente: Elaboración propia en base a los resultados del modelo no lineal ajustado.

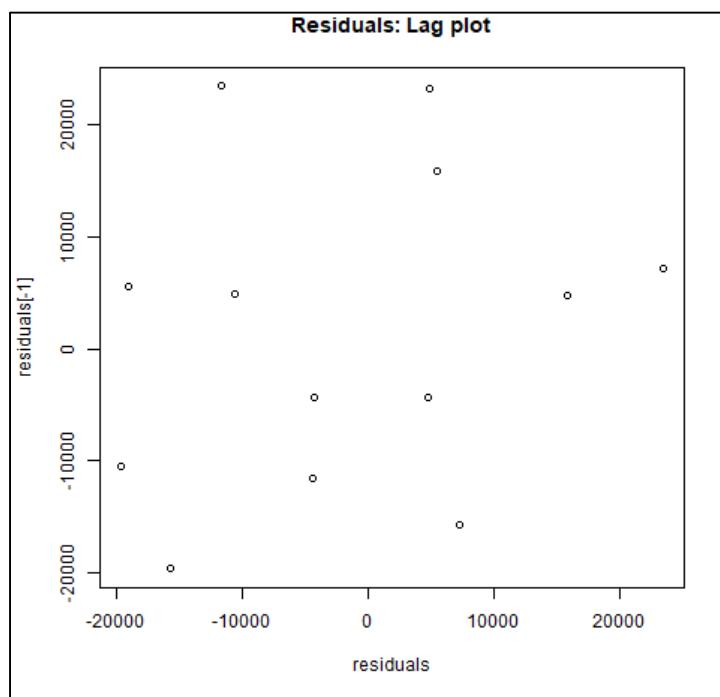
Figura 28. Supuestos del modelo de regresión no lineal: Normalidad



```
> shapiro.test(standardRES)
Shapiro-Wilk normality test
data: standardRES
W = 0.93438, p-value = 0.3511
```

Fuente: Elaboración propia en base a los resultados del modelo no lineal ajustado.

Figura 29. Supuestos del modelo de regresión no lineal: Independencia

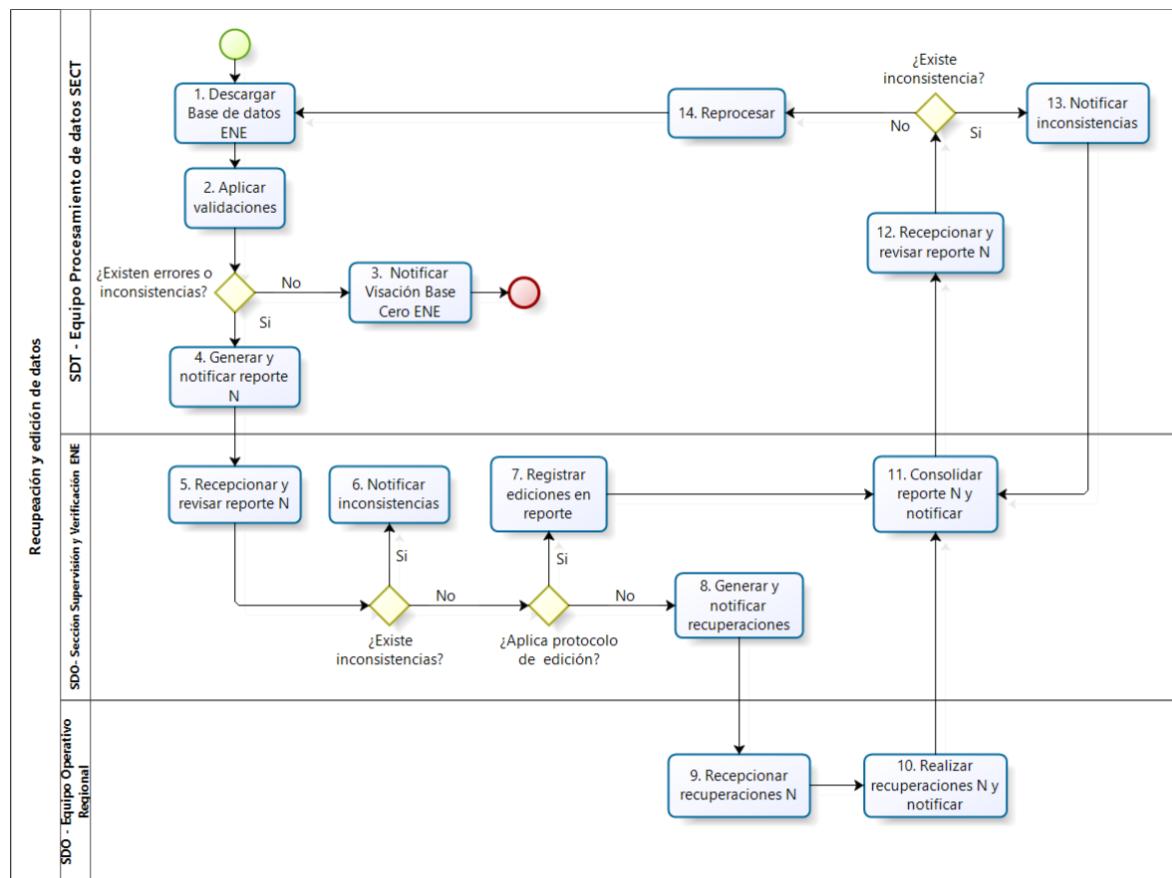


Fuente: Elaboración propia en base a los resultados del modelo no lineal ajustado.

11.9 Anexo N°9 Flujo operativo Supervisión de Gabinete

La supervisión de gabinete, mediante el proceso de recuperación y edición de datos, monitorea y asegura que la calidad de los datos recolectados, tanto en el cuestionario central como el sociodemográfico por cada miembro del hogar, permitiendo determinar si una entrevista es “completa” mediante del cumplimiento de una serie de reglas de evaluación de respuestas válidas en preguntas claves. Para comprender funcionamiento del subproceso, se detalla en la figura 30 flujo del proceso de recuperación de información y en la tabla 32 se detalla la matriz de riesgo basado en el flujo operativo:

Figura 30. Flujo proceso recuperaciones



Fuente: Subdepartamento de Supervisión y Verificación de Estadísticas Sociales

Tabla 32. Matriz de actividades flujo de recuperación de datos ENE

Nº	Actividad	Descripción	Rol
1	Descargar Base de Datos ENE	Descarga base de datos del servidor e identifica las viviendas pertenecientes a la categoría elegibles entrevistadas para generar el reporte con el código de visita 11 corresponde a aquellas de entrevista completa.	Analista de procesamiento SECT
2	Aplicar validaciones	Determinar si la información entregada por los informantes y registrada por el/la recolector/a es suficiente para clasificar y caracterizar la situación laboral de los entrevistados/as, de acuerdo al cumplimiento de validaciones definidas para tales efectos.	Analista de procesamiento SECT
◇	¿Existen errores/ Recuperaciones?	Sí: Pasa a actividad 4 No: Pasa a actividad 3	
3	Notificar Visación Base Cero ENE	Al corroborar que no existen errores/recuperaciones se notifica mediante correo electrónico a la jefatura de Departamento de Estadísticas Sociales la visación de la base cero, lo cual cierra el periodo de recolección.	Analista de procesamiento SECT
4	Generar y notificar reporte N	Para generar el reporte se aplica el set de validaciones del criterio de suficiencia determinando los casos a recuperar. El reporte N detalla casos de entrevistados/as que posee datos omitidos inconsistencias y/o fuera de rango, y se envía mediante correo electrónico. El N representa el número de reporte dentro del periodo.	Analista de procesamiento SECT
5	Repcionar y revisar reporte N	Recepción de reporte N y revisión de la completitud de las variables para determinar el flujo de corrección.	Analista SSVES
◇	¿Existe inconsistencia?	Sí: Pasa a actividad 6 No: Pasa a pregunta si aplica protocolo de edición	
6	Notificar inconsistencias	Se envía mediante correo electrónico las inconsistencias de validaciones para que el analista SECT, vuelva a ejecutar la actividad 4.	Analista SSVES
◇	¿Aplica protocolo de edición?	Sí: Pasa a actividad 7 No: Pasa a actividad 8	
7	Registrar ediciones en reporte	Identificación de aquellos casos que según protocolo son editables, realiza la corrección en el archivo de reporte N.	Analista SSVES
8	Generar y notificar recuperaciones N	Ánálisis de los casos que requieren que los datos sean recuperados en terreno, notificando al coordinador regional y enviando mediante correo electrónico el reporte N.	Analista SSVES
9	Repcionar recuperaciones	Recepción de reporte N. Se realiza la asignación al recolector/a para que gestione la recuperación de datos.	Equipo Operativo Regional
10	Realizar recuperaciones y notificar	Se lleva a cabo la recuperación y en aquellos casos en que no fue posible contactar al informante se registra una observación y se entregan los resultados al coordinador, el cual es responsable de monitorear este proceso y notificar sus resultados.	Equipo Operativo Regional
11	Consolidar reporte N y notificar	Consolidación de la edición de datos y recuperaciones realizadas por los equipos regionales en el reporte N y se notifica mediante correo electrónico los resultados.	Analista SSVES
12	Repcionar y revisar reporte N	Recepción de reporte N. Se revisa que cumpla con el nombre de archivo, completitud de estados de gestión de recuperaciones y las variables definidas que forman parte del reporte.	Analista de procesamiento SECT
◇	¿Existe inconsistencia?	Sí: Pasa a actividad 13 No: Pasa a actividad 14	

13	Notificar inconsistencias	Se envía, mediante correo electrónico, las inconsistencias para que el analista SSVES vuelva a ejecutar la actividad 11.	Analista de procesamiento SECT
14	Reprocesar	En base a la gestión de las recuperaciones se realiza el reprocesamiento de los casos, con el fin de que aquellos casos sin recuperar por rechazo no vuelvan a ser incluidos en reportes posteriores. Esta actividad termina el ciclo, volviendo al punto 1. El N, dependerá de cuantos archivos se ejecutarán durante el periodo hasta tener una visación de base cero ENE.	Analista de procesamiento SECT

11.10 Anexo N°10 Acceso a los datos

Las bases de datos de la ENE se encuentran publicadas en el portal Web del INE-Chile, en la dirección www.ine.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion. El periodo que se pone a disposición del público corresponde a la actual serie de la encuesta, que comenzó en enero-marzo de 2010 y que sigue vigente hasta la actualidad.

Cualquier requerimiento adicional de información debe ser canalizada a través del portal del Sistema Integral de Información y Atención Ciudadana (SIAC) del INE (www.ine.cl/SIAC), el que será evaluado de acuerdo a dos criterios:

- Disponibilidad de la información solicitada.
- Protección de la confidencialidad de los datos, de acuerdo al secreto estadístico establecido en la Ley 17.374.

Las bases de datos de la ENE están disponibles en tres formatos: *SPSS*, *STATA* y *comma-separated values* (.csv, por su sigla en inglés, para trabajar con otros *software*). Para seleccionar la base de datos que se requiera consultar, se deben seguir tres pasos:

1. **Paso 1.** Una vez abierta la página Web, se debe presionar la opción BASES DE DATOS. Al hacerlo se desplegarán todos los años con información disponible (Figura 31)

Figura 31. Selección de opción de descarga de base de datos en página Web institucional del INE



2. **Paso 2.** Una vez seleccionado el año, se debe seleccionar una categoría para mostrar archivos, de acuerdo a las tres alternativas de formato disponibles (Figura 32).

Figura 32. Formatos de bases de datos de la ENE disponibles



The screenshot shows a sidebar on the left with a dark blue background and white text, listing various document types: CUADROS ESTADÍSTICOS, BOLETINES, PUBLICACIONES Y ANUARIOS, DOCUMENTOS DE TRABAJO, METODOLOGÍAS, COMITÉS Y NOTAS TÉCNICAS, BASES DE DATOS, FORMULARIOS, and METADATOS. The 'BASES DE DATOS' section is highlighted with a red border. The main content area to the right is also bordered in red and contains the heading 'BASES DE DATOS' and a sub-section 'Bases De Datos > 2021' with three options: 'Formato CSV', 'Formato STATA', and 'Formato SPSS'.

3. **Paso 3.** Finalmente, una vez seleccionado el formato, se desplegarán todos los trimestres, bastando sólo con presionar el ícono  para que comience la descarga de la base requerida (Figura 33).

Figura 33. Selección de base trimestral a descargar



The screenshot shows the same sidebar and main structure as Figure 32. The 'Formato STATA' section is highlighted with a red border. To the right, under 'Bases De Datos > 2021 > Formato STATA', there is a list of four items: 'ENE 2021 01 DEF DTA, 68.99 MB' with a download icon, 'ENE 2021 02 EFM DTA, 75.40 MB' with a download icon, 'ENE 2021 03 FMA DTA, 79.98 MB' with a download icon, and 'ENE 2021 04 MAM DTA, 83.17 MB' with a download icon.

