

Instituto Nacional de Estadísticas

Documento Metodológico Encuesta Nacional de Empleo (ENE)

2020



INSTITUTO NACIONAL DE ESTADÍSTICAS

Subdirección Técnica

Departamento de Estadísticas del Trabajo

Departamento de Metodologías e Innovación Estadística

DOCUMENTO METODOLÓGICO ENCUESTA NACIONAL DE EMPLEO (ENE)

Febrero / 2020

Versión 2.0 (actualizado el 02 de marzo de 2020)

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

1	Presentación	7
2	Contexto metodológico	10
2.1.	Antecedentes de la ENE	10
2.2.	Objetivos de la ENE	11
2.2.1.	Objetivo general	11
2.2.2.	Objetivos específicos	11
3	Marco de referencia	12
3.1.	Marco conceptual	12
3.1.1.	Clasificación de la población en edad de trabajar	14
3.1.2.	Población en la fuerza de trabajo	15
3.1.3.	Población fuera de la fuerza de trabajo	20
3.1.4.	Caracterización de la Población en Edad de Trabajar	22
3.1.5.	Principales tasas e indicadores del mercado laboral.....	30
4.	Características del diseño muestral.....	32
4.1.	Población objetivo.....	32
4.2.	Cobertura geográfica y dominios de estudio	32
4.3.	Período de referencia y periodicidad	32
4.4.	Marco muestral maestro.....	33
4.4.1.	Cobertura geográfica del marco muestral.....	34
4.4.2.	Estratificación geográfica del marco muestral	34
4.4.3.	Conformación de Unidades Primarias de Muestreo	35
4.4.4.	Estratificación socioeconómica del marco muestral.....	38
4.4.5.	Áreas especiales	39
4.4.6.	Conformación definitiva del marco muestral maestro	40
4.5.	Marco muestral de la Encuesta Nacional de Empleo.....	42
4.5.1.	Estratificación geográfica	42
4.5.2.	Conformación del marco muestral.....	44
4.5.3.	Estratificación del marco muestral ENE	46
4.6.	Unidades estadísticas	47
4.6.1.	Unidad de muestreo	47
4.6.2.	Unidad de información y análisis.....	47
5	Tamaño muestral.....	48
5.1.	Estrategia muestral	48
5.1.1.	Consideraciones para el tamaño muestral.....	48
5.2.	Cálculo del tamaño muestral.....	49
5.3.	Distribución del tamaño muestral.....	51
5.4.	Cálculo de la sobremuestra	52
5.5.	Tamaño muestral de viviendas por UPM	54
5.6.	Selección de unidades muestrales	61
5.6.1.	Selección de unidades de primera etapa	61
5.6.2.	Selección de unidades de segunda etapa.....	62

5.7.	Rotación de las unidades muestrales.....	64
5.8.	Incorporación de la muestra del MMV 2017	66
5.8.1.	Incorporación de la nueva muestra ENE por turnos	67
5.8.2.	Selección de las viviendas a restar en cada mes	67
6	Desarrollo del factor de expansión.....	69
6.1.	Factor de expansión de selección de conglomerados.....	70
6.2.	Factor de expansión de selección de viviendas	70
6.2.1.	Ajuste por elegibilidad.....	71
6.2.2.	Ajuste por no respuesta	73
6.2.3.	Ajuste a stocks poblacionales (Calibración).....	74
7	Estimadores	78
7.1.	Estimadores de total para la variable Y.....	78
7.2.	Estimadores de razón	79
8	Medidas de precisión.....	81
8.1.	Varianza de los estimadores.....	81
8.2.	Errores absolutos y relativos	82
8.2.1.	Error absoluto e intervalo de confianza	82
8.2.2.	Error relativo y coeficiente de variación	83
8.3.	Representatividad de las estimaciones	84
8.3.1.	Elementos relacionados con el tamaño muestral	84
8.3.2.	Elementos relacionados con la precisión de la estimación	84
9	Alcance de la presentación de resultados.....	86
9.1.	Difusión de resultados.....	86
9.2.	Accesibilidad a los datos y marco legal	87
10	Anexos	89
	Anexo N° 1. Número de UPM, viviendas y PET según estrato geográfico de la ENE	89
	Anexo N° 2. Estratificación muestral del marco muestral de la ENE.....	95
	Anexo N° 3. Códigos de visita de la ENE	101
	Anexo N° 4. Tasas de no logro por estrato geográfico, según información de la ENE 2018	102
	Anexo N° 5. Estimaciones de población para la Encuesta Nacional de Empleo.....	106
	Anexo N° 6. Listado de marginales provinciales.....	109
	Anexo N° 7. Cálculo de los factores de expansión en el período de transición	110
	Anexo N° 8. Construcción de insumo para la calibración por población extranjera en factores de expansión de la ENE	115
11	Bibliografía.....	125

Listado de figuras

Figura 1. Clasificación de la población dentro y fuera de la fuerza de trabajo en la ENE	15
Figura 2. Criterios para clasificar como ocupado	17
Figura 3. Excepciones a los criterios para clasificar ocupación	18
Figura 4. Criterios para clasificar como desocupado	19
Figura 5. Clasificación de la población de acuerdo con su vínculo con el mercado laboral (versión extendida)	22
Figura 6: Estratificación geográfica del MMV 2017	35
Figura 7: Estratificación de entidades urbanas	36
Figura 8: Numeración por método de serpenteo	37
Figura 9: Estratificación de entidades rurales	37
Figura 10. Esquema de rotación de la muestra de viviendas.....	65
Figura 11. Incorporación de la muestra MMV 2017 a partir de las rotaciones	66
Figura 12: Esquema de configuración de marcos y muestras	110
Figura 13: Esquema actualizado de configuración de marcos y muestras.....	111
Figura 14. Resultados del modelo no lineal logístico de 4 parámetros	122

Listado de tablas

Tabla 1. Ejemplos de trabajo en la ocupación poco frecuentes (medido en la ENE) y otras formas de trabajo (actividades productivas no consideradas)	16
Tabla 2. UPM con tamaño fuera de rango por región	41
Tabla 3. Estratificación socioeconómica del MMV 2017 por región.....	41
Tabla 4. UPM especiales por región	42
Tabla 5. Cantidad de estratos geográficos por región y área.....	43
Tabla 6. Estratos que no alcanzan mínimo de UPM.....	44
Tabla 7. UPM especiales y excluidas por región, en el marco muestral de la ENE	45
Tabla 8. Cantidad de estratos de muestreo por región y área.....	46
Tabla 9. Características del parámetro de interés	48
Tabla 10. Efectos del diseño en combinaciones seleccionadas de tamaño de la muestra conglomerada y correlación intraclase	56
Tabla 11. Efectos de diseño por dominios de estudio, para distintos tamaños de viviendas promedio por UPM.....	57
Tabla 12. Errores muestrales de la tasa de desocupación para los dominios urbanos según número promedio de vivienda por UPM	58
Tabla 13. Errores muestrales de la tasa de desocupación para los dominios rurales según número promedio de vivienda por UPM	59

Tabla 15. Distribución de la muestra y resultados de simulación.....	60
Tabla 16. 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	117

Listado de gráficos

Gráfico 1. Incremento gradual de la muestra teórica durante el período de transición de la ENE.....	67
Gráfico 2. Estimación anual de población extranjera a partir del saldo migratorio de las Proyecciones de población base 2017.	118
Gráfico 3. Estimación anual de población extranjera a partir del saldo migratorio de las Proyecciones de población base 2017 y la estimación de los modelos estudiados.....	119
Gráfico 4. Estimación mensual de población extranjera según modelo logístico, y estimación anual de población extranjera a partir del saldo migratorio.	119
Gráfico 5. Estimación de población extranjera según modelo logístico, Censo 2017 y estimación de la población extranjera residente al 31 de diciembre de 2018.	120
Gráfico 6. Supuestos del modelo de regresión no lineal: función de media correcta	122
Gráfico 7. Supuestos del modelo de regresión no lineal: homocedasticidad	123
Gráfico 8. Supuestos del modelo de regresión no lineal: Normalidad.....	123
Gráfico 9. Supuestos del modelo de regresión no lineal: Independencia.....	124

1 Presentación

La Encuesta Nacional del Empleo (ENE) tiene como propósito proveer de información estadística fundamental sobre el mercado del trabajo para orientar la toma de decisiones y la evaluación de las políticas económicas y de trabajo en la ocupación a nivel nacional. El Instituto Nacional de Estadísticas (INE), responsable de su producción, elabora esta encuesta sobre el mercado laboral desde el año 1966, disponiendo de información continua y publicable a contar del trimestre diciembre-febrero de 1986 en adelante.

En la última década, la Encuesta Nacional del Empleo se ha sometido a una serie de modificaciones, tendientes a capturar de mejor manera los cambios en el mercado del trabajo, así como las fluctuaciones del mercado global, que han llevado al surgimiento de nuevas formas de ocupación y empleo, modificando, de este modo, las formas de entender el fenómeno del trabajo. Estas transformaciones de la encuesta también han tenido como objetivo la armonización de la medición de las estadísticas laborales siguiendo estándares como los desarrollados por la Organización Internacional del Trabajo (OIT) y la Organización para la Cooperación y el Desarrollo Económico (OCDE), que faciliten las comparaciones con otros países de manera armonizada.

El cambio más profundo en la encuesta se produjo en el año 2010, cuando se realizaron importantes modificaciones conceptuales y metodológicas, que llevaron a la división del estudio entre su antigua versión, aplicada entre el trimestre diciembre-febrero de 1986 hasta igual trimestre del año 2010, y la nueva ENE, iniciada en el trimestre enero-marzo de 2010 y vigente hasta la fecha. Ambas series no fueron comparables entre sí, puesto que los cambios incorporados produjeron un quiebre metodológico en el estudio. Al introducir los nuevos conceptos (más rigurosos) que redefinían a la población objetivo de la encuesta, situación que alteró la comparabilidad a nivel conceptual de los grupos de ocupados, desocupados y personas fuera de la fuerza de trabajo o inactivos. Estos lineamientos emanaron de la 19ª Conferencia Internacional de Estadísticos del Trabajo de la OIT. 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 actualizar la ENE con las más recientes recomendaciones internacionales, en línea con el desarrollo y el avance del mercado laboral a esa fecha.

A partir del año 2020, el INE incorpora nuevas modificaciones a la medición vigente, que permiten capturar las innovaciones que vive el mercado laboral en el país. Entre los principales cambios realizados, se produjo la actualización del marco muestral de la encuesta, del diseño estadístico y del método de calibración de los factores de expansión, el cual incorpora las proyecciones de población teniendo como sustento el Censo 2017.

La actualización del **marco muestral de la encuesta**, realizada a partir del PreCenso 2016. La primera fase de actualización de ese marco muestral se hizo entre noviembre de 2018 y noviembre de 2019, cuando se actualizó el número de viviendas a información de Precenso 2016. La segunda fase, está siendo aplicada a contar de enero de 2020, incluye el uso de nuevos conglomerados de muestreo construidos con información de Censo 2017 y que dentro de sus bondades se encuentra que son homogéneos en término del tamaño de viviendas.

La modernización del **diseño muestral** esta basada tanto en el marco de viviendas que se construyó con el Censo 2017 como en las nuevas proyecciones de población base 2017. Este nuevo diseño muestral utiliza un nivel de confianza del 95% para el cálculo de los errores de estimación, que es mayor al usado en el diseño que se tenía hasta ahora, que era equivalente al 90%. Este diseño se implementará durante todo el primer semestre de 2020.

La renovación del **método de calibración de los factores de expansión**, el que incorporan las proyecciones de población en base a Censo 2017 y toma en consideración las buenas prácticas internacionales.

Adicionalmente, como parte de las recomendaciones de la OIT para la medición de las estadísticas sobre las relaciones de trabajo, durante el año 2020 se estarán testeando nuevas preguntas en el cuestionario, sobre el uso de plataformas digitales, la propiedad de los medios de producción, entre otras. Estas preguntas serán monitoreadas para posteriormente definir un plan de publicación a contar del año 2021. El presente documento es resultado de un intenso trabajo de fortalecimiento de la Encuesta Nacional de Empleo, el cual contó con la asesoría de la Organización Internacional del Trabajo (OIT) y del Grupo de Especialistas en Encuestas y Empleo, conformado durante 2019 por el INE. Agradecemos el aporte de este grupo de especialistas conformado por Carolina Casas-Cordero Valencia¹, Susana Eyheramendy Duerr², Javiera Barrera Martínez³, Ricardo González⁴ y Andrés Gutiérrez Rojas⁵, cuya participación fue fundamental para impulsar este proceso de cambio metodológico. La revisión periódica y el fortalecimiento de las operaciones estadísticas constituyen buenas prácticas internacionales en el ámbito de la calidad estadística⁶, pues las metodologías, procedimientos o supuestos que utilizan las encuestas deben ser revisados con cierta periodicidad considerando el dinamismo propio de los fenómenos sociales, demográficos y económicos de los diferentes países.

Por último, el documento se encuentra organizado en nueve capítulos, en que se aborda la actualización de la metodología de la Encuesta Nacional de Empleo a partir de este año 2020.

- En el segundo y tercer capítulo se expone el contexto metodológico de la encuesta, detallando los antecedentes y objetivos del producto estadístico, además del marco conceptual que sostiene la medición del empleo.

¹ Ph.D in Survey Methodology, University of Maryland, EEUU. Investigadora asociada al Centro MIDE UC y al Centro UC de Encuestas y Estudios Longitudinales de la Pontificia Universidad Católica de Chile. Actualmente es jefa del Programa de Magíster en Diseño y Análisis de Encuestas Sociales y realiza el curso Fundamentals of Survey and Data Science, University of Maryland.

² Ph.D. in Statistics, Rutgers University, EEUU. Es profesora asociada en el Departamento de Estadística de la Pontificia Universidad Católica de Chile, y de la Facultad de Ingeniería y Ciencias de la Universidad Adolfo Ibáñez.

³ Ph.D. en Matemáticas Aplicadas de la Universidad de París Descartes y en Ciencias de la Ingeniería en Modelación Matemática de la Universidad de Chile. Actualmente es profesora asociada en la Facultad de Ingeniería y Ciencias de la Universidad Adolfo Ibáñez (UAI), subdirectora de posgrado de la misma facultad.

⁴ Ingeniero comercial y Magíster en Economía de la Pontificia Universidad Católica de Chile. También tiene estudios en metodología avanzada de encuestas en GESIS-Leibniz Institute for the Social Sciences, Colonia, Alemania. Desde septiembre de 2013, es coordinador del Área de Opinión Pública del CEP y coordinador del International Social Survey Programme (ISSP) en Chile.

⁵ Ph.D. en Estadísticas y Magíster en Ciencias en Estadísticas de la Universidad Nacional de Colombia. Actualmente se desempeña como experto regional en Estadísticas Sociales de Naciones Unidas, Comisión Económica para América Latina y el Caribe (CEPAL).

⁶ La *“Recomendación del Consejo de la OCDE sobre Buenas Prácticas Estadísticas”*, (OECD, 2015), señala en su recomendación n°8 la necesidad de revisiones periódicas y exhaustivas, en las que participen expertos externos, de las producciones estadísticas.

- El cuarto, quinto y sexto capítulo se explica el nuevo diseño estadístico de la encuesta y las características del marco muestral.
- En el séptimo y octavo capítulo se presentan los estimadores y las medidas de precisión de la nueva medición.
- El noveno capítulo aborda los alcances de la difusión de resultados.

2 Contexto metodológico

2.1. Antecedentes de la ENE

Las estadísticas sobre el mercado laboral tienen larga tradición en el INE. En 1966 el instituto elaboró por primera vez la Encuesta Nacional de Empleo, siendo una de las más antiguas en América Latina (Cecchini, 2005). Desde aquellos años, este estudio ha experimentado diversas modificaciones, entre ellas, los cambios de marcos muestrales (BCCh, 2001). De este modo, desde la década del 60 han existido tres series de datos.

En su historia reciente, como resultado de la incorporación de los estándares internacionales de medición, en el año 2010, el diseño conceptual y metodológico de la ENE incluyó, entre otros aspectos, la actualización de los conceptos de población ocupada, desocupada y fuera de la fuerza de trabajo o inactiva, y la distinción de los ocupados entre los ocupados “tradicionales” y los “no tradicionales”, así como las distintas razones de inactividad.

Manteniendo aquella estructura, las actualizaciones conceptuales fueron complementadas con la descripción de los principales indicadores de razón calculados a partir de las categorías de condición de actividad laboral de la población y con un resumen de los principales indicadores analíticos, consistente en un conjunto de indicadores complementarios que facilitan el análisis y seguimiento del mercado laboral.

Más en detalle, entre las modificaciones que se realizaron a la encuesta en el año 2010, se integraron baterías de preguntas que permiten actualizar la Población en Edad de Trabajar (PET), y se amplió la cobertura de indicadores para el análisis de la ocupación y el desocupación. Además, se elaboraron indicadores complementarios sobre la caracterización de la ocupación y respecto de las personas fuera de la fuerza de trabajo. Los nuevos temas incorporados se relacionan con la calidad del empleo de los trabajadores dependientes y con las medidas de subutilización laboral, como el trabajo a tiempo parcial involuntario y las personas inactivas potencialmente activas.

Durante el año 2017 se incorporaron preguntas que permiten medir la informalidad laboral, a través de la “Tasa de ocupación informal”, “Tasa de ocupación en el sector informal” y “Ocupados informales”, entre otros.

A partir del año 2020 se ha incorporado en la ENE una serie de mejoras tendientes a capturar de mejor manera las nuevas formas de trabajo (por ejemplo, a través de aplicaciones digitales), así como la flexibilidad laboral que actualmente experimentan trabajadores dependientes e independientes, explicitadas en mayor detalle en la Resolución I de la 20° CIET (2018). Estos últimos ajustes tienen como propósito avanzar en la adopción de las principales recomendaciones internacionales en la materia.

2.2. Objetivos de la ENE

2.2.1. Objetivo general

El objetivo de la Encuesta Nacional de Empleo es clasificar y caracterizar a la población en edad de trabajar (población de 15 años o más), con residencia habitual en Chile, según su vínculo con el mercado laboral. Esto, cumpliendo los estándares internacionales de medición de las estadísticas del trabajo, a fin de proporcionar indicadores que permitan analizar y comparar a nivel internacional las características ocupacionales del país de forma coyuntural y estructural.

2.2.2. Objetivos específicos

1. Cuantificar la población de 15 años o más, de acuerdo con su clasificación en tres grupos mutuamente excluyentes: ocupados, desocupados (ambos constitutivos de la fuerza de trabajo), y las personas fuera de la fuerza de trabajo, de acuerdo con los estándares internacionales vigentes.
2. Determinar características sobre las condiciones de ocupación y las modalidades de inserción laboral de trabajadores dependientes o independientes, formales o informales, a tiempo completo o parcial, y sobre la subutilización de la fuerza de trabajo, entre otros aspectos, permitiendo construir indicadores complementarios a los tradicionales para fines de seguimiento y análisis del mercado laboral.
3. Establecer la evolución de los indicadores tradicionales, tales como la tasa de desocupación, tasa de participación y tasa de ocupación a nivel nacional y en los correspondientes niveles de estimación definidos en el diseño muestral.
4. Caracterizar de manera sociodemográfica a la población residente en el país (se incluye nacionales y extranjeros), con especial atención en la población de 15 años y más, definida como población en edad de trabajar.

3 Marco de referencia

El presente capítulo expone una revisión de los conceptos en los que se basa la ENE para clasificar y caracterizar a la población en edad de trabajar, según su vínculo con el mercado laboral. Estos conceptos, a través de las preguntas que conforman el cuestionario, son medidos en un período de referencia específico, que corresponde al marco temporal definido y acotado en el que los encuestados son consultados sobre actividades o experiencias que sucedieron en este período. En particular, la ENE utiliza la *semana de referencia*.

3.1. Marco conceptual

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)⁷. 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, definiciones, clasificaciones y otros procedimientos metodológicos que se acuerdan como representativos, sobre estadísticas del trabajo.

En particular, el marco conceptual de la ENE considera 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)⁸; además de la Resolución sobre las estadísticas de las relaciones de trabajo del año 2018. Estos criterios son, al mismo tiempo, los mismos que adopta OCDE y EUROSTAT para todos sus países miembros.

Específicamente, 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).

⁷ 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

⁸ 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.

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,
5. Otras actividades productivas.

El **trabajo en la ocupación** establece el alcance de las actividades de referencia para las estadísticas de la fuerza de trabajo, que 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.

En relación a la expresión “a cambio de remuneración o beneficios”, la 19 CIET establece que “...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” (19° CIET, 2013). Aquel intercambio supone transacciones de mercado. La misma resolución establece respecto a la remuneración o beneficio de la ocupación que:

- a) Incluye la remuneración en efectivo o en especie, ya sea efectivamente recibida o no, y también puede incluir componentes adicionales de ingresos en efectivo o en especie.
- b) La remuneración puede ser pagada directamente a la persona que realiza el trabajo o indirectamente a un miembro del hogar o de la familia” (19° CIET, 2013).

Para la medición, 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).

Por su parte, 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.

Desde su versión del 2010, la ENE adopta los conceptos que, al año 2013, son parte de las directrices provenientes de la resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo. En concreto, la encuesta mide el trabajo en la ocupación, que comprende el trabajo realizado para terceros a cambio de remuneración o beneficios.

La incorporación de los estándares internacionales en la ENE se traduce en:

- 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 los ocupados, a los ocupados “tradicionales” de los "no tradicionales", en consideración de si el entrevistado 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 el mercado del trabajo.
- Establecer las relaciones de autoridad y riesgo económico entre los trabajadores y las unidades económicas para las que trabajan.

A partir de este marco conceptual, la ENE da cumplimiento a su objetivo de estudio. A través de él es posible clasificar a las personas en edad de trabajar que son residentes habituales en viviendas particulares ocupadas del país, como ocupados, desocupados y personas fuera de fuerza de trabajo (categorías mutuamente excluyentes) y caracterizarlos según diversas variables.

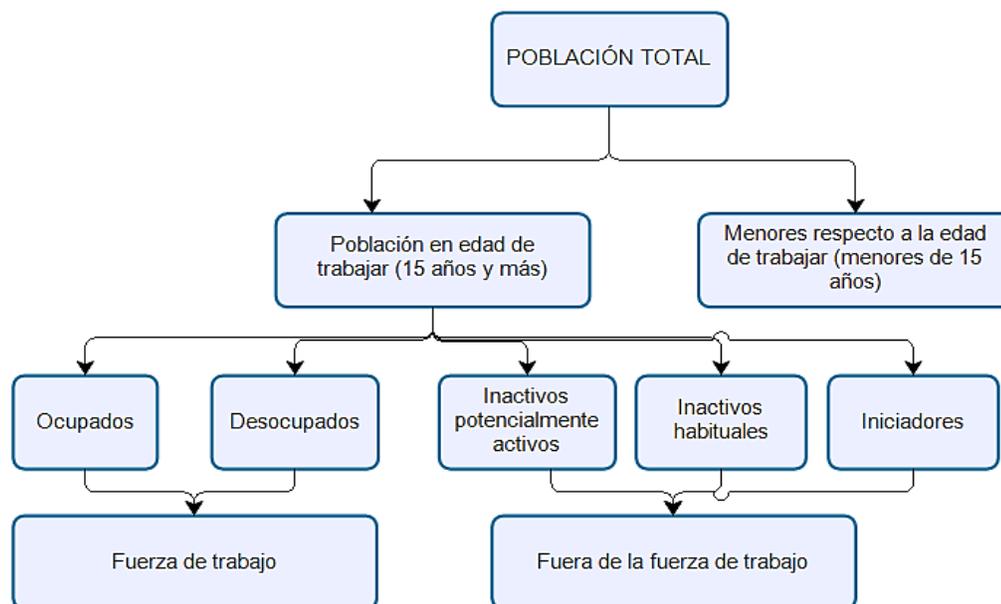
3.1.1. Clasificación de la población en edad de trabajar

La población en edad de trabajar (PET) corresponde a las personas que pueden tener un puesto de trabajo, es decir, trabajo en la ocupación. En Chile, se establece como tal a toda la población de 15 años o más.

La PET puede clasificarse según la situación en la fuerza de trabajo, aplicando un conjunto de reglas de prioridad, en las que se basa el enfoque internacional para esta medición mediante encuestas a hogares.

Para implementar la clasificación, en primer lugar, se establece una distinción entre quienes son parte de la fuerza de trabajo y quienes están fuera de ella, como se presenta en la Figura 1.

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.

3.1.2. Población en la fuerza de trabajo

La población en la fuerza de trabajo, anteriormente conocida como población económicamente activa, corresponde a todas aquellas personas de 15 años o más que, durante la semana de referencia, cumplen los requisitos para ser incluidas en la categoría de personas ocupadas o desocupadas, esto es, personas que se encuentran vinculadas con el mercado laboral.

Para establecer la situación de la población en la fuerza de trabajo (ocupada y, desocupada) o fuera de la fuerza de trabajo, se debe recurrir a dos conceptos: si recibe remuneración o beneficio por el bien producido el servicio prestado y el criterio de una hora.

En primer lugar, 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 remuneradas. 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.

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 poco frecuentes (medido en la ENE) y otras formas de trabajo (actividades productivas no consideradas)

Trabajo en la ocupación	Otras formas de trabajo
• Trabajo doméstico remunerado.	• 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 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.	• Trabajador por cuenta propia que está en trámites de inicio de actividad o negocio.
• 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.	• 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 no 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.

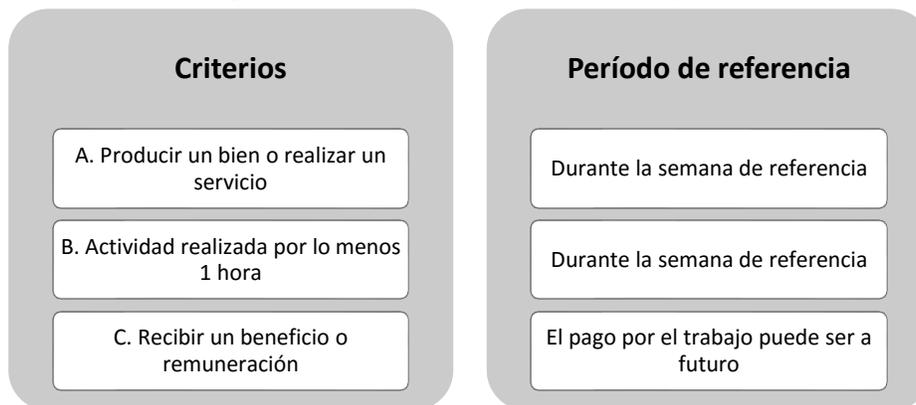
Fuente: Elaborado por el Instituto Nacional de Estadísticas.

En segundo lugar, el criterio de una hora, establecido por la OIT, es complementario al de pago en dinero o especies y alude al tiempo mínimo requerido en de la semana de referencia para la realización de una actividad económica con orientación al mercado. Estos dos conceptos permiten delimitar dentro de la fuera de trabajo a los ocupados y a los no ocupados. Estos últimos pueden diferenciarse, posteriormente, en desocupados y personas fuera de la fuerza de trabajo.

3.1.2.1. Ocupados

Los ocupados son todas las personas en edad de trabajar que, durante el período de referencia, cumplen con los siguientes criterios:

Figura 2. Criterios para clasificar como ocupado



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Entre la población ocupada, se puede hacer una distinción entre ocupados “tradicionales” y “no tradicionales”, lo que se relaciona con la concepción 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 tradicional”. Mientras que quienes no identifiquen desde un principio su actividad como un trabajo se clasifican como “ocupados no tradicionales”, mediante una mayor indagación respecto de la actividad realizada.

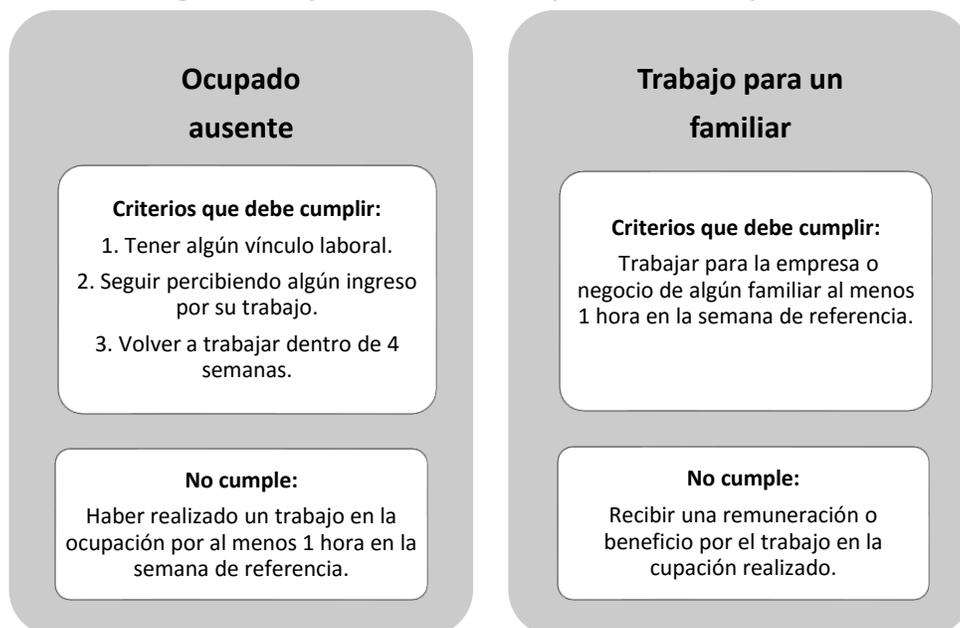
La mayor precisión en la clasificación de los ocupados permite incluir personas que desarrollan trabajos que, por sus características, ya sea en horas trabajadas, frecuencia con que lo realizan o lugar de trabajo, no habrían sido identificadas, en principio, como trabajo en la ocupación. Un ejemplo de 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.

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

- a) **Los trabajadores 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 trabajadores familiares no remunerados son las familias que trabajan en un puesto en la feria y que no tienen establecido un pago para sus integrantes.
- b) **Los ocupados 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 del entrevistado en reincorporarse o no al trabajo en un tiempo determinado y en la percepción o no de recibir una remuneración. 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: ocupados ausentes de su ocupación pero que mantienen un sueldo o ganancia.
 - Con vínculo por pronto retorno: ocupados ausentes de su ocupación pero que retornaran en un plazo de cuatro semanas.

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

Figura 3. Excepciones a los criterios para clasificar ocupación



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

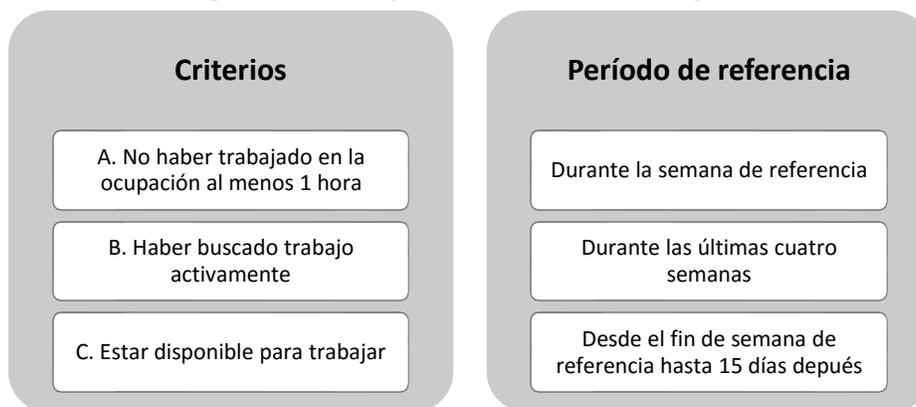
3.1.2.2. Desocupados

Son clasificados como desocupados aquellas personas de 15 años o más que durante el período de referencia estuvieron:

- a) 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.
- b) 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).
- c) 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 4 semanas.

Esto se resume en la siguiente figura:

Figura 4. Criterios para clasificar como desocupado



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

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 en el mercado laboral.

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.

Los desocupados se subdividen en:

- a) **Cesantes:** son todos los desocupados que han tenido anteriormente un trabajo en la ocupación que durara por lo menos un mes.
- b) **Busca trabajo por primera vez:** población que, habiendo cumplido con los criterios para ser clasificado como desocupado, 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, 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. Por lo tanto, en este caso no se exige el criterio de búsqueda efectiva de empleo.

Por otro lado, las personas ausentes del trabajo que no sean consideradas ocupadas, esto es, aquellas que reciben un ingreso del trabajo igual o menor al 50% de lo recibido habitualmente y que van a reincorporarse después de pasados 30 días, son desocupadas si buscaron trabajo y están disponibles para desempeñarlo.

3.1.3. Población fuera de la fuerza de trabajo

La población fuera de la fuerza de trabajo abarca a todas las personas de 15 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 ocupados o desocupados). Estos son:

- a) **Inactivos potenciales:** los inactivos potencialmente activos 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 el mercado laboral, efectivamente lo haga en algún momento.

- b) **Inactivos habituales:** son aquellas personas que no buscaron trabajo en la ocupación y no están disponibles para comenzar uno.

- c) **Iniciadores:** 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.

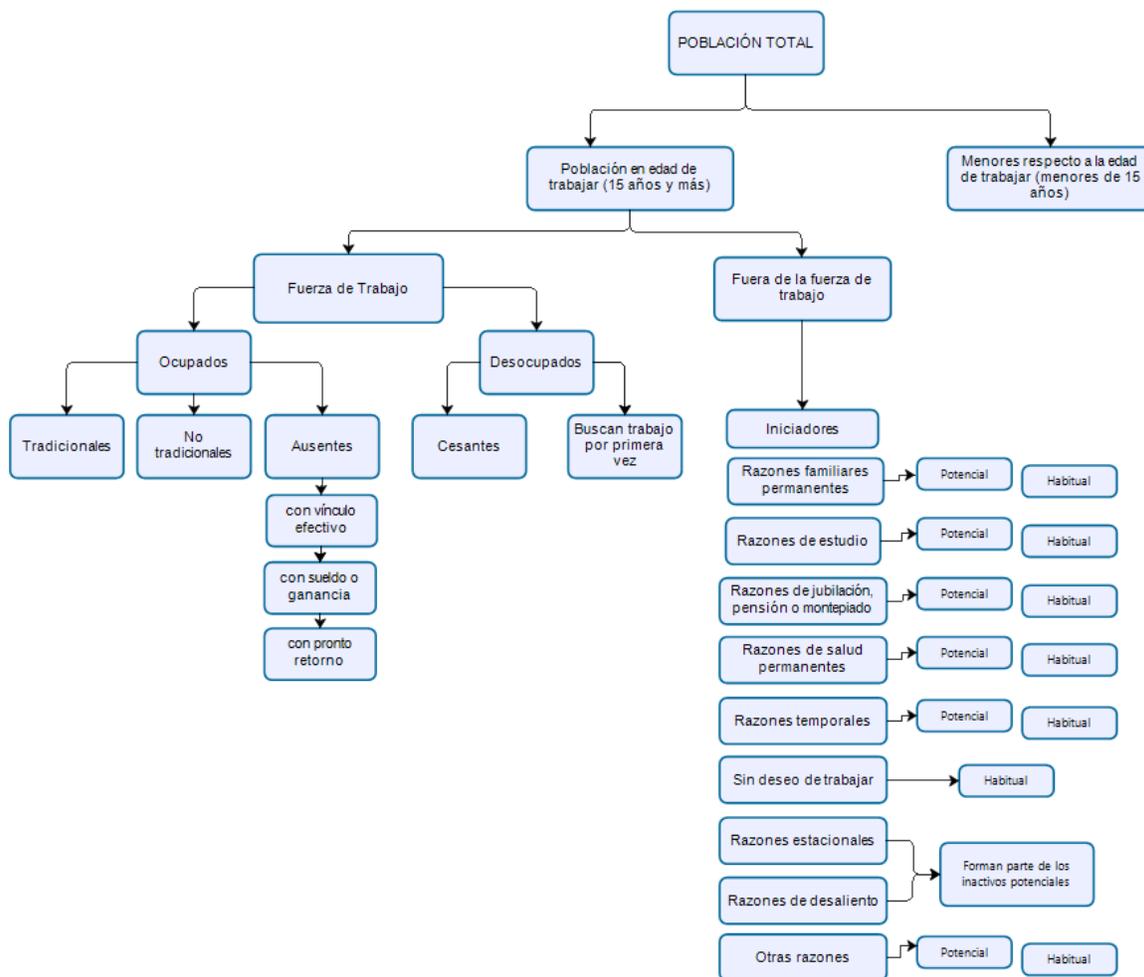
Estos últimos, si señalan estar disponibles, tanto la OIT como los países miembros de la Organización para la Cooperación y el Desarrollo Económico (OCDE), recomiendan incorporarlos dentro de los desocupados, y, por lo tanto, deben ser incluidos dentro del cálculo de la tasa de desocupación. En caso contrario, se les considera fuera de la fuerza de trabajo pues no buscaron ni estarían 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.

Por su parte, las personas fuera de la fuerza de trabajo esgrimen diversas razones de por qué se encuentran inactivas, las que pueden categorizarse según su potencialidad o habitualidad, algunas de ellas son: estacionales, de desaliento, personales temporales, familiares permanentes, de estudio, de pensión o montepiado, de jubilación, de salud permanente, sin deseo de trabajar y otras razones.

Todo lo presentado hasta este momento en materia de clasificación de la población en edad de trabajar se resume en la Figura 5.

Figura 5. Clasificación de la población de acuerdo con su vínculo con el mercado laboral (versión extendida)



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

3.1.4. Caracterización de la Población en Edad de Trabajar

Una vez identificada y clasificada la población en edad de trabajar según su vínculo con el mercado laboral, se caracteriza de acuerdo con una serie de variables sociodemográficas y económicas.

La captura de preguntas sociodemográficas permite caracterizar a la población a partir de ciertas variables básicas, segmentándola en grupos homogéneos o de interés. Estas variables se recogen en los cuestionarios sociodemográficos o tarjetas de registro del hogar y están asociadas a las personas que participan o no en la entrevista. En la ENE, se recolectan preguntas sociodemográficas de todos los integrantes del hogar, cumplan o no con el requisito de ser parte de la muestra objetivo de la encuesta, es decir, personas de 15 años o más.

En la encuesta, la caracterización sociodemográfica considera las preguntas sobre identificación del hogar y sus miembros, indagando sobre el sexo, la edad, fecha de nacimiento, el estado civil, el nivel educacional, el parentesco, el proveedor principal, la nacionalidad, migración, pueblos originarios, entre otras. Para la caracterización de la población según su nivel educativo, se utiliza la Clasificación Internacional

Normalizada de la Educación (CINE)⁹, concebida por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) en 1997. 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. Este clasificador cubre básicamente dos aspectos: niveles y campos de educación.

3.1.4.1. Caracterización de la población ocupada

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)¹⁰ del año 1993, de la Organización Internacional del Trabajo.

La CISE 93 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:

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

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 trabajadores bajo su responsabilidad; o *independiente*, trabajador sin vínculo contractual o acuerdo de trabajo, que puede tener o no trabajadores bajo su responsabilidad, asumiendo el riesgo económico inherente del ejercicio de su actividad.

Esta caracterización inicial se puede complementar, de acuerdo con las distintas necesidades nacionales de los países, permitiendo desagregarla y adaptarla al contexto nacional. De este modo, la adaptación nacional de la CISE 93, permite clasificar a los ocupados en:

- a) Empleadores: trabajador independiente que en su trabajo tienen trabajadores bajo su responsabilidad.
- b) Trabajador por cuenta propia: trabajador independiente que realiza su actividad económica sin empleador y trabaja solo o con un socio (sin trabajadores bajo su responsabilidad).
- c) Asalariado sector privado: trabajador dependiente con vínculo contractual en una unidad económica de tipo privada.
- d) Asalariado sector público: trabajador dependiente con vínculo contractual en una unidad económica de tipo pública o estatal, es decir, que trabajan en instituciones o empresas del

⁹ Para mayores detalles revisar la Clasificación Internacional Normalizada de la Educación CINE 2011, (UNESCO, 2006).

¹⁰ Para mayores detalles revisar la Clasificación Internacional de la Situación en el Empleo (CISE), (OIT, 1993).

Estado, ya sea dentro de la administración pública como en otros sectores económicos (salud, enseñanza, etc).

- e) Personal de servicio doméstico: trabajador dependiente que se desempeña en quehaceres domésticos, ya sea puertas adentro o puertas afuera.
- f) Familiar no remunerado: trabajador que, como se expuso previamente, no cumple con la condición de recibir un pago en dinero o en especie por su labor, pero debido a que su trabajo es realizado para algún familiar propietario de la unidad económica, participa indirectamente de los beneficios económicos producidos por esta unidad.

Una de las características de esta adaptación nacional es que permite desagregar a los trabajadores o personal de servicio doméstico (quienes realizan trabajo doméstico por una remuneración o beneficio), ya que en la CISE original éstos se encuentran incluidos dentro del grupo de trabajadores asalariados. También, se realiza la distinción entre los trabajadores asalariados del sector público y privado, que en la versión original son presentados como una gran categoría. Esta adaptación no considera la categoría “Miembros de cooperativas de productores”.

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)¹¹, que es una herramienta para organizar las ocupaciones en una serie de grupos definidos en función de las tareas que se desarrollan.

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, 2018a).

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, que incluye los siguientes grandes grupos:

- 1 Directores, gerentes y administradores
- 2 Profesionales, científicos e intelectuales
- 3 Técnicos y profesionales de nivel medio
- 4 Personal de apoyo administrativo

¹¹ Para mayores detalles ver el “Clasificador Chileno de Ocupaciones”, (INE, 2018a).

- 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

Adicionalmente, 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 a 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. De este modo, la CIUU hace posible establecer el sector de la unidad económica. Actualmente, y de manera oficial desde el año 2016, la ENE utiliza la Clasificación de Actividades Económicas Nacional de Encuestas Sociodemográficas (CAENES)¹², es una adaptación del CIIU4.cl 2012, desarrollado por el INE siguiendo recomendaciones internacionales. 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).

Cabe señalar que previamente, entre el año 2010 y 2016, la encuesta empleó la CIIU revisión 3 que contenía 17 secciones para la clasificación de las ramas de actividad económica.

El clasificador CAENES posee 21 secciones, a partir de las que se puede categorizar a los ocupados:

- 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

¹² Para mayores detalles ver el “Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas”, (INE, 2016a) y la separata técnica que aborda su adaptación (INE, 2016b)

- 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.

Para la aplicación de esta clasificación, es importante considerar para quién trabaja el ocupado y con quién está establecido el acuerdo contractual o de trabajo. En caso de no coincidir la unidad económica para la que se trabaja con la unidad económica con la que se tiene el acuerdo de trabajo, se está ante un caso de terciarización.

La terciarización 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 el trabajador 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 terciarización 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)

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 laboral¹³.

En el año 1993, la OIT en la 15° Conferencia Internacional de Estadísticos del Trabajo, realizó una propuesta estandarizada sobre el concepto de sector informal y los alcances de la informalidad laboral, desde la perspectiva de la unidad económica. En 1997 el Departamento de Estadísticas de Naciones Unidas creó el Grupo Delhi encargado de robustecer las normas estadísticas realizadas por OIT. Fue así como se incorporó el concepto de ocupación informal, en la 17° CIET de 2003, que introdujo la perspectiva laboral basada en las características y condiciones del puesto de trabajo que ocupa una persona. Ambos conceptos, permitieron establecer la definición de la economía informal como marco general e integrador para la medición de las estadísticas de informalidad laboral.

Posteriormente, la OIT en el año 2013, elaboró el Manual estadístico sobre el Sector Informal y el empleo informal, que constituye, principalmente, el marco conceptual para la medición de la informalidad laboral en la ENE.

En Chile se adopta la siguiente definición de sector y ocupación informal¹⁴:

- a) **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 cuasi-sociedades, ya que no poseen una contabilidad completa o simplificada que les permita realizar una efectiva separación de sus gastos, por el lado de los trabajadores 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 trabajadores asalariados.
- b) **Ocupación Informal:** Son todos aquellos asalariados o trabajadores 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 informales por definición a todos los familiares no remunerados del hogar, además de los trabajadores por cuenta propia y empleadores propietarios 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 dimensiones¹⁵. 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 2017¹⁶.

¹³ Para mayores detalles sobre el marco conceptual de la medición de Informalidad laboral, se sugiere revisar el documento "Estadísticas de informalidad laboral: Marco conceptual y manual metodológico", (INE, 2018b)

¹⁴ Para mayor información visitar <https://www.ine.cl/estadisticas/sociales/mercado-laboral/informalidad-y-condiciones-laborales>

¹⁵ A contar de julio de 2017, se inició la recolección oficial de las nuevas preguntas de informalidad laboral y, por tanto, el primer dato publicado de la serie fue el trimestre móvil de julio-septiembre 2017, siendo el 31 de enero 2018 la fecha de su lanzamiento.

¹⁶ El lanzamiento oficial de los datos sobre informalidad laboral se realizó el 31 de enero de 2018.

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.

3.1.4.2. Desocupados

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 cesantes, es decir, aquellas personas que han tenido anteriormente un empleo que duró por lo menos un mes.

Por lo tanto, se pueden clasificar de acuerdo con su última actividad en la ocupación. Para ello se utiliza un set de variables similares a los ocupados para obtener detalles como la rama de actividad económica en la que trabajó, la categoría en la ocupación, entre otras.

3.1.4.3. Fuera de la fuerza de trabajo¹⁷

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 ocupados o desocupados), determinando quienes son inactivos potencialmente activos, inactivos habituales o iniciadores.

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 al mercado laboral, 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 encontrarán, 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

¹⁷ Indentificados en el anterior marco conceptual de la encuesta como “inactivos”

estacionales, generalmente esta población se suma a los marginalmente vinculados siendo, de hecho, el grupo más numeroso entre éstos.

- **Otras razones:** es una categoría de personas fuera de la fuerza de trabajo que considera a todas aquellas razones que no califican en las otras categorías de inactividad.
- **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 los inactivos, 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 de trabajar:** Son personas que declaran no querer trabajar por lo que su potencial de participación es igual a cero.

3.1.5. Principales tasas e indicadores del mercado laboral

Del marco conceptual de la ENE, se derivan los siguientes indicadores principales para la medición.

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

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

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

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

3. *Tasa de participación*: es el porcentaje de personas que participa en el mercado laboral o componen la fuerza de trabajo dentro del total de personas que tienen edad de trabajar.

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

4. *Tasa de desocupación con iniciadores disponibles (SU1)*: es la tasa de desocupación incluyendo a los iniciadores 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 empezara una actividad laboral en el futuro próximo.

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

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

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

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 los iniciadores disponibles y la fuerza de trabajo potencial (inactivos potencialmente activos), dando cuenta de la desocupación de quienes participan del mercado del trabajo y quienes potencialmente podrían entrar a participar en el corto plazo.

$$SU3 = \frac{\text{desocupados} + \text{iniciadores disponibles} + \text{inactivos potencialmente activos}}{\text{fuerza de trabajo} + \text{iniciadores disponibles} + \text{inacivos potencialmente activos}} * 100$$

7. *Tasa de presión laboral*: es la razón entre los desocupados más los ocupados que buscan empleo más los iniciadores disponibles y la fuerza de trabajo más los iniciadores 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.

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

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 ocupados.

$$\text{Tasa de ocupación informal} = \frac{\text{ocupados informales}}{\text{ocupados totales}} * 100$$

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

$$\text{Tasa de ocupación en el sector informal} = \frac{\text{ocupados en unidades económicas informales}}{\text{ocupados totales}} * 100$$

De forma complementaria a los principales indicadores de ocupación, se sugiere la inclusión de otro tipo de indicadores, cuya evolución es relevante analizar para dar cuenta de la relación del trabajo con la economía. Estos indicadores permiten, además, comparaciones internacionales.

11. *Rama de actividad económica*: Distribución de las incidencias de las distintas ramas de actividad económica, de acuerdo con CAENES, en la población ocupada.
12. *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.
13. *Grupo ocupacional*: Distribución de las incidencias de las categorías de grupo ocupacional, de acuerdo con la CIUO 08.CL, en la población ocupada.
14. *Horas habituales y efectivas*: Distribución del volumen de trabajo en la ocupación, medido en horas efectivas trabajadas por la población ocupada. Se recomienda incluir la separación entre trabajo de tiempo parcial voluntario e involuntario, con objeto de evidenciar la subutilización de la fuerza de trabajo.

4. Características del diseño muestral

El diseño muestral de la ENE implementado a partir de enero de 2020, tiene como objetivo, generar estimaciones de los principales indicadores¹⁸ de la población respecto a su vínculo con el mercado laboral, con mejores niveles de precisión en las estimaciones de los dominios de estudio definidos, implementando mecanismo 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.

4.1. Población objetivo

Corresponde a las personas, de 15 años o más, que residen habitualmente en viviendas particulares ocupadas situadas en el territorio nacional.

4.2. Cobertura geográfica y dominios de estudio

La ENE busca ser representativa de todo el territorio nacional, abarcando 97% de las comunas de todo el territorio, por lo tanto, se utiliza como insumo casi la totalidad del marco muestral maestro¹⁹.

El diseño muestral de la ENE implementado a partir de enero de 2020 garantiza representatividad y errores de muestreo aceptables²⁰ para los dominios de estudio: nacional, nacional urbano, nacional rural, regional, el área urbana de las 16 regiones y el área rural de las regiones de O'Higgins, El Maule, La Araucanía, Los Lagos, Metropolitana, Los Ríos y Ñuble.

4.3. Período de referencia y periodicidad

Para el período de referencia de la información recogida, se adopta como norma, la semana anterior (de lunes a domingo) a la fecha en que se realiza la entrevista, en línea con la recomendación de OIT (19° CIET, 2013). Sin embargo, para las preguntas referidas a la medición de ingresos, que se levanta una vez al año en el último trimestre calendario, el período de referencia es principalmente el mes anterior incluyendo el pago devengado.

La ENE es una encuesta continua, con periodicidad de publicación mensual de los resultados obtenidos a partir de un trimestre móvil.

¹⁸ Tasa de desocupación, tasa de ocupación y tasa de participación.

¹⁹ Debido a que se realizan algunas exclusiones por motivos operativos, la ENE no cubre 0,57% del total de viviendas a nivel nacional, lo que corresponde a 35.287 viviendas. El proceso de exclusión se explica con mayor detalle en el apartado 4.5.2.

²⁰ Para más información consultar documento "Lineamientos para el uso de medidas de precisión en encuestas de hogares" (INE, 2020).

4.4. 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 geográficas²¹, 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.

La muestra de la ENE, vigente desde enero de 2020, fue seleccionada desde el Marco Muestral de Viviendas²², elaborado a partir de los resultados del Censo de Población y Vivienda 2017, y los listados de viviendas asociados al Precenso 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 ausentes²³ 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].

²¹ 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.

²² 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 Precenso 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 Precenso 2016 (MMS 2016) para los estratos RAU (Resto de Área Urbana) y rural.

²³ Se excluyen las viviendas de temporada y colectivas

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.

4.4.1. 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.

4.4.2. Estratificación geográfica del marco muestral

El MMV 2017 contiene variables que permiten la identificación de la división política administrativa de país²⁴, 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 amanzanamiento 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 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.

²⁴ Región, provincia y comuna.

Figura 6: Estratificación geográfica del MMV 2017



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

4.4.3. 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.

4.4.3.1. 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 7: 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.

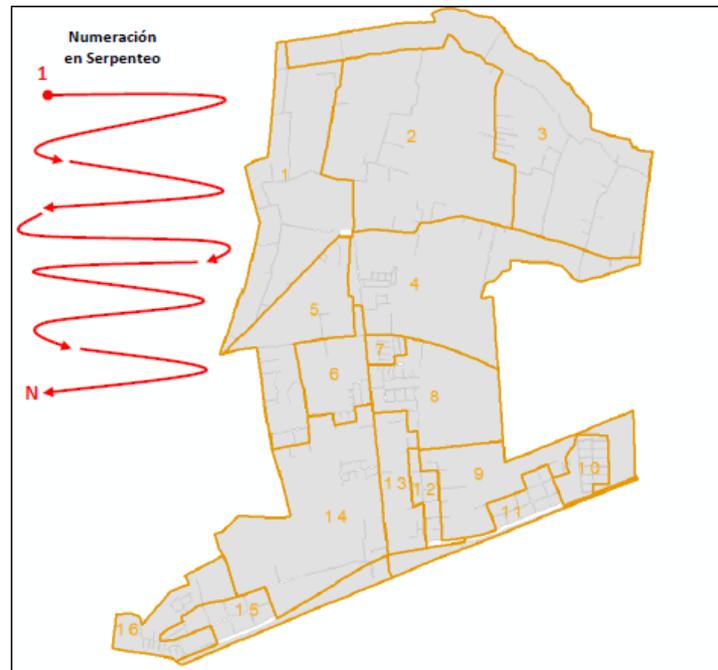
Para la construcción de las UPM, el criterio fundamental es mantener unidades compactas y homogéneas, con el objetivo de que el muestreo bietápico sea eficiente y la muestra pueda entrar en operación en forma óptima. Por esta razón, existen excepciones para el rango de tamaño de las UPM con el propósito 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 dividir las.

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 entorno²⁵.

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 8).

²⁵ 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).

Figura 8: Numeración por método de serpiente



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

4.4.3.2. 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 9: 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 exceda 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 otros²⁶.

4.4.4. 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 por 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 variable²⁷.

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 3 y 5 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:

- Tasa de ocupación por sexo
- Tasa de desocupación por sexo

²⁶ 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.

²⁷ Mayores antecedentes disponibles en el documento “Estratificación socioeconómica del marco muestral de viviendas 2017 (MMV 2017)”, (INE, 2019c).

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

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 3 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.

4.4.5. 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:

- 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 posibilite la coordinación y valorización del esfuerzo operativo con información más detallada para el trabajo de campo.

4.4.6. 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 2. UPM con tamaño fuera de rango por región

Región	Total UPM	Urbano		Total UPM	Rural	
		UPM < 160 viviendas	UPM > 240 viviendas		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 3. 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
Metropolitana	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 4. 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.

4.5. 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.

4.5.1. 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 11 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 5, mientras que los casos excepcionales de estratos con menos de 36 UPM se detallan en la Tabla 6. El listado completo de los estratos geográficos queda disponible en Anexo N° 1.

Tabla 5. Cantidad de estratos geográficos por región y área

Región	Área		Total
	Urbano	Rural	
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 6. 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.

4.5.2. 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 levantamiento, de modo que la decisión incluya también las capacidades instaladas de cobertura²⁸.

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 7 muestra la distribución regional de UPM y viviendas marco, especiales y excluidas.

²⁸ 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 7. 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.

4.5.3. 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 8.

Tabla 8. Cantidad de estratos de muestreo por región y área

Región	Área		Total
	Urbano	Rural	
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.

4.6. Unidades estadísticas

4.6.1. Unidad de muestreo

Existen dos unidades de muestreo.

- a) Unidad primaria de muestreo (UPM) definida como un conglomerado homogéneo, en términos de la cantidad de viviendas particulares²⁹ 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 4.4 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.

4.6.2. 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 15 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 cada una de las personas residentes en la vivienda seleccionada mediante el cuestionario construido para estos fines.

²⁹ Excluyendo las viviendas de temporada.

5 Tamaño muestral

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.

5.1. Estrategia muestral

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

Los estratos de muestreo corresponden a la combinación “Estrato geográfico – Estrato Socioeconómico”, en aquellos estratos donde, según lo descrito en el apartado 4.5.3, puede existir dicha combinación.

5.1.1. 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 9. 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 UPM³⁰. Asimismo, se estimó el efecto de diseño a nivel de región-área que se utiliza en el cálculo del tamaño definitivo.

5.2. 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 2017³¹, 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 desempleo³² 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.

³⁰ En el apartado 5.5 se desarrolla la metodología implementada para determinar el número óptimo de viviendas por UPM.

³¹ 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.

³² $P_{ra} = \frac{\text{Total de personas desocupadas}}{\text{Total de personas en la fuerza de trabajo}}$

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

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 (3).

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

Donde,

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

def_{ra} : Efecto de diseño a nivel de viviendas para el dominio región-área. Estimado a partir de simulaciones³³ 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 (3), con información a nivel de personas en la fuerza de trabajo³⁴, 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.

³³ Para más información consultar el Informe de Simulaciones de tamaños muestrales de la ENE 2020.

³⁴ 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 ocupados o desocupados.

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.

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:
- Todos los estratos deben contar con muestra de al menos seis³⁵ UPM.
 - 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.

5.3. 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 **¡Error! No se encuentra el origen de la referencia..** 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:

³⁵ Con la finalidad de asegurar un mínimo de 2 UPM mensuales y el aporte del estrato al cálculo de la varianza.

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

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 **¡Error! No se encuentra el origen de la referencia.**, 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 10 (ver apartado 5.5) 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 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 muestreo³⁶, 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.

5.4. 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

³⁶ Estos casos se presentaron en dos estratos urbanos de la región de Aysén.

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 logro³⁷ 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 visita³⁸ señalados en la Tarjeta de Registro del Hogar (TRH)³⁹vigente en la versión anterior de la encuesta. Esta clasificación permite categorizar la situación de las viviendas en 13 códigos diferentes y cuatro grupos.

La tasa de logro sobre viviendas gestionadas⁴⁰ (TL_{gest}) se calcula mediante la siguiente expresión:

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

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 permita compensar la no respuesta, independiente de la causa que la origine.

³⁷ 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.).

³⁸ Para conocer más detalle sobre los códigos de registro de visitas empleados, ver Anexo N° 3. Códigos de visita de la ENE.

³⁹ 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.

⁴⁰ 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.

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 posible⁴¹, 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 ronda⁴² 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 ronda⁴³ 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.

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

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

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_{hgest} : Tasa de no logro sobre viviendas gestionadas en el estrato h .

5.5. 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.

⁴¹ 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.

⁴² Entiéndase por ronda, el número de entrevistas que es realizada a las mismas unidades de análisis de manera consecutiva.

⁴³ Aunque el diseño muestral anterior consideraba seis rondas para el estrato CD, nueve para el RAU y 12 para el rural, tomando en cuenta que bajo el nuevo diseño se realizaron seis rondas en los estratos urbano y rural (ver sección 5.7), solo se analizaron las tasas de no logro para las seis primeras seis rondas.

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 intraclase 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)⁴⁴:

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

Donde,

deff : Efecto del diseño.

\tilde{n} : Número de unidades promedio de la población objetivo⁴⁵ 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 10, se puede observar cómo influyen en el efecto del diseño el tamaño medio de la muestra objetivo por UPM y la correlación intraclase.

⁴⁴ 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.

⁴⁵ 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.

Tabla 10. 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.

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 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 11, 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 conglomerados⁴⁶. Como puede apreciarse todos los efectos de diseño obtenidos son menores a 3, por lo que una muestra de 6, 8, 10 o 12 viviendas por UPM, garantiza estimaciones con un efecto de diseño aceptable.

⁴⁶ Número de viviendas del conglomerado (*m*) multiplicado por el promedio de personas en la fuerza de trabajo por vivienda (Pers. FTTT).

Tabla 11. Efectos de diseño por dominios de estudio, para distintos tamaños de viviendas promedio por UPM

Región	Urbano						Rural					
	ρ	Pers. FTT x Viv.	deff aproximado según \bar{n}				ρ	Pers. FTT x Viv.	deff aproximado según \bar{n}			
			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 tolerancia⁴⁷, que los tamaños 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 12 muestra los errores absolutos obtenidos para la tasa de desocupación y los errores relativos obtenidos para el nivel de desocupados 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 12 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 (8, 10 y 12 viviendas) y en la región de Magallanes con 10 y 12 viviendas.

⁴⁷ 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 desocupados, se fija como error máximo admisible 30%, en términos relativos.

Tabla 12. Errores muestrales de la tasa de desocupación para los dominios urbanos según número promedio de vivienda por UPM

Región	Tasa de Desocupación			Nivel de desocupados		
	Error absoluto ⁴⁸			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 8 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.

En el caso de los dominios rurales, la Tabla 13 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 10 viviendas por UPM, la mayoría de los errores absolutos resultaron mayores a 2%, por ende en los casos de 12 y 14 viviendas por UPM, los errores absolutos resultaron aún mayores.

⁴⁸ Para más información consultar el Informe de Simulaciones de tamaños muestrales de la ENE 2020.

Tabla 13. Errores muestrales de la tasa de desocupación para los dominios rurales según número promedio de vivienda por UPM

Región	Tasa de Desocupación			Nivel de desocupados		
	Error absoluto ⁴⁹			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 10 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 10 viviendas o más resultaban superiores a los límites de tolerancia.

Por último, en la **¡Error! No se encuentra el origen de la referencia.** se presentan los pasos de cálculo realizados hasta obtener el tamaño de muestra objetivo definitivo. Posteriormente, la Tabla 14 presenta los tamaños a nivel de viviendas y UPM, y los errores de simulación asociados, para todos los niveles de estimación.

⁴⁹ Para más información consultar el Informe de Simulaciones de tamaños muestrales de la ENE 2020.

Tabla 14. 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.

5.6. Selección de unidades muestrales

5.6.1. 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 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:

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

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 .

5.6.2. 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 previo⁵⁰, 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:

Se calcula el período K .

⁵⁰ 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.

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

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:

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

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.

5.7. 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 rotativo⁵¹ 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 submuestras⁵². 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 contemplan 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.

⁵¹ 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.

⁵² Este procedimiento se lleva a cabo distribuyendo equitativamente la muestra de UPM de cada estrato de muestreo ENE, entre las tres submuestras.

Figura 10. Esquema de rotación de la muestra de viviendas

		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
Panel de rotación	6	R6	R6	R6																					
	5	R5	R5	R5	R6	R6	R6																		
	4	R4	R4	R4	R5	R5	R5	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									
	1	R1	R1	R1	R2	R2	R2	R3	R3	R3	R4	R4	R4	R5	R5	R5	R6	R6	R6						
	6				R1	R1	R1	R2	R2	R2	R3	R3	R3	R4	R4	R4	R5	R5	R5	R6	R6	R6			
	5							R1	R1	R1	R2	R2	R2	R3	R3	R3	R4	R4	R4	R5	R5	R5	R6	R6	R6
	4										R1	R1	R1	R2	R2	R2	R3	R3	R3	R4	R4	R4	R5	R5	R5
	3													R1	R1	R1	R2	R2	R2	R3	R3	R3	R4	R4	R4
	2																R1	R1	R1	R2	R2	R2	R3	R3	R3
	1																			R1	R1	R1	R2	R2	R2
6																						R1	R1	R1	

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

La Figura 10 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 mensualmente⁵³. 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.

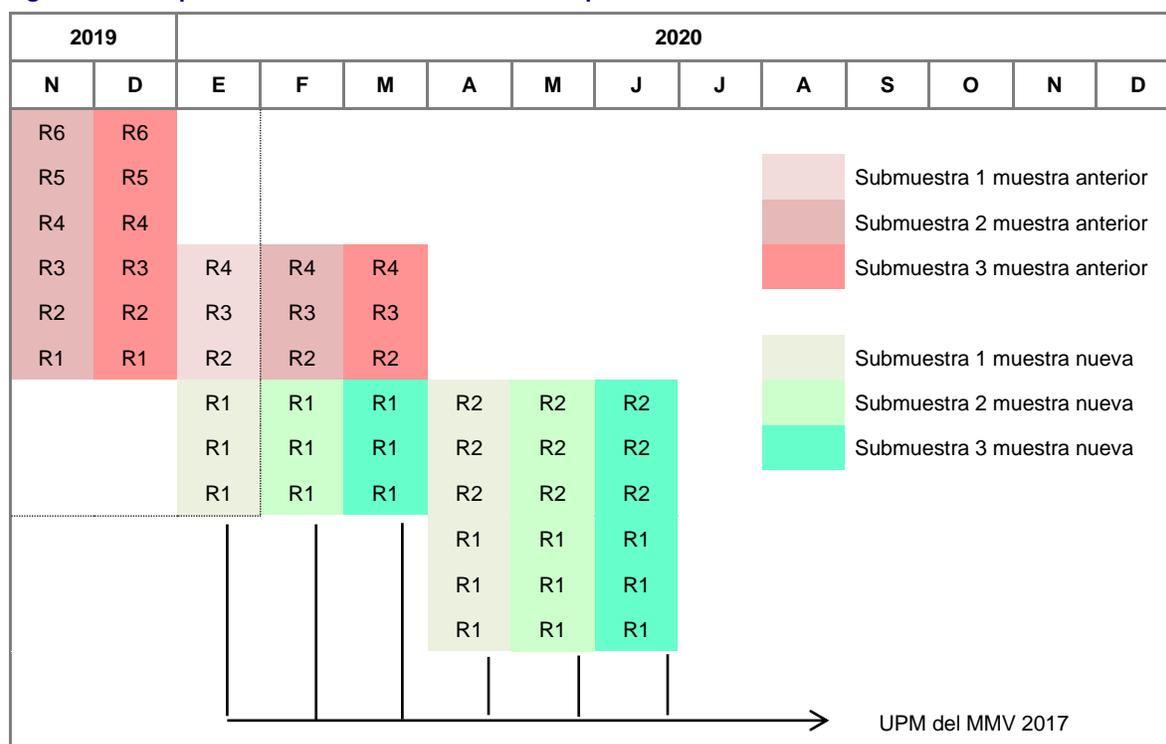
⁵³ Se debe recordar que la muestra total ENE está conformada por las tres submuestras, por lo que, de un mes a otro, la fracción que se renueva representa una dieciochoava parte de la muestra total de la ENE.

5.8. 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 (NDE2019) es incorporada una fracción de la muestra proveniente del nuevo marco, a razón de 1/6 de la muestra trimestral⁵⁴ (aproximadamente 9.000 viviendas).

Paralelamente 1/6 de la muestra proveniente del marco muestral anterior⁵⁵ (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 (AMJ2020) la muestra ENE tendrá un nuevo tamaño muestral⁵⁶ aproximado de 54.000⁵⁷ viviendas, todas provenientes del nuevo MMV 2017.

Figura 11. Incorporación de la muestra MMV 2017 a partir de las rotaciones



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

⁵⁴ Considerar que este 1/6 es calculado sobre el nuevo tamaño muestral de la ENE, el cual contempla una sobremuestra.

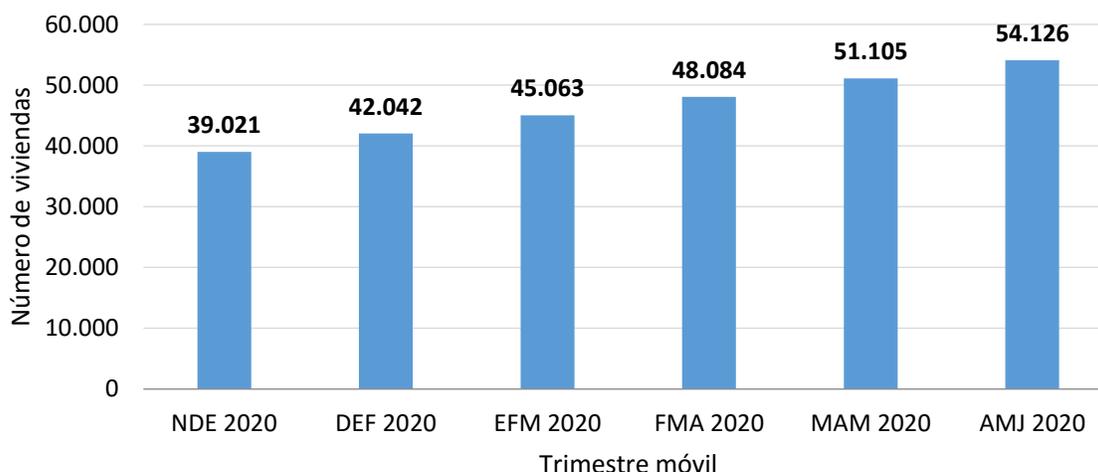
⁵⁵ Este 1/6 es calculado sobre el tamaño de muestra del diseño muestral de la ENE vigente hasta diciembre 2019.

⁵⁶ El tamaño muestral del diseño anterior era de 38.000 viviendas y no contemplaba sobremuestra.

⁵⁷ Incluyendo la sobremuestra.

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 de muestra definido para alcanzar los niveles de precisión óptimos, en el mes de junio 2020⁵⁸ (ver Gráfico 1).

Gráfico 1. Incremento gradual de la muestra teórica durante el período de transición de la ENE



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

5.8.1. 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 (10 viviendas), con excepción de los estratos urbanos de las regiones de Atacama, Aysén y Magallanes que es de 8 viviendas.

5.8.2. Selección de las viviendas a restar en cada mes

La selección de la muestra proveniente de los marcos MMS2016 y MMM2016⁵⁹ 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

⁵⁸ 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.

⁵⁹ 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 Precenso 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 Precenso 2016 (MMS 2016) para los estratos RAU (Resto de Área Urbana) y rural.

construcción de las unidades primarias de muestreo entre los marcos MMS2016, MMM2016 y MMV 2017.

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 5, 4 y hasta 3 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.

6 Desarrollo 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 unidades⁶⁰, 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

⁶⁰ 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).

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.

6.1. 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”.

6.2. 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 .
- $\pi_{hj/i}$: 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 .

6.2.1. 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.

6.2.1.1. 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:

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

Donde,

- Θ_i : Conjunto de viviendas j seleccionadas en el conglomerado i .
- $\Theta_{i,known}$: 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, $\hat{R}_{h,known}$ 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:

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

6.2.1.2. Ajuste por no elegibilidad

La población objetivo de la ENE está conformada por aquellas personas de 15 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:

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

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

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

6.2.2. 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 motivos⁶¹.

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 6.2.1.2 (corregido por no elegibilidad); y la estimación del total de viviendas de las que se obtiene 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.

⁶¹ En correspondencia con los códigos de disposición final definidos en el instrumento de recolección a partir de enero 2020.

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} .

6.2.3. 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 15 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 (Vanderhoeft, 2001). Este último es el método escogido para la calibración de la ENE.

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 $g^{-1}(w^{cal}, w^{NR})$.

Las marginales definidas⁶² son las siguientes:

- A nivel nacional: según sexo, 13 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 extranjera⁶³. Esta marginal se incluirá a partir del trimestre móvil EFM 2013. La construcción del insumo para la calibración mediante esta marginal se detalla en el Anexo N°8.

⁶² 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).

⁶³ 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)

- A nivel regional: Población total según sexo.
- A nivel regional: Población en Edad de Trabajar (PEA) según sexo.
- A nivel regional: Población total Área Urbana según sexo.
- Región Metropolitana: Población total para 6 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 conurbaciones.

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.
 - 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 4 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.

7 Estimadores

Luego de haber definido las características del diseño muestral de la ENE, se presentan a continuación las fórmulas implementadas para el cálculo de los principales estimadores.

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

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

También puede ser expresado como:

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

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.

7.1. Estimadores de total para la variable Y

La ENE genera estimaciones para los siguientes totales:

- Personas en la Fuerza de trabajo
- Ocupados
- Ocupados informales

- Desocupados
- Inactivos

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 **Y** en el estrato **h**.*

$$\hat{Y}_h = \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk} \quad (28)$$

*Total estimado de la variable **Y** en el nivel de estimación requerido.*

$$\hat{Y} = \sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk} = \sum_h \hat{Y}_h \quad (29)$$

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 **Y** en personas con la característica **C** el estrato **h**.*

$$\hat{Y}_{h_c} = \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk_c} \quad (30)$$

*Total estimado de la variable **Y** en personas con la característica **C** en el nivel de estimación requerido.*

$$\hat{Y}_C = \sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk_c} = \sum_h \hat{Y}_{h_c} \quad (31)$$

7.2. Estimadores 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.

La ENE genera estimaciones de razón para obtener los siguientes indicadores:

- Tasa de participación
- Tasa de ocupación
- Tasa de desocupación
- Tasa de desocupación con iniciadores disponibles
- Tasa combinada de desocupación y tiempo parcial involuntario
- Tasa combinada de desocupación y fuerza de trabajo potencial
- Tasa de presión laboral
- Tasa de ocupación informal
- Tasa de ocupación en el sector informal

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:

$$\hat{R}_{h_{y/x}} = \frac{\hat{Y}_h}{\hat{X}_h} = \frac{\sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk}}{\sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot X_{hijk}} \quad (32)$$

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.

$$\hat{R}_{y/x} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk}}{\sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot X_{hijk}} \quad (33)$$

8 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 muchos 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.

8.1. 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 cualquier de ellas a partir de las otras mediante la aplicación simples 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 6.2.3), 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)⁶⁴, 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

⁶⁴ Para más información ver (Särndal, Swensson, & Wretman, 1991, pág. 235).

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).

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 Raking⁶⁵.

8.2. 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.

8.2.1. 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 estudio⁶⁶).

El error absoluto permite observar cuanto 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 ν grados de libertad ($t_{1-\frac{\alpha}{2}, \nu}$), 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 ν grados de libertad con los que se calcula el valor crítico⁶⁷. Su expresión algebraica está dada por:

$$E_{abs}(\hat{y}) = t_{1-\frac{\alpha}{2}, \nu} \cdot S(\hat{y}) \quad (34)$$

⁶⁵ 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).

⁶⁶ Cuando la variable Y es un porcentaje o tasa, el error absoluto se expresa en puntos porcentuales.

⁶⁷ 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.

Alternativamente, en términos probabilísticos, se puede expresar como:

$$\Pr\left(|\hat{Y} - Y| \leq E_{abs}(\hat{Y})\right) = 1 - \alpha \quad (35)$$

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

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:

$$IC(\hat{Y})_{95\%} = \left(\hat{Y} - E_{abs}(\hat{Y}); \hat{Y} + E_{abs}(\hat{Y})\right) = \left(\hat{Y} - t_{1-\frac{\alpha}{2},v} \cdot S(\hat{Y}); \hat{Y} + t_{1-\frac{\alpha}{2},v} \cdot S(\hat{Y})\right) \quad (36)$$

8.2.2. 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:

$$E_{rel}(\hat{Y}) = t_{1-\frac{\alpha}{2},v} \cdot \frac{\hat{\sigma}_{\hat{Y}}}{\hat{Y}} = t_{1-\frac{\alpha}{2},v} \cdot CV_{\hat{Y}} \quad (37)$$

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.

8.3. 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 flujograma para publicar o suprimir estimaciones (INE, 2020), el cual contempla los siguientes elementos:

8.3.1. 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.

8.3.2. 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.

9 Alcance de la presentación de resultados

9.1. Difusión de resultados

Los resultados y el material elaborado a partir de la encuesta es diseminado al público y los usuarios disponibles, a través de la página web institucional (www.ine.cl/), el Banco de Datos ENE y la plataforma INE.STAT (<https://stat.ine.cl>).

La publicación de resultados se realiza periódicamente, a las 9:00 am horas del último día hábil de cada mes, con excepción de la plataforma INE.STAT la cual es actualizada 48 horas hábiles luego de la publicación de la cifra.

El material publicado por la encuesta corresponde a los siguientes:

- **Boletín de empleo Nacional:** Documento que detalla los principales indicadores del mercado del trabajo en el país, describe las variaciones de la situación dentro y fuera de la fuerza de trabajo en distintas categorías, cómo el grupo ocupacional, la rama de actividad económica, la categoría en la ocupación entre otros. Incluye tasas de subutilización, con ajuste estacional, informalidad laboral y comparación con registros administrativos.
- **Boletín de empleo región metropolitana y gran Santiago:** Documento que detalla los principales indicadores del mercado del trabajo en la región metropolitana y el gran Santiago, describe las variaciones de la situación dentro y fuera de la fuerza de trabajo en distintas categorías, cómo el grupo ocupacional, la rama de actividad económica, la categoría en la ocupación entre otros.
- **Nota estadística de empleo:** Archivo que contiene tablas con la información principal de la ENE (niveles y tasa de desocupación), además de tasas de respuestas, error absoluto, coeficientes de variación e intervalos de confianza, según corresponda.
- **Nota estadística de informalidad laboral:** Archivo que contiene tablas con la información principal sobre informalidad laboral (niveles de ocupados informales y tasa de ocupación informal), además de tasas de respuestas, error absoluto, coeficientes de variación e intervalos de confianza, según corresponda.
- **Boletín de informalidad laboral:** Documento de publicación trimestral que detalla los principales indicadores de ocupación informal en el trabajo en el país, describe las variaciones ocupados formales e informales en distintas categorías, cómo el grupo ocupacional, la rama de actividad económica, la categoría en la ocupación entre otros. Incluye tasas de ocupación informal y la tasa de ocupación en el sector informal.
- **Anexos coyuntura ENE:** Archivo que contiene tablas con los niveles principales de la ENE, además de su variación absoluta y relativa anual e incidencias.

- **Separata Anual:** Documento que resume la dinámica promedio de las principales clasificaciones y caracterizaciones en la ocupación ocurridas el año que terminó el mes pasado, dicha síntesis se realiza a nivel nacional y regional, destacando las principales variaciones según incidencia en las principales desagregaciones como (sexo, rama, categoría ocupacional y principales tasas).
- **Separata técnica:** Documento que resume una dinámica y/o situación en particular sobre aspectos técnicos y/o operativos relacionados con la Encuesta Nacional de Empleo. Su publicación es eventual, y generalmente cumple con el objetivo de apoyar la difusión de la información de manera precisa, simple y oportuna.
- **Base de datos:** Base de datos anonimizadas con la información trimestral de la ENE desde año 2010 en adelante.
- **Tabulados:** Series de datos sobre la información principal de la encuesta, a nivel nacional y regional, desde el año 2010 en adelante, con actualización trimestral.
- **Metodología de la Encuesta Nacional de Empleo:** Documento que sintetiza los principales aspectos metodológicos del estudio, específicamente, resume el marco conceptual de la medición del empleo y el diseño estadístico de la encuesta actualizado en el 2020.

9.2. 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:

- Ley 19.628 de protección de vida privada y datos personales.
- Ley 20.285 regula el principio de transparencia de la función pública y el derecho de acceso a la información de los órganos de la Administración del Estado.
- Artículo 29° de la Ley 17.374- El Instituto Nacional de Estadísticas, los organismos fiscales, semifiscales y Empresas del Estado, y cada uno de sus respectivos funcionarios, no podrán divulgar los hechos que se refieren a personas o entidades determinadas de que hayan tomado conocimiento en el desempeño de sus actividades.
- El estricto mantenimiento de estas reservas constituye el "Secreto Estadístico". Su infracción por cualquier persona sujeta a esta obligación hará incurrir en el delito previsto y penado por el artículo 247, del Código Penal, debiendo en todo caso aplicarse pena corporal.
- Artículo 30° de la Ley 17.374- Los datos estadísticos no podrán ser publicados o difundidos con referencia expresa a las personas o entidades a quienes directa o indirectamente se

refieran, si mediare prohibición del o los afectados.” (Extractos de la Ley Orgánica 17.374 de creación del Instituto Nacional de Estadísticas, publicada el 10-12-1970.)

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

10 Anexos

Anexo N° 1. 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 (1) (Valparaíso)	549	110.459	243.343
Valparaíso	Urbano	510021	Conurbación Valparaíso (2) (Viña del Mar, Concón)	728	148.135	310.830

Región	Área	Código de estrato geográfico	Estrato geográfico	UPM	Viviendas	PET
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, Penciahue, Curepto)	88	17.516	36.876
Maule	Urbano	710031	R.P. Talca (2) (Pelarco, Río 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, Yervas 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) (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, Yervas 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, Pitrufrquén, 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, Pitrufrquén, 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 (Futroneo, 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 (Futroneo, 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.

Anexo N° 2. 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
Valparaíso	Urbano	510021	5100211	90	18.227
Valparaíso	Urbano	510021	5100212	381	77.601

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
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
Metropolitana	Urbano	1320011	13200113	177	37.472
Metropolitana	Urbano	1320021	13200210	35	7.124

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
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	610011	6100111	116	23.91
O'Higgins	Urbano	610011	6100112	364	76.037
O'Higgins	Urbano	610011	6100113	117	24.524
O'Higgins	Urbano	610021	6100210	47	9.394
O'Higgins	Urbano	610031	6100310	255	50.988
O'Higgins	Urbano	620011	6200110	113	23.774
O'Higgins	Urbano	620021	6200210	47	9.33
O'Higgins	Urbano	630011	6300110	137	28.124
O'Higgins	Rural	610012	6100120	100	8.773
O'Higgins	Rural	610022	6100220	41	3.575
O'Higgins	Rural	610032	6100320	374	33.299
O'Higgins	Rural	620012	6200120	40	3.664
O'Higgins	Rural	620022	6200220	111	10.347
O'Higgins	Rural	630012	6300121	224	20.574
O'Higgins	Rural	630012	6300122	58	5.033
Maule	Urbano	710011	7100111	95	19.412
Maule	Urbano	710011	7100112	253	51.356
Maule	Urbano	710011	7100113	99	19.958
Maule	Urbano	710021	7100210	88	17.516

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
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
Biobío	Urbano	830021	8300210	144	29.455
Biobío	Urbano	830031	8300310	95	19.479

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
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
Los Lagos	Rural	1010022	10100220	152	14.208
Los Lagos	Rural	1010032	10100320	42	3.699

Región	Área	Código estrato geográfico	Código estrato de muestreo	UPM	Viviendas
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.

Anexo N° 3. Códigos de visita de la ENE

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.

Códigos de visita o disposición de la ENE 2020.

Código	Glosa descriptiva
Elegibles, entrevistadas	
11	Entrevista completa
12	Entrevista parcial
Elegibles, rechazadas	
21	Rechazo del hogar
Elegibles, no contactadas	
31	Vivienda ocupada, sin moradores presentes
32	Se impide el acceso a la vivienda
33	Informante no ubicable, se contacta a otra persona
34	Informante ausente en período de levantamiento de la encuesta
Elegibles, no entrevistadas por otros motivos	
41	Informante impedido/a 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
49	Otra razón de no entrevista
Elegibilidad desconocida	
51	No se envió a terreno
52	No fue posible localizar la dirección
53	Inaccesibilidad por razones climáticas o desastres naturales
54	Área peligrosa
59	Otra razón de elegibilidad desconocida
No elegibles	
61	Empresa, oficina de gobierno u otra organización
62	Institución (hospital, cárcel, asilo de ancianos, etc.)
63	Dormitorio colectivo (militar, de trabajo, internado, hostales, residenciales, etc.)
64	Vivienda en demolición, incendiada, destruida o erradicada
65	Vivienda desocupada (por ejemplo, en venta, arriendo)
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

Fuente: Elaborado por el Instituto Nacional de Estadísticas.

Anexo N° 4. 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%
Metropolitana	Urbano	1320021	36%	30%
Metropolitana	Rural	1310012	36%	30%

Región	Área	Código de estrato geográfico	Tasas de no logro	Tasas de no logro (Truncada a 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%
Maule	Urbano	710021	43%	30%
Maule	Urbano	710011	36%	30%

Región	Área	Código de estrato geográfico	Tasas de no logro	Tasas de no logro (Truncada a 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%
La Araucanía	Rural	920012	15%	15%
La Araucanía	Rural	910032	15%	15%

Región	Área	Código de estrato geográfico	Tasas de no logro	Tasas de no logro (Truncada a 30%)
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.

Anexo N° 5. Estimaciones de población para la Encuesta Nacional de Empleo

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.

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.

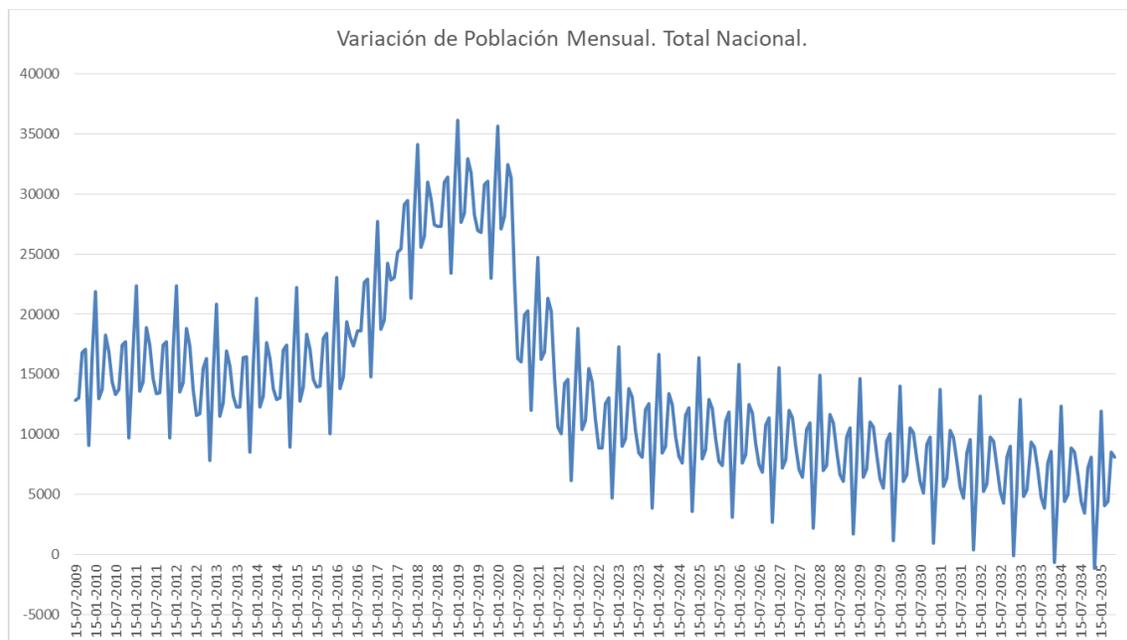
P_i : Población pivote inicial.

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.



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.



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.

Anexo N° 6. 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
Maule	71	Talca	357.162	EFM	2010
Maule	73	Curicó	249.679	EFM	2010
Maule	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.

Anexo N° 7. Cálculo de los factores de expansión en el período de transición

De acuerdo con lo expuesto en el apartado 5.8 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.

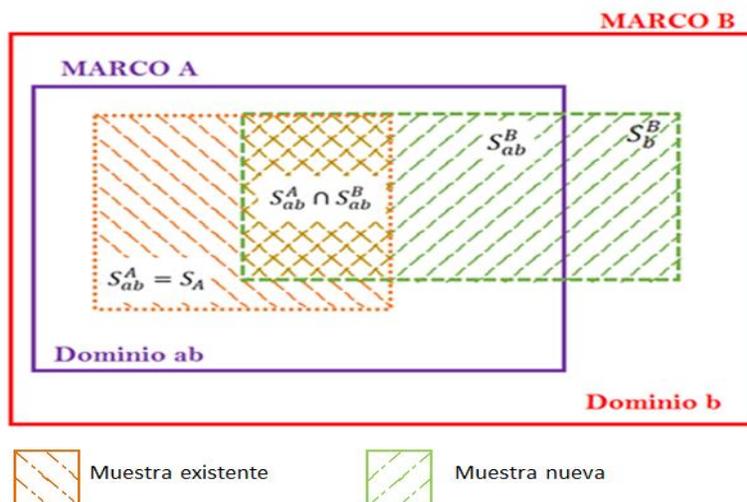
Para efectos de nomenclatura denominaremos marco “A” a la unión de los listados MMS 2016 y MMM 2016 y marco “B” al MMV 2017.

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 población⁶⁸.

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 12 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 12: Esquema de configuración de marcos y muestras



Fuente: Elaborado por el Instituto Nacional de Estadísticas.

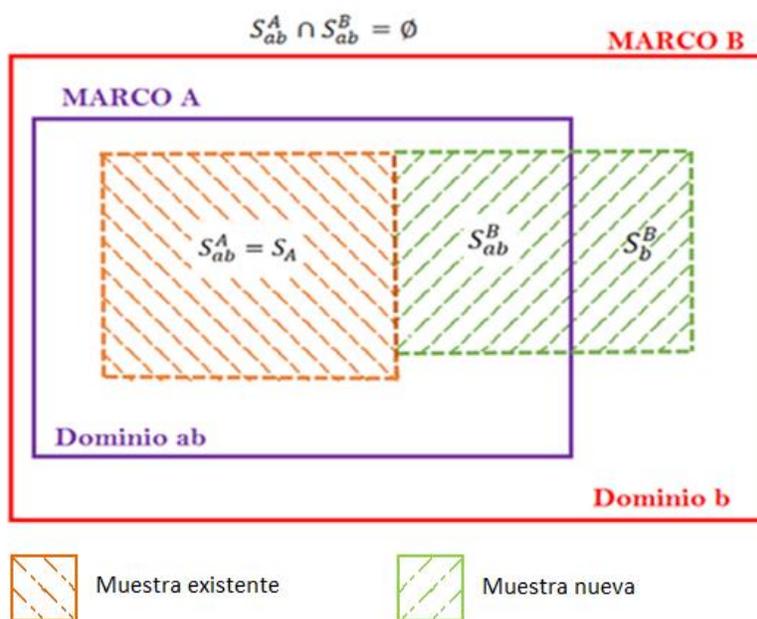
⁶⁸ Suponiendo que no haya falta de respuesta o cobertura insuficiente.

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 pérdida muestral se espera sea cubierta por el sobremuestreo⁶⁹.

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 12 se transforma al esquema de la Figura 13. **¡Error! No se encuentra el origen de la referencia.**

Figura 13: 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 .

⁶⁹ 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 muestras⁷⁰ 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 forma⁷¹:

$$w_{Lij}^{MM} = \begin{cases} w_{h_A ij}^{NR} \cdot \alpha & , \text{ si la muestra proviene del marco A (muestra anterior)} \\ w_{h_B ij}^{NR} \cdot (1 - \alpha) & , \text{ si la muestra proviene del marco B (muestra nueva)} \end{cases}$$

siendo, $\alpha = \frac{m_A}{m_A + m_B}$ y $L = \{h_A \text{ y } h_B \in \text{al dominio común } ab\}$ (38)

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).
- $w_{h_A ij}^{NR}$: Ponderador de selección de la vivienda j dentro del conglomerado i en el estrato h_A , ajustado por no respuesta de la muestra anterior.
- $w_{h_B ij}^{NR}$: 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 : 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 sobremuestreo⁷² y la nueva clasificación de códigos de disposición final entrará en vigor a partir de enero 2020.

⁷⁰ 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).

⁷¹ 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.

⁷² 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:

$$w_{h_A i} = \frac{M_{h_A}}{n_{h_A}^{sel} \cdot M_{h_A i}} \cdot \frac{M'_{h_A i}}{m_{h_A i}^{sel}} \quad (39)$$

Donde:

- M_{h_A} : Número de viviendas particulares en el estrato h_A al censo correspondiente⁷³, o al año de actualización del marco.
- $M'_{h_A i}$: Número de viviendas particulares ocupadas, según actualización de los listados de viviendas, en la unidad de primera etapa i en el estrato h_A .
- $M_{h_A i}$: 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 .
- $m_{h_A i}^{sel}$: 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.

$$Ponderador\ conglomerado = \frac{n_{h_A}^{sel}}{n_{h_A}} \quad (40)$$

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.

$$Ponderador\ vivienda = \frac{m_{h_A i}^{sel}}{m_{h_A i}} \quad (41)$$

Donde:

⁷³ 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.

$m_{h_A i}^{sel}$: Número de viviendas seleccionadas en la unidad de primera etapa i en el estrato h_A .

$m_{h_A i}$: 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:

$$w_{h_A i j} = w_{h_A i} \cdot \frac{n_{h_A}^{sel}}{n_{h_A}} \cdot \frac{m_{h_A i}^{sel}}{m_{h_A i}} = \frac{M_{h_A}}{n_{h_A}^{sel} \cdot M_{h_A i}} \cdot \frac{M'_{h_A i}}{m_{h_A i}^{sel}} \cdot \frac{n_{h_A}^{sel}}{n_{h_A}} \cdot \frac{m_{h_A i}^{sel}}{m_{h_A i}} = \frac{M_{h_A}}{n_{h_A} \cdot M_{h_A i}} \cdot \frac{M'_{h_A i}}{m_{h_A i}} \quad (42)$$

Donde,

M_{h_A} : Número de viviendas particulares en el estrato h_A al censo correspondiente, o al año de actualización del marco.

$M'_{h_A i}$: Número de viviendas particulares ocupadas, según actualización de los listados de viviendas, en la unidad de primera etapa i en el estrato h_A .

$M_{h_A i}$: 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 .

$m_{h_A i}^{sel}$: 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 .

$m_{h_A i}$: 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} ; $m_{h_A i}$) y no de las seleccionadas inicialmente ($n_{h_A}^{sel}$; $m_{h_A i}^{sel}$). Este factor lleva implícitamente los dos ajustes mencionados, análogo al ajuste de no-respuesta aplicado a los factores del Marco B (ver apartado 6). Entonces,

$$w_{h_A i j} = w_{h_A i j}^{NR} \quad (43)$$

Finalmente, el ponderador calibrado por proyecciones de población ($w_{h_A i j}^{cal}$), durante el período de transición se obtiene a partir del ponderador $w_{L i j}^{MM}$ (ponderador de selección ajustado por marcos múltiples).

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 6,6⁷⁴ 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 respectivamente⁷⁵.

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%⁷⁶ (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 mensualizada, a partir del cual sea posible incorporar este componente en la calibración de los factores de expansión de la ENE.

⁷⁴ 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.

⁷⁵ 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,

⁷⁶ 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%.

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 periodos 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-DEM⁷⁷ 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.

Para aproximar el total (*stock*) de personas extranjeras residentes en Chile, con frecuencia mensual, para el período 2012-2025⁷⁸, 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

⁷⁷ Departamento de Extranjería y Migración (DEM).

⁷⁸ 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.

los años 1992-2012 (ver Tabla 15), 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 15. 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 Chile⁷⁹.

⁷⁹ 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)*.

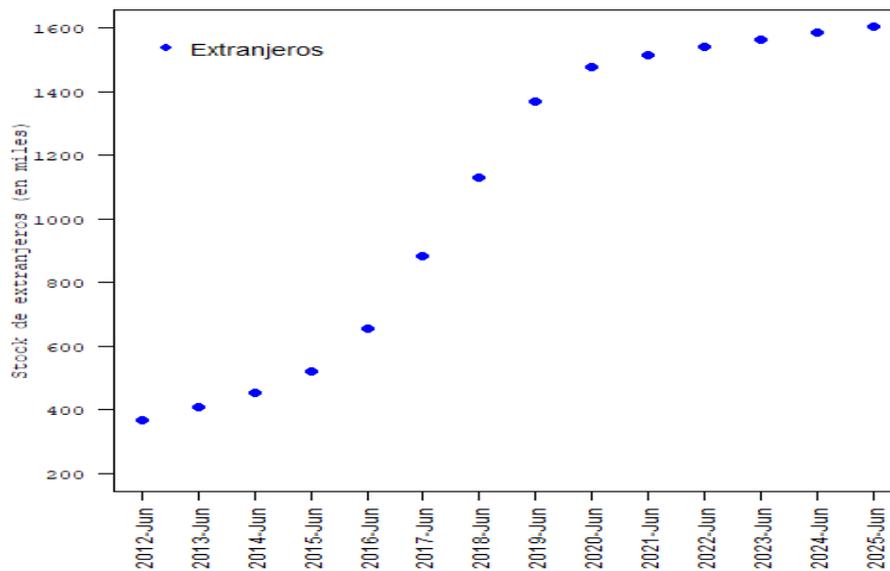
Resultados del modelo

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 el Gráfico 2 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.

Gráfico 2. Estimación anual de población extranjera a partir del saldo migratorio de las Proyecciones de población base 2017.

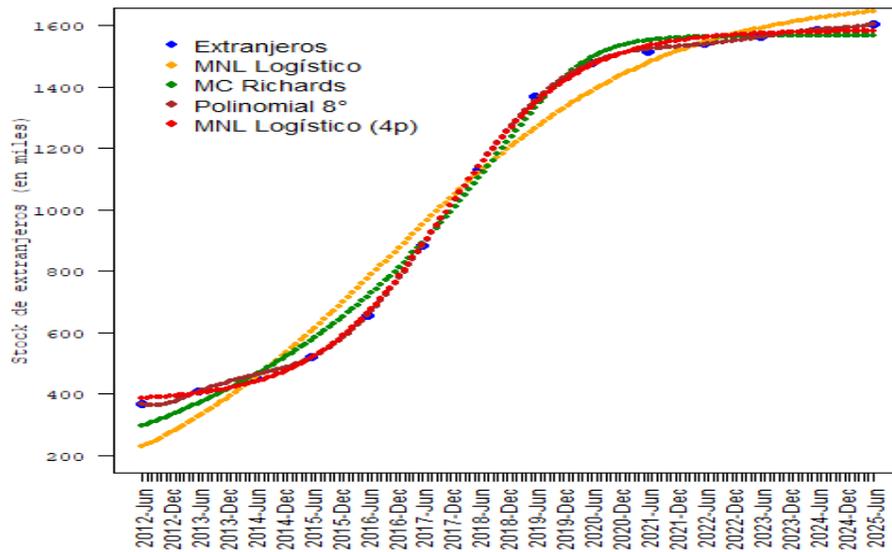


Fuente: Elaboración propia en base al componente Inmigración de las Proyecciones de Población base 2017.

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 grados⁸⁰. El Gráfico 3 recoge estos resultados.

⁸⁰ 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.

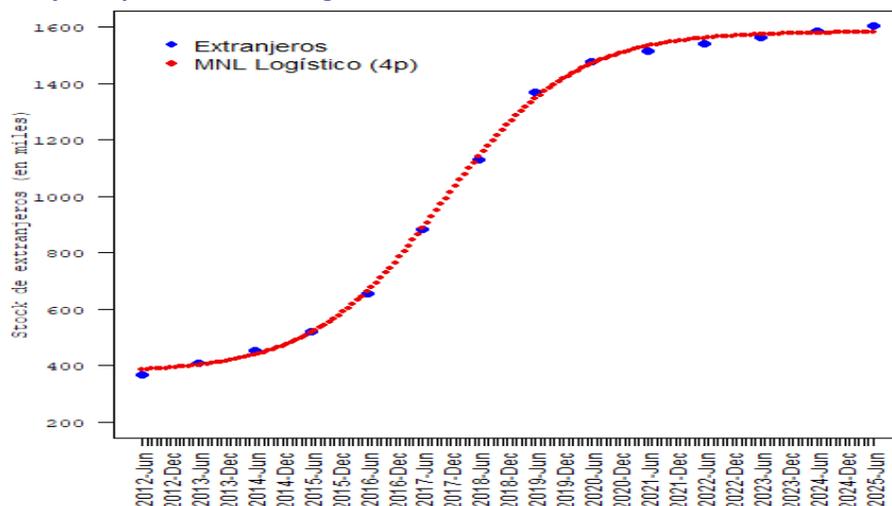
Gráfico 3. Estimación anual de población extranjera a partir del saldo migratorio de las Proyecciones de población base 2017 y la estimación de los modelos estudiados.



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. El Gráfico 4 ilustra esta serie.

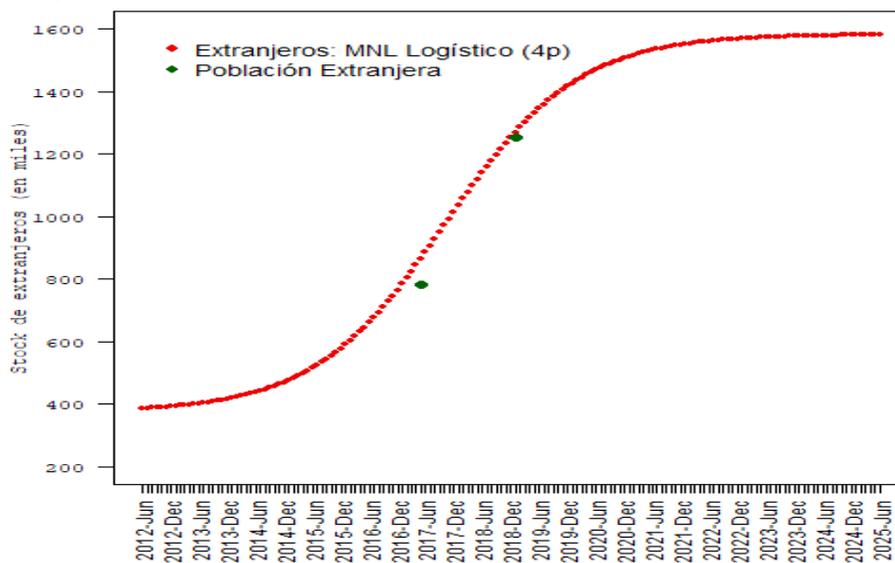
Gráfico 4. Estimación mensual de población extranjera según modelo logístico, y estimación anual de población extranjera a partir del saldo migratorio.



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.

El Gráfico 5, 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.

Gráfico 5. 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.

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.

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 detalle⁸¹:

$$f(x, A, B, xmid, scal) = A + \frac{B - A}{1 + e^{(xmid-x)/scal}} \quad (44)$$

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 14, 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.

⁸¹ 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 14. Resultados del modelo no lineal logístico de 4 parámetros

```
> summary(NLlog4_2012)

Formula: inmigracion ~ SSfpl(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

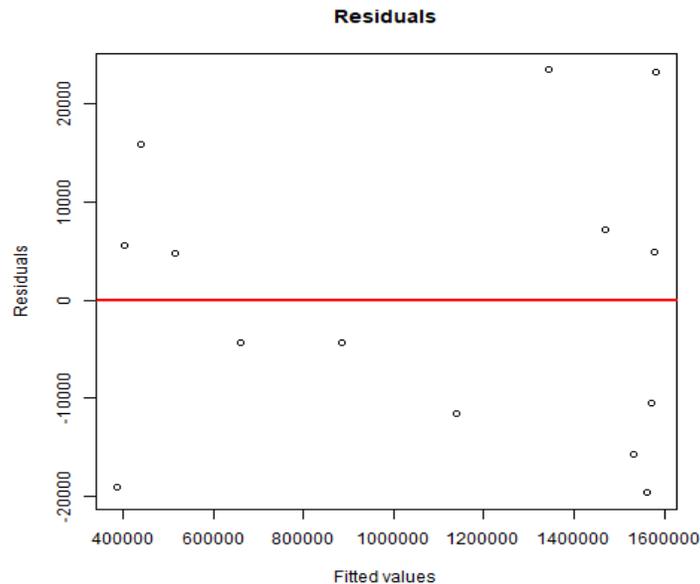
Residual standard error: 16580 on 10 degrees of freedom

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

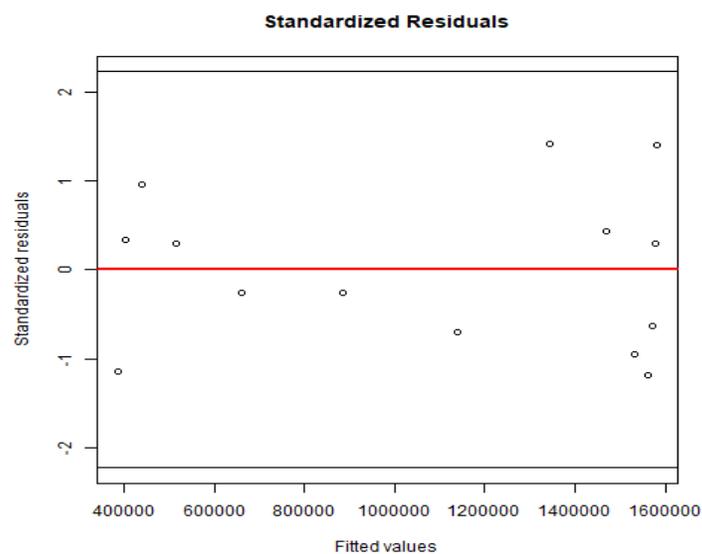
Finalmente, como señalan (Ritz & Streibig, 2008), se deben validar los supuestos detrás del modelo de regresión no lineal. Para esto, se debe verificar que: la función de media sea la correcta, homocedasticidad (con los residuos estandarizados), distribución normal de los residuos y, por último, independencia de los residuos, lo cuales se cumplen en el modelo ajustado.

Gráfico 6. Supuestos del modelo de regresión no lineal: función de media correcta



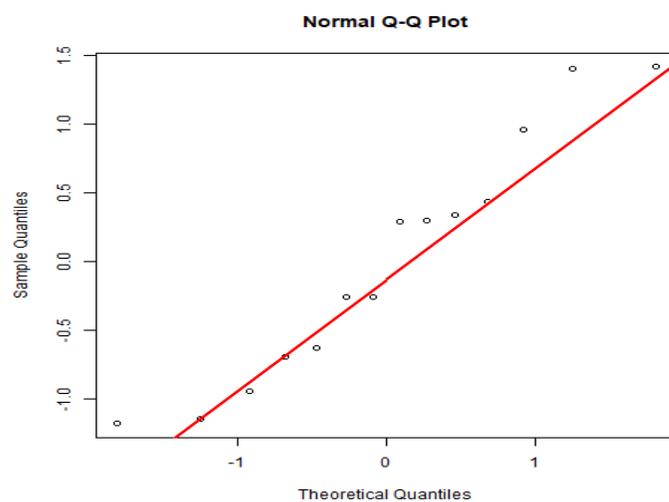
Fuente: Elaboración propia en base a los resultados del modelo no lineal ajustado.

Gráfico 7. Supuestos del modelo de regresión no lineal: homocedasticidad



Fuente: Elaboración propia en base a los resultados del modelo no lineal ajustado.

Gráfico 8. Supuestos del modelo de regresión no lineal: Normalidad



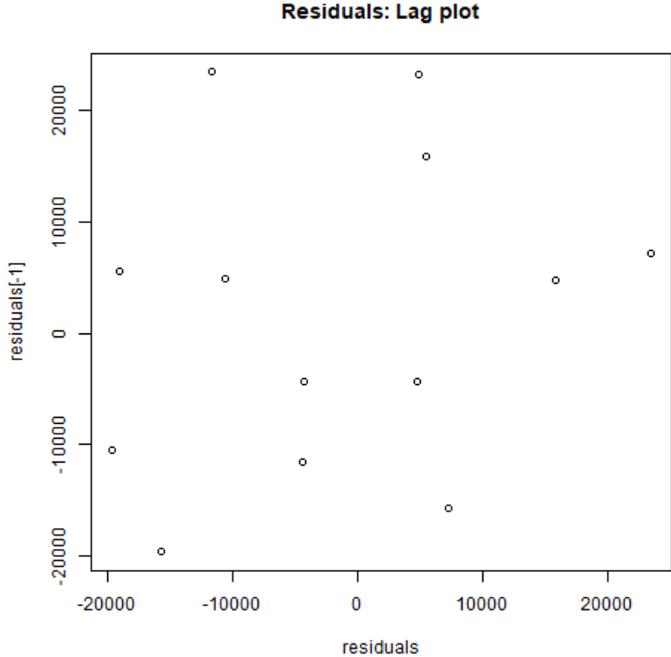
```
> shapiro.test(standardRES)
```

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.

Gráfico 9. Supuestos del modelo de regresión no lineal: Independencia



Fuente: Elaboración propia en base a los resultados del modelo no lineal ajustado.

11 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.
- 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.
- BCCh. (2001). *Indicadores Económicos y Sociales en Chile 1960 – 2000*. Santiago, Chile: Departamento Publicaciones de la Gerencia de Investigación Económica. División Estudios. Banco Central de Chile.
- 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 of 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*. Santiago, Chile.: Division de Estadística y Proyecciones Económicas. CEPAL.
- 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.
- Department 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.
- 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.iesta.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. (Marzo de 2019a). *Ciudades, pueblos, aldeas y caseríos 2019*. Obtenido de https://geoarchivos.ine.cl/File/pub/Cd_Pb_Al_Cs_2019.pdf
- 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-expansi%C3%B3n-de-la-encuesta-nacional-de-empleo>
- 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-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>
- Kish, L. (1963). Changing strata and selection probabilities. *Proceedings of the Social Statistics Section, American Statistical Association*, 124-131.

- Konijin, 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 activay 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.
- OECD. (11 de 2015). *Recomendación del Consejo de la OCDE sobre Buenas Prácticas Estadísticas*. Obtenido de <http://www.oecd.org/statistics/good-practice-toolkit/OECD-LEGAL-0417-spa.pdf>
- OIT. (Enero de 1993). *Clasificación Internacional de la Situación en el Empleo (CISE)*. Obtenido de https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087564.pdf
- OIT. (2013). *La Medición de la Informalidad: Manual Estadístico sobre el Sector Informal y el Empleo Informal*. Ginebra: Organización Internacional del Trabajo.
- 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.