



CASEN 2022

Encuesta de Caracterización Socioeconómica Nacional

Metodología de Diseño Muestral

Septiembre 2022

Informe editado por el Ministerio de Desarrollo Social y Familia, a partir del Informe Diseño Muestral Casen 2022 entregado por el Instituto Nacional de Estadísticas (INE), en el marco del convenio "Diseño Muestral de la Encuesta de Caracterización Socioeconómica Nacional Casen 2022".

ÍNDICE

I. PRESENTACIÓN	7
<i>I.1. Antecedentes y características generales del diseño de Casen 2022</i>	9
I.1.1. Objetivo	9
I.1.2. Población objetivo.....	10
I.1.3. Unidad de información.....	10
I.1.4. Marco muestral	10
I.1.5. Dominios de estudio.....	11
I.1.6. Objetivos de precisión y tamaño de muestra	12
I.1.7. Estrategia muestral	12
I.1.8. Método de selección de unidades primarias de muestreo (UPM)	12
I.1.9. Método de selección de unidades secundarias de muestreo (viviendas)	13
II. SIMULACIÓN DEL TAMAÑO MUESTRAL	14
<i>II.1. Objetivo general de las simulaciones</i>	14
<i>II.2. Objetivos específicos</i>	15
<i>II.3. Metodología de cálculo del tamaño muestral utilizando muestras complejas</i>	15
<i>II.4. Resumen del proceso de simulación</i>	18
II.4.1. Escenario N°0 (Escenario Base)	23
II.4.2. Escenario N°1	23
II.4.3. Escenarios N°2.1 y N°2.2	24
II.4.4. Escenario N°3	26
II.4.5. Escenario N°4	28
II.4.6. Escenario N°5	29
II.4.7. Escenario N°6	31
II.4.8. Comparación tamaño objetivo entre Escenario N°5 (escenario elegido) y Escenario N°6 (modalidad telefónica)	32
<i>II.5. Tamaño muestral propuesto</i>	34
II.5.1. Comparación de tamaños y distribuciones muestrales	36
II.5.2. Distribución de las unidades primarias de muestreo.....	40
III. DISEÑO MUESTRAL	43
<i>III.1. Características del marco muestral</i>	43
<i>III.2. Cobertura geográfica del marco muestral</i>	45
<i>III.3. Áreas Especiales</i>	46
<i>III.4. Estratificación Socioeconómica del Marco Muestral MMV 2020</i>	47
<i>III.5. Estratificación Socioeconómica del Marco DE selección para Casen 2022</i>	48
IV. TAMAÑO MUESTRAL EN MUESTRAS COMPLEJAS	50
<i>IV.1. Comparación del tamaño muestral entre Casen 2022 y Casen en Pandemia 2020</i>	51
<i>IV.2. Estimación del tamaño muestral para Casen 2022</i>	53
<i>IV.3. Selección de Conglomerados (o unidades primarias de muestreo)</i>	62
<i>IV.4. Selección de viviendas de la muestra</i>	65

V.	VERIFICACIÓN Y ACTUALIZACIÓN DE UNIDADES PRIMARIAS DE MUESTREO	68
	V.1. <i>Diseño del proceso de enumeración</i>.....	68
	V.1.1. Método de Priorización	68
	V.1.2. Proceso de Enumeración	71
	V.1.3. Análisis de Gabinete	73
	V.2. <i>Resultados del proceso de actualización del directorio de casen 2022</i>.....	75
	V.2.1. Enumeración en terreno	75
	V.2.2. Validación y Mejora de Directorios de Gabinete	76
	V.2.3. Elaboración del Directorio y Productos cartográficos.....	78
	V.2.3.1. Directorio de viviendas seleccionadas	78
	V.2.3.2. Productos Cartográficos.....	80
VI.	DISEÑO FACTORES DE EXPANSIÓN.....	83
	VI.1. <i>Introducción</i>	83
	VI.2. <i>Ponderador de selección de conglomerados</i>	85
	VI.2.1. Probabilidad de selección de los conglomerados	86
	VI.2.2. Ponderador de selección de conglomerados como inverso de la probabilidad de selección	87
	VI.3. <i>Ponderadores de selección de viviendas</i>	89
	VI.3.1. Probabilidad condicional de selección de viviendas	89
	VI.3.2. Ajuste por omisión de viviendas.....	89
	VI.3.3. Ponderador de selección de viviendas como inverso de la probabilidad de selección 90	
	VI.4. <i>Ajuste por elegibilidad</i>	91
	VI.4.1. Ajuste por elegibilidad desconocida.....	92
	VI.4.2. Ajuste por no elegibilidad.....	93
	VI.5. <i>Ajuste por no respuesta</i>	94
	VI.6. <i>Ponderador de calibración</i>	96
	VI.6.1. Ponderador de Calibración Regional	98
	VI.6.2. Ponderador de Calibración Provincial	100
	VI.6.3. Ponderador de Calibración Comunal.....	102
	VI.7. <i>Redondeo Probabilístico</i>	104
VII.	FACTOR DE EXPANSIÓN ORIENTACIÓN SEXUAL E IDENTIDAD DE GÉNERO.....	105
	VII.1. <i>Diseño del factor de expansión</i>.....	106
	VII.2. <i>Variables incluidas en el modelo</i>	106
	VII.3. <i>Creación de veintiles para el ajuste</i>	108
	VII.4. <i>Ponderador de calibración</i>	109
VIII.	ESTIMACIÓN DE VARIANZA COMPLEJA	111
	VIII.1. <i>Métodos convencionales para la estimación de la varianza</i>.....	111
	VIII.1.1. Métodos exactos	112
	VIII.1.2. Método del conglomerado último	112
	VIII.1.3. Métodos de linealización.....	112
	VIII.1.4. Método de replicación	113
	VIII.2. <i>Varianza en muestras complejas</i>	113
	VIII.3. <i>Algoritmo de cálculo en muestras complejas</i>.....	115

VIII.3.1. Algoritmo de cálculo de los estimadores	116
VIII.3.2. Algoritmo de cálculo de las varianzas en muestras complejas	118
VIII.4. Variables que identifican el diseño muestral complejo	120
VIII.4.1. Creación de pseudo-estratos (varstrat).....	120
VIII.4.2. Creación de pseudo-conglomerados (varunit)	121
VIII.5. Resultados de la situación de pobreza por ingresos.....	124
IX. ANEXOS	129
Anexo N°1. Escenarios trabajados.....	129
Anexo N°2. Tamaños muestrales a nivel nacional, regional y comunal por área (Urbano – Rural) para Casen 2022	146
Anexo N°3. Fuentes de información y procesos de verificación y actualización de unidades primarias de muestreo	155
Anexo N°3.1. Fuentes de Información	155
Anexo N°3.2. Descripción de procesos de enumeración	168
Anexo N°3.3. Descripción de validación y mejora de Directorios de Gabinete	175
Anexo N°4. Stock de viviendas según factor de selección	178
Anexo N°5. Covariables consideradas en el modelo propensity score	179
Anexo N°6. Stock de viviendas según factor de no respuesta	180
Anexo N°7. Stock poblacional calibrado a proyecciones de población en viviendas particulares ocupadas región - área	181
Anexo N°8: Stock poblacional calibrado a las Proyecciones de Población de 18 años o más por región-área	182
X. BIBLIOGRAFÍA	183

ÍNDICE DE TABLAS

Tabla II.1. Resumen de parámetros utilizados y tamaños y errores resultantes según escenario, Casen 2022.....	21
Tabla II.2. Tamaños muestrales y errores asociados a nivel nacional y regional, Casen 2022	30
Tabla II.3. Tamaños muestrales y errores asociados a nivel nacional y regional. Casen 2022 en modalidad de levantamiento mixto	32
Tabla II.4. Comparación de tamaños objetivo y errores asociados Escenarios N°5 y N°6	33
Tabla II.5 Comparación distribución de tamaños objetivo regionales entre ambos escenarios	33
Tabla II.6. Tamaños y errores a nivel nacional, regional, urbano-rural, Casen 2022	35
Tabla II.7. Tamaños y errores Casen en Pandemia 2020 y propuesta Casen 2022. Nacional Urbano - Rural	35
Tabla II.8. Comparación del número de viviendas entre MMV 2020 – Casen 2022 – Casen en Pandemia 2020 a nivel nacional y regional.	37

Tabla II.9. Comparación distribución porcentual de viviendas a nivel nacional y regional. MMV 2020 – Casen 2022 – Casen en Pandemia 2020.....	39
Tabla II.10. Distribución de unidades primarias de muestreo: MMV 2020 – Casen 2022	40
Tabla II.11. Distribución de tamaños muestrales a nivel nacional, regional y provincial, por área: Casen en Pandemia 2020 – MMV 2020 – Casen 2022	41
Tabla III.1. Distribución de UPM excluidas del marco de selección de Casen 2022.....	46
Tabla IV.1. Tamaños muestrales de viviendas para Casen 2022, según parámetros de Casen en Pandemia 2020 y efecto del MMV 2020 a nivel nacional y regional	59
Tabla IV.2. Comparación tamaños objetivos y distribución porcentual de Casen desde 2011 hasta 2022 a nivel nacional y regional.....	61
Tabla IV.3. Número de viviendas en el MMV 2020 y número de viviendas en el Marco de selección para Casen 2022 a nivel regional.....	66
Tabla V.1. Distribución del número de UPM priorizadas según área urbano y rural, Casen 2022	71
Tabla V.2. Total de UPM por región, según origen de actualización de directorios	75
Tabla V.3. Total de UPM enumeradas en terreno a nivel de región.....	76
Tabla V.4. Distribución regional de directorios de UPM validadas en Gabinete según área urbano/rural. Porcentajes según Total Nacional.	77
Tabla V.5. Total de viviendas del directorio Casen a nivel de región.....	79
Tabla VI.1. Estadísticos de los ponderadores de selección de UPM, Casen 2022.....	88
Tabla VI.2. Distribución viviendas según elegibilidad	93
Tabla VI.3. Veintiles Casen 2022	96
Tabla VI.4. Proyecciones de población al 30 de noviembre de 2022 según región y área	97
Tabla VII.1. Total de personas logradas en el levantamiento	106
Tabla VII.2. Total de personas de 18 años o más según la variable dependiente r_una	107
Tabla VII.3. Variables independientes seleccionadas en el modelo para el FE OSIG.....	107
Tabla VII.4. Valores mínimo, máximo y promedio de la probabilidad de respuesta estimada por veintil	108
Tabla VIII.1. Total de estratos, conglomerados, viviendas, varstrat y varunit, según región por área geográfica, para la muestra lograda en Casen 2022	123
Tabla VIII.2. Frecuencia de varunit según número de viviendas que lo componen.....	124
Tabla VIII.3. Estimación del porcentaje de la población en situación de pobreza por ingresos, según zona y región.....	125
Tabla VIII.4. Estimación del porcentaje de hogares en situación de pobreza por ingresos, según zona y región	127
Tabla IX.1. Escenario N°0 (Escenario Base)	129
Tabla IX.2. Escenario N°1.....	130
Tabla IX.3. Escenario N°2.1.....	131
Tabla IX.4. Escenario N°2.2.....	132

Tabla IX.5. Escenario N°3.....	133
Tabla IX.6. Escenario N°4.....	134
Tabla IX.7. Escenario N°5 (Escenario elegido).....	135
Tabla IX.8. Distribución de tamaños muestrales a nivel nacional, regional, comunal y por área. Escenario N°6 levantamiento mixto. Casen 2022.....	136
Tabla IX.9. Versión Cartográfica ACON, según año a nivel de región	158
Tabla IX.10. Dotación equipo operativo para la enumeración a nivel de región.....	170
Tabla IX.11. Distribución regional de UPM priorizadas según área urbano/rural. Porcentajes según total nacional.	172
Tabla IX.12. UPM de otras muestras por año.....	177

ÍNDICE DE CUADROS

Cuadro II.1. Parámetros comunes utilizados en escenarios Casen 2022.....	20
Cuadro IV.1. Comparación de criterios utilizados para el cálculo del tamaño muestral Casen 2022 – Casen en Pandemia 2020.....	52
Cuadro IV.2. Método de selección de unidades primarias de muestreo, Casen 2022	64
Cuadro V.1. Descripción de campos contenidos en el directorio Casen.....	79
Cuadro V.2. Descripción de campos del Mapa Móvil Regional.....	81
Cuadro IX.1. Variables de la metodología de priorización cartográfica en áreas urbanas	164
Cuadro IX.2. Variables de la metodología de priorización cartográfica de áreas rurales	165

I. PRESENTACIÓN¹

Las encuestas de hogares cumplen un rol fundamental en la toma de decisiones para la formulación de políticas sociales, de ahí el creciente interés por obtener estimaciones precisas y confiables que puedan representar de manera adecuada la realidad de la población de interés. Sin embargo, para lograr esta propiedad deseable en las estimaciones, es necesario tener una muestra que responda a las exigencias de los errores muestrales propuestos. En ese sentido, la muestra debe ser construida para los objetivos propios de la encuesta, requerimientos que se deben considerar al momento de determinar un tamaño muestral adecuado. Además, el tamaño muestral debe balancear en su diseño el presupuesto disponible y los errores de muestreo considerados como admisibles para fines del estudio. La conjugación de estos tres elementos (requerimientos, errores muestrales admisibles y presupuesto disponible) permite obtener tamaños muestrales óptimos².

Al igual que en sus versiones anteriores, la Encuesta de Caracterización Socioeconómica Nacional (en adelante, Casen 2022) fue planificada por el Ministerio de Desarrollo Social y Familia (en adelante, el Ministerio) quien, en enero de 2022, mediante un convenio de transferencia, encargó la elaboración del diseño muestral al Instituto Nacional de Estadísticas (en adelante, INE).

Mediante el convenio de colaboración interinstitucional entre el Ministerio y el INE denominado: **“Diseño Muestral Encuesta de Caracterización Socioeconómica Nacional Casen 2022”**, se busca dar continuidad a una nueva versión de la Encuesta Casen, para el año 2022, que permita obtener información relevante y actualizada para la implementación de políticas públicas sobre la materia y para uso de la comunidad nacional en general.

Este proceso comienza con simulaciones de tamaños muestrales, para determinar el número óptimo de unidades a encuestar de acuerdo con los requerimientos y objetivos de la encuesta para, posteriormente, seleccionar las unidades muestrales, elaborar y preparar el material cartográfico para realizar el levantamiento de la información y, una vez realizado el levantamiento, calcular los factores expansión, estimaciones y errores muestrales para el parámetro de interés.

La metodología del diseño muestral de la Encuesta, describe el conjunto de opciones metodológicas y de procedimientos estadísticos adoptados a lo largo del proceso de planificación, que son considerados los más apropiados para seleccionar una muestra probabilística, estadísticamente representativa de las

¹ Este documento fue elaborado por el Instituto Nacional de Estadísticas (INE) y editado por el Ministerio de Desarrollo Social y Familia (MDSF).

² Hablar de tamaños muestrales óptimos no denota, en este contexto un proceso matemático de maximización/minimización de una función objetivo, sino el proceso de las múltiples iteraciones para las distintas simulaciones, distintos parámetros y requerimientos. Así como también los acuerdos llegados entre el INE y la contraparte del Ministerio en las reuniones de trabajo sostenidas.

viviendas particulares de Chile y que permita estimar, con el nivel de precisión deseado, la tasa de pobreza en cada región del territorio nacional, así como también, en las áreas geográficas urbana y rural a nivel nacional y para el país en su conjunto.

Este informe es parte de los productos entregables acordados entre ambas instituciones por medio del convenio suscrito, y se organiza en las siguientes secciones:

- “Simulación del tamaño muestral”, que incluye una descripción y la generación de tablas con distintos escenarios que darán origen al tamaño muestral elegido, considerando para esto los parámetros obtenidos desde Casen 2017 (modalidad presencial) y Casen en Pandemia 2020 (modalidad mixta secuencial) y aplicando los ajustes necesarios producto del cambio de marco muestral de viviendas del INE, considerando variables de interés, cobertura geográfica, efectos de diseño y niveles de precisión establecidos.
- “Diseño Muestral”, que incluye la descripción de las propiedades del marco muestral, la definición de los estratos de muestreo de Casen 2022, un resumen del proceso de actualización de direcciones realizado para la correspondiente selección de viviendas y la metodología utilizada para la selección de las unidades muestrales (primarias o UPM y secundarias o viviendas).
- “Tamaño muestral en muestras complejas”, que incluye una comparación entre la metodología de cálculo del tamaño muestral de Casen 2022 y Casen en Pandemia 2020, así como también el desarrollo de cada uno de los pasos a seguir para la obtención del tamaño, considerando los resultados obtenidos desde Casen 2017 y de Casen en Pandemia 2020, y utilizando estrategias para plasmar en el tamaño de Casen 2022 el efecto del cambio de marco muestral, para luego obtener los errores de muestreo según los estándares de calidad utilizados por el INE. Adicionalmente, incluye la descripción de la selección de conglomerados y de viviendas, junto con una primera aproximación al diseño de enumeración/verificación de las Unidades Primarias de Muestreo.
- “Verificación y actualización de las Unidades Primarias de Muestreo”, que incluye una descripción de la actualización cartográfica del marco muestral y los distintos procedimientos utilizados en dicha labor, así como también, los insumos con los que se cuenta para cada tipo de actualización, el método de priorización de las unidades que se actualizarán en terreno y los resultados del proceso.
- “Diseño Factores de Expansión”, que describe paso a paso la metodología del cálculo de factores de expansión a nivel regional, provincial y comunal, con el objetivo de que la muestra represente al total de población en viviendas particulares ocupadas en Chile.

- “Factor de expansión orientación sexual e identidad de género”, que explica la metodología para el cálculo de factor que se aplica a personas de 18 años o más que respondieron al menos una pregunta del módulo (autoaplicado) de orientación sexual e identidad de género.
- “Estimación de la varianza compleja”, que explica los métodos para la estimación de varianza en diseños muestrales complejos, considerando la creación de pseudo-estratos y pseudo-conglomerados para una correcta estimación de dicho estadígrafo. Además, se presenta la estimación de la tasa de pobreza por ingresos, junto con sus errores muestrales, para los dominios de estudio definidos.

Previamente, se incluye un apartado introductorio con los antecedentes y características generales del Diseño Muestral de Casen 2022, que contiene una descripción de los objetivos de la encuesta, la población objetivo, el marco muestral utilizado, los niveles de estimación y la estrategia muestral, entre otros.

I.1. ANTECEDENTES Y CARACTERÍSTICAS GENERALES DEL DISEÑO DE CASEN 2022

I.1.1. Objetivo

La Encuesta Casen es el principal instrumento para realizar el seguimiento de la política pública en materia de focalización del gasto social y, para la generación de diversos análisis de la realidad social del país. Al igual que sus versiones anteriores, Casen 2022 ha sido pensada como un instrumento de diagnóstico y evaluación que permita:

1. Conocer periódicamente la situación de los hogares y de la población, especialmente de aquella en situación de pobreza y de aquellos grupos definidos como prioritarios por la política social, con relación a aspectos demográficos, de educación, salud, vivienda y entorno, trabajo e ingresos.
2. Estimar la magnitud de la pobreza por ingresos y multidimensional y la distribución del ingreso.
3. Identificar carencias y demandas de la población en las áreas señaladas.
4. Evaluar las distintas brechas que separan a los diferentes segmentos sociales y ámbitos territoriales.
5. Evaluar el impacto de la política social, en relación con la cobertura, focalización y distribución del gasto fiscal de los principales programas sociales de alcance nacional entre los hogares, según su nivel de ingreso; así como también, el impacto de este gasto en el ingreso de los hogares y en la distribución de este.

I.1.2. Población objetivo

La población objetivo de la encuesta es la población que reside en viviendas particulares ocupadas a lo largo del territorio nacional, con exclusión de algunas áreas³ previamente definidas por el INE⁴. Estas áreas corresponden a comunas completas o parte de comunas cuyas Unidades Primarias de Muestreo (UPM) fueron clasificadas como áreas especiales en el marco muestral de viviendas MMV 2020 y que, en el marco muestral utilizado para Casen en Pandemia 2020 y anteriores, fueron identificadas como áreas de difícil acceso (ADA).

Al interior de cada vivienda particular seleccionada, el objetivo es entrevistar a todos los hogares y a todas las personas que se declaran residentes habituales.

I.1.3. Unidad de información

La unidad de información de Casen 2022 es el hogar y las personas que lo conforman. Por su parte, el informante idóneo es algún integrante del hogar de 18 años o más.

I.1.4. Marco muestral⁵

El marco muestral de viviendas MMV 2020 está constituido por unidades primarias de muestreo (UPM) que corresponden a áreas geográficas homogéneas, en términos del número de viviendas particulares ocupadas que las conforman, excluyendo viviendas de temporada y colectivas.

Las UPM presentan una estratificación geográfica, dada por la división político-administrativa del país, según la conformación región-provincia-comuna y desagregada, además según la división censal, clasificándolas en áreas urbanas y rurales. La definición de urbano y rural viene dada por las especificaciones del Censo de Población y Vivienda 2017 (en adelante, Censo 2017). A esta estratificación

³ Excluye en su totalidad las comunas de: Ollagüe, Juan Fernández, Isla de Pascua, Cochamó, Chaitén, Futaleufú, Hualaihué, Palena, Guaitecas, y O'Higgins. Además, excluye algunas UPM definidas como especiales en las comunas de: General Lagos, Colchane, Lago Verde, Tortel, Laguna Blanca, Río Verde, San Gregorio, Cabo de Hornos, Primavera, Timaukel y Torres del Paine.

⁴ Estas áreas especiales fueron definidas previamente por el INE para no ser consideradas en el levantamiento del nuevo diseño de la Encuesta Nacional de Empleo (ENE). Además, la exclusión de dichas áreas, fue presentada y aprobada por el Panel de expertos y expertas de Casen 2022.

⁵ A diferencia de las versiones previas de Casen, donde se utilizó el Marco muestral con base en Censo 2002, en Casen 2022 se utiliza el Marco Muestral de Viviendas (MMV 2020) con base en Censo 2017. Dado lo anterior, y debido a las posibles implicancias de la actualización del marco muestral en los principales indicadores de la encuesta Casen (Pobreza por ingresos y pobreza multidimensional), el Ministerio, en conjunto con el Panel de Expertas y Expertos, definieron evaluar de forma previa a la recolección de Casen 2022, el efecto del cambio de marco muestral por medio de una prueba piloto (realizada por el INE). El objetivo del estudio fue estimar el efecto de utilizar el MMV 2020 sobre los principales indicadores de la encuesta Casen. Para mayor detalle revisar: [Nota técnica N°1: Principales cambios metodológicos Casen 2022](#)

se suma la clasificación socioeconómica de las UPM (con niveles bajo, medio y alto) construida a partir de un modelo estadístico cuyas variables son obtenidas desde los resultados del Censo 2017⁶.

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

A partir del MMV 2020 se elabora el marco de selección para Casen 2022, conformado por 335 comunas que fueron definidas para levantar en el nuevo diseño de la Encuesta Nacional de Empleo (ENE 2020)⁷. Las UPM del marco muestral de Casen 2022 que se encuentran estratificadas geográficamente (comunas) y por áreas (urbana-rural) presentan, además una estratificación por nivel socioeconómico (NSE) construida, a partir de la clasificación socioeconómica del MMV 2020 pero, adecuada especialmente para esta encuesta.

A través de las UPM se accede a las viviendas y hogares, cuyos integrantes son principalmente, el objeto final de estudio.

Para la selección de viviendas o unidades secundarias de muestreo (USM) se considera la actualización de las UPM seleccionadas en primera etapa, subdividiéndose el trabajo de acuerdo con los criterios de priorización establecidos en el convenio. En el 60% de las UPM la actualización corresponde a la verificación de direcciones en gabinete, a partir de los listados obtenidos desde Precenso 2016, imágenes satelitales, Certificados de Recepción Final de conjuntos habitacionales, servidores de mapas, entre otros. En el restante 40% la actualización corresponde a la enumeración en terreno donde, a partir de los listados de Precenso se recorren las UPM verificando y/o añadiendo direcciones correspondientes a nuevas edificaciones.

I.1.5. Dominios de estudio

Los niveles de estimación o dominios de estudio para los cuales la muestra fue diseñada con el objetivo de obtener estimaciones representativas son: nacional, nacional urbano, nacional rural y regional, los cuales coinciden con los dominios de estudio definidos en las versiones de Casen 2011, 2013, 2017 y Casen en Pandemia 2020. Sin embargo, la muestra tiene presencia en 335 comunas, tanto en áreas urbanas como rurales, donde aplique.

⁶ Documentos de trabajo: Estratificación socioeconómica del marco muestral de viviendas 2017 MMV 2017. Disponible en: [Estratificación socioeconómica del MMV 2017](#).

⁷ Entre las 335 comunas a levantar en Casen 2022 se consideran las comunas de: General Lagos, Colchane, Lago Verde, Tortel, Laguna Blanca, Río Verde, San Gregorio, Cabo de Hornos, Primavera, Timaukel y Torres del Paine que, hasta Casen en Pandemia 2020, no habían sido incluidas en la muestra.

I.1.6. Objetivos de precisión y tamaño de muestra

Considerando un levantamiento bajo modalidad presencial, con un tamaño objetivo ajustado de 71.028 viviendas, se espera obtener estimaciones del parámetro de interés con un error absoluto nacional de 0,4% y, un error relativo, de 3,3%⁸.

Con este tamaño los errores absolutos regionales no superan 2,1% y; los errores relativos, 30,5%.

Dados los tamaños muestrales objetivos a nivel nacional y regional, la distribución de los tamaños a nivel de estratos geográficos (comuna-área) se estableció de manera tal que, en cada combinación comuna-urbana se consideran, al menos, 32 viviendas y; en cada combinación comuna-rural se incluyen 40 viviendas. A excepción de estratos geográficos que en el marco muestral presentan entre dos y siete UPM, para los cuales, en cada combinación de comuna-urbana, se considera un tamaño de 16 viviendas y; en cada combinación de comuna-rural, un tamaño de 20 viviendas. En el caso de estratos unitarios, se seleccionan 8 viviendas en el área urbana y 10 viviendas en el área rural (lo que corresponde a seleccionar una UPM).

I.1.7. Estrategia muestral

El diseño muestral de Casen 2022 corresponde a la obtención de una muestra probabilística, estratificada y bietápica. Los estratos geográficos corresponden a la agregación comuna-área. Adicionalmente, cada UPM del MMV 2020 presenta una estratificación socioeconómica que permite realizar una estratificación por nivel socioeconómico (NSE) especialmente adecuado para el marco de selección de Casen 2022⁹ y, por tanto, los estratos de muestreo quedan conformados por la agregación comuna-área-NSE. En cada uno de estos estratos la muestra de UPM es seleccionada de forma independiente.

I.1.8. Método de selección de unidades primarias de muestreo (UPM)

Para Casen 2022 la selección de UPM se realiza al interior de cada estrato de muestreo, en forma sistemática con probabilidad proporcional al tamaño, según el número de viviendas que las conforman.

⁸ Ver ¡Error! No se encuentra el origen de la referencia. Tamaños muestrales y errores asociados a nivel nacional y regional, Casen 2022.

⁹ Para mayor detalle sobre la estratificación socioeconómica del Marco de selección para Casen 2022, revisar apartado III.5.

I.1.9. Método de selección de unidades secundarias de muestreo (viviendas)

Las viviendas al interior de cada UPM son seleccionadas en forma sistemática y con igual probabilidad, previa actualización ya sea, desde terreno (enumeración) o desde gabinete (verificación)¹⁰ según los criterios de priorización y presupuestarios establecidos.

¹⁰ Para mayor detalle sobre el proceso de enumeración, revisar apartado V.

II. SIMULACIÓN DEL TAMAÑO MUESTRAL

Entre noviembre de 2021 y enero de 2022, el INE tuvo la misión de desarrollar la estrategia de muestreo de la Encuesta Casen 2022 que consistió en proponer y elaborar, de acuerdo con los requerimientos planteados por el Ministerio, simulaciones de tamaños muestrales con los términos específicos de errores muestrales aceptables y distribución muestral, optimizando la selección de las unidades muestrales, de acuerdo con criterios de factibilidad y de presupuesto.

Durante esta etapa, previa al desarrollo de la estrategia de trabajo de campo, se estableció como hito inicial del convenio INE-Ministerio obtener un tamaño muestral óptimo para la estimación del principal parámetro de interés, con márgenes de precisión que permitan que Casen 2022 tenga representatividad a nivel nacional, nacional urbano, nacional rural, y regional. Para ello, se simularon distintos escenarios o alternativas de tamaño muestral, basados en la unificación de criterios estadísticos solicitados y discutidos con el Ministerio.

Estas simulaciones se realizaron considerando los parámetros de los resultados del trabajo de campo de Casen 2017 y Casen en Pandemia 2020 y, en la medida de que el nuevo diseño lo permita, la distribución de la muestra objetivo de esta última, resultando en la formulación de diversos escenarios, que fueron presentados en reuniones de trabajo conformadas por los equipos de la División de Observatorio Social del Ministerio y el Subdepartamento de Diseño de Marcos y Muestras del INE. Finalmente, se estableció un escenario que se ajusta a los requerimientos establecidos por el Ministerio.

En este capítulo se resumen los distintos escenarios y estrategias de muestreo que se han simulado para Casen 2022 con el fin de obtener tamaños muestrales en los niveles de estimación definidos por el Ministerio.

II.1. OBJETIVO GENERAL DE LAS SIMULACIONES

Determinar un tamaño muestral que reúna características suficientes para representar de manera adecuada a la población chilena respecto de sus componentes sociodemográficos y, además, contenga márgenes de error tolerables¹¹ a nivel nacional, nacional urbano, nacional rural, y regional, considerando distintas estrategias de muestreo, distribución de la muestra objetivo y de la muestra con sobremuestreo.

¹¹ Los márgenes de error tolerables se definen en función de cada una de las simulaciones.

II.2. OBJETIVOS ESPECÍFICOS

Calcular el tamaño muestral considerando los siguientes elementos:

- Niveles de estimación:
 - Nacional
 - Nacional urbano y nacional rural
 - Regional
- Medida poblacional o parámetro a estimar: Tasa de pobreza (porcentaje de personas en situación de pobreza por ingresos).

- Variable de diseño: Caracterización del individuo según condición de pobreza, tal que,

$$X_j = \begin{cases} 0, & \text{si el } j - \text{ésimo individuo no se encuentra en situación de pobreza por ingresos} \\ 1, & \text{si el } j - \text{ésimo individuo se encuentra en situación de pobreza por ingresos} \end{cases}$$

- Errores de muestreo tolerables para el estudio.
- Parámetros obtenidos del trabajo de campo de Casen en Pandemia 2020.
- Distribución de la población en viviendas particulares ocupadas según el marco muestral de viviendas MMV 2020.
- Tasas de no logro¹² sobre viviendas gestionadas a nivel de comuna-área, observadas en la aplicación de la encuesta Casen en los años 2017 y 2020.
- Población objetivo de la encuesta: población residente habitual en viviendas particulares ocupadas a lo largo del territorio nacional, con exclusión de áreas especiales¹³.

II.3. METODOLOGÍA DE CÁLCULO DEL TAMAÑO MUESTRAL UTILIZANDO MUESTRAS COMPLEJAS

La población objetivo, los niveles de estimación deseados y la cobertura geográfica de la encuesta Casen 2022, hacen necesaria la confección de un diseño eficiente para la selección de una muestra de viviendas, en términos de ejecución presupuestaria e implementación operativa.

¹² La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

¹³ Las áreas especiales corresponden a unidades primarias de muestreo (UPM) que, por ciertos motivos, han quedado sin posibilidad de ser seleccionadas en la muestra. Algunos de estos motivos son: climáticos, de transporte, de costo de acceso, de acceso restringido, etc.

El muestreo aleatorio simple de viviendas necesitaría, en principio, de un marco de viviendas con las direcciones de domicilios, generado por un censo de población y viviendas que, por un lado, resultaría muy costoso de mantener actualizado y, por otro, las viviendas seleccionadas de un marco de este tipo, se encontrarían tan dispersas geográficamente que, los costos asociados al levantamiento excederían la asignación presupuestaria, por tanto, el muestreo aleatorio simple iría en desmedro de la eficiencia deseada.

Por otra parte, en los diseños de muestra complejos, las unidades que finalmente componen la muestra (en este caso viviendas) se determinan después de varias etapas; en consecuencia, no es necesario mantener un marco de viviendas actualizado, lo cual reduce significativamente los costos.

La formación de conglomerados¹⁴ hace posible agregar nuevas etapas de selección¹⁵, pasando de un diseño monoetápico a uno bietápico. Adicionalmente, al conformar estratos que permitan incrementar la precisión estadística de las estimaciones, se genera la estructura de un diseño de muestra complejo. Para la encuesta Casen 2022, los conglomerados son conjuntos de viviendas que permiten que la muestra quede lo suficientemente concentrada, facilitando el levantamiento y control, dentro de un presupuesto limitado, asegurando niveles razonables de precisión de las estimaciones.

La conformación de estratos de muestreo, la conglomeración de unidades muestrales y la definición de las probabilidades de inclusión, caracterizan los diseños de muestra complejos (Heeringa, West, & Berglund, 2010, pág. 13). El diseño de Casen presenta elementos que lo complejizan, específicamente, por el hecho que sea probabilístico, estratificado y bietápico.

Mientras, algunos de estos elementos tienden a reducir la varianza, como la estratificación, otros tienden a aumentarla, como la conglomeración. Las distintas etapas de muestreo tienden a aumentar la varianza mientras que, la disminución del tamaño de los conglomerados tiende a disminuirla. También es necesario considerar otros efectos tales como la distribución de la prevalencia de la pobreza en las unidades secundarias de muestreo o la correlación de las unidades secundarias al interior de los conglomerados, respecto a la característica de interés, llamada correlación intraclásica, que afecta la varianza y es el principal factor en la medición del efecto del diseño o efecto de no considerar un muestreo aleatorio simple (Kish, 1965).

El cociente entre la estimación de la varianza bajo un diseño complejo versus la estimación bajo un diseño aleatorio simple se denomina efecto del diseño y, en términos de tamaño muestral, equivale a cuánto se debe aumentar la muestra bajo diseño complejo para alcanzar el mismo nivel de precisión que se obtendría si aplicáramos un muestreo aleatorio simple.

¹⁴ Conjuntos de unidades muestrales (viviendas).

¹⁵ El método de selección en la etapa i es independiente al de la etapa j, pudiendo realizarse con probabilidades iguales o no, perdiendo la condición de una muestra autoponderada en este último caso.

Luego, para determinar el tamaño de la muestra compleja de Casen 2022, se incorpora este factor de ajuste a partir del tamaño de una muestra aleatoria simple, lo que permite aproximarse al número de viviendas necesarias para que el diseño complejo proporcione la misma varianza (Kish, 1965).

En este contexto, la metodología de cálculo del tamaño muestral utilizada en Casen desde 2011 hasta 2020, es básicamente la misma. Primero, se inicia con la determinación de un tamaño muestral bajo un muestreo aleatorio simple, fijando previamente el error absoluto a alcanzar con base en los resultados de encuestas anteriores. Posteriormente, se realizan tres ajustes secuenciales: el primero, por el efecto del diseño (obtenido a partir del levantamiento de una encuesta o como el promedio del levantamiento de dos o más encuestas); el segundo, corresponde a un ajuste por población finita y; el tercero, se refiere al ajuste por la tasa de no logro, que también puede ser obtenida a partir de los resultados de un levantamiento previo o, como el promedio de 2 o más levantamientos.

Para la obtención de tamaños muestrales para Casen 2022, la metodología difiere levemente del proceso descrito en el párrafo anterior pues, si bien, comienza con la determinación de un tamaño muestral bajo un muestreo aleatorio simple es, a partir de este tamaño que se obtiene un error de estimación. Luego, utilizando efectos que dan cuenta del diseño complejo, se calculan tamaños necesarios que permitan obtener estimaciones con el grado de precisión calculado en el paso anterior. Los efectos de diseño utilizados para Casen 2022 son obtenidos considerando el diseño del nuevo marco muestral MMV 2020 en conjunto con los efectos obtenidos en el levantamiento de Casen en Pandemia 2020.

De este modo, para la encuesta Casen 2022, se obtuvieron tamaños de muestra regionales, asegurando representatividad a nivel nacional, nacional urbano, nacional rural y regional.

Posteriormente, operó un procedimiento mediante el cual se distribuyó el tamaño de muestra de cada región por el estrato geográfico comuna-área y, luego por niveles socioeconómicos (NSE), conformando el estrato de muestreo comuna-área-NSE donde se buscó, por un lado, resguardar coherencia entre los tamaños muestrales y aquellos de la población, en términos de proporcionalidad y, por otro, que dicha distribución no difiriera mayormente de la distribución de tamaños de Casen en Pandemia 2020.

Bajo esta distribución, los tamaños con sobremuestreo son calculados considerando el promedio ponderado entre las tasas de no logro¹⁶ obtenidas en los levantamientos de Casen 2017 y Casen en Pandemia 2020, con ciertos ajustes que se describen en capítulos posteriores.

¹⁶ La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

II.4. RESUMEN DEL PROCESO DE SIMULACIÓN

En este capítulo se presentan las simulaciones de tamaños muestrales derivadas de distintos escenarios que fueron trabajados inicialmente, hasta llegar a la propuesta elegida. La estructura del capítulo mantiene la cronología de las simulaciones realizadas, así como también, la metodología empleada en ellas.

Como base para el desarrollo de todos los escenarios trabajados se presentan los resultados del trabajo de campo de Casen en Pandemia 2020, desde donde se obtiene la estimación de la tasa de pobreza junto con sus estadísticos asociados, a saber, el error estándar, el coeficiente de variación y el efecto de diseño, tanto a nivel nacional como regional.

A partir de estos resultados, fueron analizados siete escenarios, el primero de ellos (escenario N°1) corresponde a una simulación inicial considerando efectos de diseño provenientes desde el marco muestral MMV 2020 aprovechando, por un lado, estudios realizados para obtener el tamaño muestral para la Encuesta Nacional de Empleo y por otro, los valores utilizados en la Encuesta de Presupuestos Familiares¹⁷. El tamaño muestral resultante sirvió como punto de partida para el desarrollo de los siguientes escenarios.

Los siguientes dos escenarios (N°2.1 y N°2.2) corresponden a un recálculo de tamaños muestrales considerando la adecuación de efectos de diseño para las regiones de: Arica y Parinacota, Tarapacá, Atacama, Ñuble, Los Ríos, Aysén y Magallanes, donde se utilizó un efecto de 1,4; y para la Región Metropolitana, donde se utilizó un efecto de 3,0.

La generación del escenario N°2.1 considera la exclusión de comunas-áreas donde el tamaño muestral asignado, en forma proporcional al marco muestral, resulta inferior a 16 viviendas. Por su parte, el escenario N°2.2 considera un tamaño objetivo mínimo de 8 viviendas por cada estrato geográfico (comuna-área).

Dado que el escenario N°2.2 no excluye estratos geográficos generados por la combinación de comuna-área, pero fija un tamaño objetivo mínimo de viviendas para los estratos pequeños, los siguientes escenarios se desarrollan a partir de él.

¹⁷ Los efectos de diseño utilizados en la Encuesta Nacional de Empleo toman un valor fijo de 1,7 para todas las regiones. Para mayor detalle revisar Documento Informe de Simulaciones de tamaños muestrales Encuesta Nacional de Empleo (ENE). Disponible en: [Informe de simulaciones de tamaños muestrales, ENE](#)
Los efectos de diseño utilizados en la Encuesta de Presupuestos Familiares toman valores entre 1,7 y 2,5.

El escenario N°3 considera dos ajustes a la distribución proporcional inicial del tamaño objetivo por estrato geográfico. El primero, fija un tamaño mínimo de viviendas y el segundo, controla por el número de UPM resultantes dado dicho tamaño.

En el escenario N°4 se realizan ajustes a los tamaños por estrato geográfico según el número de UPM que contienen en el marco muestral. Así, para estratos que contienen una sola UPM se les asigna una UPM en la muestra; para estratos que contienen entre dos y siete UPM, se les asignan dos UPM en la muestra y finalmente, para los estratos que contienen ocho UPM o más, se calcula un tamaño muestral de modo que se asignen, al menos, cuatro UPM en la muestra.

A partir del escenario N°4, manteniendo los tamaños objetivos obtenidos surge el escenario N°5 sobre el cual ocurrirá el trabajo operativo de Casen 2022. Este escenario considera el aumento del tamaño con sobremuestreo en algunos estratos donde es necesario resguardar su tamaño a lograr, según la experiencia del trabajo de campo de otras operaciones estadísticas, como el Estudio Piloto Cambio de Marco Casen 2022¹⁸. El aumento de tamaño se realiza al interior de las UPM, considerando seleccionar 2 viviendas adicionales a las definidas originalmente en el diseño, según el estrato geográfico (comuna – área) que se ha definido como prioritario. En consecuencia, en cada UPM se seleccionan 8, 10 o 12 viviendas para encuestar.

Adicionalmente y también a partir del escenario N°4 se genera el escenario N°6 con la finalidad de considerar una eventual modalidad de levantamiento mixto donde, es necesario, a partir del tamaño con sobremuestreo calculado en el escenario N°4, calcular un nuevo tamaño objetivo. Para obtener este tamaño se aplica la tasa de no logro¹⁹ correspondiente a Casen en Pandemia 2020 a nivel regional y luego, para la distribución a nivel de estratos de muestreo de Casen 2022, se consideran los mismos criterios de tamaño del escenario N°4. Estos tamaños objetivos son inferiores a los calculados en dicho escenario, sin embargo, mantienen su proporcionalidad.

Luego, a partir de los tamaños muestrales objetivos del escenario elegido (escenario N°5) se realiza una comparación entre su distribución regional, la distribución del marco muestral y la distribución de Casen en Pandemia 2020, con la finalidad de mostrar que no existen mayores diferencias con dichos referentes.

¹⁸ A diferencia de las versiones previas de Casen, donde se utilizó el Marco muestral con base en Censo 2002, en Casen 2022 se utiliza el Marco Muestral de Viviendas (MMV 2020) con base en Censo 2017. Dado lo anterior, y debido a las posibles implicancias de la actualización del marco muestral en los principales indicadores de la encuesta Casen (Pobreza por ingresos y pobreza multidimensional), el Ministerio, en conjunto con el Panel de Expertas y Expertos, definieron evaluar de forma previa a la recolección de Casen 2022, el efecto del cambio de marco muestral por medio de una prueba piloto (realizada por el INE). El objetivo del estudio fue estimar el efecto de utilizar el MMV 2020 sobre los principales indicadores de la encuesta Casen. Para mayor detalle revisar: [Nota técnica N°1: Principales cambios metodológicos Casen 2022](#)

¹⁹ La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

El Cuadro II.1 presenta los parámetros comunes utilizados en los escenarios considerados.

Cuadro II.1. Parámetros comunes utilizados en escenarios Casen 2022

Parámetro	Descripción
Parámetro a estimar	Tasa de pobreza por ingresos
Variable de diseño	$X_j = \begin{cases} 0, & \text{si el individuo } j \text{ no es pobre por ingresos} \\ 1, & \text{si el individuo } j \text{ es pobre por ingresos} \end{cases}$
Estimador asociado	Estimador de razón: $r = \frac{\text{Número de personas pobres por ingresos}}{\text{Número total de personas}}$

Elaborado por el Instituto Nacional de Estadísticas.

La Tabla II.1 presenta un resumen de los parámetros utilizados en cada una de las simulaciones, así como también, los tamaños y errores muestrales resultantes, por orden cronológico, comenzando con los resultados del trabajo de campo de Casen en Pandemia 2020 (escenario N°0) para luego enumerar los siguientes escenarios desde el 1 al 6.

Tabla II.1. Resumen de parámetros utilizados y tamaños y errores resultantes según escenario, Casen 2022

Nro.	N° viviendas MMV 2020	Niveles de estimación	Error Absoluto	Error Relativo	Efectos de Diseño	Tamaño Muestral Objetivo	Tamaño Muestral Objetivo (ajustado) ²⁰	Tasa de no logro ²¹	Tamaño con sobremuestreo
0		Nacional	0,4%	3,9%	2,2	62.540			
		Regional	Entre 0,7% - 2,1%	Entre 8,3% -31,0%	Entre 0,4 - 3,2				
1	6.206.626	Nacional	0,4%	3,4%	2,2	72.555	71.025	Promedio TNL Casen 2015 - 2017 Comuna	95.498
		Regional	Entre 0,7% - 3,1%	Entre 8,3% - 47,3%	Entre 1,7 - 2,5				
2.1	6.115.331	Nacional	0,4%	3,3%	2,2	69.846	70.403	Promedio Ponderado TNL Casen 2017 - 2020 Comuna U-R ²²	105.928
		Regional	Entre 0,7% - 2,1%	Entre 8,3% - 31,9%	Entre 1,4 - 3				
2.2	6.206.626	Nacional	0,4%	3,3%	2,2	69.829	71.755	Promedio Ponderado TNL Casen 2017 - 2020 Comuna U-R	107.832
		Regional	Entre 0,7% - 2,1%	Entre 8,3% - 31,3%	Entre 1,4 - 3				
3		Nacional	0,4%	3,4%	2,2	68.034	68.857	Promedio Ponderado TNL Casen 2017 - 2020	101.338
		Regional	Entre 0,7% - 2,1%	Entre 8,3% - 30,5%	Entre 1 - 3,2				

²⁰ Los errores absolutos y relativos son calculados con el tamaño objetivo ajustado.

²¹ La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

²² Se obtuvieron las tasas de no logro por estrato geográfico (comuna-área) desde los levantamientos de Casen 2017 y Casen en Pandemia 2020. En el caso de comunas nuevas incluidas en Casen 2022, las tasas de no logro fueron obtenidas como un promedio a nivel de región-área, considerando ambos levantamientos.

Nro.	N° viviendas MMV 2020	Niveles de estimación	Error Absoluto	Error Relativo	Efectos de Diseño	Tamaño Muestral Objetivo	Tamaño Muestral Objetivo (ajustado) ²⁰	Tasa de no logro ²¹	Tamaño con sobremuestreo
								Comuna U-R ajustada ²³	
4	6.206.626	Nacional	0,4%	3,3%	2,2			Promedio Ponderado TNL Casen 2017 - 2020	104.246
		Regional	Entre 0,7% - 2,1%	Entre 8,2% - 30,5%	Entre 1 - 3,2	68.034	71.028	Comuna U-R ajustada	
5	6.206.626	Nacional	0,4%	3,3%	2,2			Promedio Ponderado TNL Casen 2017 - 2020	106.856
		Regional	Entre 0,7% - 2,1%	Entre 8,2% - 30,5%	Entre 1 - 3,2	68.034	71.028	Comuna U-R ajustada ²⁴	
6	6.206.626	Nacional	0,4%	3,5%	2,2			TNL Casen en Pandemia 2020	104.246
		Regional	Entre 0,8% - 2,2%	Entre 8,5% - 34,3%	Entre 1 - 3,2	64.376			

Elaborado por el Instituto Nacional de Estadísticas.

²³ Las tasas de no logro son ajustadas cuando, al aplicarlas para obtener el tamaño con sobremuestreo y luego, calcular el número de UPM a seleccionar, el tamaño resultante excede la cantidad de UPM por estrato que contiene el marco muestral.

²⁴ Tasa de no logro ajustada según el resultado del trabajo de campo del Estudio Piloto Cambio de Marco Casen 2022.

II.4.1. Escenario N°0 (Escenario Base)

El total de viviendas logradas en el levantamiento de Casen en Pandemia 2020 fue de 62.540, desde donde se obtuvo una estimación de la tasa de pobreza nacional de 10,8%, asociada a un error absoluto de 0,4% y a un error relativo de 3,9%. El efecto de diseño (Deff) resultante es de 2,2. A nivel regional la tasa de pobreza fluctúa entre 5,7% (región de Magallanes) y 17,4% (región de La Araucanía) ambas con errores absolutos asociados de 1,8%. El efecto de diseño a nivel regional fluctúa entre 0,4 y 3,2 obteniendo el mayor valor en la región Metropolitana.

Luego, con estos parámetros se calcula un tamaño muestral, obviando el efecto diseño proveniente del marco estratificado de manzanas y secciones utilizado en Casen en Pandemia 2020, a partir del cálculo de una cuasi varianza poblacional ajustada (llamada Cuasi-Varianza Deff.Ajustado) que acota los efectos de diseño obtenidos a valores entre 1 y 5. Con lo anterior se calcula un tamaño de 40.001 viviendas provenientes de un diseño aleatorio simple, a partir del cual, es posible obtener una estimación de la tasa nacional con errores absolutos y relativos de 0,4% y 3,9%, respectivamente. A nivel regional se obtienen errores absolutos entre 0,7% y 2,1% para las regiones Metropolitana y Coquimbo y, errores relativos entre 8,3% y 31,0% para las regiones Metropolitana y Magallanes, respectivamente²⁵.

Esta simulación sirve como tamaño inicial para luego, aplicar efectos de diseño ad-hoc al levantamiento de Casen 2022, basados en el marco muestral de viviendas MMV 2020.

II.4.2. Escenario N°1

Este escenario corresponde a una simulación inicial de tamaños muestrales realizada con base en los parámetros obtenidos del trabajo de campo de Casen en Pandemia 2020 (escenario N°0) y en efectos de diseño similares a los utilizados en la Encuesta Nacional de Empleo (ENE) y en la IX Encuesta de Presupuestos Familiares (EPF) cuyos valores fueron obtenidos desde estudios previos acerca del comportamiento del marco muestral MMV 2017²⁶.

Considerando la estimación de la tasa de pobreza junto con sus parámetros asociados, resultantes del trabajo de campo de Casen en Pandemia 2020, y utilizando efectos de diseño simulados y acotados, se obtiene un tamaño objetivo de 72.555²⁷ viviendas a nivel nacional que, aplicando corrección por población finita, resulta en un tamaño de 71.025 viviendas. Este tamaño permite estimar la tasa de pobreza con errores absolutos y relativos de 0,4% y 3,4%, respectivamente.

²⁵ Más detalle en Tabla VII.1 Escenario N°0 (Escenario Base) del Anexo N°1.

²⁶ Para mayor detalle revisar: [Informe de simulaciones de tamaños muestrales, ENE](#)

²⁷ Este tamaño es obtenido considerando un tamaño bajo muestreo aleatorio simple de 39.997 viviendas, calculado a partir de la cuasivarianza poblacional, ajustada según los efectos de diseño propios de este escenario. Más detalle en Tabla VII.2 Escenario N°1 del Anexo N°1.

A nivel regional los errores absolutos fluctúan entre 0,7% y 3,1% para las regiones Metropolitana y Aysén, respectivamente y; los errores relativos, entre 8,3% y 47,3% para las mismas regiones.

Los valores considerados para el efecto del diseño (Deff) se detallan a continuación:

- Si Deff Casen en Pandemia 2020 > 2,5 → Deff Casen 2022: 2,5
- Si Deff Casen en Pandemia 2020 < 1,7 → Deff Casen 2022: 1,7
- En cualquier otro caso, Deff Casen 2022: Deff Casen en Pandemia 2020

Se observan errores relativos altos para las regiones de Aysén y Magallanes, sin embargo, a partir del error estándar se concluye que estos valores cumplen con el criterio establecido por el *Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares*²⁸.

Al distribuir los tamaños regionales entre los estratos geográficos (comuna-área) en forma proporcional al MMV 2020 y aplicando una tasa de no logro²⁹ promedio ponderada entre Casen 2015 y Casen 2017 (a nivel comunal, sin distinguir por área) se obtiene un tamaño nacional con sobremuestreo de 95.498 viviendas³⁰. Es importante mencionar que, en aquellas comunas que se incorporan al diseño de Casen 2022 se les imputa el promedio de la región.

Las tasas de no logro a nivel comunal utilizadas en el cálculo del tamaño con sobremuestreo se obtienen desde la ecuación (1):

$$tnl_p = \frac{tnl_j \cdot eg_j + tnl_k \cdot eg_k}{eg_j + eg_k} = \frac{\frac{ec_j}{eg_j} \cdot eg_j + \frac{ec_k}{eg_k} \cdot eg_k}{eg_j + eg_k} \quad (1)$$

Donde:

tnl_p : Tasa de no logro ponderada por el número de encuestas gestionadas.

$tnl_{j,k}$: Tasa de no logro del año j o del año k .

$eg_{j,k}$: Número de encuestas gestionadas el año j o el año k .

$ec_{j,k}$: Número de encuestas logradas el año j o el año k .

II.4.3. Escenarios N°2.1 y N°2.2

Con base en una revisión de efectos de diseño utilizados en el escenario N°1, surgen los escenarios N°2.1 y N°2.2 que modifican estos efectos en las regiones de: Arica y Parinacota, Tarapacá, Atacama, Ñuble, Los Ríos, Aysén y Magallanes, asignándoles un valor fijo de 1,4 y; en la Región Metropolitana, un valor de

²⁸ Para mayor detalle revisar: [Fundamentos del Estándar para la evaluación de la calidad de las estimaciones para encuestas de hogares](#)

²⁹ La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

³⁰ Más detalle en Tabla VII.2 Escenario N°1 del Anexo N°1.

3,0. El motivo por el cual se modifican los efectos en las regiones anteriormente expuestas es porque, aun cuando, en el marco muestral presentan tamaños pequeños, en término de número de viviendas, los tamaños objetivos asignados en forma proporcional, eran más altos que los considerados en versiones anteriores de la encuesta, lo que podría implicar un desafío operativo mayor en el trabajo de campo en Casen 2022, en relación con versiones anteriores.

La diferencia entre ambos escenarios se encuentra en que, el escenario N°2.1 excluye estratos geográficos (comunas-áreas) donde el tamaño muestral asignado, en forma proporcional al marco muestral, resulta inferior a 16 viviendas (similar a lo aplicado en simulación para obtener el tamaño muestral de Casen en Pandemia 2020) mientras que, el escenario N°2.2 no excluye estratos geográficos, pero fija un tamaño objetivo mínimo de 8 viviendas por cada uno de ellos.

El escenario N°2.1 con exclusión de comunas-áreas considera un marco muestral que pasa de 6.206.626 a 6.115.331 viviendas.

El tamaño calculado bajo muestreo aleatorio simple es ajustado por el efecto del diseño, resultando en un tamaño de 69.846³¹ viviendas. Aplicando un ajuste por finitud de la población se obtiene un tamaño nacional de 68.625 viviendas. Al distribuir los tamaños regionales resultantes, en forma proporcional al MMV 2020, entre los estratos geográficos (comuna-área) y, luego de ciertos ajustes para cumplir con tamaños adecuados a ese nivel, se obtiene un tamaño nacional ajustado de 70.403 viviendas que permiten obtener una estimación de la tasa de pobreza con un error absoluto de 0,4% y un error relativo de 3,3%.

A nivel regional, los errores absolutos fluctúan entre 0,7% y 2,1% para las regiones Metropolitana y Coquimbo, respectivamente. Por otro lado, los errores relativos varían entre 8,3% y 31,9% para las regiones Metropolitana y Magallanes, respectivamente.

Aplicando una tasa de no logro³² promedio ponderada entre Casen 2017 y Casen en Pandemia 2020 a nivel de comuna-área se obtiene un tamaño nacional con sobremuestreo de 105.928 viviendas³³. Se debe señalar que, la ventaja de utilizar la tasa de no logro promedio ponderada considerando además la desagregación por área (urbano-rural) se encuentra en que, no sólo es posible incorporar el efecto de los resultados del trabajo de campo de Casen en Pandemia 2020 frente a Casen 2017, sino que, además se logran observar las diferencias en el desempeño operativo entre las áreas urbanas y rurales.

³¹ Este tamaño es obtenido considerando un tamaño bajo muestreo aleatorio simple de 40.007 viviendas, calculado a partir de la cuasivarianza poblacional, ajustada según los efectos de diseño propios de este escenario.

³² La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

³³ Más detalle en Tabla VII.3. Escenario N°2.1 del Anexo N°1.

Según la asignación de tamaño explicada en párrafo anterior, en este escenario queda excluida el área urbana de algunas comunas, el área rural de otras y, tanto en su área urbana como rural, las comunas de: La Higuera, Panquehue, La Estrella, Empedrado, Contulmo, Antuco, Quilaco, San Rosendo, Curaco De Vélez, Queilén, Alhué y María Pinto.

El escenario N°2.2 considera tamaño objetivo mínimo de 8 viviendas para cada estrato geográfico (comuna-área).

El tamaño calculado bajo muestreo aleatorio simple es ajustado por el efecto del diseño, resultando en un tamaño de 69.829³⁴ viviendas. Aplicando un ajuste por finitud de la población se obtiene un tamaño nacional de 68.625 viviendas. Al distribuir los tamaños regionales resultantes, en forma proporcional al MMV 2020, entre los estratos geográficos (comuna-área) y, luego de ciertos ajustes para cumplir con tamaños adecuados a ese nivel, se obtiene un tamaño nacional ajustado de 71.755 viviendas que permiten obtener una estimación de la tasa de pobreza con un error absoluto de 0,4% y un error relativo de 3,3%.

A nivel regional los errores absolutos fluctúan entre 0,7% y 2,1% para las regiones Metropolitana y Coquimbo, respectivamente y; los errores relativos, entre 8,3% y 31,3% para las regiones Metropolitana y Magallanes.

Aplicando una tasa de no logro³⁵ promedio ponderada entre Casen 2017 y Casen en Pandemia 2020 a nivel de comuna-área se obtiene un tamaño nacional con sobremuestreo de 107.832 viviendas³⁶.

A partir de estos dos escenarios los tamaños son distribuidos entre los estratos de muestreo conformados especialmente para Casen 2022, que corresponden a comuna-área y nivel socioeconómico donde aplique.

Los siguientes escenarios, N°3 y N°4, toman como base el escenario N°2.2, fijando un tamaño objetivo mínimo de viviendas para los estratos que, por distribución proporcional inicial, se les había asignado un bajo tamaño muestral.

II.4.4. Escenario N°3

A partir de este escenario se considera distribuir el tamaño muestral de viviendas, en UPM (o unidades primarias de muestreo) definiendo encuestar 8 viviendas en cada una, tanto en áreas urbanas como rurales cumpliendo, en la medida de lo posible, con los siguientes criterios:

- En estratos que contengan ocho UPM o más en el marco muestral se seleccionan, al menos cuatro UPM, tanto para áreas urbanas como rurales.

³⁴ Este tamaño es obtenido considerando un tamaño bajo muestreo aleatorio simple de 39.996 viviendas, calculado a partir de la cuasivarianza poblacional, ajustada según los efectos de diseño propios de este escenario.

³⁵ La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

³⁶ Más detalle en Tabla VII.4. Escenario 2.2 del Anexo N°1.

- En aquellos estratos que, en el marco muestral contengan entre dos y siete UPM, se seleccionan dos UPM.
- En estratos que presentan una UPM en el marco muestral, se selecciona una UPM en la muestra.
- El tamaño objetivo debe considerar que, la fracción de muestreo, en términos de UPM, debe ser máximo 70% respecto del marco muestral, salvo en caso de estratos que presentan una o dos UPM en el marco.

Es importante destacar que, en este escenario, existen estratos que tienen asignado un menor número de viviendas que las correspondientes a los criterios descritos anteriormente, esto ocurre dado que, en la distribución inicial, que incorpora sólo la proporcionalidad, según número de viviendas del marco, les corresponde un tamaño inferior al de los criterios definidos.

Otro factor importante que considerar es que, para la distribución por estrato de muestreo, se ajustaron los efectos de diseño en algunas regiones, por un lado, para estar en línea con la estimación regional de estos, provenientes desde el levantamiento de Casen en Pandemia 2020 y, por otro, porque es necesario que el marco presente suficientes unidades primarias de muestreo para la selección.

En este escenario se ajustan (disminuyen) los efectos de diseño de las siguientes regiones: Arica y Parinacota, Tarapacá, Aysén y Magallanes. Por otro lado, se aumentan en la Región Metropolitana, O'Higgins y Biobío, según los resultados de trabajo de campo de Casen en Pandemia 2020.

El tamaño calculado bajo muestreo aleatorio simple es ajustado por el efecto del diseño, resultando en un tamaño de 68.034³⁷ viviendas. Aplicando un ajuste por finitud de la población se obtiene un tamaño nacional de 67.032 viviendas. Al distribuir los tamaños regionales resultantes, en forma proporcional al MMV 2020, entre los estratos geográficos (comuna-área) y considerando los criterios expuestos en párrafos anteriores, se obtiene un tamaño nacional ajustado de 68.857 viviendas que permiten obtener una estimación de la tasa de pobreza con un error absoluto de 0,4% y un error relativo de 3,4%.

A nivel regional los errores absolutos fluctúan entre 0,7% y 2,1% para las regiones Metropolitana y Coquimbo, respectivamente. Por su parte, los errores relativos fluctúan entre 8,3% y 30,5% para las regiones Metropolitana y Magallanes, respectivamente.

Aplicando una tasa de no logro³⁸ promedio ponderada entre Casen 2017 y Casen en Pandemia 2020 a nivel de comuna-área se obtiene un tamaño nacional con sobremuestreo de 101.338³⁹ viviendas.

³⁷ Este tamaño es obtenido considerando un tamaño bajo muestreo aleatorio simple de 39.705 viviendas, calculado a partir de la cuasivarianza poblacional, ajustada según los efectos de diseño propios de este escenario.

³⁸ La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

II.4.5. Escenario N°4

El escenario N°4 corresponde al tamaño y distribución de la muestra objetivo aprobados por Ministerio según los últimos ajustes realizados al escenario N°3. En este escenario, dada la cantidad de UPM disponibles en el marco muestral, se decide seleccionar 10 viviendas por UPM en el área rural y mantener 8 viviendas por UPM en el área urbana. A partir de los criterios de determinación de tamaños muestrales definidos en el escenario N°3, los siguientes dos criterios se vuelven imprescindibles:

- Se seleccionan en la muestra, al menos, cuatro UPM para todos los estratos que tengan en el marco muestral ocho UPM o más. Según lo anterior, en el área urbana se seleccionan como mínimo 32 viviendas por estrato y; en el área rural, 40.
- En todos los estratos que tengan en el marco muestral, entre dos y siete UPM, se seleccionan dos UPM en la muestra. Esto corresponde a seleccionar 16 viviendas en el área urbana y 20, en el área rural.

Cabe mencionar que, a pesar de la proporcionalidad según marco muestral, utilizada en la distribución del número de viviendas que se considera en el escenario N°3, en este escenario no existen excepciones respecto del número mínimo de viviendas asignadas como producto de la distribución inicial.

A partir del escenario N°3 y de los resultados del trabajo de campo de Casen en Pandemia 2020, se obtienen los valores para el cálculo de tamaño muestral y parámetros asociados para el escenario N°4.

El tamaño calculado bajo muestreo aleatorio simple es ajustado por el efecto del diseño, resultando en un tamaño de 68.034⁴⁰ viviendas. Aplicando un ajuste por finitud de la población se obtiene un tamaño nacional de 67.032 viviendas. Al distribuir los tamaños regionales resultantes, en forma proporcional al MMV 2020, entre los estratos geográficos (comuna-área) y, considerando los criterios expuestos en párrafos anteriores, se obtiene un tamaño nacional ajustado de 71.028 viviendas que permiten obtener una estimación de la tasa de pobreza con un error absoluto de 0,4% y un error relativo de 3,3%.

A nivel regional los errores absolutos fluctúan entre 0,7% y 2,1% para las regiones Metropolitana y Coquimbo, respectivamente. Por su parte, los errores relativos fluctúan entre 8,2% y 30,5% para las regiones Metropolitana y Magallanes, respectivamente.

³⁹ Más detalle en Tabla VII.5. Escenario N°3 del Anexo N°1.

⁴⁰ Este tamaño es obtenido considerando un tamaño bajo muestreo aleatorio simple de 39.705 viviendas, calculado a partir de la cuasivarianza poblacional, ajustada según los efectos de diseño propios de este escenario.

Luego, aplicando una tasa de no logro⁴¹ promedio ponderada entre Casen 2017 y Casen en Pandemia 2020 a nivel comuna-área, se obtiene un tamaño nacional con sobremuestreo de 104.246 viviendas⁴².

II.4.6. Escenario N°5

El escenario N°5, elegido finalmente, surge a partir de los tamaños objetivos obtenidos en el escenario N°4, considerando la aplicación de tasas de no logro ajustadas en algunos estratos geográficos (comuna - área) específicamente, en aquellos donde en el trabajo de campo del Estudio Piloto Cambio de Marco Casen 2022 se presentaron bajas tasas de logro de viviendas. Para el cálculo del tamaño con sobremuestreo en este escenario, se utilizaron tasas de no logro de manera tal que, en las UPM ya definidas en el escenario N°4 y pertenecientes a los estratos en cuestión, se considere seleccionar 2 viviendas adicionales, tanto en el área urbana como en el área rural, quedando conformadas las UPM por 8 o 10 viviendas en el área urbana y, 10 o 12 viviendas en el área rural.

En la Tabla II.2 se observan los tamaños muestrales y los errores asociados, junto con los tamaños con sobremuestreo definitivos a nivel nacional y regional.

⁴¹ La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

⁴² Más detalle en Tabla VII.6 Escenario N° 4 del Anexo N°1.

Tabla II.2. Tamaños muestrales y errores asociados a nivel nacional y regional, Casen 2022

Escenario N°5 Casen 2022

Región	Viviendas MMV 2020	Tasa de Pobreza	Efecto Diseño Ajustado	Tamaños Ajustados Efectos de Diseño	Error Absoluto	Error Relativo	Tamaños Ajustados por finitud	Tamaño Objetivos Ajustados	Error Absoluto Ajustado	Error Relativo Ajustado	Tamaño con Sobremuestreo
Nivel País	6.206.626	10,8%	2,2	68.034	0,4%	3,4%	67.032	71.028	0,4%	3,3%	106.856
Arica y Parinacota	73.783	11,9%	1,2	2.737	1,8%	15,5%	2.640	2.646	1,9%	15,8%	3.698
Tarapacá	110.680	14,0%	1	2.621	1,7%	12,1%	2.561	2.638	1,7%	12,1%	4.156
Antofagasta	190.638	9,3%	1,7	2.684	1,8%	19,7%	2.647	2.787	1,8%	19,3%	4.580
Atacama	106.193	9,5%	1,4	3.363	1,8%	18,9%	3.260	3.290	1,8%	19,1%	5.060
Coquimbo	285.374	11,7%	2,2	2.695	2,1%	18,3%	2.670	2.837	2,1%	17,8%	4.180
Valparaíso	677.668	11,3%	1,9	6.325	1,3%	11,0%	6.267	7.070	1,2%	10,4%	10.984
Metropolitana	2.379.740	9,0%	3,2	13.123	0,7%	8,3%	13.052	13.607	0,7%	8,2%	22.804
O'Higgins	333.576	10,0%	1,8	4.455	1,6%	16,0%	4.397	4.798	1,5%	15,4%	6.672
Maule	392.061	12,3%	1,7	4.889	1,5%	12,1%	4.829	5.107	1,5%	11,8%	7.090
Ñuble	182.744	14,7%	1,4	3.175	1,8%	12,2%	3.121	3.319	1,8%	11,9%	4.696
Biobío	558.582	13,2%	1,8	6.363	1,4%	10,7%	6.292	6.909	1,4%	10,2%	9.800
La Araucanía	359.601	17,4%	1,7	4.860	1,8%	10,4%	4.796	5.052	1,8%	10,2%	7.006
Los Ríos	144.531	12,2%	1,4	3.907	1,5%	12,6%	3.805	3.820	1,6%	12,7%	5.400
Los Lagos	308.709	11,3%	1,7	3.580	1,8%	15,5%	3.539	3.826	1,7%	15,0%	5.466
Aysén	40.838	6,6%	1	1.400	2,0%	29,9%	1.353	1.397	2,0%	29,9%	2.070
Magallanes	61.908	5,7%	1	1.857	1,8%	31,0%	1.803	1.925	1,7%	30,5%	3.194

Elaborado por el Instituto Nacional de Estadísticas.

II.4.7. Escenario N°6

Este escenario se construye a partir del tamaño con sobremuestreo obtenido en el escenario N°4 y surge de la necesidad de contar con un tamaño muestral objetivo para un eventual levantamiento mixto, pero considerando que la situación sanitaria puede cambiar y pueda ser necesario transitar desde un levantamiento mixto a una aplicación principalmente telefónica, tal como ocurrió en Casen en Pandemia 2020.

Suponiendo un levantamiento bajo modalidad mixta y considerando el tamaño con sobremuestreo calculado en el escenario N°4, se utiliza la tasa de no logro⁴³ regional de Casen en Pandemia 2020 para obtener los tamaños objetivos a este nivel y según los siguientes criterios:

- La distribución de la muestra regional en los estratos de muestreo (comuna-área-NSE) se realiza de manera proporcional al peso del estrato del escenario N°4 para modalidad presencial.
- La distribución de la muestra a nivel de los estratos geográficos (comuna-área) cumple con los siguientes mínimos establecidos para la muestra en modalidad presencial:
 - Estratos de una UPM en el marco: debe ser incluido en la muestra.
 - Estratos entre dos y siete UPM en el marco: se debe cumplir un mínimo de dos UPM en la muestra. Esto implica que en el área urbana de cada comuna se considera entrevistar, al menos, 16 viviendas y; en el área rural, 20.
 - Estratos de ocho UPM o más en el marco: se seleccionan en la muestra, al menos, cuatro UPM. Esto implica que, en el área urbana de cada comuna, se considera entrevistar, al menos, 32 viviendas y; en el área rural, al menos 40.

⁴³ La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

Tabla II.3. Tamaños muestrales y errores asociados a nivel nacional y regional. Casen 2022 en modalidad de levantamiento mixto

Escenario N°6 Casen 2022							
Región	Viviendas MMV 2020	Tasa de Pobreza	Tamaño Objetivo Ajustado	Error Absoluto Objetivo	Error Relativo Objetivo	Tasa de Logro Regional	Tamaño con Sobremuestreo
Nivel País	6.206.626	10,8%	64.376	0,4%	3,5%		104.246
Arica y Parinacota	73.783	11,9%	1.801	2,2%	18,8%	59,9%	2.970
Tarapacá	110.680	14,0%	2.522	1,7%	12,4%	60,5%	4.156
Antofagasta	190.638	9,3%	2.723	1,9%	19,9%	59,3%	4.580
Atacama	106.193	9,5%	2.603	2,0%	21,2%	61,5%	4.158
Coquimbo	285.374	11,7%	2.611	2,2%	19,1%	61,4%	4.176
Valparaíso	677.668	11,3%	6.749	1,3%	11,3%	60,6%	10.984
Metropolitana	2.379.740	9,0%	13.234	0,8%	8,5%	57,9%	22.804
O'Higgins	333.576	10,0%	4.333	1,7%	16,8%	62,6%	6.672
Maule	392.061	12,3%	4.586	1,6%	12,7%	63,0%	7.070
Ñuble	182.744	14,7%	2.966	1,9%	12,9%	61,4%	4.680
Biobío	558.582	13,2%	6.257	1,5%	11,2%	62,5%	9.792
La Araucanía	359.601	17,4%	4.430	1,9%	11,1%	61,3%	7.002
Los Ríos	144.531	12,2%	3.319	1,6%	13,5%	61,2%	5.400
Los Lagos	308.709	11,3%	3.485	1,8%	16,3%	61,8%	5.450
Aysén	40.838	6,6%	1.180	2,2%	32,5%	64,4%	1.784
Magallanes	61.908	5,7%	1.577	1,9%	34,3%	59,9%	2.568

Elaborado por el Instituto Nacional de Estadísticas.

II.4.8. Comparación tamaño objetivo entre Escenario N°5 (escenario elegido) y Escenario N°6 (modalidad telefónica)

La Tabla II.4 muestra los tamaños objetivos y los errores asociados a dichos tamaños de los escenarios N°5 (escenario elegido) y N°6 (escenario bajo levantamiento telefónico), y la comparación de la distribución regional de tamaños, en términos de la proporcionalidad.

Tabla II.4. Comparación de tamaños objetivo y errores asociados Escenarios N°5 y N°6

Región	Escenario N°5				Escenario N°6			
	Tamaño Objetivo	Error Absoluto	Error Relativo	Proporción Regional	Tamaño Objetivo	Error Absoluto	Error Relativo	Proporción Regional
Nivel País	71.028	0,4%	3,3%		64.376	0,4%	3,5%	
Arica y Parinacota	2.646	1,9%	15,8%	3,7%	1.801	2,2%	18,8%	2,8%
Tarapacá	2.638	1,7%	12,1%	3,7%	2.522	1,7%	12,4%	3,9%
Antofagasta	2.787	1,8%	19,3%	3,9%	2.723	1,9%	19,9%	4,2%
Atacama	3.290	1,8%	19,1%	4,6%	2.603	2,0%	21,2%	4,0%
Coquimbo	2.837	2,1%	17,8%	4,0%	2.611	2,2%	19,1%	4,1%
Valparaíso	7.070	1,2%	10,4%	10,0%	6.749	1,3%	11,3%	10,5%
Metropolitana	13.607	0,7%	8,2%	19,2%	13.234	0,8%	8,5%	20,6%
O'Higgins	4.798	1,5%	15,4%	6,8%	4.333	1,7%	16,8%	6,7%
Maule	5.107	1,5%	11,8%	7,2%	4.586	1,6%	12,7%	7,1%
Ñuble	3.319	1,8%	11,9%	4,7%	2.966	1,9%	12,9%	4,6%
Biobío	6.909	1,4%	10,2%	9,7%	6.257	1,5%	11,2%	9,7%
La Araucanía	5.052	1,8%	10,2%	7,1%	4.430	1,9%	11,1%	6,9%
Los Ríos	3.820	1,6%	12,7%	5,4%	3.319	1,6%	13,5%	5,2%
Los Lagos	3.826	1,7%	15,0%	5,4%	3.485	1,8%	16,3%	5,4%
Aysén	1.397	2,0%	29,9%	2,0%	1.180	2,2%	32,5%	1,8%
Magallanes	1.925	1,7%	30,5%	2,7%	1.577	1,9%	34,3%	2,4%

Elaborado por el Instituto Nacional de Estadísticas.

Tabla II.5. Comparación distribución de tamaños objetivo regionales entre ambos escenarios

Escenario N°5		Escenario N°6	
Región	Proporción Regional	Región	Proporción Regional
Nivel País		Nivel País	
Aysén	2,0%	Aysén	1,8%
Magallanes	2,7%	Magallanes	2,4%
Tarapacá	3,7%	Arica y Parinacota	2,8%
Arica y Parinacota	3,7%	Tarapacá	3,9%
Antofagasta	3,9%	Atacama	4,0%
Coquimbo	4,0%	Coquimbo	4,1%
Atacama	4,6%	Antofagasta	4,2%
Ñuble	4,7%	Ñuble	4,6%
Los Ríos	5,4%	Los Ríos	5,2%
Los Lagos	5,4%	Los Lagos	5,4%
O'Higgins	6,8%	O'Higgins	6,7%
La Araucanía	7,1%	La Araucanía	6,9%
Maule	7,2%	Maule	7,1%
Biobío	9,7%	Biobío	9,7%
Valparaíso	10,0%	Valparaíso	10,5%
Metropolitana	19,2%	Metropolitana	20,6%

Elaborado por el Instituto Nacional de Estadísticas.

Los tamaños objetivo a nivel regional del escenario N°6 son obtenidos a partir del tamaño con sobremuestreo del escenario N°4 y aplicando tasas de no logro, según los resultados del trabajo de campo de Casen en Pandemia 2020. Luego, la distribución de estos tamaños en los estratos de muestreo (comuna-área-NSE) se realiza de manera proporcional al peso del estrato del escenario N°4.

Se observa que las proporciones regionales de los tamaños entre ambos escenarios no son exactamente iguales y, esto se debe a que las tasas de no logro utilizadas en el escenario N°6 (a nivel de región) podrían presentar diferencias con las tasas de no logro utilizadas en el escenario N°4 (a nivel del estrato geográfico comuna-área) donde el tamaño con sobremuestreo a nivel regional se obtiene de la agregación de los tamaños con sobremuestreo de las comunas que componen cada región.

Según lo anterior, los tamaños objetivo de 12 de las 16 regiones del escenario N°6 comparten el mismo ordenamiento que la distribución de los tamaños objetivo del escenario N°4. El bloque de regiones que presentan diferencia en dicho ordenamiento es el compuesto por Arica y Parinacota, Tarapacá, Antofagasta y Atacama.

II.5. TAMAÑO MUESTRAL PROPUESTO

El tamaño objetivo del escenario elegido (N°5) definido para obtener estimaciones de la tasa de pobreza con suficiente grado de precisión a nivel regional, alcanza 71.028 viviendas a nivel nacional, cuyos errores absolutos y relativos asociados son de 0,4% y 3,3%, respectivamente. Aplicando las tasas de no logro a nivel de estrato geográfico (comuna-área) con ciertos ajustes (con el objetivo, por un lado, que el marco muestral cuente con suficientes unidades para seleccionar la muestra y, por otro, cuidando las tasas de respuesta de aquellos estratos que en el Estudio Piloto Cambio de Marco Casen 2022, resultaron deficientes) se obtiene el tamaño con sobremuestreo que alcanza 106.856 viviendas.

En la Tabla II.6 se observa el tamaño objetivo junto con sus errores muestrales asociados y el tamaño con sobremuestreo a nivel nacional y regional, desagregados por área urbano-rural.

Tabla II.6. Tamaños y errores a nivel nacional, regional, urbano-rural, Casen 2022

Región	Error Absoluto	Error Relativo	Tamaño Objetivo			Tamaño con Sobremuestreo		
			Total	Urbano	Rural	Total	Urbano	Rural
Nivel País	0,4%	3,3%	71.028	56.905	14.123	106.856	87.252	19.604
Arica y Parinacota	1,9%	15,8%	2.646	2.412	234	3.698	3.350	348
Tarapacá	1,7%	12,1%	2.638	2.439	199	4.156	3.856	300
Antofagasta	1,8%	19,3%	2.787	2.577	210	4.580	4.200	380
Atacama	1,8%	19,1%	3.290	2.880	410	5.060	4.510	550
Coquimbo	2,1%	17,8%	2.837	2.158	679	4.180	3.296	884
Valparaíso	1,2%	10,4%	7.070	5.777	1.293	10.984	9.104	1.880
Metropolitana	0,7%	8,2%	13.607	12.659	948	22.804	21.344	1.460
O'Higgins	1,5%	15,4%	4.798	3.291	1.507	6.672	4.712	1.960
Maule	1,5%	11,8%	5.107	3.587	1.520	7.090	5.084	2.006
Ñuble	1,8%	11,9%	3.319	2.177	1.142	4.696	3.040	1.656
Biobío	1,4%	10,2%	6.909	5.571	1.338	9.800	8.032	1.768
La Araucanía	1,8%	10,2%	5.052	3.402	1.650	7.006	4.836	2.170
Los Ríos	1,6%	12,7%	3.820	2.637	1.183	5.400	3.760	1.640
Los Lagos	1,7%	15,0%	3.826	2.631	1.195	5.466	3.888	1.570
Aysén	2,0%	29,9%	1.397	1.014	383	2.070	1.430	640
Magallanes	1,7%	30,5%	1.925	1.693	232	3.194	2.810	384

Elaborado por el Instituto Nacional de Estadísticas.

Luego, a partir de las desagregaciones de tamaños por estrato geográfico (comuna-área) y de los resultados del trabajo de campo de Casen en Pandemia 2020 se obtienen, por agregación, los tamaños y errores asociados para los niveles nacional urbano y nacional rural que se muestran en la Tabla II.7.

Tabla II.7. Tamaños y errores Casen en Pandemia 2020 y propuesta Casen 2022. Nacional Urbano - Rural

Región	Resultados Casen en Pandemia 2020				Propuesta Casen 2022			
	Viviendas Logradas	Tasa de Pobreza	Error Absoluto	Error Relativo	Tamaño Objetivo	Error Absoluto	Error Relativo	Tamaño con Sobremuestreo
Nivel País	62.540	10,8%	0,4%	3,9%	71.028	0,4%	3,3%	106.856
Urbano	52.663	10,4%	0,5%	4,8%	56.905	0,5%	4,6%	87.252
Rural	9.877	13,8%	1,2%	8,7%	14.123	1,3%	9,2%	19.604

Elaborado por el Instituto Nacional de Estadísticas.

A nivel nacional urbano, con un tamaño objetivo ajustado de 56.905 viviendas se obtiene una estimación para la tasa de pobreza con un error absoluto de 0,5% y un error relativo de 4,6%.

A nivel nacional rural, con un tamaño objetivo de 14.123 viviendas se obtiene una estimación para la tasa de pobreza con un error absoluto de 1,3% y un error relativo de 9,2%.

II.5.1. Comparación de tamaños y distribuciones muestrales

La Tabla II.8 muestra los tamaños del marco muestral MMV 2020, de la muestra objetivo y de la muestra con sobremuestreo para Casen 2022, además de los tamaños de la muestra objetivo y con sobremuestreo de Casen en Pandemia 2020 a nivel nacional y regional, donde se observa que no existen grandes diferencias entre las distribuciones de Casen 2022 y Casen en Pandemia 2020.

Tabla II.8. Comparación del número de viviendas entre MMV 2020 – Casen 2022 – Casen en Pandemia 2020 a nivel nacional y regional.

Región	MMV 2020		Casen 2022				Casen en Pandemia 2020			
	N° viviendas	% Regional	Muestra Objetivo	% Regional	Muestra con Sobremuestreo	% Regional	Muestra objetivo	% Regional	Muestra con Sobremuestreo	% Regional
Nivel País	6.206.626		71.028		106.856		76.793		103.030	
Arica y Parinacota	73.783	1,2%	2.646	3,7%	3.698	3,5%	3.009	3,9%	3.811	3,7%
Tarapacá	110.680	1,8%	2.638	3,7%	4.156	3,9%	2.946	3,8%	4.332	4,2%
Antofagasta	190.638	3,1%	2.787	3,9%	4.580	4,3%	2.607	3,4%	3.993	3,9%
Atacama	106.193	1,7%	3.290	4,6%	5.060	4,7%	2.747	3,6%	3.906	3,8%
Coquimbo	285.374	4,6%	2.837	4,0%	4.180	3,9%	3.469	4,5%	4.392	4,3%
Valparaíso	677.668	10,9%	7.070	10,0%	10.984	10,3%	7.392	9,6%	10.439	10,1%
Metropolitana	2.379.740	38,3%	13.607	19,2%	22.804	21,4%	13.755	17,9%	22.663	22,0%
O'Higgins	333.576	5,4%	4.798	6,8%	6.672	6,2%	5.529	7,2%	6.716	6,5%
Maule	392.061	6,3%	5.107	7,2%	7.090	6,6%	5.924	7,7%	6.847	6,6%
Ñuble	182.744	2,9%	3.319	4,7%	4.696	4,4%	3.096	4,0%	3.696	3,6%
Biobío	558.582	9,0%	6.909	9,7%	9.800	9,2%	7.769	10,1%	9.621	9,3%
La Araucanía	359.601	5,8%	5.052	7,1%	7.006	6,6%	5.647	7,4%	6.534	6,3%
Los Ríos	144.531	2,3%	3.820	5,4%	5.400	5,1%	3.693	4,8%	4.557	4,4%
Los Lagos	308.709	5,0%	3.826	5,4%	5.466	5,1%	4.891	6,4%	5.791	5,6%
Aysén	40.838	0,7%	1.397	2,0%	2.070	1,9%	1.959	2,6%	2.630	2,6%
Magallanes	61.908	1,0%	1.925	2,7%	3.194	3,0%	2.360	3,1%	3.102	3,0%

Elaborado por el Instituto Nacional de Estadísticas.

Nota: Es importante destacar que, el tamaño objetivo expuesto en la columna correspondiente a Casen en Pandemia 2020 se refiere a la muestra planificada inicialmente considerando un operativo presencial.

La Tabla II.9 presenta la comparación, en términos de distribución porcentual, entre la muestra objetivo de Casen 2022 y la de Casen en Pandemia 2020 donde se observa que:

- Las regiones: Metropolitana, O'Higgins, Maule, La Araucanía, Los Ríos, Los Lagos, Aysén y Magallanes, presentan el mismo ordenamiento en términos de distribución a nivel regional.
- Las regiones de: Arica y Parinacota, Tarapacá, Antofagasta, Atacama, Coquimbo y Ñuble si bien, presentan un ordenamiento algo distinto en la distribución porcentual de sus tamaños muestrales, se observa que sus proporciones se encuentran acotadas en el intervalo de porcentajes [3,4% - 4,7%]. Algo similar ocurre en las regiones de Biobío y Valparaíso cuyos valores se encuentran acotados en el intervalo de porcentajes [9,6% - 10,1%].

Las diferencias se deben, principalmente, a que los tamaños que se requieren en Casen 2022 para obtener estimaciones del parámetro de interés con los errores definidos, dado el resultado del trabajo de campo de Casen en Pandemia 2020, implican tener que ajustar las distribuciones porcentuales iniciales (según MMV 2020) modificando los tamaños en algunas regiones.

Respecto de la comparación del MMV 2020 con la muestra objetivo de Casen 2022, se observa que:

- En las regiones de: Valparaíso, Metropolitana, O'Higgins, Maule, Biobío, La Araucanía, Los Lagos, Aysén y Magallanes si bien, las distribuciones porcentuales de tamaños muestrales no son exactamente iguales, se obtiene el mismo ordenamiento.
- En las regiones de: Arica y Parinacota, Tarapacá, Antofagasta, Atacama, Coquimbo, Ñuble y Los Ríos si bien, presentan un ordenamiento algo distinto, se observa que sus valores se encuentran acotados en el intervalo de porcentaje [1,2% - 5,4%].

Tabla II.9. Comparación distribución porcentual de viviendas a nivel nacional y regional. MMV 2020 – Casen 2022 – Casen en Pandemia 2020

MMV 2020		Muestra Objetivo Casen 2022		Muestra Objetivo Casen en pandemia 2020		Sobremuestreo Casen 2022		Sobremuestreo Casen en pandemia 2020	
Región	Porcentaje	Región	Porcentaje	Región	Porcentaje	Región	Porcentaje	Región	Porcentaje
Aysén	0,7%	Aysén	2,0%	Aysén	2,6%	Aysén	1,9%	Aysén	2,6%
Magallanes	1,0%	Magallanes	2,7%	Magallanes	3,1%	Magallanes	3,0%	Magallanes	3,0%
Arica y Parinacota	1,2%	Tarapacá	3,7%	Antofagasta	3,4%	Arica y Parinacota	3,5%	Ñuble	3,6%
Atacama	1,7%	Arica y Parinacota	3,7%	Atacama	3,6%	Tarapacá	3,9%	Arica y Parinacota	3,7%
Tarapacá	1,8%	Antofagasta	3,9%	Tarapacá	3,8%	Coquimbo	3,9%	Atacama	3,8%
Los Ríos	2,3%	Coquimbo	4,0%	Arica y Parinacota	3,9%	Antofagasta	4,3%	Antofagasta	3,9%
Ñuble	2,9%	Atacama	4,6%	Ñuble	4,0%	Ñuble	4,4%	Tarapacá	4,2%
Antofagasta	3,1%	Ñuble	4,7%	Coquimbo	4,5%	Atacama	4,7%	Coquimbo	4,3%
Coquimbo	4,6%	Los Ríos	5,4%	Los Ríos	4,8%	Los Ríos	5,1%	Los Ríos	4,4%
Los Lagos	5,0%	Los Lagos	5,4%	Los Lagos	6,4%	Los Lagos	5,1%	Los Lagos	5,6%
O'Higgins	5,4%	O'Higgins	6,8%	O'Higgins	7,2%	O'Higgins	6,2%	La Araucanía	6,3%
La Araucanía	5,8%	La Araucanía	7,1%	La Araucanía	7,4%	La Araucanía	6,6%	O'Higgins	6,5%
Maule	6,3%	Maule	7,2%	Maule	7,7%	Maule	6,6%	Maule	6,6%
Biobío	9,0%	Biobío	9,7%	Valparaíso	9,6%	Biobío	9,2%	Biobío	9,3%
Valparaíso	10,9%	Valparaíso	10,0%	Biobío	10,1%	Valparaíso	10,3%	Valparaíso	10,1%
Metropolitana	38,3%	Metropolitana	19,2%	Metropolitana	17,9%	Metropolitana	21,3%	Metropolitana	22,0%

Elaborado por el Instituto Nacional de Estadísticas.

II.5.2. Distribución de las unidades primarias de muestreo

En la Tabla II.10 se observa la distribución de las unidades primarias de muestreo, según región y área para el marco muestral de viviendas MMV 2020 y para la muestra de Casen 2022.

Tabla II.10. Distribución de unidades primarias de muestreo: MMV 2020 – Casen 2022

Región	Urbano		Rural	
	N° UPM MMV 2020	N° UPM muestra	N° UPM MMV 2020	N° UPM muestra
Nivel País	26.333	10.602	8.866	1.943
Arica y Parinacota	339	335	71	29
Tarapacá	514	482	61	30
Antofagasta	905	525	68	38
Atacama	464	451	145	55
Coquimbo	1.111	412	646	88
Valparaíso	3.001	1.138	676	188
Metropolitana	11.168	2.668	919	146
O'Higgins	1.200	589	955	196
Maule	1.394	635	1.172	199
Ñuble	602	380	634	164
Biobío	2.384	1.004	754	176
La Araucanía	1.225	604	1.189	217
Los Ríos	501	470	483	164
Los Lagos	1.090	485	928	157
Aysén	150	143	114	64
Magallanes	285	281	51	32

Elaborado por el Instituto Nacional de Estadísticas.

La Tabla II.11 muestra la distribución de los tamaños muestrales en número de viviendas y número de UPM a nivel nacional, regional y provincial por área.

Tabla II.11. Distribución de tamaños muestrales a nivel nacional, regional y provincial, por área: Casen en Pandemia 2020 – MMV 2020 – Casen 2022

Región/Provincia	Casen en Pandemia 2020								MMV 2020		Casen 2022							
	Muestra Objetivo				Sobremuestra				N° UPM Urbano	N° UPM Rural	Muestra Objetivo			Sobremuestra			N° UPM Muestra	
	Total	CD	RAU	Rural	Total	CD	RAU	Rural			Total	Urbana	Rural	Total	Urbana	Rural	Urbano	Rural
Nacional	76.793	58.440	4.481	13.872	103.030	81.669	5.140	16.221	26.333	8.866	71.028	56.905	14.123	106.856	87.252	19.604	10.602	1.943
Arica y Parinacota	3.009	2.776	50	183	3.811	3.537	53	221	339	71	2.646	2.412	234	3.698	3.350	348	335	29
Arica	2.929	2.776	0	153	3.726	3.537	0	189	337	65	2.600	2.396	204	3.642	3.330	312	333	26
Parinacota	80	0	50	30	85	0	53	32	2	6	46	16	30	56	20	36	2	3
Tarapacá	2.946	2.746	50	150	4.332	4.078	64	190	514	61	2.638	2.439	199	4.156	3.856	300	482	30
Del Tamarugal	238	68	50	120	282	75	64	143	27	48	278	129	149	414	184	230	23	23
Iquique	2.708	2.678	0	30	4.050	4.003	0	47	487	13	2.360	2.310	50	3.742	3.672	70	459	7
Antofagasta	2.607	2.407	50	150	3.993	3.714	75	204	905	68	2.787	2.577	210	4.580	4.200	380	525	38
Antofagasta	1.593	1.533	0	60	2.497	2.420	0	77	591	20	1.768	1.668	100	2.724	2.584	140	323	14
El Loa	811	701	50	60	1.250	1.084	75	91	263	42	834	754	80	1.592	1.392	200	174	20
Tocopilla	203	173	0	30	246	210	0	36	51	6	185	155	30	264	224	40	28	4
Atacama	2.747	2.492	0	255	3.906	3.580	0	326	464	145	3.290	2.880	410	5.060	4.510	550	451	55
Chañaral	339	309	0	30	440	401	0	39	50	14	347	307	40	530	480	50	48	5
Copiapó	1.707	1.642	0	65	2.582	2.488	0	94	310	49	2.079	1.951	128	3.220	3.050	170	305	17
Huasco	701	541	0	160	884	691	0	193	104	82	864	622	242	1.310	980	330	98	33
Coquimbo	3.469	2.535	150	784	4.392	3.303	166	923	1.111	646	2.837	2.158	679	4.180	3.296	884	412	88
Choapa	479	244	50	185	559	289	56	214	104	150	365	207	158	514	304	210	38	21
Elqui	2.152	1.858	50	244	2.844	2.497	54	293	806	218	1.808	1.559	249	2.782	2.448	334	306	33
Limarí	838	433	50	355	989	517	56	416	201	278	664	392	272	884	544	340	68	34
Valparaíso	7.392	6.200	250	942	10.439	8.880	353	1.206	3.001	676	7.070	5.777	1.293	10.984	9.104	1.880	1.138	188
Los Andes	461	341	0	120	622	462	0	160	158	77	466	306	160	684	424	260	53	26
Marga Marga	1.215	1.095	0	120	1.555	1.405	0	150	574	71	1.239	1.099	140	1.828	1.568	260	196	26
Petorca	403	125	150	128	547	147	223	177	99	108	395	220	175	596	336	260	42	26
Quillota	722	571	0	151	818	650	0	168	304	104	777	579	198	1.036	776	260	97	26
San Antonio	957	787	50	120	1.229	1.015	62	152	274	72	731	531	200	1.100	840	260	105	26
San Felipe	670	437	50	183	911	599	68	244	199	157	636	396	240	898	568	330	71	33
Valparaíso	2.964	2.844	0	120	4.757	4.602	0	155	1.393	87	2.826	2.646	180	4.842	4.592	250	574	25
Metropolitana	13.755	13.022	100	633	22.663	21.721	114	828	11.168	919	13.607	12.659	948	22.804	21.344	1.460	2.668	146
Chacabuco	361	271	0	90	462	347	0	115	318	181	486	367	119	704	504	200	63	20
Cordillera	1.135	1.075	0	60	1.495	1.401	0	94	866	77	1.092	1.002	90	1.638	1.448	190	181	19
Maipo	834	702	0	132	1.128	949	0	179	643	195	905	738	167	1.366	1.056	310	132	31
Melipilla	468	167	100	201	592	227	114	251	192	276	472	240	232	674	344	330	43	33
Santiago	10.405	10.405	0	0	18.285	18.285	0	0	8.764	36	10.009	9.869	140	17.516	17.376	140	2.172	14

Región/Provincia	Casen en Pandemia 2020								MMV 2020		Casen 2022							
	Muestra Objetiva				Sobremuestra				N° UPM Urbano	N° UPM Rural	Muestra Objetiva			Sobremuestra			N° UPM Muestra	
	Total	CD	RAU	Rural	Total	CD	RAU	Rural			Total	Urbana	Rural	Total	Urbana	Rural	Urbano	Rural
Talagante	552	402	0	150	701	512	0	189	385	154	643	443	200	906	616	290	77	29
O'Higgins	5.529	3.401	500	1.628	6.716	4.226	589	1.901	1.200	955	4.798	3.291	1.507	6.672	4.712	1.960	589	196
Cachapoal	3.555	2.495	150	910	4.309	3.104	167	1.038	903	518	3.256	2.464	792	4.498	3.488	1.010	436	101
Cardenal Caro	585	253	150	182	743	319	193	231	47	111	373	135	238	564	224	340	28	34
Colchagua	1.389	653	200	536	1.664	803	229	632	250	326	1.169	692	477	1.610	1.000	610	125	61
Maule	5.924	3.201	800	1.923	6.847	3.859	866	2.122	1.394	1.172	5.107	3.587	1.520	7.090	5.084	2.006	635	199
Cauquenes	423	234	50	139	473	261	57	155	83	76	347	223	124	470	312	158	39	15
Curicó	1.617	846	250	521	1.913	1.051	269	593	377	318	1.427	982	445	1.876	1.308	568	163	56
Linares	1.643	778	150	715	1.796	864	160	772	330	441	1.327	829	498	1.870	1.200	670	150	67
Talca	2.241	1.343	350	548	2.665	1.683	380	602	604	337	2.006	1.553	453	2.874	2.264	610	283	61
Ñuble	3.096	1.493	600	1.003	3.696	1.844	684	1.168	602	634	3.319	2.177	1.142	4.696	3.040	1.656	380	164
Diguillín	1.809	1.175	200	434	2.220	1.473	232	515	457	282	2.106	1.634	472	3.042	2.312	730	289	73
Itata	571	100	250	221	647	113	284	250	53	117	494	214	280	706	288	418	36	41
Punilla	716	218	150	348	829	258	168	403	92	235	719	329	390	948	440	508	55	50
Biobío	7.769	6.144	400	1.225	9.621	7.762	443	1.416	2.384	754	6.909	5.571	1.338	9.800	8.032	1.768	1.004	176
Arauco	866	498	100	268	1.022	588	118	316	210	161	796	499	297	1.056	656	400	82	40
Biobío	2.178	1.180	300	698	2.455	1.347	325	783	503	463	1.907	1.196	711	2.598	1.664	934	208	93
Concepción	4.725	4.466	0	259	6.144	5.827	0	317	1.671	130	4.206	3.876	330	6.146	5.712	434	714	43
La Araucanía	5.647	3.144	650	1.853	6.534	3.784	701	2.049	1.225	1.189	5.052	3.402	1.650	7.006	4.836	2.170	604	217
Cautín	4.315	2.429	450	1.436	5.081	2.999	487	1.595	954	957	3.863	2.636	1.227	5.462	3.852	1.610	481	161
Malleco	1.332	715	200	417	1.453	785	214	454	271	232	1.189	766	423	1.544	984	560	123	56
Los Ríos	3.693	2.320	150	1.223	4.557	2.994	176	1.387	501	483	3.820	2.637	1.183	5.400	3.760	1.640	470	164
Ranco	1.001	513	50	438	1.124	571	61	492	101	181	975	536	439	1.358	728	630	91	63
Valdivia	2.692	1.807	100	785	3.433	2.423	115	895	400	302	2.845	2.101	744	4.042	3.032	1.010	379	101
Los Lagos	4.891	2.925	500	1.466	5.791	3.585	541	1.665	1.090	928	3.826	2.631	1.195	5.466	3.888	1.578	485	157
Chiloé	1.198	440	250	508	1.365	523	269	573	188	316	911	472	439	1.334	756	578	94	57
Llanquihue	2.120	1.516	50	554	2.517	1.837	53	627	571	386	1.813	1.356	457	2.604	2.024	580	253	58
Osorno	1.573	969	200	404	1.909	1.225	219	465	331	226	1.102	803	299	1.528	1.108	420	138	42
Aysén	1.959	1.450	158	351	2.630	1.911	231	488	150	114	1.397	1.014	383	2.070	1.430	640	143	64
Aysén	521	391	50	80	700	516	73	111	45	26	427	339	88	580	450	130	45	13
Capitán Prat	81	0	51	30	124	0	78	46	6	7	56	16	40	90	30	60	3	6
Coyhaique	1.193	1.059	0	134	1.572	1.395	0	177	92	54	807	643	164	1.220	920	300	92	30
General Carrera	164	0	57	107	234	0	80	154	7	27	107	16	91	180	30	150	3	15
Magallanes	2.360	2.184	73	103	3.102	2.891	84	127	285	51	1.925	1.693	232	3.194	2.810	384	281	32
Antártica Chilena									3	0	16	16	0	20	20	0	2	0
Magallanes	1.930	1.887	0	43	2.599	2.541	0	58	231	28	1.478	1.366	112	2.460	2.280	180	228	15
Tierra Del Fuego	103	0	73	30	118	0	84	34	12	9	135	75	60	216	120	96	12	8
Última Esperanza	327	297	0	30	385	350	0	35	39	14	296	236	60	498	390	108	39	9

Elaborado por el Instituto Nacional de Estadísticas.

III. DISEÑO MUESTRAL

La población objetivo de la encuesta Casen la constituyen todas las personas y hogares que residen en viviendas particulares ocupadas a lo largo del territorio nacional. La cobertura de la encuesta Casen 2022 es todo el territorio nacional, excluyendo aquellas áreas identificadas por el INE como áreas especiales según la definición en el marco muestral de viviendas MMV 2020.

El diseño muestral de la encuesta Casen 2022 se define como probabilístico y estratificado, según área geográfica y por nivel socioeconómico, tanto en el área urbana como rural. La selección de la muestra se realiza en dos etapas. Las unidades primarias de muestreo (UPM) se seleccionan en forma sistemática con probabilidad proporcional al tamaño, según el número de viviendas particulares ocupadas que posee cada UPM en el marco muestral. Las unidades secundarias de muestreo, viviendas, se seleccionan de forma sistemática con igual probabilidad, al interior de cada UPM.

Al interior de la vivienda, el objetivo es identificar a todos los hogares y las personas integrantes del hogar. La entrevista se realiza al informante idóneo, que es algún integrante del hogar de 18 años o más y, a través de él, se recolectan datos de todos los integrantes que son residentes habituales de la vivienda.

III.1. CARACTERÍSTICAS DEL MARCO MUESTRAL

Un marco muestral se define como un conjunto de unidades, inequívocamente identificables donde, es posible asociarles una probabilidad de selección distinta de cero, acorde a la estrategia de muestreo empleada.

Un marco muestral completo y preciso es requisito fundamental para asegurar la calidad de los procesos de diseño, selección muestral, recolección y análisis. Así también, es necesario contar con marcos muestrales actualizados para no afectar la confiabilidad de las estimaciones y de los productos estadísticos, además de evitar que el sesgo producido por la cobertura insuficiente se traspase a la muestra.

Sin embargo, es necesario señalar que los marcos muestrales perfectos o completamente actualizados no existen y que, lo realmente importante es que estas desactualizaciones, que son un hecho con el cual conviven todas las oficinas de estadísticas, sean minimizadas.

El marco muestral utilizado para Casen 2022 es el marco muestral de viviendas MMV 2020 conformado a partir de la información del Censo 2017, con actualización al año 2020.

El MMV 2020 es completo y preciso en su información respecto de todo el territorio nacional y de las unidades presentes en él, debido a la necesidad de contar con marcos de muestreo que posibiliten mejorar la eficiencia y eficacia a nivel técnico y operativo de los estudios realizados por el INE.

En línea con lo anterior cabe señalar que, en la necesidad de contar con un marco muestral que minimiza las desactualizaciones, el INE ha contemplado realizar actualizaciones de manera anual, de forma de acotar los problemas de completitud y precisión que presentan los marcos muestrales, resguardando también, la comparabilidad y los atributos del marco actual.

El marco muestral de viviendas MMV 2020 está constituido por unidades primarias de muestreo (UPM) que corresponden a áreas geográficas homogéneas, en términos del número de viviendas que las conforman.

En la conformación de las UPM se consideran viviendas particulares con moradores presentes o ausentes, excluyendo viviendas de temporada y colectivas. Esto 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.

Las UPM presentan una estratificación geográfica, dada por la división político-administrativa, según el ordenamiento región–provincia-comuna desagregadas, según la división censal, clasificándolas en áreas urbanas y rurales y, además poseen una estratificación socioeconómica.

La conformación de las UPM en el área urbana considera un tamaño medio de 200 viviendas, con un intervalo de 20%, es decir, la cantidad de viviendas por UPM se encuentra en el rango [160 - 240] y, en el área rural, un tamaño medio de 90 viviendas, con un intervalo de 20%, quedando su cantidad de viviendas por UPM en el intervalo [70 - 110] con algunas excepciones.

Con relación a la definición de áreas urbanas y rurales una de las tareas realizadas para preparar el Precenso 2016 (insumo fundamental para la construcción del MMV 2020) fue revisar la base territorial, es decir, la información geográfica existente respecto a los límites del territorio nacional. Lo anterior buscaba segmentar el territorio y, posteriormente, distribuir las cargas de trabajo y controlar los avances del operativo además de identificar las modificaciones geográficas. Para ello, se utilizó como punto de partida la división político-administrativa del país y la división del territorio comunal en distritos y zonas censales. Además, se redefinió el Límite Urbano Censal (LUC) que establece la división entre las áreas (urbanas y rurales) para lo cual fue necesario realizar un trabajo en terreno y en oficina en cada una de las regiones del país. A partir de estos criterios se establecieron los límites que debían ser utilizados en el Precenso 2016, los cuales constituían la base para realizar el levantamiento censal.

El LUC se define como una línea imaginaria que separa el área urbana del área rural y tiene una finalidad estadística censal. Su fijación corresponde a criterios técnicos propios y particulares del INE⁴⁴ y su delimitación tiene como objetivo principal establecer la diferencia metodológica que tendrá el área urbana y rural de cada comuna.

La definición de urbano y rural viene dada por las especificaciones del Censo 2017, según la siguiente descripción:

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

III.2. COBERTURA GEOGRÁFICA DEL MARCO MUESTRAL

Una característica del MMV 2020, en línea con las propiedades deseables que deben cumplir los marcos muestrales, es su completitud geográfica donde las UPM cubren 345 comunas de las 346 que constituyen el territorio nacional, continental e insular. La comuna que queda excluida corresponde a la Antártica Chilena perteneciente a la región de Magallanes. En ese sentido, el MMV 2020 establece una diferencia importante con los marcos derivados del Censo 2002 ya que, en estos se excluyeron en su construcción, las áreas definidas por el Censo como de difícil acceso (ADA) o alto costo (ACO) y, por tanto, sus UPM no cubren todo el territorio nacional omitiendo, en algunos casos, comunas completas.

El MMV 2020 se destaca por la incorporación de áreas de tratamiento especial, en sustitución de lo que los marcos anteriores definían como ADA y ACO donde, no se excluyen a priori estas áreas, sin embargo, es posible identificarlas y caracterizarlas dentro del marco muestral. Estas áreas son llamadas “áreas geográficas de tratamiento especial” dada la complejidad en la realización de los trabajos operativos (traducido, en la mayoría de los casos, en un aumento en el costo de levantamiento).

De las 345 comunas que componen el MMV 2020, 335 son consideradas para el levantamiento de Casen 2022, quedando excluidas aquellas comunas que, luego de la evaluación operativa de las Direcciones Regionales del INE no fueron incorporadas en la Encuesta Nacional de Empleo (ENE) por tener todas sus

⁴⁴ Para mayor información revisar: <https://www.ine.cl/herramientas/portal-de-mapas/geodatos-abiertos> y [alcances-base-cartografica-censo-2017.pdf \(ine.cl\)](#).

UPM clasificadas como “Áreas Especiales” (AE). Estas comunas son: Ollagüe, Juan Fernández, Isla de Pascua, Cochamó, Chaitén, Futaleufú, Hualaihué, Palena, Guaitecas y O'Higgins.

Así también, se excluyen además UPM consideradas como áreas especiales en las comunas de General Lagos (4), Colchane (5), Tortel (1) y Cabo de Hornos (1).

En la Tabla III.1 se presenta la distribución de UPM excluidas en el marco de selección de Casen 2022, con el número de viviendas particulares que contienen (ocupadas y desocupadas) de acuerdo con la información del MMV 2020.

Tabla III.1. Distribución de UPM excluidas del marco de selección de Casen 2022

Región	Comuna	Total de UPM MMV 2020	Total de UPM excluidas en ENE	Comunas completas excluida en ENE	Total viviendas excluidas
Nivel País		139	129		15.069
Arica y Parinacota	General Lagos	5	4	NO	368
Tarapacá	Colchane	8	5	NO	479
Antofagasta	Ollagüe	2	2	SI	152
Valparaíso	Juan Fernández	5	5	SI	405
	Isla de Pascua	15	15	SI	2.760
	Cochamó	20	20	SI	1.860
	Chaitén	18	18	SI	1.969
Los Lagos	Futaleufú	8	8	SI	1.230
	Hualaihué	35	35	SI	3.973
	Palena	7	7	SI	821
	Guaitecas	4	4	SI	633
Aysén	O'Higgins	4	4	SI	312
	Tortel	4	1	NO	61
Magallanes	Cabo de Hornos	4	1	NO	46

Elaborado por el Instituto Nacional de Estadísticas.

III.3. ÁREAS ESPECIALES

En la construcción de un marco muestral, es necesario identificar áreas susceptibles de exclusión. Estas áreas suelen estar afectadas por factores que dificultan su levantamiento, que pueden ser climáticos, de transporte, costo de acceso o de acceso restringido, entre otros. Para el caso del MMV 2020, estas áreas se definieron como “Áreas Especiales”.

Para clasificar las áreas especiales, el MMV 2020 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. Estas tipologías son:

- Tiempo de trayecto
- Clima
- Altitud
- Transporte especial
- Insularidad
- Acceso Pedestre
- Estado de los caminos
- 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. Esto permite que la exclusión se lleve a cabo a nivel de UPM, en lugar de comuna, con lo que, el MMV 2020 consigue un mayor nivel de cobertura que su predecesor (marco muestral de manzanas MMM 2016 y marco muestral de secciones MMS 2016).

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

III.4. ESTRATIFICACIÓN SOCIOECONÓMICA DEL MARCO MUESTRAL MMV 2020

Entre las novedades incorporadas en el MMV 2020 se cuenta con la estratificación por nivel socioeconómico, que se realiza a nivel de UPM y se basa en un análisis de diversas variables recopiladas en el Censo 2017. De acuerdo con el documento Estratificación socioeconómica del marco muestral de viviendas 2017⁴⁵ la clasificación de las UPM según el nivel socioeconómico permite aumentar la eficiencia de la inferencia en las encuestas de hogares que utilizan este marco (Guerrero et al., 2019). En línea con esto, el principal objetivo de realizar la estratificación es inducir mayor eficiencia y precisión en las estimaciones de las estadísticas oficiales, además de servir como fuente de información para encuestas de hogares que requieran mayor análisis de carácter socioeconómico. Al estratificar se espera que las UPM dentro de un mismo estrato sean similares (disminución de la intra-varianza) y diferentes a las de otro estrato (aumento de la inter-varianza).

A partir de los análisis realizados a las variables recopiladas en el Censo y considerando diversos algoritmos multivariados de clasificación (Principal Component Analysis (PCA), K-medias, y Principal Component Analysis by means of Alternating Least Squares (PRINCALS)) se obtuvieron 180 vectores de

⁴⁵ Para más detalle sobre la estratificación socioeconómica del marco muestral de viviendas revisar: [Estratificación socioeconómica del marco muestral de viviendas 2017](#)

estratificación, que consideraban entre 3 y 5 niveles socioeconómicos, y para evaluar su desempeño se calculó el efecto de diseño de 24 indicadores, eligiendo aquella estratificación que permitiera reducir mayormente la varianza. Los indicadores seleccionados fueron:

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

Luego de la evaluación, el método utilizado para la estratificación del MMV 2020 es el de estratificación óptima de la primera componente principal del PCA, que considera los siguientes indicadores:

- Porcentaje de personas en la educación superior.
- Tasa de ocupación.
- Porcentaje de viviendas con índice de materialidad alto.
- Indicador sobre total de hijos nacidos vivos.

Para estos indicadores, según el método elegido, se obtuvo como resultado una variable de estratificación con tres categorías, denominada nivel socioeconómico (NSE) considerando, además, las distribuciones de las UPM según el cruce de comuna y área, ya que la baja prevalencia de UPM en regiones extremas, particularmente en el área rural, hace que sea más conveniente considerar solo 3 estratos.

III.5. ESTRATIFICACIÓN SOCIOECONÓMICA DEL MARCO DE SELECCIÓN PARA CASEN 2022

Las 335 comunas presentes en Casen 2022 clasificadas según la división censal en áreas urbanas y rurales dan origen a 623 estratos geográficos. Al dividir estos estratos por los niveles socioeconómicos (NSE) presentes en el MMV 2020 (bajo-medio-alto) dan origen a 1.303 estratos. Sin embargo, considerar todas estas desagregaciones en el diseño de Casen 2022 implicaría que un porcentaje importante de estratos que presentan un bajo número de UPM se vean expuestos a la pérdida de unidades por falta de respuesta, lo que, en casos extremos podría significar la pérdida completa del estrato, finalizado el proceso de recolección.

Como antecedente, en el ejercicio inicial del proceso de estratificación socioeconómica para Casen 2022, se evidenció que, al incorporar todas las desagregaciones por NSE, es decir, considerar un total de 1.303 estratos, 34,3% corresponden a estratos que contienen de 2 a 9 UPM (448 estratos) y 21,6% corresponden a estratos con 10 a 20 UPM (282 estratos). Esto motivó a revisar como referencia los resultados del estudio realizado para la ENE⁴⁶, en donde se propuso un flujo de decisión para la subdivisión de los estratos geográficos en estratos muestrales que incluyen el nivel socioeconómico, con reglas similares a las establecidas para Casen 2022, como lo son: considerar un número mínimo de UPM en el estrato y la concentración de los NSE en el estrato geográfico para incorporar esta variable como subdivisión.

En reuniones entre los equipos del Ministerio y el INE se acordó que, para que una desagregación (comuna-área-NSE) sea considerada como un estrato de muestreo para Casen 2022, debe cumplir que:

- La cantidad de UPM en dicha desagregación sea mayor o igual a 20.
- Concentre menos de 80% dentro de la desagregación comuna-área.

Al realizar el análisis se concluye que 264 desagregaciones de comuna-área no cumplen con el primer criterio, en cuyo caso se decide no desagregar por nivel socioeconómico.

Luego, para estratos por niveles socioeconómicos que concentran 80% o más del total de UPM dentro de la desagregación comuna-área, se decide colapsarlos con estratos de los niveles socioeconómicos restantes, formando un solo estrato en dicha desagregación.

Finalmente, los estratos por niveles socioeconómicos que concentran menos de 80% del total de UPM dentro de la desagregación comuna-área y, además contienen más de 20 UPM, se consideran como estratos de muestreo para Casen 2022.

Según lo anterior, el MMV 2020 ajustado para Casen 2022 queda conformado por 767 estratos de muestreo, sin embargo, el marco de selección para Casen 2022 considera solo 764 estratos, ya que 3 contienen una sola UPM y estas no poseen viviendas particulares ocupadas⁴⁷. De los 764 estratos, 9 corresponden a estratos unitarios (conformados por una sola UPM en el marco). Estos se ubican en el área rural de las comunas de General Lagos, Alto Hospicio, María Elena, Huechuraba, La Florida, Chiguayante, Lota, San Pedro de la Paz y Río Verde.

⁴⁶ Para mayor detalle revisar Documento Metodológico Encuesta Nacional de Empleo (ENE). Disponible en: [Documento metodológico Encuesta Nacional de Empleo \(ENE\)](#)

⁴⁷ Estos corresponden al área rural de las comunas de: La Reina, Las Condes y Peñalolén.

IV. TAMAÑO MUESTRAL EN MUESTRAS COMPLEJAS

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

Para cada ubicación geográfica se pueden observar comportamientos distintos en los fenómenos de estudio y en la configuración de las unidades de primera y segunda etapa de selección, incorporando mayor complejidad al diseño muestral. Esto impacta directamente en la construcción de los factores de expansión para los distintos dominios de estudio⁴⁸, ya sea en las estimaciones obtenidas con la encuesta, así como en la estimación de sus errores de muestreo. Por ende, los parámetros iniciales deben considerar los elementos antes mencionados.

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

En Casen 2022 se busca estimar un parámetro desconocido⁴⁹ de la población objetivo de modo que, esta estimación se acerque lo más posible al parámetro poblacional⁵⁰. La distancia entre el parámetro poblacional y su estimación⁵¹ puede ser usada como una medida que cuantifica cuán lejos o cerca se está del verdadero valor. Esto se puede expresar en términos probabilísticos y tomar los márgenes de error de la estimación, como el error absoluto y relativo, con cierto nivel de confianza.

En esta sección se presentan los elementos considerados para dar inicio al cálculo del tamaño muestral para Casen 2022.

Como se menciona anteriormente, la muestra ha sido construida para el logro de los objetivos propios de la encuesta, considerando factores como la variable de interés, el estimador asociado a esa variable, los niveles de estimación, los errores de muestreo tolerables y disponer de una fuente de información para obtener estimaciones que servirán para dar inicio al cálculo del tamaño muestral (encuestas anteriores).

⁴⁸ Se denomina dominios de estudios a los niveles de desagregación, generalmente geográficos, para los cuales se requieren estimaciones confiables y precisas.

⁴⁹ Para efectos de este estudio el parámetro a estimar es la tasa de pobreza por ingresos.

⁵⁰ Las estimaciones puntuales buscan dar el mejor valor estimado individual de un parámetro, en comparación con la estimación de intervalo, que procede de especificar un rango de valores (OCDE, 2007).

⁵¹ Esta distancia o diferencia absoluta es denominada error de estimación, error de muestreo o margen de error.

Según lo anterior, se realizó un proceso iterativo donde se consideraron distintas estrategias de muestreo (escenarios) y se fueron ajustando los parámetros hasta obtener un tamaño muestral adecuado a los requerimientos del estudio. En este capítulo se describen los elementos utilizados en la metodología de cálculo del tamaño muestral, pero antes se realiza una comparación entre las metodologías usadas en Casen en Pandemia 2020 y anteriores, y Casen 2022.

IV.1. COMPARACIÓN DEL TAMAÑO MUESTRAL ENTRE CASEN 2022 Y CASEN EN PANDEMIA 2020

La metodología de cálculo del tamaño muestral utilizada en Casen desde 2011 a 2020 es básicamente la misma. Se inicia con la determinación de tamaños muestrales bajo un muestreo aleatorio simple, considerando parámetros del trabajo de campo de encuestas anteriores (número de viviendas logradas, error estándar, efecto del diseño, etc.). A partir de los tamaños bajo muestreo aleatorio simple se calculan los errores asociados (absoluto y luego, el relativo) determinando si dichos errores concuerdan con los valores máximos definidos al inicio del proceso de simulación. Posteriormente, se realizan tres ajustes secuenciales: el primero, por el efecto del diseño (obtenido de una o como el promedio de dos o más encuestas), el segundo corresponde a un ajuste por finitud de la población; y el tercero, al ajuste por la tasa de no respuesta, que también puede ser obtenida desde un levantamiento previo o el promedio de más de un levantamiento.

Para Casen 2022, esta metodología se mantiene, sin embargo, como el diseño muestral ha experimentado un cambio respecto de los diseños previos, dados por la nueva estructura del marco muestral MMV 2020 (en la estratificación y en la definición de las UPM) es necesario considerar efectos de diseño acordes a la nueva estructura, tomando como referencia los efectos de diseño, a nivel regional, utilizados en la Encuesta Nacional de Empleo, y ajustándolos para obtener tamaños y errores según lo definido al inicio de la simulación.

Obtenidos los tamaños muestrales regionales se procede a distribuirlos en los estratos de muestreo, primero a nivel de comuna-área y luego, según el diseño muestral. Para el diseño de Casen en Pandemia 2020 y anteriores, se prosigue con los estratos de muestreos conformados por la comuna, área y tamaño de las manzanas en el área urbana y; para el diseño de Casen 2022, corresponde a la comuna, área y nivel socioeconómico. En ambos diseños, la distribución inicial se realiza en forma proporcional a los tamaños del marco muestral utilizado y, a partir de estos, se incorporan ciertos ajustes para cumplir con los requerimientos de tamaño definidos.

Cabe mencionar además que, la determinación de tamaños a nivel de estratos tiene relación con la cantidad de UPM resultantes según el diseño. Para Casen en Pandemia 2020 se consideran, al menos dos UPM por estrato y, para Casen 2022, depende de la disponibilidad del marco muestral: en estratos en

que el número de UPM en el marco es mayor o igual a ocho, se consideran, al menos, cuatro UPM en la muestra; en estratos en que el número de UPM está entre dos y siete, se consideran dos UPM en la muestra; y en estratos que presentan una UPM en el marco, esta UPM es incluida en la muestra.

El Cuadro IV.1 muestra los criterios utilizados para los pasos considerados en ambos diseños para determinar el tamaño muestral.

Cuadro IV.1. Comparación de criterios utilizados para el cálculo del tamaño muestral Casen 2022 – Casen en Pandemia 2020.

	Casen 2022	Casen en Pandemia 2020
Obtención de Tamaños Muestrales	Regionales, asegurando representatividad a nivel nacional, nacional urbano, nacional rural y regional.	
Efecto del Diseño	Según estudios realizados para la ENE con ajustes (mínimo: 1 y máximo: 3,2).	Según trabajo de campo de Casen 2017.
Afijación del tamaño	Por comuna y área en forma proporcional al MMV 2020 con mínimo de 16 viviendas en el área urbana y 20 en el área rural (con excepción de estratos unitarios).	Por comuna y área en forma proporcional al MM 2016 con mínimo de 30 viviendas en las ciudades y 50 en las áreas rau – rural.
Ajustes al tamaño	Se revisa distribución en función de Casen en Pandemia 2020 y de MMV 2020.	Según distribución de Casen 2017.
Criterios de Errores admisibles	A nivel regional, error absoluto máximo de 2,1% y error relativo máximo de 30,5%. Además, se revisa que el error estándar cumpla con los criterios de calidad definido por el INE.	A nivel regional, error absoluto máximo de 4% y error relativo máximo de 30%, con excepción en casos donde la tasa es menor a 5%, donde se acepta un error relativo máximo de 35%.
Tamaño con Sobremuestreo	Según promedio ponderado entre Tasas de No Logro de Casen 2017 y Casen en Pandemia 2020, con ajustes según disponibilidad de UPM en el MMV 2020 y según resultados operativos del Estudio Piloto Cambio de Marco Casen 2022.	Según promedio ponderado entre Tasas de No Logro de Casen 2015 y Casen 2017.

Elaborado por el Instituto Nacional de Estadísticas.

IV.2. ESTIMACIÓN DEL TAMAÑO MUESTRAL PARA CASEN 2022

Al igual que en versiones anteriores, el cálculo del tamaño muestral para Casen 2022 prevé una serie de procedimientos que corresponden, en principio, al paso desde un diseño bajo muestreo aleatorio simple a un diseño complejo, seguido de otros dos ajustes, determinados por la finitud de la población y la tasa de no logro⁵², obteniendo así el tamaño con sobremuestreo.

En los párrafos siguientes se describe, en detalle, cada uno de los nueve pasos realizados para definir la muestra con el total de viviendas a encuestar, en principio, a nivel de región, luego a nivel de comuna y área y finalmente, a nivel de estratos de muestreo de Casen 2022 (comuna-área-NSE).

Obtenidos los tamaños muestrales en términos de número de viviendas, para el escenario elegido, se determinó, en principio, un tamaño fijo de viviendas a encuestar por UPM según su área geográfica. En el caso del área urbana se consideró encuestar ocho viviendas por UPM y, en el área rural, diez. Cabe señalar que, a partir de estudios realizados en la Encuesta Nacional de Empleo (ENE) donde se consideraron 3 escenarios sobre el número de viviendas a encuestar por UPM, a saber, ocho, diez y doce viviendas, se determinó que la elección de estos valores no impacta en los efectos de diseño obtenidos al momento de seleccionar muestras con distintos tamaños fijos de viviendas por UPM.⁵³

La cantidad de viviendas a encuestar por UPM determinó un número fijo de UPM a seleccionar. A partir de esta cantidad y, considerando un ajuste al tamaño con sobremuestreo calculado, motivado por los resultados del trabajo de campo del Estudio Piloto Cambio de Marco Casen 2022, en algunos estratos, se decide aumentar en 2 el número de viviendas a seleccionar por UPM, tanto en el área urbana como en el área rural, quedando las UPM muestrales conformadas por 8 o 10 viviendas a encuestar en el área urbana y, por 10 o 12, viviendas a encuestar en el área rural.

La nomenclatura para describir las fórmulas que siguen, se presenta a continuación, asumiendo que el índice r corresponde a región; ce , identifica el estrato de muestreo Casen 2022; i , corresponde a la unidad primaria de muestreo (UPM); y j , a la unidad secundaria de muestreo (viviendas).

Sean:

p : Prevalencia de la variable cualitativa de interés en el estudio, corresponde a la tasa de pobreza por ingreso en la región r .

⁵² La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

⁵³ Para mayor detalle revisar Informe de simulaciones de tamaños muestrales Encuesta Nacional de Empleo (INE, 2020, pág. 27). Disponible en: [Informe de simulaciones de tamaños muestrales, ENE](#)

- n_r : Número de conglomerados o unidades primarias de muestreo (UPM) resultantes en la región r .
- \bar{m}_r : Número promedio de viviendas a encuestar por unidad primaria de muestreo resultantes en la región r .
- $n \cdot \bar{m} = m_r$: Número de viviendas a encuestar en la región r .
- M_r : Número de viviendas en el marco muestral para la región r .
- $Def f(p)_r$: Efecto del diseño asociado a la tasa de pobreza (p) en la región r . Por definición el efecto del diseño es $Def f(p) = \sigma^2(p)_{MC} / \sigma^2(p)_{MAS}$, donde $\sigma^2(p)_{MC}$ es la varianza bajo un diseño complejo y $\sigma^2(p)_{MAS}$ es la varianza bajo un muestreo aleatorio simple.
- $SE(p)_r$: Error estándar de la estimación de la pobreza (p) en la región r . Corresponde a la raíz cuadrada de la varianza de la estimación.
- $S(p)_r^2$: Cuasivarianza poblacional de la tasa de pobreza (p) en la región r . Para variables cualitativas y en un muestreo aleatorio simple, corresponde aproximadamente a $(M_r / (M_r - 1)) \times (P \times Q)$, siendo P la tasa de pobreza por ingresos regional, con $Q = (1 - P)$ y M_r el total de viviendas en la región r .
- $t^v_{1-\alpha/2}$: Cuantil de una distribución t de Student con nivel de confianza $1-\alpha$ (para efectos de la simulación se utiliza un valor fijo $T=2$).

Paso 1

Se obtienen estimaciones, a nivel regional, de la tasa de pobreza por ingresos y sus estadísticos asociados: error estándar y efecto del diseño, usando los resultados del trabajo de campo de Casen en Pandemia 2020. La cuasivarianza poblacional $S(p)_r^2$ se deriva a partir del error estándar obtenido, utilizando la siguiente fórmula:

$$S(p)_r^2 = \frac{n \cdot \bar{m}}{Def f(p)_r} \cdot SE(p)_r^2 \quad (2)$$

Paso 2

Se obtienen, a partir del error estándar, los errores absolutos por región (d_{0r}) y luego, los errores relativos (e_{0r}) según una aproximación del cuantil de una distribución t de Student con nivel de confianza de 95% que, para efectos de la simulación, tiene un valor correspondiente a $t=2$. Con esto y la cuasivarianza de la tasa de pobreza derivada en (2) se obtienen los tamaños muestrales iniciales a nivel regional, considerando un muestreo aleatorio simple monoetápico m_{0r} , como se muestra en la fórmula (3):

$$m_{0r} = \frac{t_{1-\alpha/2}^2 \cdot S(p)_r^2}{d_{0r}^2} \quad (3)$$

Paso 3

Se estima el tamaño muestral para la nueva encuesta considerando el efecto de un diseño en dos etapas. Este efecto representa el ajuste del tamaño anterior debido al cambio que experimenta la varianza al pasar de un muestreo monoetápico a uno bietápico. Los valores que se utilizan como efectos de diseño consideran, por un lado, el comportamiento, en términos de la correlación intraclásica del nuevo marco MMV 2020 y, por otro, los efectos obtenidos en el levantamiento de Casen en Pandemia 2020; definiendo una cota mínima de 1, en regiones donde el valor obtenido fue menor que 1 (Tarapacá, Aysén y Magallanes) y, cotas máximas iguales a las obtenidas en dicho levantamiento (región Metropolitana con un valor de 3,2).

La fórmula utilizada para el cálculo del tamaño muestral m_{1r} , incorporando el efecto del diseño se presenta a continuación:

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

Donde m_{0r} representa el número de viviendas estimadas a encuestar en la región r bajo un muestreo aleatorio simple y m_{1r} representa el número de viviendas a encuestar en la región r considerando el efecto de diseño, es decir, bajo un diseño complejo.

Paso 4

A partir del tamaño bajo muestreo aleatorio simple y de la cuasivarianza poblacional obtenida en el Paso 1, se obtiene el error absoluto asociado según la siguiente fórmula:

$$d_{1r} = 2 * \sqrt{S(p)_r^2 / \sqrt{m_{0r}}} \quad (5)$$

Nota: El error relativo asociado se obtiene al dividir el error absoluto por la estimación de la tasa de pobreza.

Paso 5

Luego, el tamaño m_{1r} es corregido por un ajuste para muestreo en poblaciones finitas o corrección por finitud $(1 - m_{1r}/M_r)$ obteniéndose el tamaño m_{2r} . La fórmula utilizada se presenta a continuación:

$$m_{2r} = \frac{Z_{1-\alpha/2}^2 \cdot S(p)_r^2 \cdot Deff(p)_r}{d_{0r}^2 + Z_{1-\alpha/2}^2 \cdot Deff(p)_r \cdot S(p)_r^2 / M_r} = \frac{m_{1r}}{1 + m_{1r}/M_r} \quad (6)$$

Donde m_{2r} representa el total de viviendas a encuestar en la región r considerando un diseño muestral bietápico que incorpora el efecto del diseño y el ajuste por finitud y M_r corresponde al total de viviendas de la población en la región r .

Paso 6

A partir del tamaño regional m_{2r} se determina el tamaño para cada comuna m_{2c} , obtenido en forma proporcional según la distribución comunal del marco muestral MMV 2020. Posteriormente, los tamaños m_{2c} son distribuidos en cada área (urbana y rural) bajo el mismo procedimiento.

Luego, para cada estrato geográfico comuna–área se ajusta el tamaño según el siguiente detalle: Se asigna, para el área urbana, un tamaño mínimo de muestra objetivo de 32 viviendas (4 UPM) con excepción de aquellos estratos que, en el marco presentan entre 2 y 7 UPM, donde se asigna un tamaño de 16 viviendas, lo que corresponde a muestrear 2 UPM. Para el área rural, se asigna un tamaño mínimo objetivo de 40 viviendas (4 UPM) con excepción, al igual que en el área urbana, de aquellos estratos que, en el marco presentan entre 2 y 7 UPM, donde se asigna un tamaño de 20 viviendas, lo que corresponde a muestrear 2 UPM. Tanto para el área urbana como rural, en los casos de estratos unitarios (que presentan una UPM en el marco) se selecciona dicha UPM considerando muestrear 8 viviendas objetivo en áreas urbanas y 10, en áreas rurales.

Según lo anterior, en comparación con el diseño de Casen en Pandemia 2020 y, como una forma de estabilizar la varianza de las estimaciones, se establece un número fijo de viviendas a seleccionar por UPM (en principio, 8 en el área urbana y 10 en el área rural) y se fijan tamaños mínimos en términos de número de UPM por estratos, evitando la subestimación de esta, permitiendo, a partir del diseño, fortalecer las estimaciones de áreas pequeñas que se requieran.

A partir de lo anterior, es necesario aumentar el tamaño de muestra obtenido inicialmente en forma proporcional, para algunas áreas en determinadas comunas y; disminuirlo en otras, siempre y cuando pertenezcan a la misma región, de tal forma que se respeten los tamaños de muestra regionales definidos en el paso 5 (m_{2r}).

Obtenidos los tamaños por comuna–área se distribuyen entre los estratos de muestreo (comuna–área–NSE) definidos para Casen 2022, en forma proporcional al número de viviendas que cada uno contiene en el marco.

Paso 7

Una vez definido el total de unidades a encuestar para estimar la tasa de pobreza por ingresos, se debe tener en consideración que, es posible no lograr algunas unidades de la muestra por diversas razones, tales como rechazos, moradores ausentes, entre otras. Con la finalidad de salvaguardar la precisión de la

estimación de la tasa de pobreza, se aplica un factor de corrección a la fórmula (6) con base en la estimación de la tasa de no logro considerada para Casen 2022 que, inicialmente⁵⁴ fue obtenida como un promedio ponderado entre las tasas de no logro⁵⁵ de Casen 2017 y Casen en Pandemia 2020, a nivel de comuna-área.

Una vez aplicado el factor de corrección por tasas de no logro a nivel de estratos de muestreo, se obtiene el tamaño con sobremuestreo y, a partir de este, se calcula el número de UPM a seleccionar considerando, en principio 8 viviendas por UPM en el área urbana y, 10 en el área rural.

Al utilizar la tasa de no logro como ajuste en el cálculo del tamaño muestral permite, no sólo contar con una aproximación de la cantidad de viviendas que no responderán la encuesta sino que además, considera el hecho que, en el trabajo de campo, es posible que en algunas unidades no se conozca su estado de elegibilidad, quedando clasificadas como elegibilidad desconocida o bien, que sean edificaciones que no pertenecen al universo de estudio (edificaciones de otros usos) quedando clasificadas como no elegibles. En presencia de estos casos, la precisión de la estimación del parámetro de interés no debiera verse afectado puesto que, el tamaño con sobremuestreo ya considera una pérdida de unidades por estos motivos.

La fórmula que define el tamaño muestral final para la encuesta Casen 2022, que incorpora los ajustes conjuntos de efecto del diseño, finitud y tasa de no logro para cada estrato de muestreo de Casen 2022 (comuna-área-NSE) está dada por:

$$m_{3ce} = \frac{m_{2ce}}{1 - tnl_c} \quad (7)$$

Donde

m_{3ce} : número de viviendas a encuestar en el estrato de muestreo *ce* (comuna-área-NSE).

m_{2ce} : número de viviendas objetivo en el estrato de muestreo *ce*.

tnl_c : Tasa de no logro estimada a nivel de comuna – área.

Adicionalmente, como una forma de incorporar los efectos del MMV 2020, en el cálculo del tamaño con sobremuestreo de Casen 2022, se decide revisar los resultados del trabajo de campo del Estudio Piloto Cambio de Marco Casen 2022, aplicando un aumento de tamaño en aquellos estratos que resultaron con bajas tasas de logro (estratos prioritarios). Para llevar a cabo este aumento en cada UPM ya seleccionada en los estratos prioritarios, se modifica el número de viviendas a seleccionar, considerando 2 viviendas adicionales tanto en UPM urbanas como rurales.

⁵⁴ En algunos estratos geográficos (comuna-área) fue necesario modificar las tasas obtenidas con el objetivo de que el número de UPM que contiene el marco fuera suficiente para seleccionar el número de UPM a muestrear.

⁵⁵ Se pondera por el número de viviendas elegibles dentro de cada comuna.

Según lo anterior, en los estratos prioritarios, se seleccionan 10 viviendas por UPM en áreas urbanas y, 12 en UPM de áreas rurales.

Paso 8

Obtenidos los tamaños con sobremuestreo por comuna-área-NSE se calculan por agregación, los tamaños con sobremuestreo comunales, regionales y nacional por área urbano y rural.

Paso 9

Dado que el tamaño objetivo inicial se ajusta por finitud y por consideraciones especiales en cada estrato muestral definido para Casen 2022, se obtiene un tamaño regional algo distinto del obtenido en el Paso 3, motivo por el cual, es necesario recalcular los errores de estimación asociados a estos tamaños. Este cálculo se obtiene desde la ecuación (8).

$$d_{2r} = d_{1r} \cdot \sqrt{\frac{m_{1r}}{m_{2r}}} \quad (8)$$

Donde

d_{2r} : Error absoluto para el tamaño de muestra m_{2r} .

d_{1r} : Error absoluto para el tamaño de muestra m_{1r} .

m_{1r} : Número inicial de viviendas a encuestar en la región r .

m_{2r} : Número de viviendas objetivo en la región r , que incorpora ajuste por finitud y ajustes por estrato Casen 2022 según consideraciones especiales definidas en párrafos anteriores.

De forma análoga se estima el error relativo.

Obtenidos los errores se calcula el error estándar a partir del cociente entre el error absoluto y el cuantil de una distribución t de Student con nivel de confianza $1-\alpha$ (para efectos de la simulación se utiliza un valor fijo $T=2$). Este valor se compara con la función cuadrática definida en el documento Fundamentos del Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares⁵⁶, dada por la ecuación (9):

$$\text{máximo error estándar tolerable} = \begin{cases} \frac{\sqrt[3]{p^2}}{9} & , \text{ si } 0 < p \leq 0,50 \\ \frac{\sqrt[3]{(1-p)^2}}{9} & , \text{ si } 0,50 < p < 1 \end{cases} \quad (9)$$

Donde, p Corresponde a la Tasa de Pobreza.

⁵⁶ Para mayor información revisar: [Fundamentos del Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares](#)

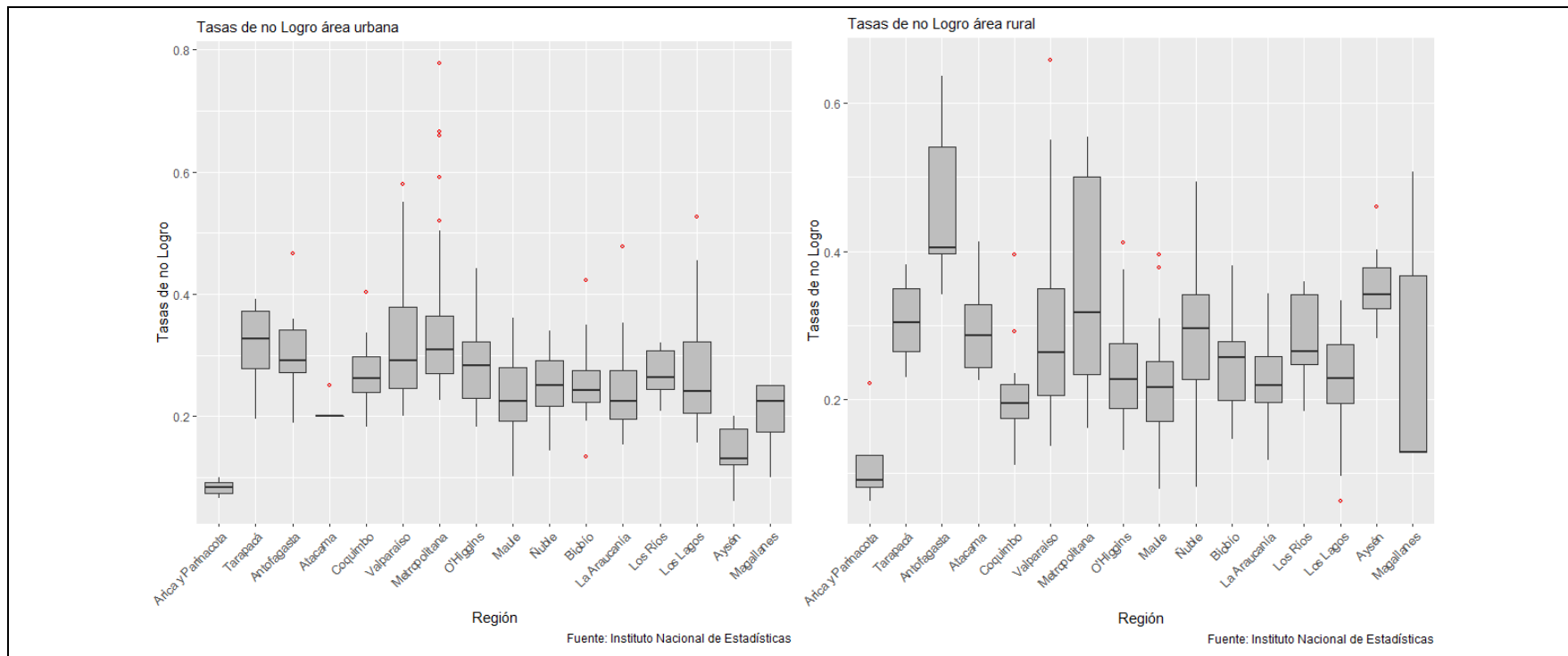
La Tabla IV.1 muestra el detalle de la obtención de los tamaños muestrales regionales para Casen 2022 en función del trabajo de campo de Casen en Pandemia 2020.

Tabla IV.1. Tamaños muestrales de viviendas para Casen 2022, según parámetros de Casen en Pandemia 2020 y efecto del MMV 2020 a nivel nacional y regional

Región	Viviendas MMV 2020	Resultados Casen en Pandemia 2020							Propuesta Casen 2022									
		Viviendas Efectivas	Tasa de Pobreza	Error Estándar	Efecto Diseño	Cuasi-Varianza Deff.Ajustado	Error Absoluto	Error Relativo	Tamaño M.A.S.	Efecto Diseño	Tamaños efecto de diseño	Tamaños finitud	Tamaño Objetivo	Error Absoluto	Error Relativo	Tamaño con Sobremuestreo	Error Estándar	Máximo EE Tolerable
Nivel País	6.206.626	62.540	10,8%	0,2%	2,2	0,13	0,4%	4,0%	39.705	2,2	68.034	67.032	71.028	0,4%	3,3%	106.856		
Arica y Parinacota	73.783	2.282	11,9%	0,9%	0,5	0,19	1,8%	15,5%	2.281	1,2	2.737	2.640	2.646	1,9%	15,8%	3.698	0,009	0,027
Tarapacá	110.680	2.621	14,0%	0,8%	0,5	0,19	1,7%	12,1%	2.621	1,0	2.621	2.561	2.638	1,7%	12,1%	4.156	0,008	0,030
Antofagasta	190.638	2.368	9,3%	0,9%	1,5	0,13	1,8%	19,7%	1.579	1,7	2.684	2.647	2.787	1,8%	19,3%	4.580	0,009	0,023
Atacama	106.193	2.402	9,5%	0,9%	0,7	0,19	1,8%	18,9%	2.402	1,4	3.363	3.260	3.290	1,8%	19,1%	5.060	0,009	0,023
Coquimbo	285.374	2.695	11,7%	1,1%	2,2	0,14	2,1%	18,3%	1.225	2,2	2.695	2.670	2.837	2,1%	17,8%	4.180	0,010	0,027
Valparaíso	677.668	6.326	11,3%	0,6%	1,9	0,13	1,3%	11,0%	3.329	1,9	6.325	6.267	7.070	1,2%	10,4%	10.984	0,006	0,026
Metropolitana	2.379.740	13.122	9,0%	0,4%	3,2	0,06	0,7%	8,3%	4.101	3,2	13.123	13.052	13.607	0,7%	8,2%	22.804	0,004	0,022
O'Higgins	333.576	4.207	10,0%	0,8%	1,7	0,16	1,6%	16,0%	2.475	1,8	4.455	4.397	4.798	1,5%	15,4%	6.672	0,008	0,024
Maule	392.061	4.314	12,3%	0,7%	1,5	0,16	1,5%	12,1%	2.876	1,7	4.889	4.829	5.107	1,5%	11,8%	7.090	0,007	0,027
Ñuble	182.744	2.268	14,7%	0,9%	0,8	0,18	1,8%	12,2%	2.268	1,4	3.175	3.121	3.319	1,8%	11,9%	4.696	0,009	0,031
Biobío	558.582	6.009	13,2%	0,7%	1,7	0,18	1,4%	10,7%	3.535	1,8	6.363	6.292	6.909	1,4%	10,2%	9.800	0,007	0,029
La Araucanía	359.601	4.003	17,4%	0,9%	1,4	0,23	1,8%	10,4%	2.859	1,7	4.860	4.796	5.052	1,8%	10,2%	7.006	0,009	0,035
Los Ríos	144.531	2.791	12,2%	0,8%	0,6	0,16	1,5%	12,6%	2.791	1,4	3.907	3.805	3.820	1,6%	12,7%	5.400	0,008	0,027
Los Lagos	308.709	3.580	11,3%	0,9%	1,7	0,16	1,8%	15,5%	2.106	1,7	3.580	3.539	3.826	1,7%	15,0%	5.466	0,008	0,026
Aysén	40.838	1.694	6,6%	0,9%	0,4	0,14	1,8%	27,2%	1.400	1,0	1.400	1.353	1.397	2,0%	29,9%	2.070	0,010	0,018
Magallanes	61.908	1.858	5,7%	0,9%	0,7	0,14	1,8%	31,0%	1.857	1,0	1.857	1.803	1.925	1,7%	30,5%	3.194	0,009	0,016

Elaborado por el Instituto Nacional de Estadísticas.

Gráfico IV.1. Distribución de las tasas de no logro comunales utilizadas en Casen 2022, según región



Elaborado por el Instituto Nacional de Estadísticas.

La Tabla IV.2 muestra los tamaños objetivos y la distribución porcentual a nivel nacional y regional de Casen 2011 – 2013 – 2015 – 2017 – en Pandemia 2020 – 2022

Tabla IV.2. Comparación tamaños objetivos y distribución porcentual de Casen desde 2011 hasta 2022 a nivel nacional y regional

Región	Casen 2011		Casen 2013		Casen 2015		Casen 2017		Casen en Pandemia 2020		Casen 2022	
	Muestra	%	Muestra	%	Muestra	%	Muestra	%	Muestra	%	Muestra	%
Nivel País	90.122	100	70.080	100	82.370	100	69.816	100	76.793	100	71.028	100
Arica y Parinacota	3.495	3,9	2.581	3,7	838	1,0	2408	3,4	3.009	4,3	2.646	3,7
Tarapacá	6.241	6,9	2.977	4,2	2.428	2,9	2.974	4,3	2.946	4,2	2.638	3,7
Antofagasta	6.108	6,8	2.513	3,6	1.965	2,4	2.511	3,6	2.607	3,7	2.787	3,9
Atacama	4.382	4,9	2.071	3,0	3.861	4,7	2.331	3,3	2.747	3,9	3.290	4,6
Coquimbo	3.780	4,2	3.018	4,3	3.738	4,5	3.028	4,3	3.469	5,0	2.837	4,0
Valparaíso	8.064	8,9	6.737	9,6	8.847	11,0	6.717	9,6	7.392	10,6	7.070	10,0
Metropolitana	11.461	13,0	12.865	18,0	17.304	21,0	12.954	18,6	13.755	19,7	13.607	19,2
O'Higgins	5.469	6,1	5.137	7,3	7.144	8,7	5.099	7,3	5.529	7,9	4.798	6,8
Maule	7.464	8,3	4.872	7,0	5.678	6,9	5.007	7,2	5.924	8,5	5.107	7,2
Ñuble	-	-	-	-	-	-	2.834	4,1	3.096	4,4	3.319	4,7
Biobío	8.652	9,6	9.826	14,0	11.402	14,0	6.901	9,9	7.769	11,1	6.909	9,7
La Araucanía	5.782	6,4	5.387	7,7	6.846	8,3	5.136	7,4	5.647	8,1	5.052	7,1
Los Ríos	6.576	7,3	3.768	5,4	3.318	4,0	3.624	5,2	3.693	5,3	3.820	5,4
Los Lagos	6.391	7,1	4.192	6,0	6.097	7,4	4.129	5,9	4.891	7,0	3.826	5,4
Aysén	3.941	4,4	2.051	2,9	1.134	1,4	1.862	2,7	1.959	2,8	1.397	2,0
Magallanes	2.315	2,6	2.085	3,0	1.770	2,1	2.301	3,3	2.360	3,4	1.925	2,7

Elaborado por el Instituto Nacional de Estadísticas.

IV.3. SELECCIÓN DE CONGLOMERADOS (O UNIDADES PRIMARIAS DE MUESTREO)

Como se mencionó en la sección II.1, el INE mantiene un marco de áreas que contiene un listado de unidades (que se denominan UPM o unidades primarias de muestreo) que sirven de base para la identificación y posterior selección de las viviendas, que son las unidades últimas de selección de la muestra Casen 2022 y que, a través de éstas, es posible acceder a los hogares y personas que constituyen el objeto final del estudio.

Las UPM corresponden a áreas geográficas homogéneas, en términos del número de viviendas particulares que las conforman excluyendo, viviendas de temporada y colectivas. Presentan una estratificación geográfica, dada por la división político-administrativa, según región–provincia-comuna y desagregada, además según la división censal, clasificándolas en áreas urbanas y rurales. A esta estratificación se suma la estratificación socioeconómica (NSE) adecuada especialmente para esta encuesta.

Para la selección de viviendas o unidades secundarias de muestreo (USM) se considera una actualización de las UPM seleccionadas en primera etapa, cuyo objetivo es verificar las direcciones ya existentes o añadir direcciones correspondientes a nuevas edificaciones.

Previo a la selección de UPM se determina la cantidad de viviendas a seleccionar en cada estrato de muestreo Casen 2022, conformado por comuna–área–NSE. Para obtener esta cantidad, a partir del tamaño objetivo calculado a nivel regional, se consideran las siguientes etapas:

- **Distribución por comuna:** La cantidad de viviendas a nivel regional es distribuida en forma proporcional al MMV 2020 en todas las comunas que conforman la muestra Casen 2022.
- **Distribución por área urbano–rural:** Obtenido el tamaño a nivel comunal, se realiza la desagregación urbano-rural, en forma proporcional a la distribución por área del MMV 2020.
- **Ajustes:** Se fijan tamaños mínimos por cada combinación comuna–área, en términos de viviendas y considerando, además un número mínimo de UPM a encuestar según la cantidad de estas en el MMV 2020.
- **Distribución por nivel socioeconómico (NSE):** Obtenido el tamaño ajustado se distribuye entre los estratos de muestreo (comuna–área–NSE) conformados para Casen 2022 y en forma proporcional al MMV 2020.

- Ajuste por Tasa de No Logro⁵⁷: Al tamaño objetivo resultante (a nivel de comuna–área–NSE) se le aplican tasas de no logro (TNL) a partir de los levantamientos de Casen 2017 y Casen en Pandemia 2020 a nivel de comuna–área, con algunos ajustes; obteniendo los tamaños con sobremuestreo según el escenario N°5.
- Ajustes a tamaños con sobremuestreo: Obtenidos los tamaños con sobremuestreo, se debe calcular la cantidad de UPM a seleccionar. Para esto, se considera que, en estratos urbanos se seleccionan 8 viviendas por UPM y; en estratos rurales, 10. Por tanto, los tamaños con sobremuestreo se deben redondear a múltiplos de 8 o 10, respectivamente. Según lo anterior y se obtiene una muestra de 106.856 viviendas a nivel nacional cuya desagregación por área corresponde a: 87.252 viviendas urbanas y 19.604 viviendas rurales.
- Determinación del número de UPM: A partir de los tamaños a nivel de estrato de muestreo se calcula la cantidad de UPM a seleccionar que, a nivel nacional, alcanza a 12.545 UPM, desagregado en 10.602 UPM en el área urbana y 1.943 UPM en el área rural.
- Una vez determinado el número de UPM a seleccionar por cada estrato de muestreo, se procede a la selección de forma sistemática y con probabilidad proporcional al tamaño, según el número de viviendas que las conforman, luego del siguiente ordenamiento: región – comuna – área – nivel socioeconómico (NSE Casen 2022) - Estratificación Socioeconómica (bajo-medio-alto según MMV 2020) – vector_Orden (ordenamiento de la UPM por comuna y área).

La selección sistemática en forma proporcional al tamaño se realiza mediante un algoritmo de acumulación de viviendas en el estrato de muestreo donde se elige un período y arranque aleatorio, como sigue:

Paso 1

Se construyen Q intervalos (equivalente al número de UPM en el estrato de muestreo) como se muestra en el Cuadro IV.2.

⁵⁷ La tasa de no logro se obtiene sobre el total de la muestra con sobremuestreo.

Cuadro IV.2. Método de selección de unidades primarias de muestreo, Casen 2022

UPM	N° Viviendas	Acumulado del número de viviendas	Límite inferior	Límite superior
i	M_i	ΣM_i	L_{inf}	L_{sup}
1	M_1	$\Sigma_1 = M_1$	1	Σ_1
2	M_2	$\Sigma_2 = M_1 + M_2$	$\Sigma_1 + 1$	Σ_2
3	M_3	$\Sigma_3 = M_1 + M_2 + M_3$	$\Sigma_2 + 1$	Σ_3
⋮	⋮	⋮	⋮	⋮
Q-1	M_{Q-1}	$\Sigma_{Q-1} = M_1 + M_2 + \dots + M_{Q-1}$	$\Sigma_{Q-2} + 1$	Σ_{Q-1}
Q	M_Q	$\Sigma_Q = M_1 + M_2 + \dots + M_Q$	$\Sigma_{Q-1} + 1$	Σ_Q

Elaborado por el Instituto Nacional de Estadísticas.

Paso 2

Se calcula el período (K) a partir de la siguiente fórmula: $K = SQ / n$, donde SQ corresponde al total de viviendas en el estrato de muestreo y n al número de UPM a seleccionar (se supone $n \leq Q$, es decir, que el número de UPM a seleccionar es menor o igual al número de UPM en el marco muestral).

Paso 3

Se genera un número aleatorio⁵⁸ entero A entre 1 y el período K .

Paso 4

Enseguida, se va sumando sucesivamente el período K para obtener distintos valores que pertenecerán a los distintos intervalos que indican la UPM seleccionada. De acuerdo con este algoritmo, las UPM seleccionadas son aquellas cuyo intervalo asociado contiene a los valores dados por: $A, A + K, A + 2K, A + 3K, \dots, A + (n - 1)K$.

El procedimiento anterior para UPM homogéneas en relación con el número de viviendas que las conforman permite que, al momento de ser seleccionada una UPM, ésta no vuelva a salir en la muestra. Por el contrario, cuando las UPM son heterogéneas, en relación con el número de viviendas, llegando a ser, por ejemplo, algunas el doble de otras, el procedimiento o algoritmo selecciona menos de las requeridas porque esencialmente, puede seleccionar dos o más veces una misma UPM.

Finalmente, se puede demostrar que para UPM homogéneas, la probabilidad de inclusión de la i -ésima UPM, que contiene M_i viviendas está dada por $n \times M_i / SQ$.

⁵⁸ Para la generación de este número aleatorio, se puede introducir una semilla con el objetivo de conservar la selección de las UPM en caso de que se replique el procedimiento.

IV.4. SELECCIÓN DE VIVIENDAS DE LA MUESTRA

Como se mencionó en capítulos precedentes, la selección de viviendas ocurre luego de la selección de UPM donde, para llevar a cabo una selección certera de viviendas, es necesario contar con UPM actualizadas en términos de: el número de viviendas particulares ocupadas que las componen y la ubicación exacta de estas, a partir de la cartografía asociada.

A partir de los resultados del Precenso realizado durante el año 2016 y del Censo de población y viviendas del año 2017, se dispone de información sobre el total de viviendas existentes en cada UPM presente en el marco muestral, junto con su cartografía asociada que incluye las direcciones y el número que identifica cada vivienda.

En Casen 2022, al igual que en Casen en Pandemia 2020, una vez seleccionada la UPM, es necesario actualizarla, ya sea verificando las direcciones de las viviendas que las componen o, añadiendo nuevas direcciones, provenientes de edificaciones construidas posteriores al Censo 2017⁵⁹.

Una vez determinadas las UPM definitivas de la muestra, se procede a la selección de las viviendas que se realiza al interior de cada UPM seleccionada y actualizada previamente, en forma sistemática y con igual probabilidad. Específicamente, al algoritmo de selección sistemático de viviendas, se introduce una semilla de aleatorización al igual que para la selección de UPM para que, independientemente de quien aplique el mismo algoritmo, bajo las mismas condiciones de ordenamiento previo⁶⁰, seleccione las mismas viviendas.

A partir del escenario N°5 se determinó que, se seleccionan 8 viviendas por UPM en el área urbana y 10 en el área rural. Sin embargo, como una forma de incluir en el diseño el posible comportamiento del MMV 2020 en el logro de la encuesta, se incorporan los resultados del trabajo de campo del Estudio Piloto Cambio de Marco Casen 2022, identificando aquellos estratos que resultaron con bajas tasas de logro. En ellos se decide seleccionar 2 viviendas adicionales por UPM, tanto en el área urbana como en el área rural, dando origen al escenario N°5 que, bajo el mismo tamaño objetivo del escenario N°4 presenta un tamaño con sobremuestreo de 106.856 viviendas, desagregadas en 87.252 viviendas urbanas y 19.604 viviendas rurales.

Dado que la selección de UPM de Casen 2022 incluye algunas UPM que han sido seleccionadas en otras encuestas encargadas al INE, ya sean continuas o bien, que impliquen un solo levantamiento. Algunas de

⁵⁹ El proceso de enumeración/verificación se detalla en capítulos posteriores.

⁶⁰ Las UPM son ordenadas según el recorrido realizado en el proceso de enumeración, que considera cada una de las manzanas que las conforman. En cada manzana se identifican primero las edificaciones (generando la variable "orden_edificación") y luego, las viviendas que las conforman (generando la variable "orden_vivienda"). Por lo tanto, para la selección de viviendas se considera el siguiente ordenamiento de la UPM: manzana – orden_edificación – orden_vivienda.

estas UPM coinciden en el periodo de trabajo de campo con Casen 2022 y en otras, no ha pasado suficiente tiempo desde la última vez que fueron levantadas. Por este motivo, es necesario realizar un ajuste al directorio de viviendas, previo a la selección que le otorgue probabilidad igual a 0 a aquellas que ya han sido seleccionadas en otras encuestas lo que se traduce en una reducción del stock de viviendas disponibles para la selección como se muestra en la Tabla IV.3 que da cuenta de que las 26.274 viviendas excluidas corresponden a un 1,1% del total de viviendas del directorio.

Tabla IV.3. Número de viviendas en el MMV 2020 y número de viviendas en el Marco de selección para Casen 2022 a nivel regional.

	Directorio completo	Directorio ajustado para Casen 2022	Viviendas excluidas	Porcentaje de viviendas excluidas
Total general	2.404.017	2.377.743	26.274	1,1%
Arica y Parinacota	66.684	65.482	1.202	1,8%
Tarapacá	94.483	93.273	1.210	1,3%
Antofagasta	111.460	110.129	1.331	1,2%
Atacama	94.429	92.912	1.517	1,6%
Coquimbo	93.759	92.539	1.220	1,3%
Valparaíso	251.291	248.662	2.629	1,0%
Metropolitana	595.762	591.581	4.181	0,7%
O'Higgins	143.429	142.034	1.395	1,0%
Maule	152.123	150.601	1.522	1,0%
Ñuble	98.917	97.693	1.224	1,2%
Biobío	229.822	227.335	2.487	1,1%
La Araucanía	148.138	146.706	1.432	1,0%
Los Ríos	111.435	109.596	1.839	1,7%
Los Lagos	117.374	115.887	1.487	1,3%
Aysén	34.277	33.635	642	1,9%
Magallanes	60.634	59.678	956	1,6%

Elaborado por el Instituto Nacional de Estadísticas.

Una vez realizado el ajuste al directorio, se realiza la selección y para ello sea M_i el número de viviendas que posee la UPM seleccionada y m_i el número de viviendas a seleccionar. Para la selección de m_i viviendas se siguen los siguientes pasos:

Paso 1

Se calcula el período $K = M_i / m_i$.

Paso 2

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

Paso 3

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

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

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 $A + (m_i - 1)K$.

V. VERIFICACIÓN Y ACTUALIZACIÓN DE UNIDADES PRIMARIAS DE MUESTREO

V.1. DISEÑO DEL PROCESO DE ENUMERACIÓN

Considerando que la encuesta Casen requiere disponer de estimaciones precisas y confiables que permitan representar de manera adecuada a la población, es que el proceso de selección de viviendas a encuestar se realiza sobre la base de información completa, precisa y actualizada de cada unidad primaria de muestreo (UPM). El resultado del proceso de selección de unidades de segunda etapa es un directorio que contiene un listado de direcciones de viviendas particulares ocupadas, siendo imperativo que dicho directorio contenga los datos necesarios para identificar inequívocamente la vivienda seleccionada en terreno al momento de la encuesta. En este sentido, la cantidad, identidad y ubicación georreferenciada de dichas viviendas, serán los aspectos fundamentales en el proceso de generación de dichos directorios.

El marco de selección se genera a partir del marco muestral de viviendas 2020 (MMV 2020) que contiene la división territorial en áreas urbanas y rurales para todas las unidades primarias de muestreo (UPM) a nivel nacional. Sobre dicha selección y previo a la selección de viviendas se generan los directorios de las UPM, a través de la actualización o verificación de gabinete, detectando mediante el método de priorización, aquellos casos donde el directorio de viviendas institucional vigente no cumple con la coherencia o completitud necesaria y requiere de una actualización in situ. En relación con lo anterior, se ha definido que todas aquellas UPM que no cumplan con las exigencias descritas, se resolverán a través del trabajo de campo, cumpliendo protocolos y procedimientos definidos por el proceso de enumeración de terreno, recolectando y actualizando la totalidad de viviendas y otros usos contenidos en cada área seleccionada.

Para llevar a cabo este proceso de recolección de información, se utilizan dispositivos móviles de captura (DMC) que contienen el formulario de enumeración, tanto para áreas urbanas y rurales. Esta recolección se define sobre el 40% de la selección de la muestra de UPM a nivel nacional que presenta dinámica territorial, siendo el 60% del directorio validado y actualizado mediante diversas fuentes de información propias de la actualización cartográfica continua (ACON), que pueden ser tanto internas como externas y que, están disponibles para el proceso de revisión y actualización en gabinete.

V.1.1. Método de Priorización

El INE, en lo que respecta a la actualización cartográfica, dispone de un proceso de actualización permanente de la cartografía (ACON) que permite disminuir la brecha de desactualización; por otra parte, se asume un menor nivel de desactualización intercensal, dado al tiempo transcurrido desde el

último censo (2017) que determina que las modificaciones cartográficas son más bien parciales y focalizadas; la información que se dispone se encuentra en general desagregada y con mayor detalle en relación con periodos censales anteriores, donde además, se disponen de Registros Administrativos (RRAA⁶¹) con una amplia cobertura y georreferencia; como también el uso e implementación de nuevas tecnologías de información geográfica que permiten una mejor gestión, captura y análisis de los datos geográficos (Plataforma Geoestadística).

Conforme a lo anterior, se plantea para los requerimientos de la muestra Casen una metodología de actualización cartográfica basada en el actual proceso de actualización continua (ACON) a partir del cual se puedan identificar y focalizar las áreas con mayor dinámica territorial y/o donde no se tenga suficiente información planimétrica, de registro de viviendas y direcciones según los insumos disponibles; por lo que estas áreas se deben actualizar y validar necesariamente en terreno. Esta modalidad se traduce en una mayor eficiencia en el uso de los recursos disponibles para la ejecución de este operativo, evitando recorrer todo el territorio para su actualización.

La priorización aplicada corresponderá al análisis realizado sobre las áreas de selección de la muestra, basados en la necesidad de identificar las UPM que presenten cambios o dinamismo de la información y que están asociadas a las variaciones o cambios de estructura de la(s) manzana(s), cantidad de viviendas contenidas en ellas, uso u ocupación de cada edificación, así como también, permitirán la identificación de aquellas que se mantienen estables en el tiempo o que requieren un proceso de actualización de sus directorios en base a información disponible a través de RRAA, siendo el objetivo de la priorización, dirigir los recursos disponibles a las áreas que realmente requieren un proceso de actualización en terreno.

Es importante indicar que, por las características de la conformación de las UPM, entendiéndolas como conglomerados homogéneos de viviendas, la priorización es generada en la totalidad de las áreas, tanto urbanas como rurales, no descartando dentro del proceso de análisis ninguna manzana o entidad dentro de la selección. Los resultados de la priorización definirán las áreas que requerirán un proceso de enumeración en terreno correspondiente al 40% de la muestra, como también, el proceso de análisis y actualización en gabinete de directorios vigentes disponibles, para el 60% restante de la selección.

El método de priorización incluye por una parte, aspectos cartográficos, incorporando variables de interés para áreas urbanas y rurales, que se cruzan con las áreas de las UPM seleccionadas y que buscan a través de la identificación, la presencia o no de dichas variables de interés, permitiendo evaluar

⁶¹ Corresponderán a direcciones del “Marco georreferenciado de Edificaciones”, que consiste en la vinculación geográfica de los catastros de direcciones georreferenciadas del Ministerio de Vivienda y Urbanismo (MINVU) con los directorios de Precenso 2016, los que se complementan con la base cartográfica de actualización continua (ACON), como también listados de direcciones de los Certificados de Recepción Final y también la disponibilidad de registros de direcciones de la Superintendencia de Servicios Sanitario (SISS).

potenciales cambios en los directorios disponibles y sus direcciones, con base en el total de viviendas y cambios en las estructuras de manzanas y entidades que conforman cada UPM; por otra parte, se encuentran las de orden metodológico, enfocado directamente a identificar la disponibilidad de directorios en completitud para áreas seleccionadas, como también de viviendas particulares ocupadas suficientes que permitan la selección de viviendas en una segunda etapa.

Con base en este análisis, también se incorpora el conocimiento regional, donde los equipos que permanentemente salen a terreno disponen de información actualizada del territorio, que puede incluir conocimiento de cambios estructurales dentro de sus áreas geográficas, como, por ejemplo: la instalación de campamentos, erradicaciones o cambios en estructuras viales, como también de situaciones contingentes que puedan incidir en la priorización.

Entre otros aspectos considerados para el análisis de la priorización están las observaciones derivadas de la construcción del marco muestral, donde se integran a la priorización todas aquellas UPM que puedan contener viviendas con ubicaciones tentativas derivadas del Censo 2017.

Finalmente es importante indicar, que dentro de la priorización enfocada al terreno, se incluirán en la enumeración, todas aquellas UPM que tendrán presente variables cartográficas, metodológicas, regionales y otras asociadas a la construcción del marco⁶², que conforman un 40% aproximado de la muestra, descartando aquellas UPM que poseen directorios actualizados en otros procesos de enumeración en terreno realizados por el INE post Censo 2017 (enumeraciones de UPM de las siguientes encuestas: Encuesta Nacional de Empleo (ENE) 2020-2022; Encuesta de Presupuestos Familiares (EPF) 2021; Estudio Piloto Cambio de Marco Casen 2022, pasando directamente dichos directorios al análisis y actualización de gabinete del 60%.

En la Tabla V.1 se indica el número de UPM según áreas urbanas y rurales, priorizadas para la enumeración de terreno, según el marco de selección Casen 2022.

⁶² Para más información revisar apartado IV.1.3 Priorización de UPM a enumerar.

Tabla V.1. Distribución del número de UPM priorizadas según área urbano y rural, Casen 2022

Región	Urbana	Rural	Total UPM
Nivel País	4.393	794	5.187
Arica y Parinacota	167	13	180
Tarapacá	208	21	229
Antofagasta	199	26	225
Atacama	181	22	203
Coquimbo	152	49	201
Valparaíso	455	75	530
Metropolitana	1.076	54	1.130
O'Higgins	221	94	315
Maule	282	73	355
Ñuble	162	56	218
Biobío	414	57	471
Los Ríos	176	78	254
La Araucanía	291	71	362
Los Lagos	195	62	257
Aysén	96	30	126
Magallanes	118	13	131

Elaborado por el Instituto Nacional de Estadísticas.

V.1.2. Proceso de Enumeración

La enumeración es una operación de registro de viviendas en un área geográfica específica, que tiene como finalidad identificar el universo en el cual se realizará la selección de viviendas de segunda etapa. Para realizar esta tarea, cada enumerador debe cumplir con las etapas y objetivos del referido estudio.

El proceso de la enumeración en terreno para la muestra Casen se realiza través de la utilización de un dispositivo móvil de captura (DMC) tanto para áreas urbanas y rurales, que contempla 3 fases de ejecución, que buscan particularmente asegurar la calidad de la información, correspondiendo a la recolección, supervisión y análisis de datos enumerados. Su objetivo principal es generar directorios actualizados de manera de permitir localizar las viviendas particulares ocupadas y sus informantes al momento del levantamiento de la encuesta.

La recolección de datos en terreno es efectuada por el enumerador de datos e implica un recorrido previo a la unidad asignada, verificando los tipos de vía y nombre de calles con la cartografía dispuesta en los DMC; posteriormente, se debe establecer contacto con las personas para determinar el uso de las edificaciones y la situación de ocupación de las viviendas.

Una vez terminada la enumeración en terreno, es necesario que el recolector revise en gabinete la información registrada, de manera tal, de corroborar la completitud y coherencia de los datos recabados. Al concluir la revisión se envían los datos a través del DMC (sincronización) e informa al

supervisor, quien a su vez coordina y asigna cargas de trabajo y establece un control de calidad de la información a través de la validación de datos.

Dentro del proceso de enumeración se definen los siguientes objetivos:

- Clasificar la carga a enumerar en terreno: La carga a verificar en el trabajo de campo, para los equipos de terreno, responde a una organización que considera diversos factores entre ellos, la muestra y el tiempo disponible para actualizar la información, la cantidad de personas que compone el equipo de levantamiento, además de las características geográficas y sociales de cada una de las UPM, que responde a una clasificación donde es posible encontrar lugares de difícil acceso, o sectores peligrosos, así como también la identificación de áreas con crecimientos formales e informales disponibles en el proceso de actualización ACON, permitiendo identificar dónde y cuánta carga se les entregará a quienes componen los equipos de terreno.
- Evitar omisión de viviendas: Corresponde a disponer de la cobertura del 100% del área de la UPM que permita recolectar viviendas y otros usos para la conformación del directorio. En cuanto a la omisión de viviendas, en el trabajo de levantamiento con DMC los equipos de supervisión verifican, entre otras cosas, con la base cartográfica e imagen satelital disponibles en las plataformas geográficas, que las edificaciones tengan puntos asociados. En caso contrario, de encontrar áreas sin registros, el equipo de supervisión solicita la verificación para asegurar que no hubiese ocurrido una omisión de viviendas.
- Aprobar en gabinete las UPM que cumplan con la identificación y ubicación de cada vivienda contenida en los registros: Previo al proceso de Análisis y una vez que se sincroniza la recolección de terreno de las UPM, pasa un primer filtro, que corresponde a una revisión en gabinete por parte del supervisor. En este sentido, uno de los puntos claves, es que las viviendas no se encuentren asociadas a información ambigua o que exista omisión de información, que dificulte y/o impida ubicar las viviendas en terreno. Es decir, todos aquellos casos que no presenten un número de domicilio o posean una numeración irregular, además de aquellos donde se haya rescatado el número domiciliario con el informante, pero este no se encuentra visible, o bien, donde exista más de una vivienda dentro del sitio. Si estos casos no presentan una observación respecto de su ubicación, se solicita una nueva visita a terreno para consolidar la información en la calidad requerida. Cabe destacar que esto es nuevamente verificado en la etapa posterior por la unidad de análisis.
- Identificar información duplicada: Corresponde a la verificación de cada una de las áreas de trabajo contenidas en la UPM para asegurar, entre otras cosas, que no existan registros duplicados, ya sea por un error en las asignaciones de las áreas de trabajo o por procedimientos mal ejecutados por el enumerador.

- Supervisión del levantamiento: mediante el cual se evalúa el desempeño del recolector, a través de una pauta de observación, la que se aplica a través del acompañamiento del enumerador en el terreno, que permite controlar su desempeño frente al informante y que este cumpla con todos los protocolos y calidad en la obtención del dato recolectado.

V.1.3. Análisis de Gabinete

El objetivo de este procedimiento es disminuir los errores no muestrales, pudiendo detectar de manera oportuna aquellas áreas de la selección que no cumplen con el estándar de calidad de la dirección de la vivienda, susceptible de ser seleccionadas para la encuesta. El análisis de Gabinete corresponde a la validación y actualización a través del análisis de directorios disponibles para cada UPM, sea derivada del proceso de enumeración como del directorio de viviendas institucional. Dentro de los objetivos del proceso se encuentran:

- Analizar y aprobar la completitud y calidad de la información: Corresponde a la revisión de los listados de direcciones con la UPM, asegurando correspondencia al área y completitud con base en su recorrido, verificando la coherencia en el listado de unidades correspondientes a su identificación, número de edificación y orden de la vivienda, según los criterios definidos para cada uno de los conceptos. En esta etapa, se valida la correspondencia de la dirección, que incluye la revisión de registros administrativos disponibles por la institución, para validar y actualizar dichas direcciones, en caso de no disponer de otras fuentes de datos, esta etapa incluye agregar para todas las direcciones irregulares, una observación asociada a la ubicación de la vivienda, en caso de no contar con nombres de calles o numeraciones de dichas viviendas que permitan su clara identificación.
- Corrección de direcciones mal registradas: Las direcciones tienen una coherencia lógica, aquellas que se escapan de la norma son catalogadas como numeración irregular y deben traer una observación asociada, si ese no es el caso, podría tratarse de una dirección mal registrada y es necesario que los equipos de terreno validen o corrijan dicha información. Se agrega a ello, no solo la numeración, sino que también el nombre y la nomenclatura de la vía, además de las otras variables como número de torre o block, número de piso, número de casa o departamento –si corresponde–. Estos datos deben ser verificados en caso de presentar inconsistencias.
- Clasificación de vivienda: Consiste en validar la correcta definición de las viviendas del directorio, considerando el concepto de vivienda como a toda edificación que haya sido construida, convertida o dispuesta para el alojamiento permanente o temporal de personas, abarcando cualquier clase de albergue, fijo o móvil, que sea ocupado como lugar de residencia, no importando los materiales que se hayan empleado en su construcción. Para ser considerada

como vivienda particular, esta debe encontrarse dentro de una propiedad y tener acceso independiente desde la calle, esto implica que no se interrumpa la privacidad de moradores de otras viviendas. Es en estos casos cuando podemos encontrar más de una vivienda dentro de un sitio y que debe tener una observación que dé cuenta de su ubicación dentro del sitio. La vivienda particular ocupada cumple con todo esto, y, además, se encuentra habitada al momento de la enumeración o verificación, si esta no se encontrase habitada, correspondería a una vivienda desocupada. En el ámbito de las viviendas particulares, se encuentran las ocupadas (morador presente – morador ausente) que tendrán siempre un número de hogar y las desocupadas, que no se encuentran habitadas al momento de la enumeración/verificación y, por lo tanto, no lleva asociado un número de hogar.

La vivienda de temporada, por su parte, corresponde a aquellas viviendas utilizadas solo en algunas épocas del año, ya sea por motivos recreacionales, de trabajo, u otros, en este sentido, aquellas personas que se encuentran en una vivienda de temporada tienen, necesariamente, residencia habitual en otra vivienda. Las viviendas colectivas, por su parte, son aquellas utilizadas como lugar de alojamiento por un conjunto de personas sin relación de parentesco, que comparten la vivienda o parte de ella, y que hacen vida en común por razones de salud, trabajo, religión, entre otras. Ejemplo de estas son los hostales, hospitales, residenciales, cárceles, etc.

- Descripción de viviendas sin identificación domiciliaria. Para todas aquellas viviendas que no dispongan de una identificación domiciliaria visible en terreno, es necesario darle una ubicación en la UPM, según su recorrido. En este sentido, es necesario utilizar de referencia las numeraciones domiciliarias de las viviendas aledañas o RRAA disponibles.
- Omisión y cobertura: En este proceso también se registra una revisión de los registros con base en la cobertura en el levantamiento de la UPM o del contenido del directorio, el cual se revisa a través de las plataformas de visualización geográfica disponibles para cada proceso, ya sea terreno o gabinete. En caso de detectar inconsistencias que deriven del terreno, que no pueden ser solucionadas desde un trabajo en gabinete por el equipo de analistas, se procede a rechazar la UPM devolviéndose a terreno para su corrección, pasando nuevamente por el proceso de análisis, hasta que la coherencia y completitud de los datos cumpla con los estándares requeridos para este proyecto.

Una vez finalizado el proceso de análisis de gabinete, se envían todos los registros analizados a la unidad encargada de consolidar los directorios para posteriormente, ser enviados a selección de la muestra de viviendas y con ello, construir el directorio final que será enviado a terreno para la realización de la encuesta.

V.2. RESULTADOS DEL PROCESO DE ACTUALIZACIÓN DEL DIRECTORIO DE CASEN 2022

Una de las principales definiciones del origen de los directorios es la determinación de 40% y 60% de UPM que se extraerán de un proceso de actualización en terreno y una validación en gabinete, el resultado se indica en la Tabla V.2, donde se define que 55% de UPM de la muestra, son obtenidas de directorios de Precenso 2016; 5% de UPM de otras muestras levantadas por el INE, que han sido obtenidas de directorios de enumeraciones de terreno posterior al año 2020; y, 40% de UPM que han sido enumeradas en terreno durante los meses de marzo a julio del año 2022 como un proceso de actualización de la muestra Casen 2022.

Tabla V.2. Total de UPM por región, según origen de actualización de directorios

Región	Muestra Casen		Otras Muestras		Directorios		Enumeración Terreno	
	Total UPM	%		%	Precenso 2016	%		%
Nivel País	12.545	100	625	5	6.933	55	4.987	40
Arica y Parinacota	364	2,9	43	0,3	141	1,1	180	1,4
Tarapacá	512	4,1	50	0,4	233	1,9	229	1,8
Antofagasta	563	4,5	44	0,4	295	2,4	224	1,8
Atacama	506	4,0	35	0,3	270	2,2	201	1,6
Coquimbo	500	4,0	34	0,3	265	2,1	201	1,6
Valparaíso	1.326	10,6	35	0,3	850	6,8	441	3,5
Metropolitana	2.814	22,4	55	0,4	1.671	13,3	1.088	8,7
O'Higgins	785	6,3	36	0,3	434	3,5	315	2,5
Maule	834	6,6	25	0,2	454	3,6	355	2,8
Ñuble	544	4,3	13	0,1	313	2,5	218	1,7
Biobío	1.180	9,4	49	0,4	662	5,3	469	3,7
La Araucanía	821	6,5	39	0,3	447	3,6	335	2,7
Los Ríos	634	5,1	40	0,3	340	2,7	254	2,0
Los Lagos	642	5,1	32	0,3	357	2,8	253	2,0
Aysén	207	1,7	43	0,3	69	0,6	95	0,8
Magallanes	313	2,5	52	0,4	132	1,1	129	1,0

Elaborado por el Instituto Nacional de Estadísticas.

En la construcción de los directorios para la selección de viviendas, existe una serie de actividades que fueron desarrolladas desde el inicio del proceso de enumeración hasta la generación de productos finales entre el mes de marzo y agosto del año 2022. Dentro de ellas, se encuentran:

V.2.1. Enumeración en terreno

El proceso de actualización de directorios en terreno, se desarrolló a partir de la enumeración digital de viviendas y otros usos, a través de la utilización de un dispositivo móvil de captura (DMC) que contiene el formulario de enumeración, donde se registran los datos que se levantan en terreno para todas las áreas

contenidas en las UPM priorizadas para la muestra, tanto para las UPM urbanas como rurales que se hayan definido previamente; teniendo como principal ventaja, la precisión en la ubicación de las edificaciones y/o viviendas capturadas. Cabe señalar que esta recolección requiere que los/as supervisores (as) de cada grupo de trabajo, efectúen de manera constante un proceso de supervisión y control de las cargas gestionadas, actividad que se realiza en el sistema informático Visualizador Web.

El proceso de actualización se generó en un total de 4.987 UPM correspondiente al 40% de la muestra, su distribución se indica a nivel de región en la Tabla V.3.

Tabla V.3. Total de UPM enumeradas en terreno a nivel de región

Región	UPM/Terreno	Nivel País	Porcentaje (%)
Total país	4.987	12.545	40
Arica y Parinacota	180	364	49
Tarapacá	229	512	45
Antofagasta	224	563	40
Atacama	201	506	40
Coquimbo	201	500	40
Valparaíso	441	1.326	33
Metropolitana	1.088	2.814	39
O'Higgins	315	785	40
Maule	355	834	43
Ñuble	218	544	40
Biobío	469	1.180	40
La Araucanía	335	821	41
Los Ríos	254	634	40
Los Lagos	253	642	39
Aysén	95	207	46
Magallanes	129	313	41

Elaborado por el Instituto Nacional de Estadísticas

El desarrollo del proceso de enumeración integró varias actividades que se describen en el Anexo N°3.2.

V.2.2. Validación y Mejora de Directorios de Gabinete

Este proceso de validación y mejoramiento requiere de la revisión de los directorios institucionales, en conjunto con los sistemas informáticos y plataformas geográficas disponibles para ese fin. Actualmente se disponen de los directorios generados en el proceso precensal del año 2016, asociados a unidades geográficas del Censo 2017 y aquellos generados sobre procesos de enumeración para otras muestras enumeradas por el INE a partir del año 2020 y que se encuentran vinculados a enumeraciones sobre unidades primarias de muestreo del Marco Muestral de Viviendas (MMV) que utiliza la muestra Casen 2022.

La Tabla V.4 muestra la distribución de UPM correspondiente a 60% (7.558 UPM) que fueron validadas en el proceso de revisión en gabinete. De estas, 91,7% (6.933 UPM) corresponden a Directorios mejorados de precenso y 8,3 (625) a Directorios generados en terreno para otras muestras enumeradas por el INE. Los porcentajes regionales obtenidos, son calculados con respecto al total nacional de UPM correspondientes al 60%.

Tabla V.4. Distribución regional de directorios de UPM validadas en Gabinete según área urbano/rural. Porcentajes según Total Nacional.

Región	Directorio Gabinete		Directorio Gabinete Precenso 2016				Directorio Gabinete Otras Muestras							
	Total UPM (60%)	%	UPM Urbana	%	UPM Rural	%	Total UPM	%	UPM Urbana	%	UPM Rural	%	Total UPM	%
Nivel País	7.558	100	5.749	76,1	1.184	15,7	6.933	91,7	573	7,6	52	0,7	625	8,3
Arica y Parinacota	184	2,4	128	1,7	13	0,2	141	1,9	40	0,5	3	0,0	43	0,6
Tarapacá	283	3,7	228	3,0	5	0,1	233	3,1	46	0,6	4	0,1	50	0,7
Antofagasta	339	4,5	288	3,8	7	0,1	295	3,9	38	0,5	6	0,1	44	0,6
Atacama	305	4,0	235	3,1	35	0,5	270	3,6	35	0,5	0	0,0	35	0,5
Coquimbo	299	4,0	230	3,0	35	0,5	265	3,5	30	0,4	4	0,1	34	0,4
Valparaíso	885	11,7	717	9,5	133	1,8	850	11,2	32	0,4	3	0,0	35	0,5
Metropolitana	1.726	22,8	1.570	20,8	101	1,3	1.671	22,1	50	0,7	5	0,1	55	0,7
O'Higgins	470	6,2	336	4,4	98	1,3	434	5,7	32	0,4	4	0,1	36	0,5
Maule	479	6,3	334	4,4	120	1,6	454	6,0	19	0,3	6	0,1	25	0,3
Ñuble	326	4,3	206	2,7	107	1,4	313	4,1	12	0,2	1	0,0	13	0,2
Biobío	711	9,4	545	7,2	117	1,5	662	8,8	46	0,6	3	0,0	49	0,6
La Araucanía	486	6,4	276	3,7	171	2,3	447	5,9	37	0,5	2	0,0	39	0,5
Los Ríos	380	5,0	256	3,4	84	1,1	340	4,5	38	0,5	2	0,0	40	0,5
Los Lagos	389	5,1	262	3,5	95	1,3	357	4,7	31	0,4	1	0,0	32	0,4
Aysén	112	1,5	25	0,3	44	0,6	69	0,9	37	0,5	6	0,1	43	0,6
Magallanes	184	2,4	113	1,5	19	0,3	132	1,7	50	0,7	2	0,0	52	0,7

Elaborado por el Instituto Nacional de Estadísticas.

De este 60% de UPM, 91,7% corresponderá a directorios mejorados de Precenso 2016 y, 8,3% corresponderán a directorios generados en terreno para otras muestras enumeradas por el INE.

Esta mejora de la validación de directorios se enfocó en dos ámbitos, primeramente, en la estructura de la dirección a través del tipo de vía, nombre, letra y/o número de domicilio, nombre/letra del block o edificio, números de piso, letra o número de departamento, así como también de las observaciones que se puedan complementar con información más actualizada o mejorada.

Para una mejor comprensión de lo que involucra cada validación, éstas se describen en el Anexo N°3.3.

V.2.3. Elaboración del Directorio y Productos cartográficos

Los productos cartográficos generados para la encuesta Casen hacen relación directamente con disponer de información del directorio georreferenciado, con bases de información actualizadas para ser utilizados como insumos o herramientas que permitan el acceso e individualización de las viviendas seleccionadas para el proceso de levantamiento de la encuesta. Dentro de los productos que se generan, se encuentran los siguientes:

V.2.3.1. Directorio de viviendas seleccionadas

El resultado de la generación del directorio de viviendas seleccionadas para la Encuesta Casen 2022, corresponde a un total de 106.856, distribuidas según región, como indica la Tabla V.5. El directorio, además trae una serie de campos que definen la ubicación de la vivienda a través de su identificación geográfica, identificación de la dirección de la vivienda y sus coordenadas, las que posteriormente se pueden visualizar en los productos cartográficos asociados. El detalle de los campos integrados se encuentra descrito en el Cuadro V.1.

Tabla V.5. Total de viviendas del directorio Casen a nivel de región

Región	Total de Viviendas
Nivel País	106.856
Arica y Parinacota	3.698
Tarapacá	4.156
Antofagasta	4.580
Atacama	5.060
Coquimbo	4.180
Valparaíso	10.984
Metropolitana	22.804
O'Higgins	6.672
Maule	7.090
Ñuble	4.696
Biobío	9.800
La Araucanía	7.006
Los Ríos	5.400
Los Lagos	5.466
Aysén	2.070
Magallanes	3.194

Elaborado por el Instituto Nacional de Estadísticas.

Cuadro V.1. Descripción de campos contenidos en el directorio Casen

Campo	Descripción
ID_VIVIENDA	Identificador Único de la Vivienda
COD_REGION	Código de la región
GLOSA_REGION	Nombre de la región
COD_PROVINCIA	Código de la provincia
GLOSA_PROVINCIA	Nombre de la provincia
COD_COMUNA	Código de la comuna
GLOSA_COMUNA	Nombre de la comuna
COD_AREA	Código de área (1) urbano (2) rural
GLOSA_AREA	Nombre del área (urbano) (rural)
ID_UPM	Código innominado
TIPO_VIAS	Tipo de vía
NOMBRE_DIRECCIÓN	Nombre de la calle o vía
N_DOMICILIO	Número de domicilio (vivienda)
NUM_TERRENO	Número de terreno o sitio
N_LETRA_VIV	Letra de la vivienda
LETRA_BLOCK	Letra o número de block de la edificación
N_PISO	Número de piso de la edificación
N_LETRA_PIEZA_DPTO	Número o letra de la pieza o departamento
OBSERVACIÓN	Descripción u observación de la vivienda para casos donde no existe numeración
LONGITUD	Coordenada posición X (longitud) de la ubicación de la vivienda
LATITUD	Coordenada posición Y (latitud) de la ubicación de la vivienda

Elaborado por el Instituto Nacional de Estadísticas.

V.2.3.2. Productos Cartográficos

Los productos cartográficos generados para la encuesta Casen 2022 hacen relación directamente a disponer de información del directorio de viviendas georreferenciado, con bases de información actualizadas para ser utilizados como insumos o herramientas que permitan el acceso e individualización de las viviendas seleccionadas para el proceso de levantamiento de la encuesta. Dentro de los productos que se generan, se encuentran los siguientes:

- **Georreferenciación de viviendas seleccionadas:** Esta actividad consiste en asignar coordenadas de mapa y ubicación espacial a cada una de las viviendas que han sido seleccionadas para la encuesta para posteriormente, ser visualizados en un dispositivo móvil de captura. El método utilizado se define en dos etapas, uno semiautomático y otro manual, derivado específicamente del origen desde dónde provienen los directorios y de los insumos de viviendas georreferencias disponibles a la fecha.

Para la georreferenciación semiautomática existen 5 etapas relevantes definidas. En orden secuencial, estas son: la normalización y segmentación de la información, que corresponde a comparar nombres de calles similares para realizar búsquedas; la ejecución de una etapa automática que se enfoca en la búsqueda de las direcciones seleccionadas en el MGE; la revisión de enumeraciones previas, en específico, de todas aquellas enumeraciones post-censo que disponen de coordenadas en su origen, como es el caso de las áreas rurales; luego están todas aquellos puntos rurales de la enumeración de Precenso 2016 que también disponen de coordenadas en su origen y; finalmente se chequean direcciones georreferenciadas en muestras anteriores, desde donde ya se dispone de la coordenada de viviendas georreferenciada en una muestra anterior y que podrían ser coincidentes con la selección de viviendas Casen 2022.

Una vez terminado este proceso semiautomático, se disponen de un número menor de direcciones donde no se logran obtener sus coordenadas, para estos casos, se pone a disposición a los analistas geoespaciales y a nivel central una plataforma geográfica web donde, a partir de todos los insumos disponibles, como son croquis y formularios de enumeración, se localizan sobre la plataforma dichas direcciones. En el caso de no lograr una ubicación exacta de la dirección buscada, el punto queda localizado, a lo menos, en la cara de manzana o entidad a la que corresponde la vivienda. Luego se consolidan ambos métodos para la obtención de una capa de puntos con el 100% de viviendas seleccionadas de la encuesta (106.856). Esta información es incorporada a la base Geodatabase y Shapefile a nivel nacional que contiene la totalidad de unidades primarias de muestreo 12.545 UPM y los puntos de viviendas seleccionadas para cada una de ellas.

- **Mapas Móviles:** El Mapa móvil, es un instrumento de sistema de captura de información geográfica que está basado en un mapa georreferenciado, que permite que un usuario mediante un dispositivo móvil interactúe en un espacio geográfico definido.

Para la construcción de los mapas móviles para la encuesta Casen 2022, se utiliza como información base, la base ACON y todos sus componentes, como son ejes de calles, manzanas, áreas verdes, equipamientos entre otros. Para la generación de productos, se utilizó una versión más actualizada de la base ACON disponible, al momento de la priorización, incorporando 110 comunas con actualización por CRF hasta diciembre del 2021 y que en su mayoría se incluyen actualizaciones por imagen actualizadas en año 2022.

El mapa móvil contiene información cartográfica complementaria, orientada a ayudar al encuestador. En ella, estarán contenidas las UPM que muestran las áreas de selección, sobre las que se despliegan los puntos de viviendas georreferenciados y seleccionados para la encuesta, los cuales se pueden consultar y visualizar sus atributos a través de ventanas emergentes.

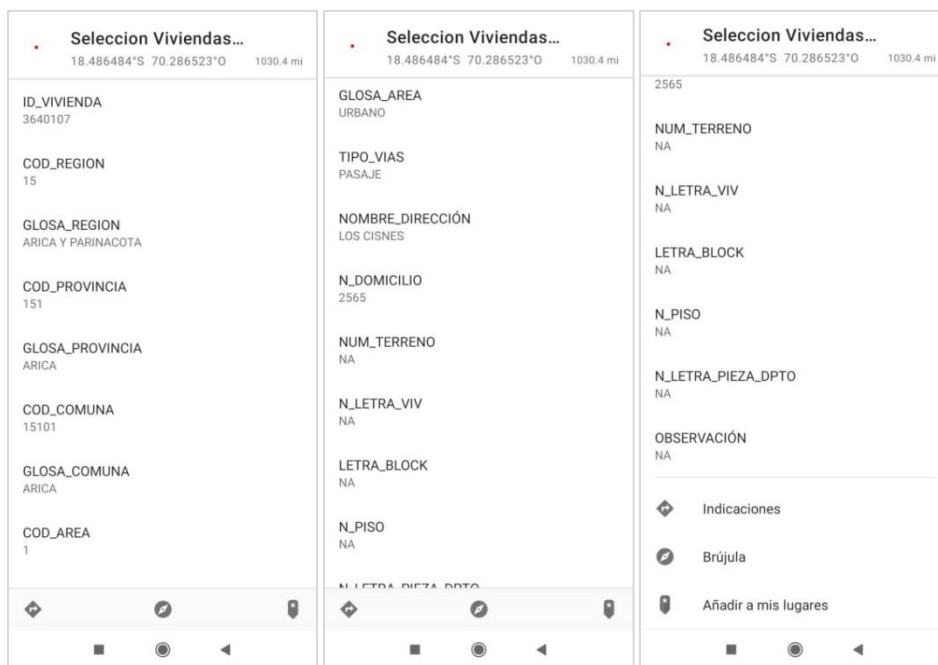
El resultado es la construcción de 16 mapas móviles a nivel regional, la descripción de los campos se encuentra en el Cuadro V.2, en conjunto con la visualización de los campos, áreas de trabajo y viviendas seleccionadas sobre DMC (Imágenes V.1 y V.2):

Cuadro V.2. Descripción de campos del Mapa Móvil Regional

Campo	Descripción
ID_VIVIENDA	Código único de la vivienda
COD_REGIÓN	Código de la Región
GLOSA_REGIÓN	Nombre de la Región
COD_PROVINCIA	Código de la Provincia
GLOSA_PROVINCIA	Nombre de la Provincia
COD_COMUNA	Código de Comuna (Código Único Territorial)
GLOSA_COMUNA	Nombre de la Comuna
COD_AREA	Código de Área (1) Urbano (2) Rural
GLOSA_AREA	Nombre del Área (Urbano/Rural)
TIPO_VIAS	Tipo de vía
NOMBRE_DIRECCIÓN	Nombre de la calle o vía
N_DOMICILIO	Número de Domicilio
NUM_TERRENO	Número del terreno o sitio
N_LETRA_VIV	Letra de la vivienda
N_LETRA_BLOCK	Letra o número de block de la edificación
N_PISO	Número de piso de la edificación
N_LETRA_PIEZA_DPTO	Número o letra de la pieza o departamento
OBSERVACIÓN	Descripción de la vivienda

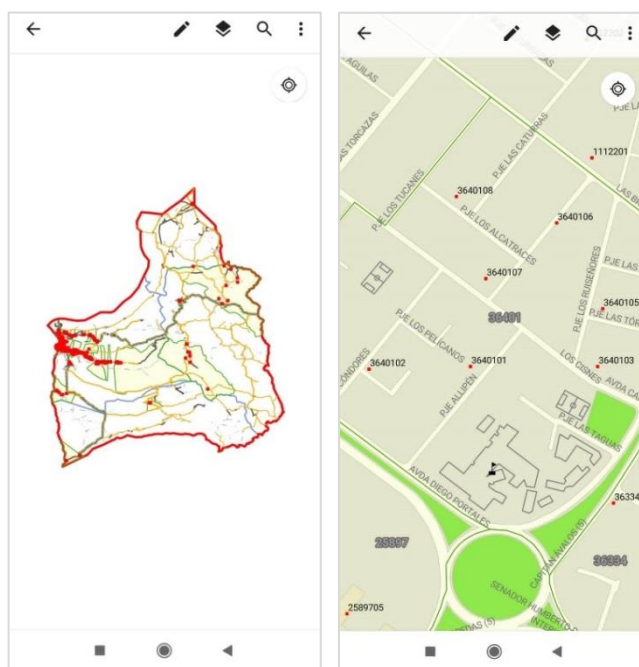
Elaborado por el Instituto Nacional de Estadísticas.

Imagen V.1. Visualización de campos de Mapa Móvil en DMC



Elaborado por el Instituto Nacional de Estadísticas.

Imagen V.2. Visualización de Mapa Móvil con UPM de la muestra y selección de viviendas Casen 2022



Elaborado por el Instituto Nacional de Estadísticas.

VI. DISEÑO FACTORES DE EXPANSIÓN

VI.1. INTRODUCCIÓN

Los análisis y estudios basados en encuestas de hogares con diseño muestral probabilístico utilizan un ponderador en la estimación de parámetros de interés para que estos tengan validez sobre la población objetivo. Este ponderador guarda relación con las probabilidades de selección de las distintas unidades de muestreo y da cuenta del número de personas de la población que representa cada individuo que participa en dicha encuesta. Este ponderador es conocido como factor de expansión.

Para Casen 2011, 2013, 2015, 2017 y Casen en Pandemia 2020 se desarrollaron factores para expandir a la población regional, provincial y comunal, lo que se mantiene para Casen 2022 donde el factor de expansión comunal y provincial es calculado solo para fines de investigación, dado que los dominios de estudio para los cuales la muestra es representativa son: nacional, nacional urbano, nacional rural y regional. Para los diferentes dominios de estudio, el factor de expansión fue calculado a nivel hogar donde todos los integrantes de un hogar cuentan con el mismo factor de expansión. Sin embargo, se debe contemplar que como efecto del redondeo probabilístico y, además con el objeto de que los factores de expansión alcancen, con una mayor precisión, el stock de las proyecciones de población a nivel nacional y regional, es probable que existan hogares donde un integrante presente un factor con una unidad mayor o menor que el resto de sus integrantes.

La construcción de los factores de expansión para Casen 2022 responde al diseño muestral consensuado entre el INE y el Ministerio que contempla una estratificación, dada por la agregación de comuna – área – NSE, donde aplique, y una conglomeración dada por las Unidades Primarias de Muestreo (UPM) provenientes del marco muestral basado en el Censo 2017. Esto último indica que las unidades finales de muestreo son seleccionadas en dos etapas donde, primero se selecciona un conglomerado o UPM y, luego se seleccionan las viviendas que se entrevistarán. La selección de las UPM se realiza en forma sistemática con probabilidad proporcional al tamaño (según el número de viviendas que contienen). La selección de viviendas se realiza con igual probabilidad al interior de cada UPM seleccionada.

Al factor de selección de unidades se le aplican una serie de ajustes dados por el resultado del trabajo de campo y adicionalmente, el factor resultante es calibrado a totales poblacionales según metodología de raking⁶³.

Así, en detalle, la metodología de cálculo de los factores de expansión consiste en la obtención secuencial de cuatro ponderadores:

⁶³ Para más detalles sobre esta metodología revisar: https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/documentos-de-trabajo/documentos-de-trabajo/calibracion-factores-ene-metodologia.pdf?sfvrsn=cc1f6a74_8.

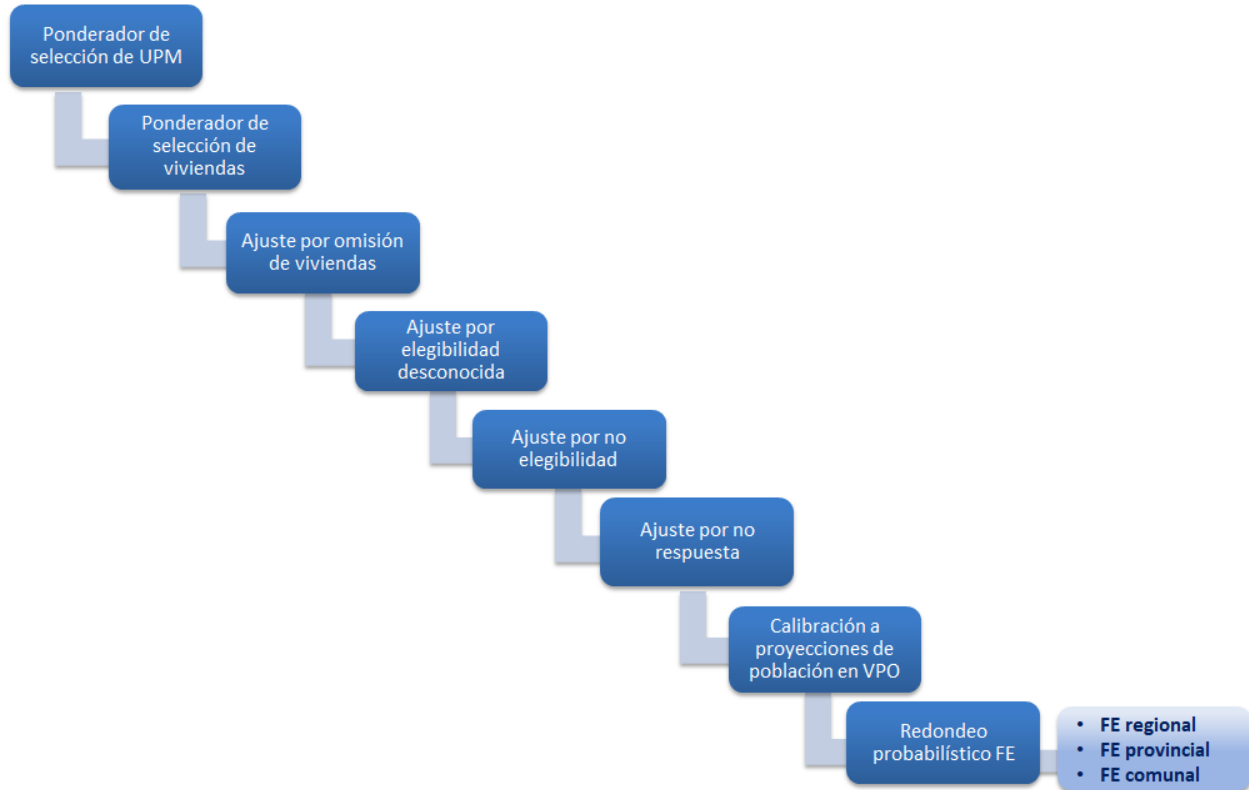
1. Ponderador de selección de unidades (de primera y segunda etapa de muestreo) que refleja cuantas viviendas en la población representa la vivienda seleccionada en la muestra. En términos generales, se calcula como el inverso de las probabilidades de selección de ambas unidades de muestreo.
2. Ponderador corregido por elegibilidad que busca corregir las fallas que llevan a la inclusión en la muestra de viviendas no elegibles y aquellas que terminan clasificadas como de elegibilidad desconocida.
3. Ponderador corregido por no respuesta que se aplica con el objetivo de reducir el sesgo asociado a la falla en conseguir la cooperación de los hogares residentes en las viviendas seleccionadas en la muestra. La técnica para implementar en Casen 2022, es la misma utilizada en Casen en Pandemia 2020 y anteriores, con la cual se busca desarrollar grupos homogéneos (en relación con la probabilidad de responder la encuesta) al interior de los cuales se calcula un ajuste de razón que permite corregir los ponderadores de selección.
4. Ponderador de calibración que busca cuadrar las estimaciones poblacionales obtenidas de la encuesta con las estimaciones poblacionales desarrolladas por el INE a partir de la información del Censo 2017, y que son utilizadas actualmente en todas las encuestas de hogares del INE, mediante la corrección de fallas en la cobertura de subpoblaciones de interés. Este ajuste se conoce también como corrección de población total.

Este ajuste se realiza en 3 niveles: comunal, provincial y regional. El ajuste “comunal”, se obtiene como la razón entre el total de personas en la comuna según las estimaciones demográficas del INE y el total de personas residentes en viviendas particulares estimadas para cada comuna a partir de la encuesta. Asimismo, el ajuste “provincial”, se obtiene como la razón entre el total de personas en este grupo poblacional según las estimaciones demográficas del INE y el total de personas residentes en viviendas particulares estimado para cada provincia a partir de la encuesta.

Por último, el ajuste “regional”, se realiza a través del método de calibración. Para esto, se utiliza la técnica por Raking que será descrita en los próximos apartados y que permite controlar por diversas variables simultáneamente y mantiene el diseño muestral ya que disminuye el sesgo que se produce por errores no muestrales, provocados por la no-respuesta y/o la imperfección de los marcos muestrales y, además, mejora la precisión de las estimaciones. En el proceso de calibración, se utiliza la información auxiliar proveniente de las Proyecciones de Población, base Censo 2017, ajustadas a la población residente en viviendas particulares ocupadas según sexo y edad y para distintos niveles de desagregación.

La Figura VI.1 muestra el flujo del proceso del cálculo de los factores de expansión para Casen 2022. Luego, a lo largo del capítulo se detalla el proceso de construcción de los factores de expansión de la muestra. Cada apartado hace referencia a las rutinas correspondientes a cada ponderador mencionado.

Figura VI.1 Diagrama de procesos de desarrollo de factor de expansión Casen 2022



Elaborado por el Instituto Nacional de Estadísticas⁶⁴.

VI.2. PONDERADOR DE SELECCIÓN DE CONGLOMERADOS

En Casen 2022 las unidades muestrales se obtienen desde el MMV 2020 a partir de un diseño probabilístico, estratificado y bietápico, donde los estratos muestrales están conformados por la combinación Comuna-Área-NSE, donde aplique y los conglomerados son seleccionados en forma sistemática con probabilidad proporcional al tamaño, según el número de viviendas que los conforman. Por su parte, las viviendas o unidades secundarias de muestreo son seleccionadas en forma aleatoria y con igual probabilidad al interior de cada UPM elegida.

A continuación, se describen los ponderadores y ajustes asociados a cada etapa del factor.

⁶⁴ VPO corresponde a Viviendas Particulares Ocupadas.

VI.2.1. Probabilidad de selección de los conglomerados

La selección de conglomerados o UPM i al interior de cada estrato de muestreo h se realiza con probabilidad proporcional al tamaño, según el número de viviendas particulares que contienen en el marco muestral. La forma de cálculo se observa en la ecuación (10):

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

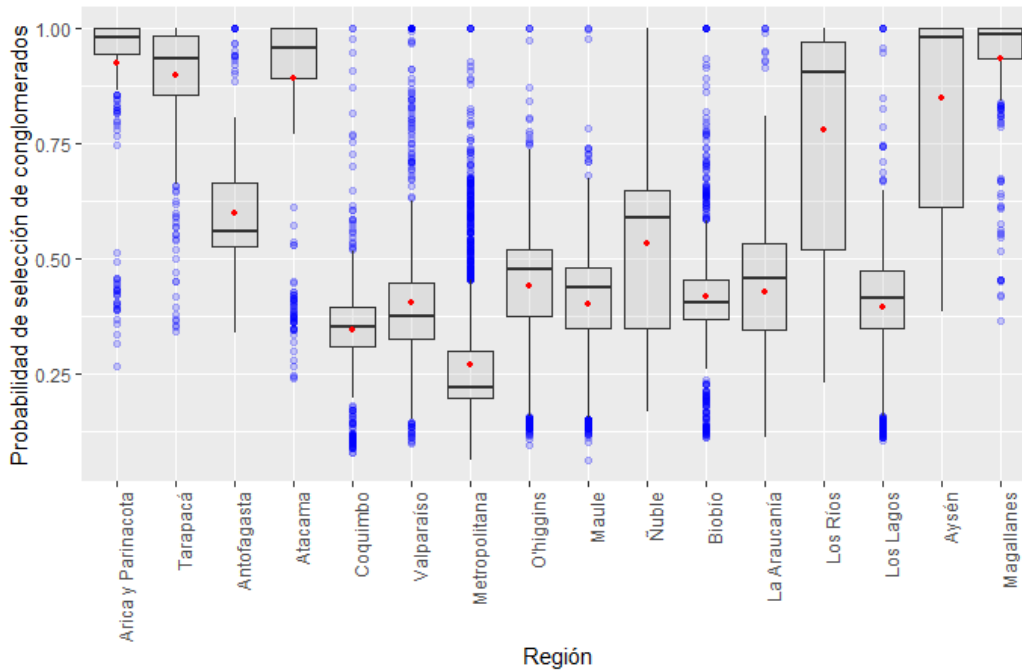
Donde,

- i : Índice de la unidad primaria de muestreo.
- h : Índice del estrato de muestreo.
- n_h : Número de UPM seleccionadas del estrato h .
- M_{hi} : Número de viviendas en la UPM i del estrato h según el MMV 2020.
- M_h : Número de viviendas en el estrato h según el MMV 2020.

En selección de unidades con probabilidad proporcional al tamaño es posible obtener probabilidades mayores que 1 en estratos pequeños o en estratos donde exista una UPM conformada por una cantidad de viviendas considerablemente superior al resto. En estos casos se puede aplicar un truncamiento a las probabilidades originales de selección, asignándoles un valor de 1, considerando con esto, una inclusión forzosa de dicha unidad (Särndal, et.al. 1992).

El Gráfico VI.1 presenta la distribución regional de probabilidades de selección de UPM. Al respecto se observa que, en gran parte de las regiones las probabilidades de selección igual a 1 corresponden casos atípicos, con excepción de Arica y Parinacota, Atacama, Aysén y Magallanes.

Gráfico VI.1: Distribución de las probabilidades de selección de UPM a nivel regional



Elaborado por el Instituto Nacional de Estadísticas.

Las regiones que presentan las menores medianas respecto a la probabilidad de selección de UPM son: Metropolitana (0,22) y Coquimbo (0,35). Por otro lado, las regiones que presentan las medianas más altas son: Magallanes (0,99) y las regiones de Arica y Parinacota, y Aysén (0,98). La mediana nacional para las probabilidades de selección se ubica en 0,42 y la media en 0,48.

VI.2.2. Ponderador de selección de conglomerados como inverso de la probabilidad de selección

Obtenidas las probabilidades de selección de los conglomerados (UPM), se obtiene el ponderador de selección como el inverso de la probabilidad de selección, tal como se indica en la ecuación (11):

$$w_{hi} = \frac{1}{P_h(i)} = \frac{M_h}{n_h \cdot M_{hi}} \quad (11)$$

Donde,

w_{hi} : Ponderador de selección base de la unidad primaria de muestreo o conglomerado.

El ponderador de selección de conglomerados se interpreta como el total de unidades en el marco muestral que representa una unidad seleccionada en la muestra.

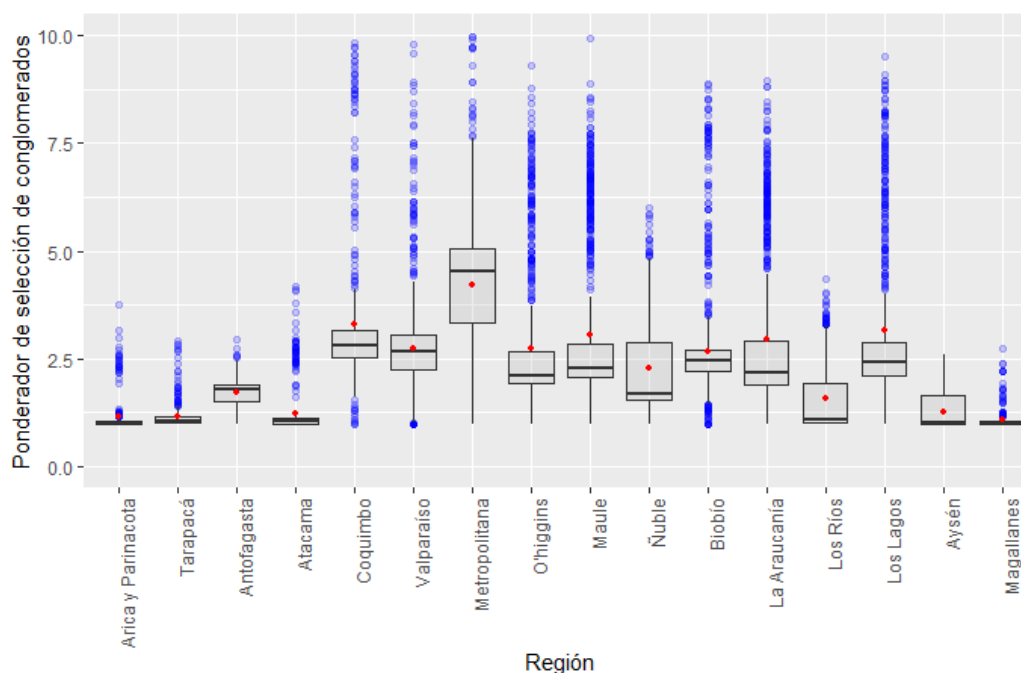
Tabla VI.1. Estadísticos de los ponderadores de selección de UPM, Casen 2022

Región	Mínimo	Máximo	Media	Mediana
Total	1,00	16,15	2,81	2,36
Arica y Parinacota	1,00	3,76	1,15	1,02
Tarapacá	1,00	2,92	1,15	1,07
Antofagasta	1,00	2,95	1,73	1,79
Atacama	1,00	4,17	1,23	1,05
Coquimbo	1,00	13,00	3,52	2,83
Valparaíso	1,00	10,03	2,76	2,68
Metropolitana	1,00	15,73	4,29	4,52
O'Higgins	1,00	10,61	2,75	2,10
Maule	1,00	16,15	3,08	2,28
Ñuble	1,00	6,01	2,27	1,70
Biobío	1,00	8,88	2,68	2,47
La Araucanía	1,00	8,94	2,93	2,19
Los Ríos	1,00	4,36	1,56	1,10
Los Lagos	1,00	9,51	3,15	2,42
Aysén	1,00	2,59	1,28	1,02
Magallanes	1,00	2,74	1,10	1,01

Elaborado por el Instituto Nacional de Estadísticas.

La Tabla VI.1 y el Gráfico VI.2 presentan la distribución regional del ponderador de selección de UPM, la mediana del ponderador a nivel nacional es de 2,36 y a nivel regional la mayor y menor mediana se presentan en la Región Metropolitana (4,52) y en la Región de Magallanes (1,01), respectivamente.

Gráfico VI.2: Distribución de los ponderadores de selección de las UPM a nivel regional



Elaborado por el Instituto Nacional de Estadísticas.

VI.3. PONDERADORES DE SELECCIÓN DE VIVIENDAS

El paso anterior describe la probabilidad de selección de las UPM. Sin embargo, la última unidad de selección de la muestra Casen 2022 es la vivienda al interior de cada UPM seleccionada.

En los apartados siguientes se detallan las probabilidades de selección de las viviendas, así como también los ajustes y calibraciones necesarios que permiten representar a toda la población objetivo.

VI.3.1. Probabilidad condicional de selección de viviendas

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

Luego, la probabilidad $P_{hi}(j|i)$ de incluir la j -ésima vivienda, condicional a que el i -ésimo conglomerado (UPM) fue seleccionado en la muestra en el estrato h , está dada por la ecuación (12).

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

Donde,

- j : Índice de la vivienda.
- M'_{hi} : Número de viviendas particulares ocupadas en la UPM i , del estrato h disponibles para seleccionar, según enumeración o verificación previa a la recolección de datos.
- m_{hi} : Número de viviendas seleccionadas en la UPM i , del estrato h .

VI.3.2. Ajuste por omisión de viviendas

Como se indica en el apartado sobre la selección de viviendas para Casen 2022, en la selección de UPM fueron incluidas algunas que ya han sido seleccionadas en otras encuestas encargadas al INE. En algunos casos el trabajo de campo de estas encuestas coincide con el trabajo de campo de Casen 2022 y, en otras, no ha pasado suficiente tiempo desde la última vez que las viviendas seleccionadas fueron levantadas. Por este motivo, es necesario realizar un ajuste al directorio de viviendas de Casen 2022, previo a la selección que le otorgue probabilidad igual a 0 a aquellas que ya han sido seleccionadas en otras encuestas. Este ajuste afecta las probabilidades originales de selección de las viviendas, lo que implica que el stock se reduce. Por lo anterior, para conservar el stock original de viviendas se estima una razón entre la cantidad total de viviendas particulares ocupadas desde el directorio original y la cantidad de viviendas disponibles de dicho directorio, es decir, con probabilidad de selección mayor que cero. Esta razón se observa en la ecuación (13).

$$\hat{R}_i = \frac{M'_{hi}}{\text{Total viviendas en la UPM}} \quad (23)$$

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

$$P_{hi}(j|i)_{aj} = P_{hi}(j|i) \cdot \hat{R}_i \quad (34)$$

Luego, la probabilidad condicional de selección de la vivienda j ubicada en el conglomerado (UPM) i del estrato h ($P_{hi}(j)$) se obtiene desde la ecuación (15).

$$P_{hi}(j) = P_h(i) \cdot P_{hi}(j|i)_{aj} \quad (45)$$

Donde,

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

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

VI.3.3. Ponderador de selección de viviendas como inverso de la probabilidad de selección

El ponderador de selección de viviendas es calculado como el inverso de la probabilidad $P_{hi}(j)$ de selección de la vivienda j ubicada en el conglomerado i (UPM) del estrato h . Esto es:

$$w_{hij} = \frac{1}{P_h(i) \cdot P_{hi}(j|i)_{aj}} = \frac{1}{P_h(i)} \cdot \frac{1}{P_{hi}(j|i) \cdot \hat{R}_i} = w_{hi} \cdot \frac{1}{P_{hi}(j|i)_{aj}} \quad (56)$$

Donde,

h, i, j : Índice del estrato, conglomerado (UPM) y vivienda, respectivamente.

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

$P_{hi}(j|i)_{aj}$: Probabilidad de selección ajustada de la vivienda j ubicada en el conglomerado (UPM) i del estrato h en la muestra de Casen 2022.

P_{hi} : Probabilidad de selección del conglomerado i (UPM) dentro del estrato h .

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

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

Esta expresión se puede escribir como:

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

Donde se agrega la nomenclatura:

- M_h : : Número de viviendas en el estrato h en el marco de selección de Casen 2022.
- n_h : : Número de UPM seleccionadas en el estrato h .
- \hat{R}_i : : Razón de ajuste por omisión de viviendas
- M_{hi} : : Número de viviendas en la UPM i del estrato h según el MMV 2020.
- M'_{hi} : : Número de viviendas particulares ocupadas en la UPM i , del estrato h , según enumeración/verificación previa a la recolección de datos.

Este ponderador puede ser interpretado como el número de viviendas en la población que representan las viviendas seleccionadas en la muestra de Casen 2022. Al interior de cada conglomerado, la ponderación es la misma para todas las viviendas.

VI.4. AJUSTE POR ELEGIBILIDAD

La población objetivo de la Encuesta Casen y, por lo tanto, a la que apunta el proceso de inferencia, son las personas que residen en forma habitual en viviendas particulares ocupadas. Estas viviendas son obtenidas desde un marco de selección para Casen, cuyas unidades han sido actualizadas en terreno o utilizando registros administrativos para su actualización en gabinete.

Sin embargo, a pesar de todos los esfuerzos que se realicen por mantener marcos muestrales actualizados y válidos para cubrir la población objeto de estudio, siempre existe la posibilidad de encontrar errores de inclusión (o de exclusión) al momento utilizar el marco muestral.

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

Por esta razón, es importante ajustar las probabilidades de selección de las viviendas seleccionadas para incorporar el hecho de que una proporción de las viviendas podría resultar como no elegible y, en otros casos, podría desconocerse su estado de elegibilidad.

Al término del trabajo de campo, todas las viviendas seleccionadas inicialmente por el INE son clasificadas en tres grandes grupos. Estas son:

1. **Elegibles:** Corresponde a edificaciones en que el encuestador pudo determinar que se trataban de viviendas particulares ocupadas (incluye entrevistas completas, entrevistas parciales y no entrevistadas).
2. **No elegibles:** Son edificaciones identificadas como negocios, viviendas colectivas, viviendas deshabitadas, viviendas de veraneo, viviendas destruidas, etc. En síntesis, edificaciones que no pertenecen al universo de estudio.
3. **Elegibilidad desconocida:** Corresponde a edificaciones en que no se pudo determinar su estado. Este es el caso, por ejemplo, de unidades que nunca fueron enviadas a terreno, viviendas a las cuales no se pudo llegar o encontrar, y otros casos similares.

El ponderador de selección de viviendas tiene valores válidos para las viviendas elegibles, no elegibles y de elegibilidad desconocida. En lo que sigue del proceso, sólo se dejarán valores válidos para las viviendas elegibles. A continuación, se describen los ajustes aplicados.

VI.4.1. Ajuste por elegibilidad desconocida

Inicialmente, todas las viviendas tienen un estatus conocido de elegibilidad: son o no viviendas particulares ocupadas. La muestra de viviendas seleccionadas incluye, en principio solo viviendas elegibles, según información disponible con posterioridad al proceso de enumeración y verificación en el caso de Casen 2022.

Durante el trabajo de campo, sin embargo, este estado no se puede comprobar en algunas viviendas, las que terminan clasificadas como de elegibilidad desconocida. Los pesos de las edificaciones consideradas en un principio, como viviendas con elegibilidad conocida (elegibles y no elegibles), pero que finalmente terminan siendo clasificadas como edificaciones de elegibilidad desconocida, son redistribuidos proporcionalmente entre las viviendas con elegibilidad conocida, al interior de cada estrato de muestreo.

La proporción de los pesos a redistribuir $R_{h,known}$ se calcula como la razón entre dos estimaciones realizadas con base en el ponderador de selección de viviendas.

El numerador corresponde a la estimación del total de viviendas en la población y el denominador corresponde a la estimación del total de viviendas con elegibilidad conocida (viviendas elegibles y no elegibles). Esto se observa en la ecuación (18):

$$\hat{R}_{h, know} = \frac{\sum_{i \in \Omega_h} \sum_{j \in \theta_i} w_{hij}}{\sum_{i \in \Omega_h} \sum_{j \in \theta_{i, know}} w_{hij}} \quad (18)$$

Donde,

- θ_i : : Conjunto de viviendas j seleccionadas en la UPM i .
- $\theta_{i, know}$: Conjunto de viviendas j seleccionadas en la en la UPM i y clasificadas como elegibles o no elegibles.
- Ω_h : Conjunto de estratos h (áreas geográficas de las comunas).
- w_{hij} : Ponderador de selección de la vivienda j en el conglomerado i del estrato h .

La razón antes descrita es la misma para todas las unidades muestrales pertenecientes a un mismo estrato de muestreo. Luego, el ponderador de selección de vivienda corregido por elegibilidad desconocida está dado por la ecuación (19):

$$w'_{hij} = \hat{R}_{h, know} \cdot w_{hij} \quad (69)$$

Donde w_{hij} es el ponderador de selección de la vivienda j desde el conglomerado i seleccionado del estrato h .

VI.4.2. Ajuste por no elegibilidad

El objetivo analítico de la Encuesta Casen es producir inferencias hacia la población que reside en viviendas particulares ocupadas (elegibles). Por lo tanto, a partir de esta etapa no se consideran, para fines analíticos, aquellas viviendas que no conforman la población objetivo de la encuesta (viviendas no elegibles, tales como oficinas de empresas, viviendas abandonadas, viviendas de veraneo, viviendas demolidas, etc.).

Este ajuste consiste en asignar un valor blanco (“missing”) en el ponderador de selección de viviendas corregido por elegibilidad, a las viviendas con clasificación no elegible.

La Tabla VI.2 presenta la distribución de las viviendas seleccionadas de acuerdo con su estado de elegibilidad según el trabajo de campo de Casen 2022.

Tabla VI.2. Distribución viviendas según elegibilidad

Elegibilidad	Frecuencia	Porcentaje
Total	106.856	100,0%
Elegible	100.632	94,2%
Elegibilidad desconocida	2.408	2,2%
No elegible	3.816	3,6%

Elaborado por el Instituto Nacional de Estadísticas.

VI.5. AJUSTE POR NO RESPUESTA

En esta etapa solo son consideradas las viviendas elegibles, pues son estas las que cumplen con las características necesarias para participar en la encuesta, es decir, ser viviendas particulares ocupadas. Sin embargo, aun cuando dichas viviendas cumplen con los atributos necesarios, es posible que los residentes de algunas de ellas no deseen participar o simplemente no puedan ser contactados. En Casen 2022, de las 100.632 viviendas que resultaron elegibles, 70.751 fueron entrevistadas lo que significa que la tasa de respuesta es de 70,3%.

Utilizar un ponderador de selección sin tomar en cuenta el fenómeno de la no respuesta a la unidad, significa obtener estimaciones representativas solo de aquellas unidades que participan en la encuesta y, por lo tanto, una parte de la población no sería cubierta.

Si no hay relación entre el “patrón de no respuesta” y la variable de interés, entonces es posible que la no respuesta solo afecte la precisión (varianza) de las estimaciones de interés. Sin embargo, si existe alguna relación entre la no respuesta y la variable de interés, se podrían obtener estimaciones sesgadas a partir del resultado de la encuesta.

Por ejemplo, si existe evidencia de que la no respuesta es mayor en alguna subpoblación con características especiales (por ejemplo, en el nivel socioeconómico alto) no ajustar por no respuesta, podría generar una subestimación de algunos estadísticos de interés, ya que en la encuesta estaría subrepresentada dicha subpoblación.

Para minimizar los problemas ocasionados por la ausencia de respuesta a la unidad, se realiza un ajuste para compensar la pérdida de las unidades, denominado *propensity score*. Este método de ajuste consiste en estimar la probabilidad de responder que tienen los hogares que son parte de las viviendas de la muestra. Como no se cuenta con información de los residentes de las viviendas donde no se realiza la entrevista, ni tampoco de las características de estas, sino solo su ubicación geográfica (región, provincia, comuna, área, etc.) resulta imposible realizar el ajuste a nivel de viviendas, por lo cual, se realiza a partir de la estimación de algún estadístico de la variable respuesta que, en caso de Casen es la tasa de respuesta⁶⁵ por unidad primaria de muestreo.

El ajuste consiste, al igual que en Casen en Pandemia 2020 y anteriores, en implementar un modelo de regresión lineal para predecir la participación en la encuesta (tasa de respuesta) de las UPM utilizando un conjunto de variables auxiliares con información a ese mismo nivel, así como también, información de la comuna, proveniente de otras encuestas o incluso de registros administrativos, con el objeto de predecir la tasa de respuesta al interior de las unidades para posteriormente, ordenarlas en orden

⁶⁵ La tasa de respuesta se obtiene como el cociente entre el número de viviendas entrevistadas y el total de viviendas elegibles.

creciente. Finalmente, todas las unidades de muestreo son agrupadas en veintiles (20 grupos de igual tamaño) siendo las viviendas pertenecientes a la misma unidad asignadas a un mismo veintil.

En resumen, la determinación de las celdas de ajuste, también denominados “estratos” de ajuste, se realiza en tres pasos:

1. Selección del modelo de predicción de la tasa de respuesta de las UPM⁶⁶.
2. Predicción de la tasa de respuesta para cada UPM.
3. Ordenamiento de las tasas predichas, de menor a mayor, y creación de veintiles.

Si el modelo de predicción es correcto, al interior de las 20 celdas de ajuste, las viviendas son similares en sus características de interés y, por consiguiente, los entrevistados pueden representar a los no entrevistados.

Bajo este supuesto, los ponderadores asociados a las viviendas elegibles que no respondieron deben ser redistribuidos entre las viviendas elegibles que si respondieron. Para estos efectos, se debe calcular una razón de ajuste.

La determinación de la razón de ajuste viene dada por la ecuación (20) y se realiza al interior de cada veintil utilizando el ponderador de selección de viviendas corregido por elegibilidad.

En el numerador se incluye la estimación del total de viviendas elegibles, y en el denominador la estimación del total de viviendas elegibles que responde.

$$\hat{R}_{g,R} = \frac{\sum_{i \in \Phi_g} \sum_{j \in \theta_{i,eleg}} w'_{hij}}{\sum_{i \in \Phi_g} \sum_{j \in \theta_{i,eleg,R}} w'_{hij}} \quad (20)$$

Donde,

- g : : Índice del grupo para la corrección de no respuesta (veintil).
- Φ_g : : Conjunto de UPM asignadas al veintil g .
- $\theta_{g,eleg}$: : Conjunto de viviendas pertenecientes al veintil g y catalogadas como elegibles.
- $\theta_{g,eleg,R}$: : Conjunto de viviendas pertenecientes al veintil g , catalogadas como elegibles y que responden la encuesta.
- w'_{hij} : : Ponderador de selección de viviendas corregido por elegibilidad.

Una vez estimados los factores de ajuste, el ponderador de selección corregido por no respuesta w_{hij}^{NR} puede ser expresado como:

⁶⁶ En el anexo N°4 se presentan las variables que componen el modelo seleccionado y su fuente.

$$w_{hij}^{NR} = \hat{R}_{g,R} \cdot w'_{hij} \quad (21)$$

Este ponderador se asigna a las 70.751 viviendas elegibles entrevistadas. A las viviendas elegibles no entrevistadas se asigna un valor blanco (“missing”).

La Tabla VI.3 presenta la distribución de los veintiles y la tasa de respuesta de cada uno de ellos.

Tabla VI.3. Veintiles Casen 2022

Veintil	Viviendas elegibles	Viviendas que responden	Viviendas que no responden	TR veintil
Total país	100.632	70.751	29.881	0,703
1	4.804	1.844	2.960	0,384
2	4.803	1.974	2.829	0,411
3	4.846	2.405	2.441	0,496
4	4.918	2.697	2.221	0,548
5	5.002	3.003	1.999	0,600
6	4.996	3.171	1.825	0,635
7	5.063	3.435	1.628	0,678
8	4.994	3.536	1.458	0,708
9	5.015	3.623	1.392	0,722
10	4.956	3.672	1.284	0,741
11	4.933	3.704	1.229	0,751
12	4.970	3.835	1.135	0,772
13	4.961	3.883	1.078	0,783
14	4.956	3.967	989	0,800
15	4.948	3.983	965	0,805
16	4.983	4.017	966	0,806
17	5.107	4.175	932	0,818
18	5.292	4.439	853	0,839
19	5.505	4.624	881	0,840
20	5.580	4.764	816	0,854

Elaborado por el Instituto Nacional de Estadísticas.

VI.6. PONDERADOR DE CALIBRACIÓN

La Encuesta Casen se realiza a partir de una muestra de viviendas y, dado el diseño muestral los factores de expansión calculados sirven para llevar estas unidades muestrales a sus respectivos stocks poblacionales (totales de viviendas). Sin embargo, como las principales unidades de interés analítico de la encuesta son las personas que residen en las viviendas seleccionadas, se requiere incorporar a los factores de expansión información actualizada a este nivel, dada por las proyecciones de población generadas por el INE a partir del Censo 2017, con actualización al periodo de realización de la encuesta.

Estas proyecciones son utilizadas en los diseños de todas las encuestas de hogares encargadas al INE. En el caso de Casen se utilizan con fecha 30 de noviembre de 2022, que coincide con el punto medio del periodo en que se realiza el trabajo de campo y para los niveles comunales, provinciales y regionales, junto con desagregaciones combinada con sexo y edad.

La Tabla VI.4 muestra las proyecciones de población, según región, utilizadas comúnmente en diseños encargados al INE y aquellas ajustadas a población residente en viviendas particulares utilizadas en la Encuesta Casen.

Tabla VI.4. Proyecciones de población al 30 de noviembre de 2022 según región y área

Región	Estimación de población total			Estimación de población residente en viviendas particulares		
	Urbano	Rural	Total	Urbano	Rural	Total
Total	17.629.090	2.254.455	19.883.545	17.624.792	2.253.781	19.878.573
Arica y Parinacota	237.462	21.125	258.587	237.133	20.929	258.062
Tarapacá	379038	19.699	398737	379.008	19.691	398.699
Antofagasta	688.045	23.472	711.517	687.535	23.425	710.960
Atacama	291.352	27.071	318.423	291.323	27.036	318.359
Coquimbo	709.268	153.771	863.039	709.119	153.726	862.845
Valparaíso	1.836.479	165.491	2.001.970	1.835.838	165.447	2.001.285
Metropolitana	8.026.487	308.137	8.334.624	8.025.026	308.047	8.333.073
O'Higgins	765.555	247.359	1.012.914	765.502	247.312	1.012.814
Maule	864.344	292.699	1.157.043	864.170	292.660	1.156.830
Ñuble	367.987	150.047	518.034	367.956	150.046	518.002
Biobío	1.494.066	184.370	1.678.436	1.493.749	184.357	1.678.106
La Araucanía	738.133	287.588	1.025.721	737.917	87.583	1.025.500
Los Ríos	298.689	111.561	410.250	298.582	111.541	410.123
Los Lagos	676.188	228.339	904.527	676.006	228.312	904.318
Aysén	88.566	19.569	108.135	88.536	19.525	108.061
Magallanes	167.431	14.157	181.588	167.392	14.144	181.536

Elaborado por el Instituto Nacional de Estadísticas.

Las proyecciones de población son estimaciones desarrolladas a partir de modelos matemáticos y demográficos que consideran la población total, sin diferenciar entre aquellos que residen en viviendas particulares y aquellos que se encuentran en viviendas colectivas. Para la encuesta Casen 2022 se realiza un ajuste a las proyecciones con el objetivo de apuntar sólo a personas pertenecientes a viviendas particulares ocupadas, que son el objeto de estudio de la encuesta. Este ajuste se realizó también en Casen en Pandemia 2020 y anteriores.

En la Tabla VI.4 se observan las estimaciones de población total y población residente en viviendas particulares a partir del Censo 2017 para el cruce región y área, la importancia de esta última estimación se explica en que la población de inferencia de la Encuesta Casen está restringida a las personas que residen en viviendas particulares ocupadas y, por tanto, las estimaciones de población elaboradas por

INE se ajustan por un factor que representa la proporción de personas que reside en viviendas particulares sobre el total de la población. En Casen 2015 y 2017, este factor se calculó con base en los resultados del Censo 2002, para cada estrato (Comuna-Área) mientras que, para Casen en Pandemia 2020 y Casen 2022, se calculó con base en información del Censo 2017. En el caso de Casen 2022 la población en viviendas particulares ocupadas a nivel de región se obtiene como la suma de los valores calculados para la desagregación región, área, sexo y grupo edad (máxima desagregación al combinar las marginales de calibración). Los niveles provinciales y comunales se obtienen directamente de multiplicar la población por el ponderador de viviendas particulares ocupadas.

Por ejemplo, según las proyecciones generadas por el INE, la estimación de población nacional al 30 de noviembre de 2022 es de 19.883.545 personas⁶⁷. Según el Censo 2017, la proporción de personas residentes en viviendas particulares es de 99,97% aproximadamente, por lo cual se estima que el total de la población que reside en viviendas particulares es 19.878.573 personas⁶⁸.

Las estimaciones poblacionales al 30 de noviembre de 2022 permiten construir ponderadores regionales, provinciales y comunales, sin embargo, dado que Casen es representativa a nivel nacional y regional, el ponderador con el cual se calculan los indicadores oficiales es el regional, considerando los ponderadores provinciales y comunales para fines de investigación.

A continuación, se describen los procedimientos para la última etapa de desarrollo de los factores de expansión de Casen 2022, que es la calibración por proyecciones de población.

VI.6.1. Ponderador de Calibración Regional

En las etapas previas de este documento se consignó un factor de expansión referido a la probabilidad de seleccionar una vivienda al interior de cada estrato de muestreo. No obstante, y a razón de contar con estimaciones de la población, en concordancia con fuentes externas, como lo son las Proyecciones de Población, se requiere obtener un factor de expansión ajustado a dichas cantidades.

La calibración es el último ajuste que se realiza al factor de expansión. Para esto, se utiliza el método *Raking* que permite incorporar información auxiliar relacionadas con la variable de interés de Casen y tiene tres propósitos, el primero de ellos es que la encuesta, a partir de los pesos calibrados, sea capaz de reproducir totales poblacionales conocidos, el segundo objetivo es alcanzar mejoras en la precisión de los parámetros de interés por medio de la obtención de estimaciones con un menor error de muestreo y

⁶⁷ Con fines analíticos, el departamento de Demografía del INE genera estimaciones poblacionales regionales por sexo y por edad y comunales por zona al 30 de noviembre del año correspondiente a la encuesta Casen, que envía al Ministerio.

⁶⁸ Las estimaciones poblacionales de residentes en viviendas particulares ocupadas se obtienen realizando el ajuste por un factor que representa la proporción de personas que residen en viviendas particulares ocupadas sobre el total de la población del Censo 2017.

finalmente, se pretende obtener una mayor precisión en las estimaciones (Alvarado, Pizarro & Guarda, 2020). Esta nueva metodología se implementa tras la realización del estudio realizado en 2020 y 2021 realizado en conjunto con la CEPAL respecto a la actualización de la metodología del factor de expansión⁶⁹.

Para obtener un factor de expansión para cada una de las personas que participan en Casen 2022, se asigna a cada una de ellas el factor de expansión de no respuesta correspondiente a la vivienda a la que pertenecen. La suma de dicho factor a nivel nacional alcanza 17,5 millones de personas, cifra estimada antes de realizar el proceso de calibración que, luego de calibrar, debe coincidir con el stock de personas totales provenientes de las proyecciones de población, ajustadas a viviendas particulares ocupadas al mes y año del levantamiento de la encuesta.

En Casen 2022 es elaborado un ponderador de Calibración Regional, también conocido como Factor de Expansión Regional (expr). Este factor se encuentra asociado a cada persona presente en la muestra y es el mismo para los diferentes integrantes de un mismo hogar. Sin embargo, se debe contemplar que como efecto del redondeo probabilístico y, además con el objeto de que los factores de expansión alcancen, con una mayor precisión, el stock de las proyecciones de población a nivel nacional y regional, es probable que existan hogares donde un integrante presente un factor con una unidad mayor o menor que el resto de sus integrantes.

En el proceso de calibración, se utiliza la información auxiliar proveniente de las proyecciones de población, según área, sexo y edad y para distintos niveles de desagregación. El periodo de levantamiento de Casen 2022 ocurre entre los meses de noviembre de 2022 y enero de 2023, por lo tanto, la fecha a la que se consideran las proyecciones corresponde al 30 de noviembre de 2022. La información auxiliar se utiliza como un vector de totales poblacionales, llamados marginales de calibración, y para Casen 2022 fueron consideradas las siguientes desagregaciones:

- Población Total, según Región por Área (Urbano/Rural).
- Población Total, según Región por Sexo.
- Población Total, según 17 grupos de edad por Sexo: [0-4], [5-9], [10-14], [15-19], [20-24], [25-29], [30-34], [35-39], [40-44], [45-49], [50-54], [55-59], [60-64], [65-69], [70-74], [75-79], [80+].

Antes de calibrar el factor de expansión, se procede a revisar los coeficientes de variación de cada una de las marginales, criterio inicial de calidad estadística, donde se espera obtener valores menores a 10% para todas las categorías. En Casen 2022 fueron revisadas dichas marginales donde se observó el mayor valor a nivel regional en la Región de Antofagasta rural, con 7,8%, explicado por la baja cantidad de

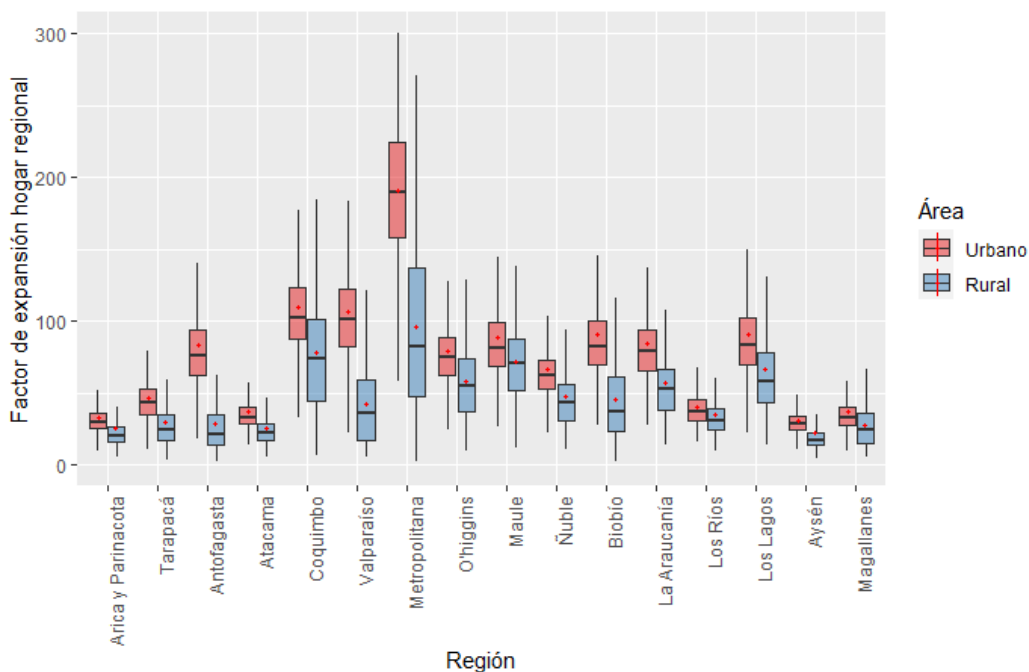
⁶⁹ Para mayor información revisar el documento Resultados de la Nueva Metodología de Calibración por Raking de los Factores de Expansión de la Encuesta Casen. Disponible en: [Nota técnica N 8: Nueva Metodología de calibración Casen](#)

unidades logradas en esa subpoblación. Ya que ningún coeficiente de variación es mayor a 10% no es necesario colapsar el nivel con otro de similares características.

A partir del proceso de calibración se obtiene una estimación de 19.878.573 personas a nivel nacional, cifra coincidente con las Proyecciones de Población, cuyo detalle por región-área se presenta en el Anexo N°7.

El Gráfico VI.3 muestra la distribución regional de los ponderadores de calibración a nivel hogar.

Gráfico VI.3. Factor de expansión hogar calibración Región – Área



Elaborado por el Instituto Nacional de Estadísticas.

VI.6.2. Ponderador de Calibración Provincial

El criterio para desarrollar el factor provincial es que se asumen válidos para todas las provincias que no fueron declaradas áreas especiales en el MMV 2020 y por lo tanto en Casen 2022. De las 56 provincias del país, 2 son consideradas áreas especiales, las cuales son: Isla de Pascua y Palena. Por otro lado, existe la provincia de la Antártica Chilena que cuenta con una sola comuna en la muestra de Casen 2022 y corresponde a Cabo de Hornos que asume la representación total de la provincia.

El Ponderador de Calibración Provincial, denominado Factor de Expansión Provincial (expp), también se asocia a cada persona en la muestra. El procedimiento de cálculo de este ponderador se realiza con base en la agregación de todas las comunas que conforman la provincia. El procedimiento se resume en cuatro pasos:

1. Ajustar las estimaciones poblacionales y estimar el total de personas en hogares particulares (N_p) en cada provincia con presencia en la muestra 2022. Las estimaciones fueron ajustadas, según la proporción de personas residentes en viviendas particulares ocupadas reportadas por el Censo de Población y Vivienda del año 2017.
2. Estimar el total de personas en hogares particulares a partir de Casen 2022. En cada vivienda entrevistada se multiplicó el ponderador corregido por no respuesta por el número de personas en la vivienda para obtener una estimación del total de personas residentes en viviendas particulares ocupadas en cada provincia en la muestra 2022.
3. Construir el ajuste a población total. Se construyó la razón \hat{R}_p , a partir del cociente entre la estimación del total de personas residentes en viviendas particulares ocupadas desarrollado, a partir de las estimaciones poblacionales (N_p , en paso 1) y el número de personas estimadas en la Provincia por la encuesta Casen 2022 (paso 2),

$$\hat{R}_p = \frac{N_p}{\sum_{h \in P} \sum_{i \in \Omega_h} \sum_{j \in \theta_{i, eleg, R}} N_{hij} \cdot w_{hij}^{NR}} \quad (22)$$

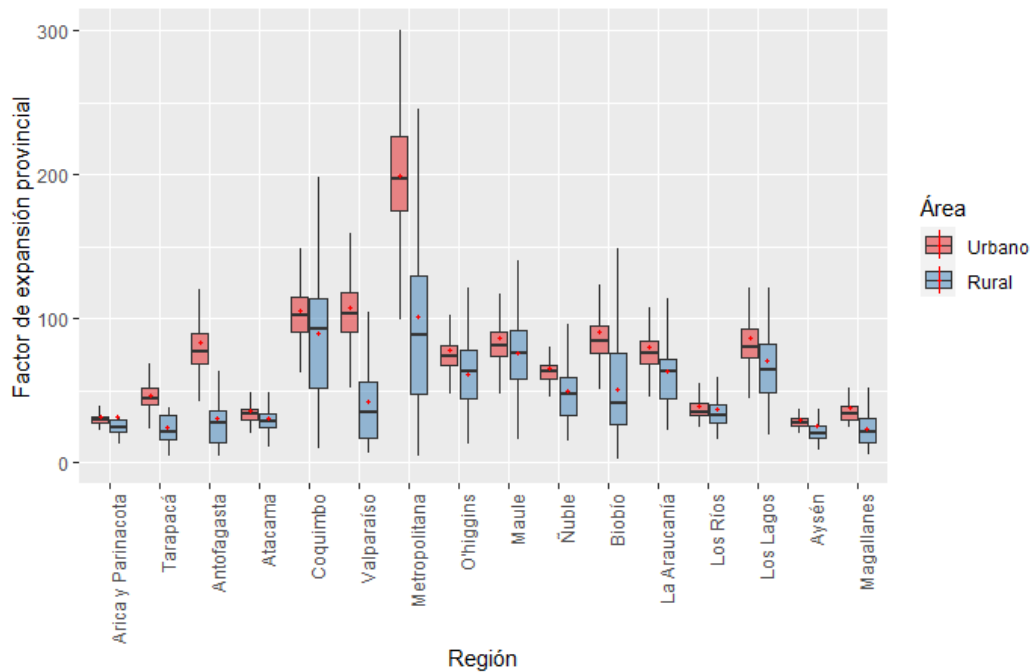
4. Construir el Ponderador de Calibración provincial. Finalmente se multiplica el Ponderador corregido por no respuesta con el ajuste de población total (paso 3).

El factor de expansión provincial tiene la siguiente expresión:

$$w_{hijk}^P = \hat{R}_p \cdot w_{hij}^{NR} \quad (237)$$

El Gráfico VI.4 muestra la distribución regional de los ponderadores de calibración provincial para el área urbana y rural donde se observa que la mayor variabilidad se encuentra tanto en el área urbana como rural de la Región Metropolitana y la menor variabilidad, en las regiones de Arica y Parinacota, Atacama y Aysén.

Gráfico VI.4. Factor de expansión calibración provincial Región – Área



Elaborado por el Instituto Nacional de Estadísticas.

VI.6.3. Ponderador de Calibración Comunal

El Ponderador de Calibración Comunal, también conocido como “Factor de Expansión Comunal” (exp_c) también se asocia a cada persona en la muestra. El procedimiento de cálculo de este ponderador puede ser resumido en cuatro pasos:

1. Ajustar las estimaciones poblacionales y estimar el total de personas en viviendas particulares ocupadas. Se utilizan las estimaciones poblacionales para las 335 comunas presentes en la muestra de Casen 2022. Las estimaciones fueron ajustadas, según la proporción de personas residentes en viviendas particulares ocupadas reportadas por el Censo de Población y Vivienda del año 2017 y estas alcanzan un total de 19.842.685. De este modo se estimó el total de personas residentes en viviendas particulares ocupadas (N_C) en las 335 comunas con presencia en la muestra 2022.
2. Estimar el total de personas en hogares particulares a partir de Casen 2022. En cada vivienda entrevistada se multiplicó el ponderador corregido por no respuesta por el número de personas en la vivienda para obtener una estimación del total de personas residentes en viviendas particulares ocupadas en cada comuna en la muestra.
3. Construir el ajuste a población total. Se construyó la razón \hat{R}_C , a partir del cociente entre la estimación del total de personas residentes en viviendas particulares ocupadas desarrollado a

partir de las estimaciones poblacionales (N_C , en paso 1) y el número de personas estimadas en la comuna por la encuesta Casen 2022 (paso 2),

$$\hat{R}_C = \frac{N_C}{\sum_{h \in C} \sum_{i \in \Omega_h} \sum_{j \in \theta_{i, eleg, R}} N_{hij} \cdot w_{hij}^{NR}} \quad (24)$$

Donde,

N_C : Es el número de personas residentes en viviendas particulares ocupadas, en la comuna C

N_{hij} : Es el número de personas encuestadas en la vivienda j , del conglomerado i , en el estrato h

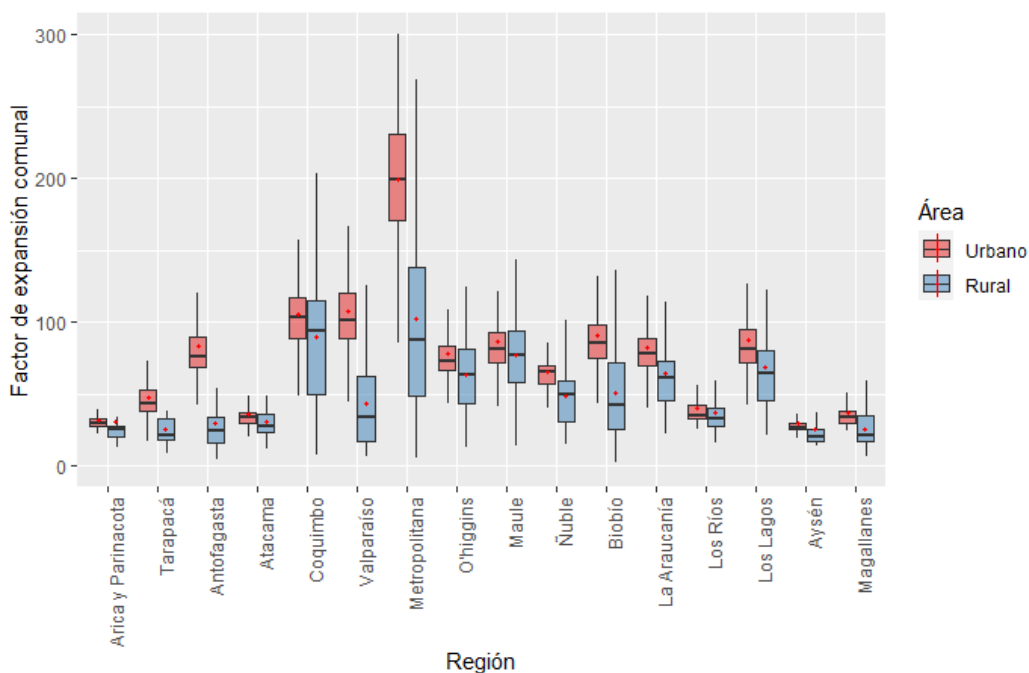
4. Construir el Ponderador de Calibración comunal. Finalmente se multiplica el Ponderador corregido por no respuesta con el ajuste de población total (paso 3).

El factor de expansión comunal tiene la siguiente expresión:

$$w_{hijk}^C = \hat{R}_C \cdot w_{hij}^{NR} \quad (258)$$

El Gráfico VI.5 muestra la distribución regional de los ponderadores de calibración comunal a nivel regional, en el área urbana y rural, para las personas de Casen 2022. En él se observa que existe una mayor variabilidad en el área rural, principalmente en las regiones Metropolitana y Valparaíso y la menor variabilidad se presenta en las regiones de Arica y Parinacota y Aysén, tanto en el área urbana como el área rural.

Gráfico VI.5. Factor de expansión calibración comunal Región – Área



Elaborado por el Instituto Nacional de Estadísticas.

VI.7. REDONDEO PROBABILÍSTICO

Finalmente, para obtener un número entero en cada uno de los factores calculados y calibrados, es decir, al factor de expansión regional, provincial y comunal, se aplica un redondeo de forma probabilística con el objetivo de que los nuevos pesos calibrados no produzcan un sesgo en las estimaciones de interés.

El redondeo probabilístico distribuye los excesos de la parte decimal aleatoriamente con base en una distribución de probabilidades entre los factores a partir de semillas de aleatorización diferentes para cada factor (regional, provincial y comunal) que permiten replicar el procedimiento y alcanzar el stock poblacional a nivel nacional; a diferencia del redondeo determinístico que aproxima el valor de los factores al entero más cercano.

Al tratarse de una distribución aleatoria de los excesos de la parte decimal es posible que existan casos donde los integrantes de un hogar que tienen el mismo factor de expansión, luego del redondeo presenten un factor de expansión diferente⁷⁰.

Por último, cabe señalar que, en el caso del factor regional, para el redondeo son utilizadas las marginales de calibración región-área, región-sexo y grupo edad-sexo, sin embargo, esto no garantiza que se alcance el stock poblacional para cada una de las categorías de las marginales y solo es posible aproximarse al total nacional como fue señalado previamente.

⁷⁰ Esta situación, también se presenta, debido a que se requiere que los factores de expansión alcancen, con una mayor precisión, el stock de las proyecciones de población a nivel nacional y regional.

VII. FACTOR DE EXPANSIÓN ORIENTACIÓN SEXUAL E IDENTIDAD DE GÉNERO

El factor de orientación sexual e identidad de género (OSIG) surge en la Encuesta Casen 2015 como respuesta a las demandas planteadas desde la sociedad civil y también, con el propósito de avanzar en el ejercicio y reconocimiento pleno de derechos y libertades individuales. Para ello, fueron incluidas, por primera vez preguntas que buscan caracterizar, de forma anónima y confidencial, la identidad de género y orientación sexual de personas de 18 años o más (MDSF, 2017).

La Encuesta Casen 2022 incluye el Módulo “Orientación sexual e identidad de género” que contiene preguntas sobre orientación sexual e identidad de género para todas las personas de 18 años o más presentes al momento de aplicar la entrevista. A diferencia del resto del cuestionario, las respuestas a estas preguntas son auto-reportadas por la persona con el objetivo de resguardar su confidencialidad. La pregunta sobre orientación sexual busca conocer la atracción afectiva o sexual que una persona puede tener hacia personas del sexo opuesto, del mismo sexo, de ambos sexos u otra alternativa distinta de las anteriores. La pregunta sobre identidad de género busca conocer cómo las personas se sienten o identifican respecto de las características que la sociedad asigna a hombres y mujeres, es decir, si se identifican con el género masculino, femenino, transmasculino, transfemenino, no binario u otro género distinto a los anteriores⁷¹.

Si se asume que todas las personas que forman parte del universo estudiado no tienen la misma probabilidad de contestar estas preguntas, es posible considerar que se podría introducir un sesgo en las estimaciones obtenidas, lo que se conoce como sesgo de selección. Esto quiere decir que si las personas de 18 años o más que no responden este módulo difieren de las que sí lo hacen, la distribución de respuestas y promedios estimados solo darán cuenta de los atributos de quienes responden y no de quienes no lo hacen.

A partir de lo expuesto en el párrafo anterior y, para corregir posibles sesgos, se construye un factor de expansión específico para las preguntas sobre orientación sexual e identidad de género⁷², que busca controlar por diferencias de características entre las personas que responden y aquellas que no lo hacen para que así, las estimaciones obtenidas reflejen adecuadamente la distribución de atributos en toda la población de 18 años o más a nivel nacional y regional, no siendo factible su uso para producir estimaciones en otros dominios de representatividad (MDSF, 2017).

⁷¹ Las preguntas sobre orientación sexual e identidad de género corresponden a variables ‘os1’ y ‘genero’ en la base de datos, respetivamente.

⁷² En la base de datos el factor de expansión de orientación sexual e identidad de género corresponde a la variable ‘expr_osig’.

VII.1. DISEÑO DEL FACTOR DE EXPANSIÓN

El cálculo de este factor de expansión se inicia con la construcción de modelos probit que determinan la probabilidad de responder las preguntas sobre orientación sexual e identidad de género. De estos modelos se obtienen nuevas probabilidades para cada individuo que corrigen el factor de expansión regional que se encuentra a nivel hogar (expr). Esto quiere decir que se genera una nueva variable en la base de datos (expr_osig) que corresponde a los pesos asignados a cada persona de 18 años o más presente al momento de aplicar la entrevista y que se aplica específicamente para el análisis de las preguntas del módulo de orientación sexual e identidad de género.

La Tabla VII.1 muestra el total de personas logradas en el levantamiento, así como el total de personas de 18 años o más y el total de menores de 18 años.

Tabla VII.1. Total de personas logradas en el levantamiento

	Casen 2022
Total levantamiento	202.231
Personas de 18 años o más	158.509
Menores de 18 años	43.722

Elaborado por el Instituto Nacional de Estadísticas.

De acuerdo con la documentación del Ministerio (2017) el objetivo de este factor de expansión es controlar el efecto de un conjunto de variables socioeconómicas y demográficas sobre la probabilidad de estar presente y responder al momento de la encuesta. A partir del supuesto de que esta probabilidad no se distribuye aleatoriamente, el factor de expansión asignará una ponderación condicionada a cada observación válida en la muestra, es decir, personas de 18 años o más presentes al momento de la encuesta que responden estas preguntas, de modo que la información recolectada pueda representar adecuadamente al total de población en el universo en estudio (personas de 18 años o más).

VII.2. VARIABLES INCLUIDAS EN EL MODELO

Para la creación del modelo solo se consideran como válidas las observaciones de personas de 18 años o más, y la variable dependiente utilizada corresponde a una variable dicotómica denominada “r_una” que posee valores 1 y 0 (es decir, responde o no responde). Para la conformación de esta variable se consideran las respuestas a las preguntas sobre orientación sexual (os1) e identidad de género (genero), asignándose valor 1 a quienes responden al menos una de estas dos preguntas, considerando como respuestas válidas las categorías de respuestas y también la alternativa “No sabe”. A partir de lo anterior la Tabla VII.2 presenta el número de personas que responden o no responden (con valores 1 y 0, respectivamente en la variable “r_una”).

Tabla VII.2. Total de personas de 18 años o más según la variable dependiente r_una

	Casen 2022
Total personas de 18 años o más	158.509
Responde al menos una pregunta del módulo (r_una = 1)	92.461
No responde el módulo (r_una = 0)	66.048

Elaborado por el Instituto Nacional de Estadísticas.

Los modelos evaluados están compuestos por variables demográficas y socioeconómicas que inciden en la presencia o ausencia de una persona informante, al momento de la entrevista. Estas variables se encuentran en el cuestionario de Casen 2022 o son construidas a partir de este y la Tabla VII.3 muestra y describe aquellas que fueron seleccionadas.

Tabla VII.3. Variables independientes seleccionadas en el modelo para el FE OSIG.

Variable	Descripción
Jh	Variable dicotómica que toma el valor 1 si la persona es jefa de hogar y 0 si no lo es.
Sexo	Variable dicotómica que toma el valor 1 si la persona es hombre y 2 si es mujer.
Uniper	Variable dicotómica que toma el valor 1 si el hogar es unipersonal y 0 si no lo es.
Tam_hogar_prom	Variable dicotómica que toma el valor 1 si el tamaño del hogar es mayor o igual al promedio y 0 si es menor.
Pareja	Variable dicotómica que toma el valor 1 si la persona tiene pareja y 0 si no la tiene.
Pareja_jh	Variable dicotómica que toma el valor 1 si la persona es esposo/a o pareja de la persona jefa de hogar y 0 si no lo es.
Region	Variable tipo factor que identifica las 16 regiones del país.
Participa	Variable dicotómica que toma el valor 1 si la persona ha participado en alguna organización o grupo organizado en los últimos 12 meses y 0 si no lo ha hecho.
Edad_50	Variable dicotómica que toma el valor 1 si la persona tiene 50 o más años y 0 si es menor.
Asiste	Variable dicotómica que toma el valor 1 si la persona asiste a un establecimiento educacional y 0 si no lo hace.
Ocupado	Variable dicotómica que toma el valor 1 si la condición de actividad de la persona es ocupada y 0 si no.
Area	Variable dicotómica que toma el valor 1 el área es urbana y 2 si es rural.
Menor_hogar	Variable dicotómica que toma el valor 1 si en la vivienda residen personas de 15 o menos años y 0 si no.

Elaborado por el Instituto Nacional de Estadísticas.

VII.3. CREACIÓN DE VEINTILES PARA EL AJUSTE

Luego de la construcción del modelo se calcula su predicción (o propensión a responder la encuesta), que toma valores en el intervalo entre 0 y 1. Estas predicciones son agrupadas en veintiles que tienen la virtud de que las personas que los componen son similares en sus características de interés y por lo tanto, los entrevistados pueden representar a los no entrevistados, al interior de cada veintil. La tabla VII.4 muestra las distribuciones de la probabilidad de respuesta a nivel de veintil, junto con la cantidad de personas en cada uno de ellos.

Tabla VII.4. Valores mínimo, máximo y promedio de la probabilidad de respuesta estimada por veintil

Veintil	Mínimo	Máximo	Promedio	n
1	0,100	0,181	0,154	7.926
2	0,181	0,227	0,203	7.926
3	0,227	0,285	0,255	7.926
4	0,285	0,332	0,309	7.926
5	0,332	0,378	0,355	7.926
6	0,378	0,423	0,401	7.926
7	0,423	0,472	0,447	7.926
8	0,472	0,523	0,496	7.926
9	0,523	0,565	0,544	7.926
10	0,565	0,605	0,584	7.925
11	0,605	0,640	0,624	7.925
12	0,640	0,667	0,654	7.925
13	0,667	0,702	0,683	7.925
14	0,702	0,743	0,724	7.925
15	0,743	0,785	0,762	7.925
16	0,785	0,825	0,806	7.925
17	0,825	0,857	0,842	7.925
18	0,857	0,898	0,876	7.925
19	0,898	0,981	0,957	7.925
20	0,981	0,996	0,989	7.925

Elaborado por el Instituto Nacional de Estadísticas.

Posteriormente, los valores obtenidos son incorporados a la base a nivel de personas donde se realiza un ajuste al ponderador de no respuesta (w_{hij}^{NR}) utilizado para la creación del factor de expansión regional original (expr). Para ello, el ponderador asociado a las personas de 18 años o más que no respondieron, debe ser redistribuido entre las personas que sí lo hicieron y por lo tanto, se calcula una razón de ajuste que viene dada por la ecuación (26). Este cálculo se realiza al interior de cada veintil utilizando el ponderador de no respuesta original a nivel de vivienda, multiplicándolo por la estimación del total de

personas de 18 años o más y luego, dividiéndolo por la estimación del total de personas de 18 años o más que responde el módulo:

$$\hat{R}_{g,R} = \frac{\sum_{J \in \theta_{g,18}} w_{hij}^{NR}}{\sum_{J \in \theta_{g,18,R}} w_{hij}^{NR}} \quad (26)$$

Donde,

- g : : Índice del grupo para la corrección de no respuesta (veintil).
- $\theta_{g,18}$: : Conjunto de personas de 18 años o más en el veintil g .
- $\theta_{g,18,R}$: : Conjunto de personas de 18 años o más en el veintil g y que responden el módulo.
- w_{hij}^{NR} : : Ponderador de no respuesta a nivel vivienda.

Una vez estimados los factores de ajuste, el ponderador de selección de orientación sexual (w_{hij}^{OS}) se expresa según la ecuación (27):

$$w_{hij}^{OS} = \hat{R}_{g,R} \cdot w_{hij}^{NR} \quad (27)$$

Este ponderador se asigna a las personas de 18 años o más que respondieron, al menos, una de las preguntas de orientación sexual e identidad de género. A las personas de 18 años o más no entrevistadas se asigna un valor blanco (“missing”).

VII.4. PONDERADOR DE CALIBRACIÓN

Luego de la obtención de los veintiles es posible obtener un factor de expansión ajustado a la posibilidad de responder las preguntas de orientación sexual e identidad de género. Este factor se calcula para cada una de las personas de 18 años y más, integrantes de los hogares y que responden el módulo.

Como todas las personas, según sus características, tienen una probabilidad distinta de responder, este factor podría ser distinto para todas las personas que componen un mismo hogar, más aún, en el proceso de calibración se debe asumir el hecho de que, según las marginales utilizadas, los pesos de personas con características distintas deben llegar a stock distintos. En ese sentido, se debe calcular un factor calibrado a nivel de persona que reproduzca las diferencias que pueden existir entre las personas, manteniendo el peso que tienen estas en la población, a diferencia del factor original, que les asigna un mismo peso a todas las personas de un mismo hogar, independiente de sus características individuales.

En este caso para la calibración se utilizó el método Raking que permite controlar diversas variables simultáneamente y mantener el diseño muestral. Para obtener el factor de expansión para cada una de las personas que respondieron este módulo, se asigna a cada una de ellas el factor de expansión de respuesta del módulo de orientación sexual e identidad de género donde, la suma para este factor a

nivel nacional es de 13.753.208. Estas cifras se obtienen antes de realizar la calibración ya que, una vez realizada, el factor de expansión debe coincidir con el stock de personas provenientes de las proyecciones de población, ajustadas a viviendas particulares ocupadas.

En el proceso de calibración, se utiliza la información auxiliar proveniente de las proyecciones de población al 30 de noviembre de 2022 para distintos niveles de desagregación. La información auxiliar corresponde a las siguientes desagregaciones:

- Población Total, según región por área (Urbano/Rural).
- Población Total, según región por sexo.
- Población Total, según 13 grupos de edad por sexo: [18-24], [25-29], [30-34], [35-39], [40-44], [45-49], [50-54], [55-59], [60-64], [65-69], [70-74], [75-79], [80+].

Antes de realizar la calibración, se procede a revisar los coeficientes de variación de cada una de las marginales donde se espera obtener valores menores a 10% para todas las categorías. En este caso el coeficiente de variación más alto se encuentra en la marginal de región por área donde en el área rural de la región de Antofagasta alcanza un 11,7%.

A partir del proceso de calibración se obtiene una estimación de 15.408.912 personas a nivel nacional, cifra coincidente con las Proyecciones de Población de personas de 18 años o más, cuyo detalle a nivel de región y área se presenta en el Anexo N°7. El ponderador de Calibración o factor de expansión específico para las preguntas sobre orientación sexual e identidad de género (variable “expr_osig” en la base de datos Casen 2022) se asocia a cada persona en la muestra que respondió al menos una de las preguntas sobre identidad de género y orientación sexual. Quienes no respondieron ninguna de estas preguntas tienen un valor blanco o “missing”.

Al igual que en el caso del factor de expansión regional, provincial y comunal se aplica un redondeo probabilístico donde se utilizan las marginales de calibración región-área, región-sexo y grupo edad-sexo. Sin embargo, es importante considerar que, tal como se mencionó en el caso del factor de expansión regional, esto no es garantía de que se alcance el stock poblacional para cada una de las categorías de las marginales y solo es posible aproximarse al total nacional.

VIII. ESTIMACIÓN DE VARIANZA COMPLEJA

Los métodos exactos de estimación de la varianza en diseños de muestras estándares, cuando son aplicables, son la mejor forma de estimarla. Sin embargo, los diseños muestrales empleados en la mayoría de las encuestas de hogares, como es el caso de Casen, son más complejos que el muestreo aleatorio simple o estratificado. A esto se le agrega que se utilizan estimadores de razón, que no son funciones lineales simples de los valores observados, por lo que la varianza de muestreo, no siempre puede expresarse por una fórmula de forma cerrada, como por ejemplo la estimación de la media de la muestra en un muestreo aleatorio simple o estratificado.

Actualmente, distintos softwares estadísticos, como R, SPSS, STATA, incorporan módulos estadísticos para el cálculo de estimaciones, varianzas y coeficientes de variación, incorporando los típicos efectos de diseño cuando se trata de muestreo complejo, por lo general polietápico.

En los apartados siguientes se exponen distintos métodos para obtener estimaciones aproximadas de la varianza, tanto convencionales como métodos más prácticos donde se busca que dicha estimación no sea subestimada, en caso de estratos donde se logró un solo conglomerado, o sobreestimada, en caso de conglomerados demasiado heterogéneos en cuanto al número de viviendas logradas.

VIII.1. MÉTODOS CONVENCIONALES PARA LA ESTIMACIÓN DE LA VARIANZA

Existen principalmente cuatro métodos convencionales para estimar las varianzas o errores muestrales de estimaciones de parámetros de interés basadas en datos de una encuesta⁷³. Estos son:

1. Métodos exactos;
2. Método del conglomerado último;
3. Aproximaciones por linealización;
4. Técnicas de replicación;

Para la estimación de parámetros de interés, así como de sus estadísticos asociados, tales como la varianza o error muestral, en general se utiliza una combinación de métodos. A continuación, se describe brevemente en qué consiste cada uno de ellos.

⁷³ Para más información ver Kish y Frankel (1974), Wolter (1985) y Lehtonen y Pahkinen (1995).

VIII.1.1. Métodos exactos

Como se ha mencionado, los métodos exactos de estimación de la varianza, cuando son aplicables (en diseños muestrales simples) son la mejor forma de estimarla. Sin embargo y, como se mencionó en párrafos anteriores, el diseño muestral empleado en Casen 2022 corresponde a un diseño complejo lo que implica que las estimaciones de parámetros de interés no son funciones lineales simples de los valores observados, por lo tanto, la varianza de muestreo no es fácil de implementar acudiendo a estos métodos.

VIII.1.2. Método del conglomerado último

El método del conglomerado último para estimar la varianza (Hansen, Hurwitz & Madow, 1953) se emplea para estimar la varianza de estimaciones de parámetros de interés basadas en una muestra obtenida a partir de un diseño muestral complejo, generalmente polietápico o polifásico.

Según este método, el conglomerado último consiste en considerar a la unidad primaria de muestreo (UPM) como la unidad de información última que conglojera la información de todas las unidades de muestreo de etapas posteriores dentro de ella.

Por ejemplo, si la unidad primaria de muestreo corresponde a un conglomerado que contiene unidades secundarias de muestreo como las viviendas y, a su vez, estas unidades secundarias de muestreo contienen unidades terciarias, como las personas, y se quiere estimar algún parámetro de interés relacionado con la ocupación, entonces cada UPM tendrá asociado el total de ocupados, como la suma de todas las personas pertenecientes a todas las viviendas contenidas en la UPM. Bajo este método, la estimación de la varianza se calcula utilizando sólo totales entre UPM, sin tener que calcular los componentes de la varianza en cada etapa de selección.

VIII.1.3. Métodos de linealización

El método de linealización es ampliamente utilizado porque puede aplicarse a casi todos los diseños muestrales y a cualquier parámetro que pueda linealizarse o expresarse como una función lineal de parámetros comunes, como medias o totales, cuyos coeficientes se extraen de derivadas parciales necesarias para la expansión en series de Taylor.

La linealización se aplica generalmente para la estimación de razones o cocientes de dos variables que se desarrollan en una serie de Taylor infinita centrada en el valor esperado (estimado) del numerador y el valor esperado (estimado) del denominador.

Seguidamente, el estimador no lineal se somete a aproximación algebraica conservando sólo los primeros términos de la serie infinita de Taylor, lo que produce una expresión que es una función lineal

de los datos muestrales, es decir, el estimador no lineal (de razón, cociente, entre otros) se ha “linealizado”. Obtenido el estimador linealizado, es posible calcular directamente la varianza estimada de la función linealizada (incluidos los términos de covarianza pertinentes) a partir de métodos exactos descritos anteriormente (Lohr, 1999).

VIII.1.4. Método de replicación

Este método consiste en tomar submuestras o réplicas de la muestra completa. Primero, se estima con toda la muestra el parámetro de interés $\hat{\theta}_0$, de acuerdo con el diseño probabilístico. Enseguida, se generan réplicas o se seleccionan submuestras, de manera que cada réplica refleje el plan de muestreo, los ajustes y procedimientos del diseño original, obteniendo distintas estimaciones del parámetro de interés.

Existen varias técnicas para implementar este método y, en algunos softwares, se incluyen los módulos de selección de réplicas, donde los más utilizados son los grupos aleatorios; la replicación repetida equilibrada (BRR); la replicación jackknife (JK1, JK2, y JK_n) y bootstrap.

Por ejemplo, si se crean K réplicas a partir de una muestra, cada una con estimaciones $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ de un parámetro $\hat{\theta}$ y además, la estimación basada en la muestra completa es $\hat{\theta}_0$. La estimación de la varianza basada en la replicación por bootstrap viene dada por: $V(\hat{\theta}) = \frac{1}{K} \sum_{r=1}^K (\hat{\theta}_r - \hat{\theta}_0)^2$, donde K es el número de réplicas.

Actualmente, la mayoría de los softwares estadísticos tienen módulos especiales para el tratamiento de los datos enfocados a muestras complejas, entendiendo por muestra compleja, a aquella que incorpora variables de ponderación, de estratificación y de conglomeración para las estimaciones de los principales estadísticos asociados a algún parámetro de interés y junto con la estimación de sus varianzas.

VIII.2. VARIANZA EN MUESTRAS COMPLEJAS

Un diseño de muestreo complejo se origina cuando existe más de una etapa de selección. Esta selección en más de una etapa suele ser, además estratificada, lo que añade otro grado de complejidad.

Bajo la estratificación, en un muestreo bietápico, las unidades muestrales se suelen llamar conglomerados o unidades de conglomeración, que corresponden a conglomeraciones de las últimas unidades de muestreo, sobre las cuales se obtienen las estimaciones de los parámetros de interés.

La encuesta Casen 2022, como también sus predecesoras, presenta un diseño muestral que se define como probabilístico, estratificado y bietápico. En el muestreo estratificado la estimación de la varianza

total se obtiene mediante la suma de las varianzas estimadas en cada estrato y la varianza en cada estrato se estima a partir de la variabilidad de la característica de interés entre los conglomerados.

Una aproximación para la estimación de varianza comúnmente utilizada en encuestas complejas, de muestreo polietápico por conglomerados como Casen, corresponde al método de estimación de varianza por conglomerado último (EVCU) o por reemplazo (WR por sus siglas en inglés). Este método permite capturar cierto grado de covarianza por medio de la correlación intraclásica de las unidades al interior de los conglomerados, aun cuando las estimaciones se basen en muestras sin traslape. También es posible la captura de covarianza a través de la correlación intraclásica de las unidades al interior de los estratos o dominios comunes (INE, 2022).

El muestreo complejo en las etapas posteriores queda automáticamente cubierto de manera adecuada con la aproximación “con reemplazo” para la primera etapa de muestreo presentada previamente. De hecho, pocos paquetes de software tienen la capacidad de incluir por separado todas las etapas de muestreo en la estimación de varianzas, en los casos en los que no se practica la aproximación con reemplazo en la primera etapa.

Estas estimaciones con EVCU, pueden efectuarse indistintamente por linealización mediante el desarrollo en Series de Taylor o con una técnica de replicación. Con este enfoque, las UPM se tratan como si se hubieran seleccionado con reemplazo entre los estratos de la primera etapa. En tal caso, cada elemento o individuo solo tiene que identificarse por el estrato de la primera etapa y por la UPM (dentro del estrato) de la que se haya seleccionado.

Para efectos de la estimación de varianzas no se necesita información sobre las etapas de muestreo por debajo del nivel de las UPM. De este modo la descripción del plan de muestreo, necesario para la estimación de varianzas con diseños complejos, se simplifica de forma que se asemeja a un muestreo por conglomerados estratificados en una etapa, esto es, una muestra estratificada de conglomerados finales completamente enumerados. Este enfoque por conglomerados finales produce una buena aproximación para la estimación de la varianza, siempre que sea razonable partir de la hipótesis de una primera etapa con reemplazo.

Así, cuando el plan de muestreo se describe como EVCU, solo se necesitan tres variables del diseño de la encuesta para la estimación de varianzas:

1. La variable de ponderación de la muestra o factor de expansión designado “expr”, “expc” o “expp”, según si el nivel territorial de expansión es regional, comunal o provincial, respectivamente.
2. La variable de estratificación (o variable de pseudo-estratificación) designada como varstrat en Casen 2022.
3. La variable UPM (o pseudo-conglomerado) designada como varunit en Casen 2022.

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

VIII.3. ALGORITMO DE CÁLCULO EN MUESTRAS COMPLEJAS

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

Si una vivienda se incluye en la muestra con una probabilidad $P_{hi}(j)$, su ponderación básica, viene dada por $w_{hij} = 1/P_{hi}(j)$, es decir, el inverso de la probabilidad de selección y se denomina factor de expansión teórico, de acuerdo al diseño probabilístico. Este ponderador es igual para todas las viviendas dentro del conglomerado (UPM) y, por lo tanto, para todas las personas dentro de ellas.

Como se presenta en el capítulo VI, este factor se ajusta por elegibilidad desconocida y también por no-respuesta de las viviendas, además se calibra para alcanzar los stocks poblacionales de personas según las proyecciones de población en viviendas particulares ocupadas.

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

$z_{hij} = w_{hij} \cdot y_{hij}$ es el valor de la variable y observada en la vivienda j expandida a la población.

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

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

n_h es el número de UPM seleccionadas en el estrato h en la muestra Casen 2022.

N_h es el número de UPM en el estrato h , contenidas en el marco muestral utilizado en Casen 2022.

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

$S_h^2 = \frac{n_h}{n_h-1} \sum_{i=1}^{n_h} (z_{hi} - \bar{z}_h)^2$ es la cuasivarianza del estrato h que corresponde a la variabilidad entre los conglomerados de la variable observada expandida en ellos. Asumiendo muestreo con reemplazo, se le agrega el factor de corrección por finitud $(1 - f_h) = (1 - n_h/N_h)$. Para muestreo sin reemplazo, se utiliza una expresión más complicada, que se puede estudiar en Yates y Grundy (1953).

VIII.3.1. Algoritmo de cálculo de los estimadores

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

El total estimado en el estrato h de la variable de interés Y es:

$$\hat{Y}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} z_{hij} = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \quad (28)$$

La estimación total a nivel agregado de la variable Y , es decir, la estimación agregada sobre todos los estratos y que forman algún nivel de estimación es:

$$\hat{Y} = \sum_{h=1}^H \hat{Y}_h \quad (29)$$

La estimación del número de unidades en la población en el estrato h es:

$$\hat{N}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \quad (30)$$

La estimación del número de unidades en la población en el nivel de estimación requerido, como agregación sobre todos los estratos de muestreo se expresa:

$$\hat{N} = \sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \quad (31)$$

La media estimada de la variable de interés Y , en el estrato h es:

$$\hat{\bar{Y}}_h = \frac{1}{\hat{N}_h} \hat{Y}_h \quad (32)$$

La media estimada de la variable de interés Y , en el nivel de estimación requerido es:

$$\hat{Y} = \frac{1}{\hat{N}} \hat{Y} \quad (33)$$

El estimador de razón entre las variables Y y X , en el estrato h es:

$$\hat{R}_h = \frac{\hat{Y}_h}{\hat{X}_h} = \frac{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij}}{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot x_{hij}} = \frac{\hat{Y}_h}{\hat{X}_h} \quad (34)$$

Por lo tanto,

$$R_h = \frac{Y_h}{X_h} = \frac{\bar{Y}_h}{\bar{X}_h} \quad (35)$$

El estimador de razón entre las variables Y y X , en el nivel de estimación requerido del estudio es:

$$\hat{R} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij}}{\sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot x_{hij}} = \frac{\hat{Y}}{\hat{X}} \quad (36)$$

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

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

$$\hat{Y}_{R_h} = \hat{R}_h \cdot X_h \text{ y también } \hat{Y}_{R_h} = \hat{R}_h \cdot \hat{X}_h \quad (37)$$

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

$$\hat{Y}_R = \hat{R} \cdot X \text{ y también } \hat{Y}_R = \hat{R} \cdot \hat{X} \quad (38)$$

Para las fórmulas (34), (35), (36), (37) y (38), se asume que el total X o la media \bar{X} son conocidas o dadas como fijas⁷⁴. Sin embargo, generalmente se desconoce el valor poblacional y entonces son estimados por \hat{X} y $\hat{\bar{X}}$ en la expresión (35), o son estimadas por \hat{X}_h y $\hat{\bar{X}}_h$ en la expresión (36).

⁷⁴ Por ejemplo, en la calibración a las proyecciones de población de personas, los stocks poblacionales son estimaciones obtenidas mediante modelos son considerados fijos, como los verdaderos valores en la población. Además, si la variable es dicotómica, X coincide con N como asimismo \hat{X} coincide con \hat{N} .

VIII.3.2. Algoritmo de cálculo de las varianzas en muestras complejas

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

Definida la estratificación en los estratos h en la encuesta Casen, la varianza del estimador del total se puede expresar como:

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \hat{V}(\hat{Y}_h) = \sum_{h=1}^H S_h^2 = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (z_{hi} - \bar{z}_h)^2 \quad (39)$$

Desglosando z_{hi} y \bar{z}_h por las expresiones definidas previamente, queda:

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\sum_{j=1}^{m_{hi}} z_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} z_{hi} \right)^2 \quad (40)$$

Reemplazando $z_{hij} = w_{hij} \cdot y_{hij}$ y $z_{hi} = \sum_{j=1}^{m_{hi}} z_{hij}$, estas expresiones pueden finalmente ser escritas como:

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

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

La varianza o cuasivarianza estimada de la variable de interés expandida en el estrato h es:

$$\hat{V}(\hat{Y}_h) = \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \right)^2 \quad (42)$$

La estimación de la varianza del Total de la variable de interés Y , en el nivel de estimación requerido es:

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \right)^2 \quad (43)$$

La estimación de la varianza en el estrato h , de la Media de la variable de interés Y es:

$$\hat{V}(\hat{Y}_h) = \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left[\left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \right) / \hat{N}_h \right]^2 \quad (44)$$

La estimación de la varianza de la media de la variable de interés Y , según el nivel de estimación requerido es:

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left[\left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} \right) / \hat{N} \right]^2 \quad (45)$$

La varianza del estimador de razón entre las variables Y y X , en el estrato h es:

$$\hat{V}(\hat{R}_h) = \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left[\left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \hat{R}_h \cdot \sum_{j=1}^{m_{hi}} w_{hij} \cdot x_{hij} \right) / \hat{X}_h \right]^2 \quad (46)$$

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

La varianza del estimador de razón entre las variables Y y X , en el nivel de estimación requerido es:

$$\hat{V}(\hat{R}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left[\left(\sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij} - \hat{R} \cdot \sum_{j=1}^{m_{hi}} w_{hij} \cdot x_{hij} \right) / \hat{X} \right]^2 \quad (47)$$

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

La estimación de la varianza del total y la media de la variable Y , en el estrato h , mediante el estimador de razón es:

$$\hat{V}(\hat{Y}_{R_h}) = \hat{V}(\hat{R}_h \cdot X_h) = \hat{V}(\hat{R}_h) \cdot \hat{X}_h^2 \quad ; \quad \hat{V}(\hat{Y}_{\hat{R}_h}) = \hat{V}(\hat{R}_h \cdot \bar{X}_h) = \hat{V}(\hat{R}_h) \cdot \hat{X}_h^2 \quad (48)$$

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

$$\hat{V}(\hat{Y}_R) = \hat{V}(\hat{R} \cdot X) = \hat{V}(\hat{R}) \cdot \hat{X}^2 \quad ; \quad \hat{V}(\hat{Y}_{\hat{R}}) = \hat{V}(\hat{R} \cdot \bar{X}) = \hat{V}(\hat{R}) \cdot \hat{X}^2 \quad (49)$$

VIII.4. VARIABLES QUE IDENTIFICAN EL DISEÑO MUESTRAL COMPLEJO

En la muestra de la encuesta Casen 2022, la variable que identifica los estratos utilizados en la selección de la muestra se denomina "estrato" y la variable que identifica los conglomerados seleccionados en la muestra se denomina "cod_upm". En total, en la muestra seleccionada existen 763 estratos⁷⁵ y 12.545 conglomerados.

Para que sea factible la estimación de la varianza en un muestreo estratificado, se deben seleccionar como mínimo 2 conglomerados por estrato. Para evitar estimaciones de varianza volátiles, se recomienda seleccionar un número mínimo de unidades muestrales (viviendas) por conglomerado. El diseño de Casen 2022 establece, para la muestra, un mínimo de 2 conglomerados por estrato⁷⁶ y una cantidad de viviendas por UPM entre 8 y 10 para el área urbana y; entre 10 y 12 para el área rural.

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

Siguiendo las mejores prácticas en la materia, se optó por esta última opción, con la finalidad de poner a disposición de los investigadores variables que permitan identificar las características del diseño complejo de Casen 2022 y producir estimaciones de varianza más estables.

Las estimaciones oficiales de la varianza compleja de Casen 2022 han sido desarrolladas utilizando pseudo-estratos y pseudo-conglomerados. A continuación, se describen los procedimientos seguidos para la construcción de estas variables.

VIII.4.1. Creación de pseudo-estratos (varstrat)

La creación de pseudo-estratos de varianza, que en la base de datos corresponde a la variable varstrat, tiene como objetivo corregir el problema de estimación de varianza que generan los estratos unitarios (1

⁷⁵ Durante la selección de viviendas la cantidad de estratos se redujo de 764 a 763 ya que la UPM presente en San José de Maipo rural no contaba con suficientes viviendas disponibles para ser seleccionadas.

⁷⁶ En la muestra de Casen 2022 existen 9 estratos unitarios definidos desde el diseño ya que presentan una UPM en el marco.

conglomerado por estrato), el cual se traduce en una subestimación de la varianza de cualquier parámetro de interés.

De los 763 estratos originales⁷⁷ 9 poseen solo una UPM y se ubican en el área rural, por lo que se hace necesario agruparlos con otros estratos de similares características. Luego de esta agrupación se mantienen 748 estratos originales y el resto sufre modificaciones ya que los 9 estratos unitarios, se agrupan con otros estratos según algunos criterios, manteniendo previamente el ordenamiento jerárquico en cuanto a la división político administrativa y el nivel socioeconómico. El procedimiento es como sigue:

1. Si el estrato tiene 2 o más UPM no se modifica.
2. Si es un estrato unitario, este se combina con un estrato de la misma provincia, área y nivel socioeconómico.
3. En caso de que lo anterior no sea posible, el estrato unitario se combina con un estrato de la misma provincia, área y diferente nivel socioeconómico⁷⁸.

A partir de un total de 763 estratos muestrales originales se construyeron 755 “varstrat”, de forma tal que cada uno contuviera al menos dos conglomerados. El solo agrupar estratos cuando se logra aplicar la encuesta en únicamente un conglomerado, permite reflejar de forma más adecuada el diseño original de la encuesta. En la Tabla VIII.1 se observa el total de estratos y varstrat por región.

VIII.4.2. Creación de pseudo-conglomerados (varunit)

La creación de pseudo-conglomerados, que en la base de datos corresponde a la variable varunit, tiene como objetivo estabilizar las estimaciones de varianza, la cual tiende a incrementarse cuando las estimaciones se basan en conglomerados con muy pocas unidades muestrales. En versiones previas de Casen tanto por diseño como por falta de respuesta existen conglomerados pequeños. Sin embargo, en esta versión, dada la implementación del MMV 2020, no existen conglomerados pequeños por diseño. Los pseudo-conglomerados, en adelante varunit, son una re-agrupación de los 12.463⁷⁹ conglomerados en los que se lograron viviendas, es decir, al menos una vivienda al interior de estos respondió la encuesta Casen 2022.

⁷⁷ Los estratos están conformados por la agrupación de comuna-área-nivel socioeconómico donde aplique. Los 763 estratos tienen al menos una UPM con viviendas logradas.

⁷⁸ Existe un varstrat creado luego de la construcción de los varunit. El estrato de Lo Barnechea rural NSE medio-alto (compuesto por 2 UPM) posee una UPM donde solo existe una vivienda lograda por lo que fue necesario crear un varunit uniéndola a la otra UPM del estrato conformándose un nuevo estrato unitario. Para subsanar ese problema, se combinó este estrato con un varstrat formado previamente, el que finalmente quedó compuesto por 3 estratos de la misma provincia, área, pero diferente NSE.

⁷⁹ El número original de conglomerados seleccionados en la muestra Casen 2022 es de 12.545 UPM.

Respecto a la creación de los varunit, en esta versión de Casen no se establece un mínimo y máximo de viviendas como sí ocurrió en versiones anteriores. En Casen 2022 solo se agrupan aquellas UPM que tienen una vivienda lograda y para ello se utilizan los siguientes criterios:

1. Se une a la UPM más cercana geográficamente dentro del mismo varstrat;
2. Si existen dos UPM igualmente cercanas, se une a aquella que posee menos viviendas logradas;
3. Cuando existen UPM con 1 vivienda lograda contiguas, estas se unen formándose varunit de hasta 4 viviendas.

En estos tres casos se procura respetar el número máximo de viviendas por conglomerado establecido en el diseño.

A partir de un total de 12.463 conglomerados en los que se lograron viviendas de la muestra Casen, se crearon 12.062 varunit, los cuales agrupan entre 2 y 12 viviendas entrevistadas. Esto, al igual que en el caso de los varstrat, refleja de forma más adecuada el diseño original de la encuesta.

Tabla VIII.1. Total de estratos, conglomerados, viviendas, varstrat y varunit, según región por área geográfica, para la muestra lograda en Casen 2022

Región	Área	N° estratos Casen 2022	N° de UPM	N° de viviendas	Varstrat	Varunit
Nivel País		763	12.463	70.751	755	12.062
Total País	Urbano	444	10.526	55.957	444	10.140
	Rural	319	1.937	14.794	311	1.922
Arica y Parinacota	Urbano	4	335	2.311	4	330
	Rural	4	29	262	3	29
Tarapacá	Urbano	8	477	2.490	8	456
	Rural	7	30	245	6	29
Antofagasta	Urbano	11	520	2.643	11	487
	Rural	8	38	300	7	37
Atacama	Urbano	11	447	2.631	11	433
	Rural	9	55	450	9	54
Coquimbo	Urbano	18	412	2.186	18	399
	Rural	16	88	723	16	88
Valparaíso	Urbano	50	1.135	5.923	50	1.103
	Rural	35	188	1.365	35	188
Metropolitana	Urbano	99	2.634	11.873	99	2.445
	Rural	28	143	958	26	139
O'Higgins	Urbano	36	589	3.422	36	585
	Rural	34	195	1.535	34	195
Maule	Urbano	42	633	3.590	42	623
	Rural	31	199	1.514	31	197
Ñuble	Urbano	24	376	2.093	24	370
	Rural	22	164	1.197	22	163
Biobío	Urbano	48	997	5.558	48	979
	Rural	33	175	1.456	31	174
La Araucanía	Urbano	38	597	3.180	38	577
	Rural	35	217	1.700	35	217
Los Ríos	Urbano	15	469	2.642	15	463
	Rural	13	164	1.200	13	164
Los Lagos	Urbano	30	481	2.638	30	471
	Rural	27	156	1.268	27	155
Aysén	Urbano	5	143	1.058	5	143
	Rural	8	64	385	8	61
Magallanes	Urbano	5	281	1.719	5	276
	Rural	9	32	236	8	32

Elaborado por el Instituto Nacional de Estadísticas.

La Tabla VIII.2 muestra que la cantidad de viviendas por varunit se encuentra entre 2 y 12, concentrándose 60,9% de ellos entre 2 y 6 viviendas.

Tabla VIII.2. Frecuencia de varunit según número de viviendas que lo componen

N° viviendas que responden	Frecuencia	Porcentaje	Porcentaje acumulado
2	607	5,0%	5,0%
3	947	7,9%	12,9%
4	1.421	11,8%	24,7%
5	2.022	16,8%	41,4%
6	2.344	19,4%	60,9%
7	2.285	18,9%	79,8%
8	1.520	12,6%	92,4%
9	511	4,2%	96,6%
10	381	3,2%	99,8%
11	14	0,1%	99,9%
12	10	0,1%	100,0%

Elaborado por el Instituto Nacional de Estadísticas.

VIII.5. RESULTADOS DE LA SITUACIÓN DE POBREZA POR INGRESOS

En Casen 2022, al igual que en versiones anteriores, las estimaciones y márgenes de error de la situación de pobreza por ingresos, fueron determinadas utilizando los parámetros varstrat, varunit y expr (factor de expansión regional) siendo estos implementados mediante cálculos realizados en paralelo utilizando los programas R y STATA.

La Tabla VIII.3 muestra la estimación del porcentaje de población en situación de pobreza por ingresos a nivel nacional, por área geográfica urbana y rural y a nivel regional, junto con sus principales estadísticos (error estándar, intervalo de confianza y coeficiente de variación).

Tabla VIII.3. Estimación del porcentaje de la población en situación de pobreza por ingresos, según zona y región

Región	Situación de pobreza por ingresos	Estimación	Error estándar	95% Intervalo de confianza		Coeficiente de variación
				Inferior	Superior	
Nacional	Pobreza extrema	2,0%	0,1%	1,9%	2,2%	3,8%
	Pobreza no extrema	4,5%	0,1%	4,3%	4,7%	2,6%
	Total pobreza	6,5%	0,1%	6,2%	6,8%	2,2%
	No pobre	93,5%	0,1%	93,2%	93,8%	0,2%
Urbano	Pobreza extrema	1,9%	0,1%	1,7%	2,0%	4,4%
	Pobreza no extrema	4,2%	0,1%	4,0%	4,5%	3,1%
	Total pobreza	6,1%	0,2%	5,8%	6,4%	2,5%
	No pobre	93,9%	0,2%	93,6%	94,2%	0,2%
Rural	Pobreza extrema	3,0%	0,2%	2,7%	3,4%	6,5%
	Pobreza no extrema	6,9%	0,3%	6,3%	7,5%	4,5%
	Total pobreza	9,9%	0,4%	9,2%	10,7%	3,9%
	No pobre	90,1%	0,4%	89,3%	90,8%	0,4%
Arica y Parinacota	Pobreza extrema	2,7%	0,5%	2,0%	3,8%	16,4%
	Pobreza no extrema	6,4%	0,7%	5,2%	7,9%	10,5%
	Total pobreza	9,2%	0,9%	7,5%	11,2%	10,2%
	No pobre	90,8%	0,9%	88,8%	92,5%	1,0%
Tarapacá	Pobreza extrema	4,5%	1,1%	2,8%	7,1%	23,7%
	Pobreza no extrema	6,5%	0,6%	5,5%	7,8%	8,9%
	Total pobreza	11,0%	1,3%	8,8%	13,8%	11,5%
	No pobre	89,0%	1,3%	86,2%	91,2%	1,4%
Antofagasta	Pobreza extrema	2,8%	0,4%	2,1%	3,7%	14,9%
	Pobreza no extrema	4,9%	0,5%	4,0%	5,9%	9,7%
	Total pobreza	7,6%	0,7%	6,4%	9,0%	8,7%
	No pobre	92,4%	0,7%	91,0%	93,6%	0,7%
Atacama	Pobreza extrema	2,8%	0,4%	2,1%	3,8%	14,3%
	Pobreza no extrema	5,4%	0,6%	4,3%	6,7%	11,4%
	Total pobreza	8,2%	0,8%	6,8%	9,8%	9,2%
	No pobre	91,8%	0,8%	90,2%	93,2%	0,8%
Coquimbo	Pobreza extrema	2,5%	0,4%	1,9%	3,3%	14,1%
	Pobreza no extrema	5,4%	0,5%	4,5%	6,5%	9,7%
	Total pobreza	7,9%	0,6%	6,8%	9,3%	8,0%
	No pobre	92,1%	0,6%	90,7%	93,2%	0,7%
Valparaíso	Pobreza extrema	1,9%	0,2%	1,6%	2,4%	11,2%
	Pobreza no extrema	4,6%	0,4%	4,0%	5,4%	7,8%
	Total pobreza	6,6%	0,4%	5,8%	7,5%	6,6%
	No pobre	93,4%	0,4%	92,5%	94,2%	0,5%
Metropolitana	Pobreza extrema	1,3%	0,1%	1,1%	1,6%	9,6%
	Pobreza no extrema	3,1%	0,2%	2,7%	3,6%	7,0%
	Total pobreza	4,4%	0,2%	4,0%	4,9%	5,6%
	No pobre	95,6%	0,2%	95,1%	96,0%	0,3%
O'Higgins	Pobreza extrema	2,3%	0,3%	1,8%	2,9%	12,9%

Región	Situación de pobreza por ingresos	Estimación	Error estándar	95% Intervalo de confianza		Coeficiente de variación
				Inferior	Superior	
	Pobreza no extrema	4,8%	0,4%	4,1%	5,5%	7,8%
	Total pobreza	7,0%	0,5%	6,2%	8,0%	6,6%
	No pobre	93,0%	0,5%	92,0%	93,8%	0,5%
Maule	Pobreza extrema	2,6%	0,3%	2,0%	3,3%	11,9%
	Pobreza no extrema	6,1%	0,4%	5,3%	7,0%	6,9%
	Total pobreza	8,7%	0,5%	7,7%	9,7%	5,9%
	No pobre	91,3%	0,5%	90,3%	92,3%	0,6%
Ñuble	Pobreza extrema	4,2%	0,4%	3,4%	5,2%	10,6%
	Pobreza no extrema	7,9%	0,6%	6,7%	9,2%	8,0%
	Total pobreza	12,1%	0,8%	10,6%	13,7%	6,4%
	No pobre	87,9%	0,8%	86,3%	89,4%	0,9%
Biobío	Pobreza extrema	2,3%	0,2%	2,0%	2,8%	8,7%
	Pobreza no extrema	5,2%	0,3%	4,6%	5,9%	6,7%
	Total pobreza	7,5%	0,4%	6,8%	8,4%	5,3%
	No pobre	92,5%	0,4%	91,6%	93,2%	0,4%
La Araucanía	Pobreza extrema	3,3%	0,3%	2,8%	4,0%	9,4%
	Pobreza no extrema	8,3%	0,5%	7,3%	9,3%	6,4%
	Total pobreza	11,6%	0,6%	10,4%	12,9%	5,5%
	No pobre	88,4%	0,6%	87,1%	89,6%	0,7%
Los Ríos	Pobreza extrema	1,3%	0,2%	0,9%	1,7%	16,7%
	Pobreza no extrema	4,6%	0,4%	3,9%	5,5%	9,1%
	Total pobreza	5,9%	0,5%	5,0%	6,9%	8,2%
	No pobre	94,1%	0,5%	93,1%	95,0%	0,5%
Los Lagos	Pobreza extrema	2,1%	0,3%	1,6%	2,8%	14,0%
	Pobreza no extrema	4,8%	0,5%	3,9%	5,9%	10,5%
	Total pobreza	7,0%	0,6%	5,9%	8,2%	8,3%
	No pobre	93,0%	0,6%	91,8%	94,1%	0,6%
Aysén	Pobreza extrema	1,1%	0,3%	0,6%	2,0%	32,0%
	Pobreza no extrema	3,0%	0,8%	1,7%	5,0%	26,7%
	Total pobreza	4,0%	0,8%	2,7%	6,0%	20,8%
	No pobre	96,0%	0,8%	94,0%	97,3%	0,9%
Magallanes	Pobreza extrema	1,0%	0,3%	0,6%	1,9%	30,8%
	Pobreza no extrema	2,4%	0,7%	1,3%	4,3%	29,7%
	Total pobreza	3,4%	0,9%	2,1%	5,6%	24,8%
	No pobre	96,6%	0,9%	94,4%	97,9%	0,9%

Elaborado por el Instituto Nacional de Estadísticas.

En la Tabla VIII.4 se observa la estimación del porcentaje de pobreza por ingresos en los hogares a nivel nacional, por área geográfica urbana y rural y a nivel regional, junto con sus principales estadísticos (error estándar, intervalo de confianza y coeficiente de variación).

Tabla VIII.4. Estimación del porcentaje de hogares en situación de pobreza por ingresos, según zona y región

Región	Situación de pobreza por ingresos	Estimación	Error estándar	95% Intervalo de confianza		Coeficiente de variación
				Inferior	Superior	
Nacional	Pobreza extrema	1,9%	0,1%	1,7%	2,0%	3,5%
	Pobreza no extrema	3,8%	0,1%	3,6%	4,0%	2,4%
	Total pobreza	5,6%	0,1%	5,9%	5,4%	2,0%
	No pobre	94,4%	0,1%	94,1%	94,6%	0,1%
Urbano	Pobreza extrema	1,8%	0,1%	1,6%	1,9%	4,0%
	Pobreza no extrema	3,5%	0,1%	3,3%	3,7%	2,8%
	Total pobreza	5,3%	0,1%	5,5%	5,0%	2,3%
	No pobre	94,7%	0,1%	94,5%	95,0%	0,1%
Rural	Pobreza extrema	2,6%	0,2%	2,3%	2,9%	6,0%
	Pobreza no extrema	5,8%	0,2%	5,3%	6,3%	4,2%
	Total pobreza	8,4%	0,3%	9,0%	7,8%	3,6%
	No pobre	91,6%	0,3%	91,0%	92,2%	0,3%
Arica y Parinacota	Pobreza extrema	2,3%	0,3%	1,7%	3,1%	14,6%
	Pobreza no extrema	5,1%	0,5%	4,2%	6,1%	9,3%
	Total pobreza	7,4%	0,6%	8,8%	6,3%	8,7%
	No pobre	92,6%	0,6%	91,2%	93,7%	0,7%
Tarapacá	Pobreza extrema	3,8%	0,8%	2,6%	5,6%	20,2%
	Pobreza no extrema	5,6%	0,5%	4,7%	6,6%	9,0%
	Total pobreza	9,4%	1,1%	11,7%	7,5%	11,6%
	No pobre	90,6%	1,1%	88,3%	92,5%	1,2%
Antofagasta	Pobreza extrema	2,5%	0,3%	1,9%	3,2%	13,8%
	Pobreza no extrema	4,3%	0,4%	3,5%	5,3%	10,3%
	Total pobreza	6,8%	0,6%	8,1%	5,7%	9,0%
	No pobre	93,2%	0,6%	91,9%	94,3%	0,7%
Atacama	Pobreza extrema	2,7%	0,4%	2,1%	3,5%	12,8%
	Pobreza no extrema	4,8%	0,5%	3,8%	6,0%	11,3%
	Total pobreza	7,5%	0,7%	9,1%	6,3%	9,4%
	No pobre	92,5%	0,7%	90,9%	93,7%	0,8%
Coquimbo	Pobreza extrema	2,5%	0,3%	2,0%	3,2%	12,4%
	Pobreza no extrema	4,2%	0,4%	3,5%	5,0%	9,3%
	Total pobreza	6,7%	0,5%	7,8%	5,8%	7,7%
	No pobre	93,3%	0,5%	92,2%	94,2%	0,6%
Valparaíso	Pobreza extrema	1,9%	0,2%	1,5%	2,3%	10,0%
	Pobreza no extrema	4,0%	0,3%	3,4%	4,6%	7,2%
	Total pobreza	5,9%	0,4%	6,6%	5,2%	6,0%
	No pobre	94,1%	0,4%	93,4%	94,8%	0,4%
Metropolitana	Pobreza extrema	1,2%	0,1%	1,0%	1,5%	9,1%
	Pobreza no extrema	2,5%	0,2%	2,2%	2,8%	6,3%
	Total pobreza	3,7%	0,2%	4,1%	3,4%	5,2%
	No pobre	96,3%	0,2%	95,9%	96,6%	0,2%
O'Higgins	Pobreza extrema	1,9%	0,2%	1,5%	2,4%	11,6%

	Pobreza no extrema	4,0%	0,3%	3,4%	4,6%	7,4%
	Total pobreza	5,9%	0,4%	6,6%	5,2%	6,2%
	No pobre	94,1%	0,4%	93,4%	94,8%	0,4%
Maule	Pobreza extrema	2,4%	0,3%	1,9%	2,9%	10,9%
	Pobreza no extrema	5,0%	0,3%	4,4%	5,7%	6,4%
	Total pobreza	7,4%	0,4%	8,2%	6,6%	5,5%
	No pobre	92,6%	0,4%	91,8%	93,4%	0,4%
Ñuble	Pobreza extrema	3,6%	0,4%	2,9%	4,4%	10,1%
	Pobreza no extrema	6,4%	0,5%	5,5%	7,4%	7,4%
	Total pobreza	10,0%	0,6%	11,2%	8,8%	6,1%
	No pobre	90,0%	0,6%	88,8%	91,2%	0,7%
Biobío	Pobreza extrema	2,5%	0,2%	2,1%	2,9%	8,5%
	Pobreza no extrema	4,7%	0,3%	4,2%	5,3%	6,0%
	Total pobreza	7,2%	0,4%	7,9%	6,5%	4,9%
	No pobre	92,8%	0,4%	92,1%	93,5%	0,4%
La Araucanía	Pobreza extrema	3,0%	0,3%	2,5%	3,5%	8,8%
	Pobreza no extrema	7,0%	0,4%	6,2%	7,9%	6,3%
	Total pobreza	10,0%	0,5%	11,0%	9,0%	5,3%
	No pobre	90,0%	0,5%	89,0%	91,0%	0,6%
Los Ríos	Pobreza extrema	1,1%	0,2%	0,8%	1,4%	15,4%
	Pobreza no extrema	3,8%	0,3%	3,2%	4,5%	8,4%
	Total pobreza	4,8%	0,4%	5,6%	4,2%	7,5%
	No pobre	95,2%	0,4%	94,4%	95,8%	0,4%
Los Lagos	Pobreza extrema	1,8%	0,2%	1,4%	2,3%	12,5%
	Pobreza no extrema	4,1%	0,4%	3,4%	5,0%	9,6%
	Total pobreza	5,9%	0,4%	6,9%	5,1%	7,6%
	No pobre	94,1%	0,4%	93,1%	94,9%	0,5%
Aysén	Pobreza extrema	1,0%	0,3%	0,6%	1,6%	27,2%
	Pobreza no extrema	2,4%	0,5%	1,6%	3,7%	21,6%
	Total pobreza	3,4%	0,6%	4,7%	2,4%	16,9%
	No pobre	96,6%	0,6%	95,3%	97,6%	0,6%
Magallanes	Pobreza extrema	1,1%	0,3%	0,6%	1,9%	29,4%
	Pobreza no extrema	1,6%	0,4%	1,1%	2,6%	22,7%
	Total pobreza	2,7%	0,6%	4,2%	1,8%	22,1%
	No pobre	97,3%	0,6%	95,8%	98,2%	0,6%

Elaborado por el Instituto Nacional de Estadísticas.

IX. ANEXOS

ANEXO N°1. ESCENARIOS TRABAJADOS

Tabla IX.1. Escenario N°0 (Escenario Base)

Escenario N°0. Resultados Casen en Pandemia 2020. Nacional Urbano - Rural y Regional										
Región	Muestra Viviendas Logradas	Tasa de Pobreza	Error Estándar	Coefficiente de Variación	Efecto Diseño	Cuasi-Varianza	Cuasi-Varianza Deff.Ajustado	Error Absoluto	Error Relativo	Tamaño obtenido bajo (M.A.S.)
Nivel País	62.540	10,8%	0,2%	2,0%	2,2	0,13	0,13	0,4%	3,9%	40.001
Urbano	52.663	10,4%	0,2%	2,2%	1,1	0,19	0,19	0,5%	4,8%	30.640
Rural	9.877	13,8%	0,6%	4,1%	1,1	0,34	0,34	1,2%	8,7%	9.361
Arica y Parinacota	2.282	11,9%	0,9%	7,7%	0,5	0,41	0,19	1,8%	15,5%	2.282
Tarapacá	2.621	14,0%	0,8%	6,1%	0,5	0,37	0,19	1,7%	12,1%	2.621
Antofagasta	2.368	9,3%	0,9%	9,8%	1,5	0,13	0,13	1,8%	19,7%	1.579
Atacama	2.402	9,5%	0,9%	9,4%	0,7	0,27	0,19	1,8%	18,9%	2.402
Coquimbo	2.695	11,7%	1,1%	9,1%	2,2	0,14	0,14	2,1%	18,3%	1.225
Valparaíso	6.326	11,3%	0,6%	5,5%	1,9	0,13	0,13	1,3%	11,0%	3.329
Metropolitana	13.122	9,0%	0,4%	4,2%	3,2	0,06	0,06	0,7%	8,3%	4.101
O'Higgins	4.207	10,0%	0,8%	8,0%	1,7	0,16	0,16	1,6%	16,0%	2.475
Maule	4.314	12,3%	0,7%	6,0%	1,5	0,16	0,16	1,5%	12,1%	2.876
Ñuble	2.268	14,7%	0,9%	6,1%	0,8	0,21	0,18	1,8%	12,2%	2.268
Biobío	6.009	13,2%	0,7%	5,3%	1,7	0,17	0,18	1,4%	10,7%	3.535
La Araucanía	4.003	17,4%	0,9%	5,2%	1,4	0,23	0,23	1,8%	10,4%	2.859
Los Ríos	2.791	12,2%	0,8%	6,3%	0,6	0,29	0,16	1,5%	12,6%	2.791
Los Lagos	3.580	11,3%	0,9%	7,8%	1,7	0,16	0,16	1,8%	15,5%	2.106
Aysén	1.694	6,6%	0,9%	13,6%	0,4	0,39	0,14	1,8%	27,2%	1.694
Magallanes	1.858	5,7%	0,9%	15,5%	0,7	0,21	0,14	1,8%	31,0%	1.858

(*) Efectos de Diseño nivel Nacional Urbano - Rural que permiten llegar a 40.001 viviendas bajo muestreo aleatorio simple a nivel nacional.

Elaborado por el Instituto Nacional de Estadísticas.

Tabla IX.2. Escenario N°1

Escenario N°1 Casen 2022										
Región	Viviendas MMV 2020	Tasa de Pobreza	Efecto Diseño Ajustado	Tamaños ajustados efecto diseño	Error Absoluto	Error Relativo	Tamaños ajustados por finitud	Error Absoluto finitud	Error Relativo finitud	Tamaño con Sobremuestreo
Nivel País	6.206.626	10,8%	2,2	72.555	0,4%	3,3%	71.025	0,4%	3,4%	95.498
Arica y Parinacota	73.783	11,9%	1,7	3.879	2,7%	22,6%	3.686	2,8%	23,2%	4.678
Tarapacá	110.680	14,0%	1,7	4.456	2,4%	16,9%	4.284	2,4%	17,3%	6.268
Antofagasta	190.638	9,3%	1,7	2.684	1,8%	19,7%	2.647	1,9%	19,9%	4.088
Atacama	106.193	9,5%	1,7	4.083	2,1%	22,6%	3.932	2,2%	23,0%	5.600
Coquimbo	285.374	11,7%	2,2	2.695	2,1%	18,3%	2.670	2,1%	18,4%	3.417
Valparaíso	677.668	11,3%	1,9	6.325	1,3%	11,2%	6.267	1,3%	11,2%	8.912
Metropolitana	2.379.740	9,0%	2,5	10.253	0,7%	8,3%	10.210	0,7%	8,3%	16.747
O'Higgins	333.576	10,0%	1,7	4.208	1,6%	15,9%	4.156	1,6%	16,0%	5.073
Maule	392.061	12,3%	1,7	4.889	1,5%	12,1%	4.829	1,5%	12,2%	5.623
Ñuble	182.744	14,7%	1,7	3.856	1,9%	13,2%	3.777	2,0%	13,4%	4.556
Biobío	558.582	13,2%	1,7	6.010	1,4%	10,6%	5.947	1,4%	10,6%	7.410
La Araucanía	359.601	17,4%	1,7	4.860	1,8%	10,4%	4.796	1,8%	10,5%	5.611
Los Ríos	144.531	12,2%	1,7	4.745	2,0%	16,6%	4.595	2,1%	16,9%	5.730
Los Lagos	308.709	11,3%	1,7	3.580	1,8%	15,6%	3.539	1,8%	15,7%	4.230
Aysén	40.838	6,6%	1,7	2.878	3,0%	45,7%	2.688	3,1%	47,3%	3.620
Magallanes	61.908	5,7%	1,7	3.154	2,1%	37,8%	3.002	2,2%	38,7%	3.935

Elaborado por el Instituto Nacional de Estadísticas.

Tabla IX.3. Escenario N°2.1

Escenario N°2.1 Casen 2022											
Región	Viviendas MMV 2020	Tasa de Pobreza	Efecto Diseño Ajustado	Tamaños ajustados efecto diseño	Error Absoluto	Error Relativo	Tamaños ajustados por finitud	Tamaño Objetivo Ajustado	Error Absoluto ajustado	Error Relativo ajustado	Tamaño con Sobremuestreo
Nivel País	6.115.331	10,8%	2,2	69.846	0,4%	3,4%	68.625	70.403	0,4%	3,3%	105.928
Arica y Parinacota	73.682	11,9%	1,4	3.195	1,8%	15,5%	3.063	3.104	1,9%	15,7%	4.528
Tarapacá	110.311	14,0%	1,4	3.669	1,7%	12,1%	3.551	3.599	1,7%	12,3%	5.664
Antofagasta	188.537	9,3%	1,7	2.684	1,8%	19,7%	2.647	2.681	1,8%	19,7%	4.400
Atacama	106.193	9,5%	1,4	3.363	1,8%	18,9%	3.260	3.291	1,8%	19,1%	5.064
Coquimbo	281.955	11,7%	2,2	2.695	2,1%	18,3%	2.670	2.753	2,1%	18,1%	4.072
Valparaíso	659.018	11,3%	1,9	6.327	1,3%	11,0%	6.267	6.461	1,2%	10,9%	10.064
Metropolitana	2.357.708	9,0%	3,0	12.303	0,7%	8,3%	12.240	12.336	0,7%	8,3%	20.744
O'Higgins	325.023	10,0%	1,7	4.209	1,6%	16,0%	4.156	4.374	1,6%	15,7%	6.104
Maule	387.296	12,3%	1,7	4.889	1,5%	12,1%	4.829	5.163	1,4%	11,7%	7.136
Ñuble	178.350	14,7%	1,4	3.177	1,8%	12,2%	3.121	3.186	1,8%	12,2%	4.472
Biobío	544.665	13,2%	1,7	6.013	1,4%	10,7%	5.947	6.058	1,4%	10,6%	8.624
La Araucanía	357.000	17,4%	1,7	4.862	1,8%	10,4%	4.796	5.182	1,8%	10,0%	7.176
Los Ríos	144.531	12,2%	1,4	3.907	1,5%	12,6%	3.805	3.859	1,5%	12,6%	5.648
Los Lagos	299.915	11,3%	1,7	3.582	1,8%	15,5%	3.539	3.651	1,7%	15,4%	5.216
Aysén	40.615	6,6%	1,4	2.370	1,8%	27,2%	2.240	2.241	1,8%	28,0%	3.360
Magallanes	60.532	5,7%	1,4	2.601	1,8%	31,0%	2.494	2.464	1,8%	31,9%	3.656

Elaborado por el Instituto Nacional de Estadísticas.

Tabla IX.4. Escenario N°2.2

Escenario N°2.2 Casen 2022											
Región	Viviendas MMV 2020	Tasa de Pobreza	Efecto Diseño Ajustado	Tamaños ajustados efecto diseño	Error Absoluto	Error Relativo	Tamaños ajustados por finitud	Tamaño Objetivo Ajustado	Error Absoluto ajustado	Error Relativo ajustado	Tamaño con Sobremuestreo
Nivel País	6.206.626	10,8%	2,2	69.829	0,4%	3,4%	68.625	71.755	0,4%	3,3%	107.832
Arica y Parinacota	73.783	11,9%	1,4	3.195	1,8%	15,5%	3.063	3.112	1,9%	15,7%	4.536
Tarapacá	110.680	14,0%	1,4	3.668	1,7%	12,1%	3.551	3.623	1,7%	12,2%	5.688
Antofagasta	190.638	9,3%	1,7	2.684	1,8%	19,7%	2.647	2.729	1,8%	19,5%	4.472
Atacama	106.193	9,5%	1,4	3.363	1,8%	18,9%	3.260	3.291	1,8%	19,1%	5.064
Coquimbo	285.374	11,7%	2,2	2.695	2,1%	18,3%	2.670	2.793	2,1%	18,0%	4.120
Valparaíso	677.668	11,3%	1,9	6.325	1,3%	11,0%	6.267	6.717	1,2%	10,7%	10.480
Metropolitana	2.379.740	9,0%	3,0	12.303	0,7%	8,3%	12.240	12.504	0,7%	8,3%	20.976
O'Higgins	333.576	10,0%	1,7	4.208	1,6%	16,0%	4.156	4.518	1,5%	15,4%	6.304
Maule	392.061	12,3%	1,7	4.889	1,5%	12,1%	4.829	5.235	1,4%	11,6%	7.232
Ñuble	182.744	14,7%	1,4	3.175	1,8%	12,2%	3.121	3.290	1,8%	12,0%	4.624
Biobío	558.582	13,2%	1,7	6.010	1,4%	10,7%	5.947	6.282	1,4%	10,4%	8.936
La Araucanía	359.601	17,4%	1,7	4.860	1,8%	10,4%	4.796	5.230	1,7%	10,0%	7.248
Los Ríos	144.531	12,2%	1,4	3.907	1,5%	12,6%	3.805	3.859	1,5%	12,6%	5.648
Los Lagos	308.709	11,3%	1,7	3.580	1,8%	15,5%	3.539	3.771	1,7%	15,1%	5.376
Aysén	40.838	6,6%	1,4	2.369	1,8%	27,2%	2.240	2.257	1,8%	27,9%	3.376
Magallanes	61.908	5,7%	1,4	2.598	1,8%	31,0%	2.494	2.544	1,8%	31,3%	3.752

Elaborado por el Instituto Nacional de Estadísticas.

Tabla IX.5. Escenario N°3

Escenario N°3 Casen 2022											
Región	Viviendas MMV 2020	Tasa de Pobreza	Efecto Diseño Ajustado	Tamaños ajustados efecto diseño	Error Absoluto	Error Relativo	Tamaños ajustados por finitud	Tamaño Objetivo Ajustado	Error Absoluto ajustado	Error Relativo ajustado	Tamaño con Sobremuestreo
Nivel País	6.206.626	10,8%	2,2	68.034	0,4%	3,4%	67.032	68.857	0,4%	3,4%	101.338
Arica y Parinacota	73.783	11,9%	1,2	2.737	1,8%	15,5%	2.640	2.646	1,9%	15,8%	2.970
Tarapacá	110.680	14,0%	1,0	2.621	1,7%	12,1%	2.561	2.618	1,7%	12,1%	4.126
Antofagasta	190.638	9,3%	1,7	2.684	1,8%	19,7%	2.647	2.747	1,8%	19,4%	4.490
Atacama	106.193	9,5%	1,4	3.363	1,8%	18,9%	3.260	3.280	1,8%	19,1%	4.148
Coquimbo	285.374	11,7%	2,2	2.695	2,1%	18,3%	2.670	2.749	2,1%	18,1%	4.076
Valparaíso	677.668	11,3%	1,9	6.325	1,3%	11,0%	6.267	6.610	1,2%	10,8%	10.340
Metropolitana	2.379.740	9,0%	3,2	13.123	0,7%	8,3%	13.052	13.311	0,7%	8,3%	22.350
O'Higgins	333.576	10,0%	1,8	4.455	1,6%	16,0%	4.397	4.530	1,6%	15,9%	6.332
Maule	392.061	12,3%	1,7	4.889	1,5%	12,1%	4.829	4.972	1,5%	11,9%	6.910
Ñuble	182.744	14,7%	1,4	3.175	1,8%	12,2%	3.121	3.194	1,8%	12,1%	4.500
Biobío	558.582	13,2%	1,8	6.363	1,4%	10,7%	6.292	6.560	1,4%	10,5%	9.342
La Araucanía	359.601	17,4%	1,7	4.860	1,8%	10,4%	4.796	4.885	1,8%	10,3%	6.782
Los Ríos	144.531	12,2%	1,4	3.907	1,5%	12,6%	3.805	3.819	1,6%	12,7%	5.404
Los Lagos	308.709	11,3%	1,7	3.580	1,8%	15,5%	3.539	3.643	1,7%	15,4%	5.210
Aysén	40.838	6,6%	1,0	1.400	2,0%	29,9%	1.353	1.378	2,0%	30,1%	1.810
Magallanes	61.908	5,7%	1,0	1.857	1,8%	31,0%	1.803	1.915	1,7%	30,5%	2.548

Elaborado por el Instituto Nacional de Estadísticas.

Tabla IX.6. Escenario N°4

Escenario N°4 Casen 2022											
Región	Viviendas MMV 2020	Tasa de Pobreza	Efecto Diseño Ajustado	Tamaños ajustados efecto diseño	Error Absoluto	Error Relativo	Tamaños ajustados por finitud	Tamaño Objetivo Ajustado	Error Absoluto ajustado	Error Relativo ajustado	Tamaño con Sobremuestreo
Nivel País	6.206.626	10,8%	2,2	68.034	0,4%	3,4%	67.032	71.028	0,4%	3,3%	104.246
Arica y Parinacota	73.783	11,9%	1,2	2.737	1,8%	15,5%	2.640	2.646	1,9%	15,8%	2.970
Tarapacá	110.680	14,0%	1,0	2.621	1,7%	12,1%	2.561	2.638	1,7%	12,1%	4.156
Antofagasta	190.638	9,3%	1,7	2.684	1,8%	19,7%	2.647	2.787	1,8%	19,3%	4.580
Atacama	106.193	9,5%	1,4	3.363	1,8%	18,9%	3.260	3.290	1,8%	19,1%	4.158
Coquimbo	285.374	11,7%	2,2	2.695	2,1%	18,3%	2.670	2.837	2,1%	17,8%	4.176
Valparaíso	677.668	11,3%	1,9	6.325	1,3%	11,0%	6.267	7.070	1,2%	10,4%	10.984
Metropolitana	2.379.740	9,0%	3,2	13.123	0,7%	8,3%	13.052	13.607	0,7%	8,2%	22.804
O'Higgins	333.576	10,0%	1,8	4.455	1,6%	16,0%	4.397	4.798	1,5%	15,4%	6.672
Maule	392.061	12,3%	1,7	4.889	1,5%	12,1%	4.829	5.107	1,5%	11,8%	7.070
Ñuble	182.744	14,7%	1,4	3.175	1,8%	12,2%	3.121	3.319	1,8%	11,9%	4.680
Biobío	558.582	13,2%	1,8	6.363	1,4%	10,7%	6.292	6.909	1,4%	10,2%	9.792
La Araucanía	359.601	17,4%	1,7	4.860	1,8%	10,4%	4.796	5.052	1,8%	10,2%	7.002
Los Ríos	144.531	12,2%	1,4	3.907	1,5%	12,6%	3.805	3.820	1,6%	12,7%	5.400
Los Lagos	308.709	11,3%	1,7	3.580	1,8%	15,5%	3.539	3.826	1,7%	15,0%	5.450
Aysén	40.838	6,6%	1,0	1.400	2,0%	29,9%	1.353	1.397	2,0%	29,9%	1.784
Magallanes	61.908	5,7%	1,0	1.857	1,8%	31,0%	1.803	1.925	1,7%	30,5%	2.568

Elaborado por el Instituto Nacional de Estadísticas.

Tabla IX.7. Escenario N°5 (Escenario elegido)

Escenario N°5 Casen 2022											
Región	Viviendas MMV 2020	Tasa de Pobreza	Efecto Diseño Ajustado	Tamaños ajustados efecto diseño	Error Absoluto	Error Relativo	Tamaños ajustados por finitud	Tamaño Objetivo Ajustado	Error Absoluto ajustado	Error Relativo ajustado	Tamaño con Sobremuestreo
Nivel País	6.206.626	10,8%	2,2	68.034	0,4%	3,4%	67.032	71.028	0,4%	3,3%	106.856
Arica y Parinacota	73.783	11,9%	1,2	2.737	1,8%	15,5%	2.640	2.646	1,9%	15,8%	3.698
Tarapacá	110.680	14,0%	1,0	2.621	1,7%	12,1%	2.561	2.638	1,7%	12,1%	4.156
Antofagasta	190.638	9,3%	1,7	2.684	1,8%	19,7%	2.647	2.787	1,8%	19,3%	4.580
Atacama	106.193	9,5%	1,4	3.363	1,8%	18,9%	3.260	3.290	1,8%	19,1%	5.060
Coquimbo	285.374	11,7%	2,2	2.695	2,1%	18,3%	2.670	2.837	2,1%	17,8%	4.180
Valparaíso	677.668	11,3%	1,9	6.325	1,3%	11,0%	6.267	7.070	1,2%	10,4%	10.984
Metropolitana	2.379.740	9,0%	3,2	13.123	0,7%	8,3%	13.052	13.607	0,7%	8,2%	22.804
O'Higgins	333.576	10,0%	1,8	4.455	1,6%	16,0%	4.397	4.798	1,5%	15,4%	6.672
Maule	392.061	12,3%	1,7	4.889	1,5%	12,1%	4.829	5.107	1,5%	11,8%	7.090
Ñuble	182.744	14,7%	1,4	3.175	1,8%	12,2%	3.121	3.319	1,8%	11,9%	4.696
Biobío	558.582	13,2%	1,8	6.363	1,4%	10,7%	6.292	6.909	1,4%	10,2%	9.800
La Araucanía	359.601	17,4%	1,7	4.860	1,8%	10,4%	4.796	5.052	1,8%	10,2%	7.006
Los Ríos	144.531	12,2%	1,4	3.907	1,5%	12,6%	3.805	3.820	1,6%	12,7%	5.400
Los Lagos	308.709	11,3%	1,7	3.580	1,8%	15,5%	3.539	3.826	1,7%	15,0%	5.466
Aysén	40.838	6,6%	1,0	1.400	2,0%	29,9%	1.353	1.397	2,0%	29,9%	2.070
Magallanes	61.908	5,7%	1,0	1.857	1,8%	31,0%	1.803	1.925	1,7%	30,5%	3.194

Elaborado por el Instituto Nacional de Estadísticas.

La Tabla IX.8 muestra la distribución de los tamaños muestrales del escenario N°6, considerando un levantamiento bajo modalidad mixta, a nivel nacional, regional y comunal, desagregados por área urbano-rural donde, a partir del tamaño con sobremuestreo calculado en el escenario N°4 y aplicando tasas de no logro correspondientes a un levantamiento telefónico, se obtiene un tamaño objetivo nacional de 64.376 viviendas.

Tabla IX.8. Distribución de tamaños muestrales a nivel nacional, regional, comunal y por área. Escenario N°6 levantamiento mixto. Casen 2022

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
Nivel País		64.376	51.120	13.256	104.246	84.816	19.430
	Total	1.801	1.627	174	2.970	2.680	290
Arica y Parinacota	Arica	1.735	1.611	124	2.904	2.664	240
	Camarones	20	0	20	20	0	20
	General Lagos	10	0	10	10	0	10
	Putre	36	16	20	36	16	20
	Total	2.522	2.325	197	4.156	3.856	300
Tarapacá	Camíña	20	0	20	30	0	30
	Colchane	20	0	20	30	0	30
	Huara	56	16	40	84	24	60
	Pica	55	35	20	86	56	30
	Pozo Almonte	119	72	47	184	104	80
	Alto Hospicio	759	749	10	1.178	1.168	10
	Iquique	1.493	1.453	40	2.564	2.504	60
	Total	2.723	2.513	210	4.580	4.200	380
Antofagasta	Antofagasta	1.549	1.529	20	2.468	2.448	20
	Mejillones	65	45	20	84	64	20
	Sierra Gorda	40	0	40	70	0	70
	Taltal	72	52	20	102	72	30
	Calama	745	705	40	1.442	1.352	90
	San Pedro de Atacama	70	30	40	150	40	110
	María Elena	40	30	10	58	48	10
	Tocopilla	142	122	20	206	176	30
	Total	2.603	2.239	364	4.158	3.608	550
Atacama	Chañaral	119	99	20	180	160	20
	Diego De Almagro	160	140	20	254	224	30
	Caldera	180	141	39	264	224	40
	Copiapó	1.336	1.296	40	2.158	2.088	70
	Tierra Amarilla	120	80	40	188	128	60
	Alto Del Carmen	59	0	59	110	0	110

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
	Freirina	71	34	37	106	56	50
	Huasco	121	81	40	186	136	50
	Vallenar	437	368	69	712	592	120
	Total	2.611	1.958	653	4.176	3.296	880
Coquimbo	Canela	56	16	40	84	24	60
	Illapel	102	64	38	146	96	50
	Los Vilos	100	60	40	154	104	50
	Salamanca	89	49	40	130	80	50
	Andacollo	52	32	20	60	40	20
	Coquimbo	699	657	42	1.148	1.088	60
	La Higuera	56	16	40	74	24	50
	La Serena	713	654	59	1.326	1.216	110
	Paiguano	40	0	40	50	0	50
	Vicuña	88	51	37	120	80	40
	Combarbalá	65	30	35	80	40	40
	Monte Patria	102	45	57	144	64	80
	Ovalle	339	254	85	520	400	120
	Punitaqui	70	30	40	90	40	50
Río Hurtado	40	0	40	50	0	50	
	Total	6.749	5.456	1.293	10.984	9.104	1.880
Valparaíso	Calle Larga	74	34	40	116	56	60
	Los Andes	231	191	40	340	280	60
	Rinconada	70	30	40	130	40	90
	San Esteban	75	35	40	98	48	50
	Limache	159	119	40	210	160	50
	Olmué	80	40	40	132	72	60
	Quilpué	526	486	40	888	768	120
	Villa Alemana	411	391	20	598	568	30
	Cabildo	77	37	40	98	48	50
	La Ligua	119	84	35	170	120	50
	Papudo	50	30	20	68	48	20
	Petorca	72	32	40	98	48	50
	Zapallar	70	30	40	162	72	90
	Calera	190	150	40	266	216	50
Hijuelas	74	34	40	98	48	50	
La Cruz	100	60	40	148	88	60	

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
	Nogales	94	54	40	122	72	50
	Quillota	284	246	38	402	352	50
	Algarrobo	82	42	40	148	88	60
	Cartagena	111	71	40	188	128	60
	El Quisco	76	56	20	132	112	20
	El Tabo	69	49	20	140	120	20
	San Antonio	293	253	40	394	344	50
	Santo Domingo	70	30	40	98	48	50
	Catemu	70	30	40	90	40	50
	Llaillay	95	55	40	122	72	50
	Panquehue	56	16	40	74	24	50
	Putendo	70	30	40	98	48	50
	San Felipe	257	217	40	406	336	70
	Santa María	72	32	40	108	48	60
	Casablanca	96	56	40	130	80	50
	Concón	175	135	40	296	256	40
	Puchuncaví	96	56	40	162	112	50
	Quintero	128	88	40	198	128	70
	Valparaíso	989	969	20	1.672	1.632	40
	Viña Del Mar	1.188	1.188	0	2.384	2.384	0
	Total	13.234	12.290	944	22.804	21.344	1.460
	Colina	228	187	41	346	256	90
	Lampa	177	140	37	268	208	60
	Tiltil	70	30	40	90	40	50
	Pirque	70	30	40	154	64	90
	Puente Alto	924	914	10	1.346	1.336	10
	San José De Maipo	70	30	40	138	48	90
	Buin	182	142	40	268	208	60
	Calera De Tango	70	30	40	118	48	70
	Paine	121	75	46	210	120	90
	San Bernardo	510	470	40	770	680	90
	Alhué	56	16	40	74	24	50
	Curacaví	75	35	40	128	48	80
	María Pinto	72	32	40	98	48	50
	Melipilla	221	151	70	324	224	100
	San Pedro	40	0	40	50	0	50

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
	Cerrillos	132	132	0	192	192	0
	Cerro Navia	202	202	0	272	272	0
	Conchalí	200	200	0	288	288	0
	El Bosque	254	254	0	344	344	0
	Estación Central	308	308	0	488	488	0
	Huechuraba	165	155	10	250	240	10
	Independencia	194	194	0	288	288	0
	La Cisterna	166	166	0	280	280	0
	La Florida	661	651	10	1.026	1.016	10
	La Granja	186	186	0	256	256	0
	La Pintana	266	266	0	376	376	0
	La Reina	160	160	0	312	312	0
	Las Condes	622	622	0	1.568	1.568	0
	Lo Barnechea	181	161	20	508	488	20
	Lo Espejo	142	142	0	200	200	0
	Lo Prado	156	156	0	208	208	0
	Macul	228	228	0	392	392	0
	Maipú	894	854	40	1.368	1.328	40
	Ñuñoa	488	488	0	1.016	1.016	0
	Pedro Aguirre Cerda	158	158	0	240	240	0
	Peñalolén	377	377	0	560	560	0
	Providencia	372	372	0	1.144	1.144	0
	Pudahuel	404	364	40	632	592	40
	Quilicura	352	332	20	508	488	20
	Quinta Normal	206	206	0	312	312	0
	Recoleta	266	266	0	416	416	0
	Renca	231	231	0	320	320	0
	San Joaquín	167	167	0	320	320	0
	San Miguel	227	227	0	432	432	0
	San Ramón	126	126	0	184	184	0
	Santiago	1.059	1.059	0	2.040	2.040	0
	Vitacura	168	168	0	776	776	0
	El Monte	91	51	40	122	72	50
	Isla De Maipo	85	45	40	122	72	50
	Padre Hurtado	136	96	40	196	136	60
	Peñaflor	179	139	40	280	200	80

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
	Talagante	139	99	40	186	136	50
	Total	4.333	2.898	1.435	6.672	4.712	1.960
O'Higgins	Codegua	70	30	40	98	48	50
	Coinco	56	16	40	74	24	50
	Coltauco	85	45	40	114	64	50
	Doñihue	107	67	40	172	112	60
	Graneros	156	116	40	218	168	50
	Las Cabras	101	37	64	154	64	90
	Machalí	241	221	20	406	376	30
	Malloa	69	30	39	90	40	50
	Mostazal	122	82	40	170	120	50
	Olivar	76	36	40	106	56	50
	Peumo	86	46	40	122	72	50
	Pichidegua	79	37	42	116	56	60
	Quinta De Tilcoco	72	32	40	90	40	50
	Rancagua	1.056	1.016	40	1.754	1.704	50
	Rengo	234	183	51	376	296	80
	Requínoa	103	59	44	158	88	70
	San Vicente	193	106	87	280	160	120
	La Estrella	56	16	40	84	24	60
	Litueche	56	16	40	84	24	60
	Marchihue	56	16	40	84	24	60
	Navidad	38	0	38	60	0	60
	Paredones	56	16	40	74	24	50
	Pichilemu	102	62	40	178	128	50
	Chépica	76	36	40	98	48	50
	Chimbarongo	145	84	61	218	128	90
	Lolol	56	16	40	84	24	60
	Nancagua	85	45	40	132	72	60
	Palmilla	56	16	40	74	24	50
	Peralillo	72	32	40	90	40	50
	Placilla	56	16	40	74	24	50
Pumanque	40	0	40	50	0	50	
San Fernando	317	273	44	524	464	60	
Santa Cruz	160	95	65	266	176	90	
	Total	4.586	3.178	1.408	7.070	5.080	1.990
Maule	Cauquenes	182	142	40	284	224	60
	Chanco	70	30	40	80	40	40

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
	Pelluhue	70	30	40	98	48	50
	Curicó	592	527	65	940	840	100
	Hualañé	70	30	40	82	32	50
	Licantén	70	30	40	90	40	50
	Molina	184	144	40	284	224	60
	Rauco	70	30	40	72	32	40
	Romeral	68	30	38	82	32	50
	Sagrada Familia	73	32	41	100	40	60
	Teno	111	37	74	148	48	100
	Vichuquén	56	16	40	66	16	50
	Colbún	85	38	47	134	64	70
	Linares	365	299	66	596	496	100
	Longaví	118	35	83	196	56	140
	Parral	175	128	47	302	232	70
	Retiro	85	32	53	120	40	80
	San Javier	182	125	57	298	208	90
	Villa Alegre	82	42	40	114	64	50
	Yerbas Buenas	76	30	46	110	40	70
	Constitución	189	149	40	308	248	60
	Curepto	56	16	40	74	24	50
	Empedrado	36	16	20	54	24	30
	Maule	204	161	43	300	240	60
	Pelarco	56	16	40	74	24	50
	Pencahue	56	16	40	74	24	50
	Río Claro	70	30	40	90	40	50
	San Clemente	170	82	88	260	120	140
	San Rafael	70	30	40	90	40	50
	Talca	895	855	40	1.550	1.480	70
	Total	2.966	1.913	1.053	4.680	3.040	1.640
Ñuble	Bulnes	117	74	43	184	104	80
	Chillán	1.025	938	87	1.694	1.544	150
	Chillán Viejo	182	142	40	290	240	50
	El Carmen	73	31	42	110	40	70
	Pemuco	70	30	40	100	40	60
	Pinto	69	32	37	118	48	70
	Quillón	107	60	47	204	104	100

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
	San Ignacio	91	37	54	146	56	90
	Yungay	120	80	40	196	136	60
	Cobquecura	56	16	40	84	24	60
	Coelemu	97	57	40	140	80	60
	Ninhue	56	16	40	64	24	40
	Portezuelo	56	16	40	74	24	50
	Quirihue	99	59	40	148	88	60
	Ránquil	56	16	40	104	24	80
	Treguaco	56	16	40	84	24	60
	Coihueco	139	46	93	230	80	150
	Ñiquén	74	16	58	114	24	90
	San Carlos	297	185	112	450	280	170
	San Fabián	56	16	40	64	24	40
	San Nicolás	70	30	40	82	32	50
	Total	6.257	4.954	1.303	9.792	8.032	1.760
	Arauco	130	93	37	186	136	50
	Cañete	123	70	53	174	104	70
	Contulmo	56	16	40	74	24	50
	Curanilahue	142	102	40	204	144	60
	Lebu	121	81	40	180	120	60
	Los Álamos	106	66	40	154	104	50
	Tirúa	56	16	40	84	24	60
	Alto Biobío	40	0	40	70	0	70
	Antuco	56	16	40	74	24	50
	Cabrero	117	77	40	170	120	50
	Laja	98	58	40	138	88	50
	Los Ángeles	725	542	183	1.132	872	260
	Mulchén	121	81	40	170	120	50
	Nacimiento	120	80	40	170	120	50
	Negrete	70	30	40	90	40	50
	Quilaco	56	16	40	74	24	50
	Quilleco	70	30	40	100	40	60
	San Rosendo	36	16	20	44	24	20
	Santa Bárbara	72	32	40	90	40	50
	Tucapel	90	50	40	130	80	50
	Yumbel	87	47	40	142	72	70

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
	Chiguayante	295	285	10	442	432	10
	Concepción	908	868	40	1.746	1.696	50
	Coronel	425	385	40	636	576	60
	Florida	70	30	40	90	40	50
	Hualpén	324	304	20	484	464	20
	Hualqui	107	67	40	164	104	60
	Lota	160	150	10	234	224	10
	Penco	181	161	20	270	240	30
	San Pedro De La Paz	479	469	10	826	816	10
	Santa Juana	75	35	40	108	48	60
	Talcahuano	520	500	20	812	792	20
	Tomé	221	181	40	330	280	50
	Total	4.430	2.930	1.500	7.002	4.832	2.170
	Carahue	104	55	49	158	88	70
	Cholchol	51	16	35	74	24	50
	Cunco	79	37	42	136	56	80
	Curarrehue	56	16	40	84	24	60
	Freire	102	32	70	158	48	110
	Galvarino	52	16	36	56	16	40
	Gorbea	85	45	40	114	64	50
	Lautaro	152	109	43	246	176	70
	Loncoche	113	73	40	154	104	50
	Melipeuco	56	16	40	74	24	50
	Nueva Imperial	130	71	59	202	112	90
	Padre Las Casas	294	182	112	462	272	190
	Perquenco	56	16	40	74	24	50
	Pitrufquén	111	69	42	164	104	60
	Pucón	122	75	47	238	168	70
	Saavedra	60	16	44	94	24	70
	Temuco	1.215	1.138	77	2.182	2.072	110
	Teodoro Schmidt	74	30	44	100	40	60
	Toltén	70	30	40	90	40	50
	Vilcún	118	65	53	176	96	80
	Villarrica	247	156	91	422	272	150
	Angol	235	195	40	338	288	50
	Collipulli	115	75	40	180	120	60

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
	Curautín	99	59	40	138	88	50
	Ercilla	70	30	40	100	40	60
	Lonquimay	52	16	36	74	24	50
	Los Sauces	70	30	40	90	40	50
	Lumaco	70	30	40	90	40	50
	Purén	74	34	40	98	48	50
	Renaico	56	36	20	86	56	30
	Traiguén	102	62	40	146	96	50
	Victoria	140	100	40	204	144	60
	Total	3.319	2.286	1.033	5.400	3.760	1.640
	Futrono	135	78	57	228	128	100
	La Unión	329	223	106	534	344	190
	Lago Ranco	92	16	76	144	24	120
	Río Bueno	289	149	140	452	232	220
	Corral	56	16	40	84	24	60
Los Ríos	Lanco	148	103	45	230	160	70
	Los Lagos	167	79	88	258	128	130
	Máfil	77	37	40	106	56	50
	Mariquina	173	75	98	252	112	140
	Paillaco	167	99	68	272	152	120
	Panguipulli	304	128	176	496	216	280
	Valdivia	1.382	1.283	99	2.344	2.184	160
	Total	3.485	2.356	1.129	5.450	3.880	1.570
	Ancud	161	109	52	314	224	90
	Castro	174	130	44	260	200	60
	Chonchi	70	30	40	100	40	60
	Curaco De Vélez	56	16	40	76	16	60
	Dalcahue	70	30	40	90	40	50
	Puqueldón	40	0	40	50	0	50
Los Lagos	Queilén	56	16	40	74	24	50
	Quellón	107	67	40	220	160	60
	Quemchi	56	16	40	74	24	50
	Quinchao	56	16	40	64	24	40
	Calbuco	130	60	70	204	104	100
	Fresia	72	32	40	90	40	50
	Frutillar	91	51	40	148	88	60

Región	Comuna	Muestra Objetivo Total	Muestra Objetivo Urbana	Muestra Objetivo Rural	Muestra Sobremuestreo Total	Muestra Sobremuestreo Urbana	Muestra Sobremuestreo Rural
	Llanquihue	87	47	40	130	80	50
	Los Muermos	72	32	40	100	40	60
	Mauñín	70	31	39	90	40	50
	Puerto Montt	929	825	104	1.556	1.416	140
	Puerto Varas	174	126	48	286	216	70
	Osorno	601	549	52	952	872	80
	Puerto Octay	56	16	40	66	16	50
	Purranque	91	51	40	132	72	60
	Puyehue	70	30	40	90	40	50
	Río Negro	70	30	40	100	40	60
	San Juan De La Costa	56	16	40	84	24	60
	San Pablo	70	30	40	100	40	60
	Total	1.180	840	340	1.784	1.144	640
	Aysén	270	230	40	366	296	70
	Cisnes	89	49	40	124	64	60
	Cochrane	36	16	20	54	24	30
	Tortel	20	0	20	30	0	30
	Coyhaique	647	529	118	1.006	736	270
	Lago Verde	20	0	20	30	0	30
	Chile Chico	56	16	40	94	24	70
	Río Ibáñez	42	0	42	80	0	80
	Total	1.577	1.357	220	2.568	2.248	320
	Cabo De Hornos	16	16	0	16	16	0
	Laguna Blanca	20	0	20	20	0	20
	Punta Arenas	1.142	1.092	50	1.924	1.824	100
	Río Verde	10	0	10	10	0	10
	San Gregorio	20	0	20	20	0	20
	Porvenir	80	60	20	136	96	40
	Primavera	20	0	20	20	0	20
	Timaukel	20	0	20	20	0	20
	Natales	229	189	40	382	312	70
	Torres Del Paine	20	0	20	20	0	20

Elaborado por el Instituto Nacional de Estadísticas.

ANEXO N°2. TAMAÑOS MUESTRALES A NIVEL NACIONAL, REGIONAL Y COMUNAL POR ÁREA (URBANO – RURAL) PARA CASEN 2022

En la siguiente Tabla se observa la distribución de los tamaños muestrales del escenario elegido (escenario N°5), a nivel nacional, regional y comunal desagregados por área urbano-rural, pasando desde el tamaño ajustado por finitud, tamaño ajustado por número de UPM y mínimos por estrato y tamaño ajustado por tasas de no logro corregidas, lo que da origen a los tamaños con sobremuestreo.

Región / Comuna	Distribución Inicial Comuna - Área			Primer ajuste: Tamaño mínimo 16 viviendas (con excepciones)			Segundo ajuste: Tamaño Objetivo Mínimo 8 o 10 viviendas. Controla por N° de UPM			Tasas de No Logro		Tamaño con Sobremuestreo				
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	N° Viviendas			N° UPM	
Nacional	67.032	56.415	10.617	68.461	56.946	11.515	71.028	56.905	14.123			106.856	87.252	19.604	10.602	1.943
Arica y Parinacota	2.640	2.412	228	2.670	2.426	244	2.646	2.412	234			3.698	3.350	348	335	29
Arica	2.580	2.396	184	2.580	2.396	184	2.580	2.396	184	10,0%	22,1%	3.618	3.330	288	333	24
Camarones	25	-	25	30	-	30	20	0	20	0,0%	6,2%	24	0	24	0	2
General Lagos	4	-	4	10	-	10	10	0	10	0,0%	9,3%	12	0	12	0	1
Putre	31	16	15	50	30	20	36	16	20	6,5%	8,8%	44	20	24	2	2
Tarapacá	2.561	2.435	126	2.608	2.439	169	2.638	2.439	199			4.156	3.856	300	482	30
Camiña	13	-	13	20	-	20	20	0	20	0,0%	25,7%	30	0	30	0	3
Colchane	7	-	7	10	-	10	20	0	20	0,0%	22,9%	30	0	30	0	3
Huara	33	12	21	46	16	30	56	16	40	19,5%	35,8%	84	24	60	3	6
Pica	49	37	12	57	37	20	57	37	20	37,2%	32,1%	86	56	30	7	3
Pozo Almonte	125	76	49	125	76	49	125	76	49	27,8%	38,2%	184	104	80	13	8
Alto Hospicio	788	786	2	796	786	10	796	786	10	32,6%	0,0%	1.178	1.168	10	146	1
Iquique	1.546	1.524	22	1.554	1.524	30	1.564	1.524	40	39,2%	28,6%	2.564	2.504	60	313	6
Antofagasta	2.647	2.566	81	2.697	2.577	120	2.787	2.577	210			4.580	4.200	380	525	38
Antofagasta	1.575	1.569	6	1.579	1.569	10	1.589	1.569	20	35,9%	0,0%	2.468	2.448	20	306	2
Mejillones	49	46	3	56	46	10	66	46	20	28,4%	0,0%	84	64	20	8	2
Sierra Gorda	7	-	7	10	-	10	40	0	40	0,0%	39,7%	70	0	70	0	7
Taltal	59	53	6	63	53	10	73	53	20	26,0%	40,5%	102	72	30	9	3
Calama	752	724	28	754	724	30	764	724	40	46,6%	54,1%	1.442	1.352	90	169	9
San Pedro de Atacama	50	25	25	60	30	30	70	30	40	18,9%	63,7%	150	40	110	5	11
María Elena	25	24	1	40	30	10	40	30	10	32,3%	0,0%	58	48	10	6	1
Tocopilla	130	125	5	135	125	10	145	125	20	29,1%	34,2%	206	176	30	22	3
Atacama	3.260	2.872	388	3.272	2.872	400	3.290	2.880	410			5.060	4.510	550	451	55
Chañaral	145	127	18	147	127	20	147	127	20	20,0%	0,0%	220	200	20	20	2
Diego De Almagro	198	180	18	200	180	20	200	180	20	20,0%	41,3%	310	280	30	28	3
Caldera	220	181	39	220	181	39	220	181	39	20,0%	0,0%	320	280	40	28	4
Copiapó	1.692	1.643	49	1.692	1.643	49	1.716	1.667	49	20,0%	33,7%	2.680	2.610	70	261	7
Tierra Amarilla	143	103	40	143	103	40	143	103	40	20,0%	32,0%	220	160	60	16	6
Alto Del Carmen	76	-	76	76	-	76	76	0	76	0,0%	28,6%	110	0	110	0	11
Freirina	89	52	37	89	52	37	81	44	37	20,0%	22,5%	120	70	50	7	5
Huasco	126	104	22	134	104	30	144	104	40	25,0%	22,7%	220	170	50	17	5

Región / Comuna	Distribución Inicial Comuna - Área			Primer ajuste: Tamaño mínimo 16 viviendas (con excepciones)			Segundo ajuste: Tamaño Objetivo Mínimo 8 o 10 viviendas. Controla por N° de UPM			Tasas de No Logro		Tamaño con Sobremuestreo					
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	
Vallenar	571	482	89	571	482	89	563	474	89	20,0%	25,9%	860	740	120	74	12	
Coquimbo	2.670	2.123	547	2.723	2.142	581	2.837	2.158	679			4.180	3.296	884	412	88	
Canela	40	8	32	40	8	32	56	16	40	25,2%	29,2%	84	24	60	3	6	
Illapel	109	71	38	109	71	38	109	71	38	23,9%	21,9%	146	96	50	12	5	
Los Vilos	83	66	17	86	66	20	106	66	40	33,7%	20,3%	154	104	50	13	5	
Salamanca	94	54	40	94	54	40	94	54	40	29,2%	18,8%	130	80	50	10	5	
Andacollo	40	35	5	45	35	10	55	35	20	18,3%	17,4%	64	40	24	5	2	
Coquimbo	775	728	47	775	728	47	775	728	47	33,1%	18,8%	1.148	1.088	60	136	6	
La Higuera	19	5	14	28	8	20	56	16	40	25,2%	23,5%	74	24	50	3	5	
La Serena	789	724	65	789	724	65	789	724	65	40,3%	39,6%	1.326	1.216	110	152	11	
Paiguano	18	-	18	20	-	20	40	0	40	0,0%	13,4%	50	0	50	0	5	
Vicuña	93	56	37	93	56	37	93	56	37	27,4%	11,1%	120	80	40	10	4	
Combarbalá	58	23	35	65	30	35	65	30	35	26,2%	19,4%	80	40	40	5	4	
Monte Patria	113	50	63	113	50	63	113	50	63	18,3%	17,5%	144	64	80	8	8	
Ovalle	376	282	94	376	282	94	376	282	94	29,7%	22,2%	520	400	120	50	12	
Punitaqui	42	21	21	60	30	30	70	30	40	19,8%	16,0%	90	40	50	5	5	
Río Hurtado	21	-	21	30	-	30	40	0	40	0,0%	21,0%	50	0	50	0	5	
Valparaíso	6.267	5.708	559	6.480	5.761	719	7.070	5.777	1.293			10.984	9.104	1.880	1.138	188	
Calle Larga	50	36	14	56	36	20	76	36	40	37,0%	36,0%	116	56	60	7	6	
Los Andes	216	203	13	223	203	20	243	203	40	27,3%	34,9%	340	280	60	35	6	
Rinconada	33	26	7	40	30	10	70	30	40	26,4%	54,1%	130	40	90	5	9	
San Esteban	63	37	26	67	37	30	77	37	40	23,6%	21,8%	98	48	50	6	5	
Limache	148	126	22	156	126	30	166	126	40	20,1%	26,3%	210	160	50	20	5	
Olmué	63	42	21	72	42	30	82	42	40	41,1%	33,3%	132	72	60	9	6	
Quilpué	525	516	9	526	516	10	556	516	40	31,9%	65,8%	888	768	120	96	12	
Villa Alemana	419	415	4	425	415	10	435	415	20	27,2%	35,1%	598	568	30	71	3	
Cabildo	63	39	24	69	39	30	79	39	40	21,4%	20,8%	98	48	50	6	5	
La Ligua	124	89	35	124	89	35	124	89	35	25,3%	26,8%	170	120	50	15	5	
Papudo	22	20	2	40	30	10	50	30	20	33,6%	0,0%	68	48	20	6	2	
Petorca	37	15	22	46	16	30	72	32	40	32,5%	24,4%	98	48	50	6	5	
Zapallar	27	18	9	40	30	10	70	30	40	58,0%	55,0%	162	72	90	9	9	
Calera	166	160	6	170	160	10	200	160	40	24,9%	14,7%	266	216	50	27	5	
Hijuelas	57	36	21	66	36	30	76	36	40	24,5%	19,4%	98	48	50	6	5	
La Cruz	73	64	9	74	64	10	104	64	40	27,3%	27,1%	148	88	60	11	6	
Nogales	69	57	12	77	57	20	97	57	40	24,2%	16,2%	122	72	50	9	5	
Quillota	300	262	38	300	262	38	300	262	38	24,6%	18,2%	402	352	50	44	5	
Algarrobo	58	45	13	65	45	20	85	45	40	48,8%	32,9%	148	88	60	11	6	
Cartagena	82	75	7	85	75	10	115	75	40	40,3%	27,8%	188	128	60	16	6	
El Quisco	64	60	4	70	60	10	80	60	20	46,3%	0,0%	132	112	20	14	2	
El Tabo	55	52	3	62	52	10	72	52	20	55,0%	0,0%	140	120	20	15	2	
San Antonio	287	269	18	289	269	20	309	269	40	21,3%	19,4%	394	344	50	43	5	
Santo Domingo	39	22	17	50	30	20	70	30	40	31,5%	20,3%	98	48	50	6	5	
Catemu	46	24	22	60	30	30	70	30	40	22,4%	23,4%	90	40	50	5	5	
Llaillay	80	58	22	88	58	30	98	58	40	22,6%	13,7%	122	72	50	9	5	
Panquehue	22	12	10	26	16	10	56	16	40	27,7%	23,8%	74	24	50	3	5	

Región / Comuna	Distribución Inicial Comuna - Área			Primer ajuste: Tamaño mínimo 16 viviendas (con excepciones)			Segundo ajuste: Tamaño Objetivo Mínimo 8 o 10 viviendas. Controla por N° de UPM			Tasas de No Logro		Tamaño con Sobremuestreo					
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	
Putendo	56	22	34	64	30	34	70	30	40	36,3%	20,9%	98	48	50	6	5	
San Felipe	253	230	23	260	230	30	270	230	40	30,0%	38,0%	406	336	70	42	7	
Santa María	51	32	19	52	32	20	72	32	40	29,7%	32,6%	108	48	60	6	6	
Casablanca	91	59	32	91	59	32	99	59	40	23,7%	17,4%	130	80	50	10	5	
Concón	152	143	9	153	143	10	183	143	40	44,4%	0,0%	296	256	40	32	4	
Puchuncaví	70	59	11	79	59	20	99	59	40	45,2%	22,2%	162	112	50	14	5	
Quintero	110	93	17	113	93	20	133	93	40	28,6%	46,1%	198	128	70	16	7	
Valparaíso	1.033	1.029	4	1.039	1.029	10	1.049	1.029	20	37,1%	44,8%	1.672	1.632	40	204	4	
Viña Del Mar	1.263	1.263	-	1.263	1.263	-	1.263	1.263	0	47,0%	0,0%	2.384	2.384	0	298	0	
Metropolitana	13.052	12.589	463	13.243	12.635	608	13.607	12.659	948			22.804	21.344	1.460	2.668	146	
Colina	235	193	42	235	193	42	235	193	42	22,6%	53,2%	346	256	90	32	9	
Lampa	181	144	37	181	144	37	181	144	37	27,9%	31,7%	268	208	60	26	6	
Tiltil	35	23	12	50	30	20	70	30	40	28,1%	25,8%	90	40	50	5	5	
Pirque	44	19	25	60	30	30	70	30	40	52,0%	55,4%	154	64	90	8	9	
Puente Alto ⁸⁰	943	942	1	952	942	10	952	942	10	29,6%	0,0%	1.336	1.336	0	167	0	
San José De Maipo	32	20	12	50	30	20	70	30	40	32,3%	54,6%	148	48	100	6	10	
Buín	169	146	23	176	146	30	186	146	40	26,3%	30,2%	268	208	60	26	6	
Calera De Tango	41	19	22	60	30	30	70	30	40	31,8%	39,2%	118	48	70	6	7	
Paine	125	78	47	125	78	47	125	78	47	33,2%	41,7%	210	120	90	15	9	
San Bernardo	493	484	9	494	484	10	524	484	40	28,9%	53,0%	770	680	90	85	9	
Alhué	13	5	8	18	8	10	56	16	40	27,5%	21,2%	74	24	50	3	5	
Curacaví	61	36	25	66	36	30	76	36	40	26,6%	47,7%	128	48	80	6	8	
María Pinto	25	12	13	36	16	20	72	32	40	33,9%	16,0%	98	48	50	6	5	
Melipilla	228	156	72	228	156	72	228	156	72	29,4%	31,5%	324	224	100	28	10	
San Pedro	20	-	20	30	-	30	40	0	40	0,0%	22,7%	50	0	50	0	5	
Cerrillos	136	136	-	136	136	-	136	136	0	29,0%	0,0%	192	192	0	24	0	
Cerro Navia	208	208	-	208	208	-	208	208	0	22,8%	0,0%	272	272	0	34	0	
Conchalí	206	206	-	206	206	-	206	206	0	27,7%	0,0%	288	288	0	36	0	
El Bosque	262	262	-	262	262	-	262	262	0	23,5%	0,0%	344	344	0	43	0	
Estación Central	317	317	-	317	317	-	317	317	0	35,6%	0,0%	488	488	0	61	0	
Huechuraba	160	159	1	169	159	10	169	159	10	31,8%	0,0%	250	240	10	30	1	
Independencia	200	200	-	200	200	-	200	200	0	31,1%	0,0%	288	288	0	36	0	
La Cisterna	171	171	-	171	171	-	171	171	0	38,2%	0,0%	280	280	0	35	0	
La Florida	672	671	1	681	671	10	681	671	10	34,1%	0,0%	1.026	1.016	10	127	1	
La Granja	192	192	-	192	192	-	192	192	0	24,9%	0,0%	256	256	0	32	0	
La Pintana	274	274	-	274	274	-	274	274	0	25,5%	0,0%	376	376	0	47	0	
La Reina	165	165	-	165	165	-	165	165	0	46,9%	0,0%	312	312	0	39	0	
Las Condes	641	641	-	641	641	-	641	641	0	59,1%	0,0%	1.568	1.568	0	196	0	
Lo Barnechea	168	166	2	176	166	10	186	166	20	66,0%	0,0%	508	488	20	61	2	
Lo Espejo	146	146	-	146	146	-	146	146	0	26,3%	0,0%	200	200	0	25	0	

⁸⁰ Para el área rural de Puente Alto, en el diseño se había establecido la selección de 1 UPM, sin embargo, al momento de la selección de viviendas, dicha UPM no contaba con viviendas elegibles. Por este motivo se selecciona una UPM adicional en el área rural de la comuna de San José de Maipo, sin alterar los tamaños a nivel de región – área.

Región / Comuna	Distribución Inicial Comuna - Área			Primer ajuste: Tamaño mínimo 16 viviendas (con excepciones)			Segundo ajuste: Tamaño Objetivo Mínimo 8 o 10 viviendas. Controla por N° de UPM			Tasas de No Logro		Tamaño con Sobremuestreo				
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural
Lo Prado	161	161	-	161	161	-	161	161	0	23,1%	0,0%	208	208	0	26	0
Macul	235	235	-	235	235	-	235	235	0	39,5%	0,0%	392	392	0	49	0
Maipú	885	880	5	890	880	10	920	880	40	33,4%	0,0%	1.368	1.328	40	166	4
Nuñoa	503	503	-	503	503	-	503	503	0	50,4%	0,0%	1.016	1.016	0	127	0
Pedro Aguirre Cerda	163	163	-	163	163	-	163	163	0	30,9%	0,0%	240	240	0	30	0
Peñalolén	388	388	-	388	388	-	388	388	0	31,4%	0,0%	560	560	0	70	0
Providencia	383	383	-	383	383	-	383	383	0	66,6%	0,0%	1.144	1.144	0	143	0
Pudahuel	382	375	7	385	375	10	415	375	40	37,2%	0,0%	632	592	40	74	4
Quilicura	343	342	1	352	342	10	362	342	20	29,3%	0,0%	508	488	20	61	2
Quinta Normal	213	213	-	213	213	-	213	213	0	30,3%	0,0%	312	312	0	39	0
Recoleta	274	274	-	274	274	-	274	274	0	33,9%	0,0%	416	416	0	52	0
Renca	238	238	-	238	238	-	238	238	0	26,1%	0,0%	320	320	0	40	0
San Joaquín	172	172	-	172	172	-	172	172	0	46,1%	0,0%	320	320	0	40	0
San Miguel	234	234	-	234	234	-	234	234	0	45,2%	0,0%	432	432	0	54	0
San Ramón	130	130	-	130	130	-	130	130	0	28,9%	0,0%	184	184	0	23	0
Santiago	1.091	1.091	-	1.091	1.091	-	1.091	1.091	0	46,5%	0,0%	2.040	2.040	0	255	0
Vitacura	173	173	-	173	173	-	173	173	0	77,8%	0,0%	776	776	0	97	0
El Monte	64	53	11	73	53	20	93	53	40	25,4%	17,8%	122	72	50	9	5
Isla De Maipo	63	46	17	66	46	20	86	46	40	34,8%	23,9%	122	72	50	9	5
Padre Hurtado	112	99	13	119	99	20	139	99	40	28,0%	37,2%	196	136	60	17	6
Peñaflor	155	143	12	163	143	20	183	143	40	26,3%	52,5%	280	200	80	25	8
Talagante	127	102	25	132	102	30	142	102	40	26,0%	18,4%	186	136	50	17	5
O'Higgins	4.397	3.264	1.133	4.554	3.325	1.229	4.798	3.291	1.507			6.672	4.712	1.960	589	196
Codegua	56	30	26	60	30	30	70	30	40	36,5%	25,0%	98	48	50	6	5
Coinco	35	16	19	50	30	20	56	16	40	36,5%	23,0%	74	24	50	3	5
Coltauco	94	52	42	94	52	42	94	52	42	18,3%	21,1%	114	64	50	8	5
Doñihue	100	77	23	107	77	30	117	77	40	32,5%	28,8%	172	112	60	14	6
Graneros	149	133	16	153	133	20	173	133	40	21,5%	17,8%	218	168	50	21	5
Las Cabras	116	43	73	116	43	73	116	43	73	33,9%	21,4%	154	64	90	8	9
Machalí	258	253	5	263	253	10	273	253	20	32,0%	41,2%	406	376	30	47	3
Malloa	63	24	39	69	30	39	69	30	39	23,4%	24,2%	90	40	50	5	5
Mostazal	115	94	21	124	94	30	134	94	40	19,9%	25,3%	170	120	50	15	5
Oliver	60	41	19	61	41	20	81	41	40	24,1%	13,0%	106	56	50	7	5
Peumo	64	53	11	73	53	20	93	53	40	22,4%	16,0%	122	72	50	9	5
Pichidegua	90	42	48	90	42	48	90	42	48	22,5%	23,8%	116	56	60	7	6
Quinta De Tilcoco	60	33	27	63	33	30	73	33	40	20,9%	26,6%	90	40	50	5	5
Rancagua	1.193	1.167	26	1.197	1.167	30	1.207	1.167	40	31,7%	18,8%	1.754	1.704	50	213	5
Rengo	269	210	59	269	210	59	299	210	59	29,2%	21,9%	376	296	80	37	8
Requinoa	119	68	51	119	68	51	119	68	51	23,1%	27,6%	158	88	70	11	7
San Vicente	222	122	100	222	122	100	222	122	100	22,6%	16,3%	280	160	120	20	12
La Estrella	17	6	11	28	8	20	56	16	40	26,0%	32,1%	84	24	60	3	6
Litueche	37	18	19	50	30	20	56	16	40	39,0%	30,0%	84	24	60	3	6
Marchihue	40	19	21	60	30	30	56	16	40	30,9%	37,5%	84	24	60	3	6
Navidad	38	-	38	38	-	38	38	0	38	38,6%	30,5%	60	0	60	0	6
Paredones	36	10	26	46	16	30	56	16	40	30,2%	22,6%	74	24	50	3	5

Región / Comuna	Distribución Inicial Comuna - Área			Primer ajuste: Tamaño mínimo 16 viviendas (con excepciones)			Segundo ajuste: Tamaño Objetivo Mínimo 8 o 10 viviendas. Controla por N° de UPM			Tasas de No Logro		Tamaño con Sobremuestreo					
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	
Pichilemu	92	71	21	101	71	30	111	71	40	44,3%	15,5%	178	128	50	16	5	
Chépica	72	41	31	72	41	31	81	41	40	19,5%	25,7%	98	48	50	6	5	
Chimbarongo	166	96	70	166	96	70	166	96	70	24,2%	18,5%	218	128	90	16	9	
Lolol	34	12	22	46	16	30	56	16	40	30,8%	32,9%	84	24	60	3	6	
Nancagua	83	52	31	83	52	31	92	52	40	27,3%	27,6%	132	72	60	9	6	
Palmilla	57	15	42	58	16	42	58	16	42	27,2%	17,5%	74	24	50	3	5	
Peralillo	56	33	23	63	33	30	73	33	40	23,9%	20,5%	90	40	50	5	5	
Placilla	40	11	29	46	16	30	56	16	40	29,3%	20,7%	74	24	50	3	5	
Pumanque	19	-	19	20	-	20	40	0	40	0,0%	22,4%	50	0	50	0	5	
San Fernando	364	313	51	364	313	51	364	313	51	31,5%	20,3%	524	464	60	58	6	
Santa Cruz	183	109	74	183	109	74	183	109	74	36,7%	17,8%	266	176	90	22	9	
Maule	4.829	3.515	1.314	4.958	3.593	1.365	5.107	3.587	1.520			7.090	5.084	2.006	635	199	
Cauquenes	207	163	44	207	163	44	207	163	44	27,3%	25,7%	284	224	60	28	6	
Chanco	41	18	23	60	30	30	70	30	40	26,8%	9,7%	88	40	48	5	4	
Pelluhue	41	22	19	50	30	20	70	30	40	35,9%	22,2%	98	48	50	6	5	
Curicó	680	605	74	679	605	74	679	605	74	28,0%	22,4%	940	840	100	105	10	
Hualañé	49	29	20	60	30	30	70	30	40	14,3%	15,0%	82	32	50	4	5	
Licantén	33	25	8	40	30	10	70	30	40	19,1%	15,1%	90	40	50	5	5	
Molina	206	165	41	206	165	41	206	165	41	27,6%	27,7%	284	224	60	28	6	
Rauco	49	25	25	60	30	30	70	30	40	10,2%	7,9%	80	32	48	4	4	
Romeral	68	30	38	68	30	38	68	30	38	13,9%	19,8%	82	32	50	4	5	
Sagrada Familia	81	34	47	81	34	47	81	34	47	16,4%	16,8%	100	40	60	5	6	
Teno	127	42	85	127	42	85	127	42	85	14,3%	13,0%	148	48	100	6	10	
Vichuquén	24	6	18	28	8	20	56	16	40	15,0%	16,5%	70	20	50	2	5	
Colbún	98	44	54	98	44	54	98	44	54	27,1%	26,7%	134	64	70	8	7	
Linares	419	343	76	419	343	76	419	343	76	30,2%	22,4%	596	496	100	62	10	
Longaví	135	40	95	135	40	95	135	40	95	25,5%	30,9%	196	56	140	7	14	
Parral	201	147	54	201	147	54	201	147	54	36,1%	20,4%	302	232	70	29	7	
Retiro	94	33	61	94	33	61	94	33	61	22,6%	23,4%	120	40	80	5	8	
San Javier	209	144	65	209	144	65	209	144	65	29,4%	30,7%	298	208	90	26	9	
Villa Alegre	80	48	32	80	48	32	88	48	40	23,9%	18,1%	114	64	50	8	5	
Yerbas Buenas	80	27	53	83	30	53	83	30	53	21,7%	22,9%	110	40	70	5	7	
Constitución	211	171	40	211	171	40	211	171	40	31,7%	37,8%	308	248	60	31	6	
Curepto	50	17	33	63	30	33	56	16	40	20,8%	20,0%	74	24	50	3	5	
Empedrado	19	13	6	26	16	10	36	16	20	29,3%	21,4%	54	24	30	3	3	
Maule	233	184	49	233	184	49	233	184	49	22,2%	21,7%	300	240	60	30	6	
Pelarco	37	11	26	46	16	30	56	16	40	20,0%	16,9%	74	24	50	3	5	
Pencahue	39	15	24	46	16	30	56	16	40	20,8%	17,1%	74	24	50	3	5	
Río Claro	62	19	43	73	30	43	73	30	43	15,9%	19,8%	90	40	50	5	5	
San Clemente	195	94	101	195	94	101	195	94	101	19,7%	27,8%	260	120	140	15	14	
San Rafael	41	21	20	60	30	30	70	30	40	19,4%	23,4%	90	40	50	5	5	
Talca	1.020	980	40	1.020	980	40	1.020	980	40	33,8%	39,5%	1.550	1.480	70	185	7	
Ñuble	3.121	2.135	986	3.186	2.169	1.017	3.319	2.177	1.142			4.696	3.040	1.656	380	164	
Bulnes	135	85	50	135	85	50	135	85	50	17,1%	39,6%	184	104	80	13	8	
Chillán	1.186	1.085	101	1.186	1.085	101	1.186	1.085	101	29,7%	34,2%	1.694	1.544	150	193	15	

Región / Comuna	Distribución Inicial Comuna - Área			Primer ajuste: Tamaño mínimo 16 viviendas (con excepciones)			Segundo ajuste: Tamaño Objetivo Mínimo 8 o 10 viviendas. Controla por N° de UPM			Tasas de No Logro		Tamaño con Sobremuestreo					
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	
Chillán Viejo	187	164	23	194	164	30	204	164	40	31,1%	21,8%	290	240	50	30	5	
El Carmen	79	31	48	79	31	48	79	31	48	27,8%	34,0%	110	40	70	5	7	
Pemuco	53	25	28	60	30	30	70	30	40	24,4%	32,0%	100	40	60	5	6	
Pinto	72	35	37	72	35	37	72	35	37	22,0%	42,9%	118	48	70	6	7	
Quillón	123	69	54	123	69	54	123	69	54	32,1%	44,1%	204	104	100	13	10	
San Ignacio	105	43	62	105	43	62	105	43	62	23,9%	32,6%	146	56	90	7	9	
Yungay	125	92	33	125	92	33	132	92	40	31,3%	28,2%	196	136	60	17	6	
Cobquecura	39	10	29	46	16	30	56	16	40	28,4%	27,3%	84	24	60	3	6	
Coelemu	98	66	32	98	66	32	106	66	40	19,4%	34,4%	140	80	60	10	6	
Ninhue	36	10	26	46	16	30	56	16	40	25,0%	8,1%	72	24	48	3	4	
Portezuelo	33	12	21	46	16	30	56	16	40	21,7%	18,2%	74	24	50	3	5	
Quirihue	85	68	17	88	68	20	108	68	40	19,6%	27,6%	148	88	60	11	6	
Ránquil	41	11	30	46	16	30	56	16	40	18,4%	49,3%	104	24	80	3	8	
Treguaco	39	12	27	46	16	30	56	16	40	25,2%	30,4%	84	24	60	3	6	
Coihueco	161	53	108	161	53	108	161	53	108	33,9%	29,6%	230	80	150	10	15	
Ñiquén	75	8	67	75	8	67	83	16	67	29,2%	22,7%	114	24	90	3	9	
San Carlos	344	214	130	344	214	130	344	214	130	24,0%	24,8%	450	280	170	35	17	
San Fabián	30	12	18	36	16	20	56	16	40	26,2%	9,3%	72	24	48	3	4	
San Nicolás	75	30	45	75	30	45	75	30	45	14,4%	11,5%	82	32	50	4	5	
Biobío	6.292	5.524	768	6.476	5.577	899	6.909	5.571	1.338			9.800	8.032	1.768	1.004	176	
Arauco	143	106	37	143	106	37	143	106	37	22,5%	26,3%	186	136	50	17	5	
Cañete	139	79	60	139	79	60	139	79	60	26,1%	18,6%	174	104	70	13	7	
Contulmo	27	13	14	36	16	20	56	16	40	25,8%	16,9%	74	24	50	3	5	
Curanilahue	125	115	10	125	115	10	155	115	40	19,6%	32,5%	204	144	60	18	6	
Lebu	102	92	10	102	92	10	132	92	40	22,3%	35,6%	180	120	60	15	6	
Los Álamos	84	75	9	85	75	10	115	75	40	24,7%	23,9%	154	104	50	13	5	
Tirúa	42	16	26	60	30	30	56	16	40	30,5%	30,7%	84	24	60	3	6	
Alto Biobío	22	-	22	30	-	30	40	0	40	13,3%	38,1%	70	0	70	0	7	
Antuco	19	9	10	26	16	10	56	16	40	30,6%	22,1%	74	24	50	3	5	
Cabrero	118	87	31	118	87	31	127	87	40	24,7%	25,9%	170	120	50	15	5	
Laja	94	66	28	96	66	30	106	66	40	21,8%	22,7%	138	88	50	11	5	
Los Ángeles	819	612	207	819	612	207	819	612	207	29,5%	19,8%	1.132	872	260	109	26	
Mulchén	117	92	25	122	92	30	132	92	40	23,1%	25,9%	170	120	50	15	5	
Nacimiento	108	90	18	110	90	20	130	90	40	22,3%	25,0%	170	120	50	15	5	
Negrete	38	23	15	50	30	20	70	30	40	30,3%	25,8%	90	40	50	5	5	
Quilaco	18	8	10	18	8	10	56	16	40	20,7%	24,6%	74	24	50	3	5	
Quilleco	41	23	18	50	30	20	70	30	40	26,7%	27,0%	100	40	60	5	6	
San Rosendo	15	12	3	26	16	10	36	16	20	19,3%	18,2%	48	24	24	3	2	
Santa Bárbara	57	32	25	62	32	30	72	32	40	21,1%	17,8%	90	40	50	5	5	
Tucapel	68	56	12	76	56	20	96	56	40	28,2%	21,1%	130	80	50	10	5	
Yumbel	97	53	44	97	53	44	97	53	44	22,6%	33,6%	142	72	70	9	7	
Chiguayante	324	323	1	333	323	10	333	323	10	24,9%	0,0%	442	432	10	54	1	
Concepción	998	980	18	1.000	980	20	1.020	980	40	42,2%	14,6%	1.746	1.696	50	212	5	
Coronel	448	434	14	454	434	20	474	434	40	24,3%	37,7%	636	576	60	72	6	
Florida	47	19	28	60	30	30	70	30	40	23,2%	17,4%	90	40	50	5	5	

Región / Comuna	Distribución Inicial Comuna - Área			Primer ajuste: Tamaño mínimo 16 viviendas (con excepciones)			Segundo ajuste: Tamaño Objetivo Mínimo 8 o 10 viviendas. Controla por N° de UPM			Tasas de No Logro		Tamaño con Sobremuestreo					
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	
Hualpén	346	343	3	353	343	10	363	343	20	24,9%	0,0%	484	464	20	58	2	
Hualqui	93	76	17	96	76	20	116	76	40	24,0%	27,0%	164	104	60	13	6	
Lota	171	170	1	180	170	10	180	170	10	22,9%	0,0%	234	224	10	28	1	
Penco	184	182	2	192	182	10	202	182	20	23,8%	27,8%	270	240	30	30	3	
San Pedro De La Paz	531	530	1	540	530	10	540	530	10	35,0%	0,0%	826	816	10	102	1	
Santa Juana	61	39	22	69	39	30	79	39	40	22,1%	30,0%	108	48	60	6	6	
Talcahuano	565	564	1	574	564	10	584	564	20	28,9%	15,0%	816	792	24	99	2	
Tomé	231	205	26	235	205	30	245	205	40	27,5%	25,6%	330	280	50	35	5	
La Araucanía	4.796	3.371	1.425	4.945	3.472	1.473	5.052	3.402	1.650			7.006	4.836	2.170	604	217	
Carahue	123	65	58	123	65	58	123	65	58	22,4%	19,3%	158	88	70	11	7	
Cholchol	54	19	35	65	30	35	51	16	35	20,4%	25,6%	74	24	50	3	5	
Cunco	93	43	50	93	43	50	93	43	50	17,0%	34,3%	136	56	80	7	8	
Curarrehue	39	12	27	46	16	30	56	16	40	22,4%	28,2%	84	24	60	3	6	
Freire	119	36	83	119	36	83	119	36	83	18,7%	26,0%	158	48	110	6	11	
Galvarino	55	19	36	66	30	36	52	16	36	15,3%	14,8%	60	20	40	2	4	
Gorbea	79	53	26	83	53	30	93	53	40	17,4%	21,8%	114	64	50	8	5	
Lautaro	180	129	51	180	129	51	180	129	51	27,7%	25,6%	246	176	70	22	7	
Loncoche	128	86	42	128	86	42	128	86	42	18,5%	22,3%	154	104	50	13	5	
Melipeuco	32	14	18	36	16	20	56	16	40	33,0%	21,7%	74	24	50	3	5	
Nueva Imperial	154	84	70	154	84	70	154	84	70	25,7%	21,0%	202	112	90	14	9	
Padre Las Casas	346	214	132	346	214	132	346	214	132	20,0%	30,0%	462	272	190	34	19	
Perquenco	37	19	18	50	30	20	56	16	40	28,2%	19,7%	74	24	50	3	5	
Pitrufuquén	131	81	50	131	81	50	131	81	50	21,6%	19,3%	164	104	60	13	6	
Pucón	144	88	56	144	88	56	144	88	56	47,8%	21,7%	238	168	70	21	7	
Saavedra	68	16	52	82	30	52	68	16	52	33,3%	22,1%	94	24	70	3	7	
Temuco	1.432	1.341	91	1.432	1.341	91	1.432	1.341	91	35,3%	21,4%	2.182	2.072	110	259	11	
Teodoro Schmidt	81	29	52	82	30	52	82	30	52	19,4%	14,0%	100	40	60	5	6	
Toltén	52	21	31	61	30	31	70	30	40	27,4%	22,4%	90	40	50	5	5	
Vilcún	138	76	62	138	76	62	138	76	62	21,9%	20,8%	176	96	80	12	8	
Villarrica	291	184	107	291	184	107	291	184	107	33,0%	26,2%	422	272	150	34	15	
Angol	253	229	24	259	229	30	269	229	40	21,9%	22,2%	338	288	50	36	5	
Collipulli	120	88	32	120	88	32	128	88	40	24,5%	30,2%	180	120	60	15	6	
Curacautín	96	70	26	100	70	30	110	70	40	19,9%	12,9%	138	88	50	11	5	
Ercilla	36	18	18	50	30	20	70	30	40	23,4%	28,4%	100	40	60	5	6	
Lonquimay	56	20	36	66	30	36	52	16	36	34,9%	31,9%	74	24	50	3	5	
Los Sauces	38	22	16	50	30	20	70	30	40	26,6%	14,9%	90	40	50	5	5	
Lumaco	50	22	28	60	30	30	70	30	40	23,9%	15,8%	90	40	50	5	5	
Purén	61	40	21	70	40	30	80	40	40	15,5%	11,8%	98	48	50	6	5	
Renaico	51	42	9	52	42	10	62	42	20	19,6%	22,1%	86	56	30	7	3	
Traiguén	94	73	21	103	73	30	113	73	40	23,6%	25,0%	146	96	50	12	5	
Victoria	165	118	47	165	118	47	165	118	47	19,4%	21,7%	204	144	60	18	6	
Los Ríos	3.805	2.652	1.153	3.817	2.654	1.163	3.820	2.637	1.183			5.400	3.760	1.640	470	164	
Futrono	148	82	66	148	82	66	156	90	66	31,3%	34,4%	228	128	100	16	10	
La Unión	381	258	123	381	258	123	381	258	123	24,4%	36,0%	534	344	190	43	19	
Lago Ranco	116	28	88	118	30	88	104	16	88	30,5%	28,4%	144	24	120	3	12	

Región / Comuna	Distribución Inicial Comuna - Área			Primer ajuste: Tamaño mínimo 16 viviendas (con excepciones)			Segundo ajuste: Tamaño Objetivo Mínimo 8 o 10 viviendas. Controla por N° de UPM			Tasas de No Logro		Tamaño con Sobremuestreo					
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	
Río Bueno	334	172	162	334	172	162	334	172	162	26,0%	25,3%	452	232	220	29	22	
Corral	55	35	20	65	35	30	56	16	40	20,9%	35,2%	84	24	60	3	6	
Lanco	163	111	52	163	111	52	171	119	52	26,6%	20,7%	230	160	70	20	7	
Los Lagos	193	91	102	193	91	102	193	91	102	27,7%	24,0%	258	128	130	16	13	
Máfil	73	43	30	73	43	30	83	43	40	25,0%	25,0%	106	56	50	7	5	
Mariquina	200	87	113	200	87	113	200	87	113	23,7%	18,4%	252	112	140	14	14	
Paillaco	193	114	79	193	114	79	193	114	79	24,3%	34,0%	272	152	120	19	12	
Panguipulli	351	148	203	351	148	203	351	148	203	31,5%	26,9%	496	216	280	27	28	
Valdivia	1.598	1.483	115	1.598	1.483	115	1.598	1.483	115	32,0%	26,1%	2.344	2.184	160	273	16	
Los Lagos	3.539	2.564	975	3.619	2.607	1.012	3.826	2.631	1.195			5.466	3.888	1.578	485	157	
Ancud	183	124	59	183	124	59	183	124	59	45,6%	30,8%	314	224	90	28	9	
Castro	198	148	50	198	148	50	198	148	50	27,2%	9,6%	260	200	60	25	6	
Chonchi	69	24	45	75	30	45	75	30	45	29,1%	22,3%	100	40	60	5	6	
Curaco De Vélez	18	5	13	28	8	20	56	16	40	17,0%	33,3%	80	20	60	2	6	
Dalcahue	61	29	32	62	30	32	70	30	40	22,4%	12,0%	90	40	50	5	5	
Puqueldón	18	-	18	20	-	20	40	0	40	0,0%	19,4%	50	0	50	0	5	
Queilén	25	10	15	36	16	20	56	16	40	25,2%	23,5%	74	24	50	3	5	
Quellón	121	76	45	121	76	45	121	76	45	52,6%	26,5%	220	160	60	20	6	
Quemchi	41	12	29	46	16	30	56	16	40	21,7%	16,7%	74	24	50	3	5	
Quinchao	38	15	23	46	16	30	56	16	40	23,0%	6,3%	72	24	48	3	4	
Calbuco	148	68	80	148	68	80	148	68	80	32,0%	20,2%	204	104	100	13	10	
Fresia	58	32	26	62	32	30	72	32	40	19,4%	20,0%	90	40	50	5	5	
Frutillar	86	58	28	88	58	30	98	58	40	35,3%	30,8%	148	88	60	11	6	
Llanquihue	68	53	15	73	53	20	93	53	40	35,4%	16,9%	130	80	50	10	5	
Los Muermos	79	34	45	79	34	45	79	34	45	17,9%	26,8%	100	40	60	5	6	
Mauñín	70	31	39	70	31	39	70	31	39	19,6%	22,5%	90	40	50	5	5	
Puerto Montt	1.055	937	118	1.055	937	118	1.055	937	118	33,8%	16,7%	1.556	1.416	140	177	14	
Puerto Varas	198	143	55	198	143	55	198	143	55	33,1%	22,8%	286	216	70	27	7	
Osorno	682	623	59	682	623	59	682	623	59	28,0%	27,6%	952	872	80	109	8	
Puerto Octay	39	8	31	39	8	31	56	16	40	16,7%	20,7%	70	20	50	2	5	
Purranque	86	58	28	88	58	30	98	58	40	22,1%	27,5%	132	72	60	9	6	
Puyehue	50	20	30	60	30	30	70	30	40	15,6%	25,8%	90	40	50	5	5	
Río Negro	63	30	33	63	30	33	70	30	40	23,1%	30,6%	100	40	60	5	6	
San Juan De La Costa	38	7	31	39	8	31	56	16	40	25,5%	30,9%	84	24	60	3	6	
San Pablo	47	19	28	60	30	30	70	30	40	20,9%	27,2%	100	40	60	5	6	
Aysén	1.353	1.014	339	1.368	1.014	354	1.397	1.014	383			2.070	1.430	640	143	64	
Aysén	304	256	48	304	256	48	328	280	48	6,0%	33,7%	440	370	70	37	7	
Cisnes	84	53	31	84	53	31	99	59	40	12,0%	34,7%	140	80	60	8	6	
Cochrane	54	41	13	61	41	20	36	16	20	20,0%	37,0%	60	30	30	3	3	
Tortel	7	-	7	10	-	10	20	0	20	0,0%	28,2%	30	0	30	0	3	
Coyhaique	763	619	144	763	619	144	787	643	144	13,0%	46,0%	1.190	920	270	92	27	
Lago Verde	16	-	16	20	-	20	20	0	20	0,0%	28,2%	30	0	30	0	3	
Chile Chico	74	45	29	75	45	30	56	16	40	18,0%	40,2%	100	30	70	3	7	
Río Ibáñez	51	-	51	51	-	51	51	0	51	0,0%	33,8%	80	0	80	0	8	
Magallanes	1.803	1.671	132	1.845	1.683	162	1.925	1.693	232			3.194	2.810	384	281	32	

Región / Comuna	Distribución Inicial Comuna - Área			Primer ajuste: Tamaño mínimo 16 viviendas (con excepciones)			Segundo ajuste: Tamaño Objetivo Mínimo 8 o 10 viviendas. Controla por N° de UPM			Tasas de No Logro		Tamaño con Sobremuestreo					
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural	Urbano	Rural	N° Viviendas			N° UPM		
Cabo De Hornos	18	18	-	30	30	-	16	16	0	10,0%	0,0%	20	20	0	2	0	
Laguna Blanca	5	-	5	10	-	10	20	0	20	0,0%	12,9%	24	0	24	0	2	
Punta Arenas	1.420	1.358	62	1.420	1.358	62	1.428	1.366	62	25,0%	36,7%	2.400	2280	120	228	10	
Río Verde	3	-	3	10	-	10	10	0	10	0,0%	12,9%	12	0	12	0	1	
San Gregorio	6	-	6	10	-	10	20	0	20	0,0%	12,9%	24	0	24	0	2	
Porvenir	76	67	9	77	67	10	95	75	20	20,0%	50,7%	168	120	48	12	4	
Primavera	9	-	9	10	-	10	20	0	20	0,0%	12,9%	24	0	24	0	2	
Timaukel	4	-	4	10	-	10	20	0	20	0,0%	12,9%	24	0	24	0	2	
Natales	258	228	30	258	228	30	276	236	40	25,0%	41,4%	474	390	84	39	7	
Torres Del Paine	4	-	4	10	-	10	20	0	20	0,0%	12,9%	24	0	24	0	2	

Elaborado por el Instituto Nacional de Estadísticas.

ANEXO N°3. FUENTES DE INFORMACIÓN Y PROCESOS DE VERIFICACIÓN Y ACTUALIZACIÓN DE UNIDADES PRIMARIAS DE MUESTREO

Anexo N°3.1. Fuentes de Información

Las fuentes de información utilizadas en el proceso de verificación y actualización de unidades primarias de muestreo son utilizadas de manera transversal en procesos de enumeración en terreno y validación de directorios institucionales en gabinete, siendo el objetivo principal, que las direcciones y observaciones dispuestas en ella, contengan la calidad requerida para el proceso de selección de viviendas permitiendo su individualización, de manera que pueda ser ubicada y reconocida en terreno al momento del levantamiento de la encuesta. Las principales fuentes de información utilizadas en la generación de directorios se describen a continuación:

Actualización Cartográfica Continua (ACON)

Corresponde al proceso de actualización continua del Marco Geoestadístico Institucional a través de la incorporación de datos provenientes de registros administrativos, información capturada en terreno e imágenes satelitales. Dicha actualización se realiza tanto a nivel planimétrico como de información de viviendas, edificaciones, infraestructura pública y direcciones asociadas. Este insumo es utilizado como componente para definir la priorización cartográfica, la que nos permite identificar dinamismo en estructura de manzanas para áreas urbanas y entidades de población para áreas rurales, como cantidad de viviendas por cada una o parte de ellas contenidas en las Unidades Primarias de Muestreo (UPM).

Las fuentes principales utilizadas en el proceso de ACON para la obtención de direcciones formales, derivan en su mayoría del proceso de recolección desde las Direcciones de Obras Municipales (DOM), dentro de las que se encuentran:

- **Certificado de Recepción Final de Obras (CRF):** Corresponde al documento oficial emitido por las DOM, que aprueba en forma definitiva la utilización de una o más viviendas dentro de un conjunto habitacional, entendiendo como conjunto habitacional un proyecto horizontal o vertical, cuyo plano original, asociado a un Permiso de Edificación (PE), contempla dos o más viviendas. Las certificaciones pueden ser finales o parciales, dependiendo de si las viviendas recepcionadas abarcan o cierran el proyecto final o aún quedan elementos pendientes por recepcionar. En caso de no existir DOM, el CRF se obtiene a través de la Secretaría Regional Ministerial del Ministerio de Vivienda y Urbanismo (Seremi MINVU). El documento certifica la recepción de obras de nuevas construcciones, modificaciones o ampliaciones, cualquiera sea su destino, para que puedan ser utilizadas o habitadas. Existen varios tipos de CRF, siendo el utilizado en la ACON el que recepciona las obras destinadas para uso habitacional. La información que se obtiene de este

documento corresponde a la cantidad de nuevas viviendas construidas, permitiendo identificar su ubicación y distribución. Los CRF de todo destino, incluyendo los con destino habitacional, se capturan mensualmente por parte de la Unidad de Edificaciones dependiente del Subdepartamento de Estadísticas de Industrias, en todas las regiones del país, que los envía al Subdepartamento de Geografía, ambos pertenecientes al INE, desde donde se extraen aquellos certificados que tienen la información que se utilizará para la ACON: viviendas nuevas asociadas a conjuntos habitacionales, viviendas no pertenecientes a conjuntos habitacionales (individuales) y sus direcciones. Semestralmente se capturan los planos de loteos asociados a los certificados clasificados, generando un listado de conjuntos recepcionados.

- **Planos de Loteo:** Si bien la información que contienen los CRF son RRAA que permiten saber dónde están construidas las nuevas viviendas y cuántas son, otro insumo que tiene relevancia para la actualización continua son los planos de loteo asociados a los CRF. En los planos se especifica cómo se distribuyen las viviendas, graficadas generalmente en una planimetría que detalla lo que involucra la construcción de nuevas viviendas: modificación de manzanas, creación de manzanas nuevas, creación de nueva vialidad, nombres de calles, áreas verdes construidas, entre otros. La información asociada al CRF es una buena fuente para actualizar la base cartográfica, definida por el plano de loteo que, valida su calidad métrica, siendo el requisito más relevante para el proceso de actualización del directorio la correcta geolocalización y la identificación de manera individual de cada una de las viviendas, a través del listado de direcciones asociados al CRF. Finalmente, el Subdepartamento de Geografía captura los CRF semestralmente, a través de sus Direcciones Regionales, siendo productos disponibles para la actualización, el documento oficial de CRF emitido por la DOM y el plano de loteo asociado a dicha recepción. Cada Dirección Regional genera como producto final los planos de loteo con las áreas recepcionadas demarcadas y la capa digital (geodatabase) con el polígono georreferenciado del área recepcionada, información que es el insumo fundamental para el proceso de actualización del directorio de viviendas.
- **Listados de Direcciones:** Corresponde a la asignación oficial de la numeración domiciliaria, emitida por el municipio a cada una de las viviendas recepcionadas. La cantidad de viviendas recepcionadas debe coincidir con las declaradas en el CRF al que corresponden. La captura de esta información es semestral y paralela a la captura de los CRF. El documento oficial que entregan las municipalidades puede ser certificado de número (o listado municipal), certificado de copropiedad, información contenida en el mismo certificado de recepción, numeración oficial incluida en el plano de loteo. Como medio alternativo de captura, se consultan fuentes externas disponibles en portales de servicios básicos (servicios sanitarios, servicios eléctricos, entre otros), o bien en el buscador de direcciones asociadas a un predio del Servicio de Impuestos Internos (SII). El resultado de esta captura es un listado de direcciones recolectadas en el semestre. Este

listado es un insumo elemental para la actualización del directorio de viviendas de cada manzana, pues su objetivo es relacionar cada uno de los registros con el plano de loteo y su vivienda correspondiente, permitiendo dar un orden de recorrido coherente con la altura domiciliar que se observa en terreno. Es importante indicar que este listado de direcciones debe cumplir con el estándar de direcciones que dispone el directorio institucional, el que incluye el tipo de vía, nombre, número, letra, piso, torre/block, y número/letra de departamento, si corresponde.

- **Imágenes de libre disposición:** El proceso también utiliza barridos sobre imágenes satelitales de libre disposición, asequibles a través de servidores de aplicaciones de mapas en la web, los cuales ofrecen imágenes de mapas desplazables, como son fotografías por satélite del mundo, historial de imágenes con distintas fechas de adquisición e incluso la ruta entre diferentes ubicaciones, como también imágenes a pie de calle a través de Google Earth y Google Street View, entre los más utilizados. Esto permite observar cambios en las estructuras de manzanas o áreas de interés, las que son utilizadas como referencia y apoyo en la localización de crecimientos formales como también aquellos que aún no están disponibles en procesos formales de recolección de datos. Estos son considerados como alertas de áreas de crecimiento y/o decrecimientos, que son incorporados posteriormente como parte de las variables de priorización, a través de la identificación de áreas que se identifican como capas de información geográficas.

Para la ejecución del proyecto de enumeración Casen 2022 la base de información disponible de ACON contempla 215 comunas con actualización hasta junio de 2021 y 130 comunas con cartografía (ACON) con actualización hasta diciembre de 2020, como indica en detalle la Tabla IX.9:

Tabla IX.9. Versión Cartográfica ACON, según año a nivel de región

Región	Comunas Cartografía (ACON) 2020	Comunas Cartografía (ACON) JUNIO 2021	Total Comunas
Nivel País	130	215	345
Arica y Parinacota	2	2	4
Tarapacá	3	4	7
Antofagasta	4	5	9
Atacama	4	5	9
Coquimbo	4	11	15
Valparaíso	11	27	38
Metropolitana	0	52	52
O'Higgins	12	21	33
Maule	19	11	30
Ñuble	14	7	21
Biobío	15	18	33
La Araucanía	11	21	32
Los Ríos	4	8	12
Los Lagos	16	14	30
Aysén	5	5	10
Magallanes	6	4	10

Elaborado por el Instituto Nacional de Estadísticas.

Marco Georreferenciado de Edificaciones (MGE)

Otro de los insumos relevantes dentro del proceso de actualización de directorios es el Marco Georreferenciado de Edificaciones (MGE), que corresponde a la vinculación geográfica de los catastros de direcciones georreferenciadas del Ministerio de Vivienda y Urbanismo (MINVU) con los directorios de Precenso 2016, los que se complementan con la base ACON. Se encuentra conformado por áreas geoestadísticas y administrativas, además de elementos estructurales del territorio que permiten su identificación operativa. Para comprender el proceso de actualización derivado de la utilización del MGE es necesario conocer los insumos y procedimientos utilizados en su generación, los que se pasan a describir.

El insumo principal utilizado para la construcción del MGE corresponde a las direcciones georreferenciadas de edificaciones dispuestas por MINVU, que contienen una ubicación precisa en el espacio geográfico con información de tipo de vía, nombre de vía y altura domiciliaria de cada dirección. Considera dos tipos de insumos al ámbito geográfico del INE: a) Caracterización de predios y vías con georreferenciación de direcciones, b) Plataforma DOM en Línea. El proceso de construcción del MGE se constituyó en el traspaso y ajuste de las coordenadas de las edificaciones hacia el Marco Geoestadístico Institucional, que permitió la vinculación al maestro de direcciones y de calles disponible. Esta información corresponde a edificaciones con información de direcciones que no necesariamente

representan viviendas. De lo anterior, se pueden identificar edificaciones con destino industrial o de comercio, las cuales corresponden a edificaciones con direcciones válidas y existentes. Otro ejemplo, corresponde a las construcciones en altura, donde una edificación puede contener cientos de viviendas, donde la edificación permite vincular esas unidades a la dirección principal.

En la construcción del MGE, se integraron las direcciones georreferenciadas al Marco Geoestadístico Institucional dentro de un contexto espacial de comuna, manzana y ejes de calles para áreas urbanas y a nivel de comuna sobre áreas rurales. Para ello, requirió de cuatro insumos esenciales que contribuyeron a su construcción, dentro de ellas, están:

- **Cartografía MINVU:** Corresponde a un catálogo de capas cartográficas dispuestas por MINVU, de las cuales se extraen edificaciones con direcciones y coordenadas. Esta base cartográfica contiene capas de información para 146 comunas de 15 regiones del país, cabe destacar que la distribución de direcciones dispuestas en el catálogo MINVU se concentra en un 99% en las áreas urbanas de cada comuna. En la construcción del MGE, se consolidan todas las direcciones urbanas dispuestas a nivel comunal en una sola capa de información geográfica de puntos, para así facilitar la administración y seguimiento de toda la base de edificaciones. Las regiones Metropolitana y de Valparaíso concentran más de la mitad de las direcciones MINVU disponibles.
- **Certificados de recepción final (CRF):** De este insumo se extraen nombres de calles oficiales y direcciones georreferenciadas de viviendas recepcionadas. Esto se realiza a partir del listado de direcciones provenientes de los CRF, los cuales son georreferenciados e integrados al MGE.
- **Centroide de manzana de Precenso 2016:** Corresponde a la capa geográfica de punto que corresponde a una abstracción del polígono de manzana de Precenso 2016 que permite dar completitud territorial al MGE, pero con coordenadas menos precisas. El operativo de recolección precensal 2016 fue una etapa previa de preparación al último Censo de población y vivienda 2017. En dicho Precenso, se recolectaron datos censales con información geográfica, direcciones de edificaciones y número de viviendas, entre otros. Si bien estos datos censales tienen un contexto geográfico a través de la manzana, estos son poco precisos porque quedan circunscritos al interior de ella, sin una especificación espacial exacta de su posición o del frente de manzana en el cual se encuentran ubicados. Lo anterior permite dar completitud territorial al marco de edificaciones donde no se disponga de información del MINVU, puesto que estas áreas son actualizadas con otros RRAA disponibles o enumeraciones de otras muestras enumeradas por el INE. Con lo anterior, es que se establece el centro geométrico de cada manzana (centroide) como insumo para la construcción del marco de edificaciones como alternativa para georreferenciar el maestro de direcciones en aquellas manzanas y/o comunas donde no se disponga de direcciones provenientes del MINVU o CRF. Cabe destacar que este insumo no constituye una dirección, ni menos una vivienda, solo es una representación geométrica abstracta de una manzana que

permite dar una georreferenciación aproximada de la base de datos del maestro de direcciones, en caso de aquellas comunas donde no se dispongan de fuentes de mayor precisión.

- **Actualización Cartográfica Continua (ACON):** De aquí se extrae todo el set de capas geográficas disponibles, en particular manzanas y ejes de calles actualizados. Dentro de la información principal utilizada se encuentra:
 - Manzana: Capa de polígono que se utiliza como contenedor geográfico, que permite contextualizar y ubicar espacialmente en el territorio una dirección georreferenciada.
 - Ejes viales: Capa de líneas que representa la red caminera del país y que contiene ejes de calles y caminos con información de tipos de vías, nombres, comuna y ciudad a la que pertenecen.
 - CRF: Capa de polígonos que representan la ubicación y delimitación de nuevas construcciones con certificado de recepción final, generado por las direcciones de obras municipales. Contiene información relevante como el año y la cantidad de viviendas recepcionadas, entre otros.
 - Direcciones Georreferenciadas CRF: Capa de punto que contiene direcciones georreferenciadas asociadas a listados de direcciones, que a partir del año 2020 son parte del insumo de CRF entregado por las direcciones de obras municipales y gestionadas por la institución.

El proceso de mejora de direcciones incorporado en el MGE incorpora definiciones, implementación de estándares, conceptualizaciones, tratamiento y administración de información geoespacial.

Para conceptualizar un MGE, es necesario destacar que existe una estrecha relación con el concepto de *dirección*, dado que una edificación puede ser representada a través de una *dirección georreferenciada*. Bajo este contexto, se puede indicar que la dirección se comprende como un objeto geoespacial que permite identificar y posicionar una edificación (o lugar) dentro de un contexto general y diferenciarlo de otra edificación vecina.

El concepto principal de esta especificación de datos es que toda dirección tiene cuando menos un localizador, que facilita a los usuarios distinguir una edificación respecto de unidades vecinas, y una posición geográfica. Lo que, en su conjunto, permite potenciar el uso y administración de estos datos espaciales en las aplicaciones de los SIG y de geolocalización.

Para identificar una edificación de forma inequívoca, en un contexto más amplio, la dirección de la edificación debe asociarse a un número de componentes de dirección que definen su ubicación dentro de un área geográfica determinada. Cada una de las componentes representa un Identificador espacial,

como, por ejemplo: el nombre de calle, comuna, código postal, región o país, los cuales en su conjunto permiten dar mayor precisión a la ubicación del elemento.

Como se mencionó en los párrafos anteriores, para la construcción del MGE se considera como insumo base las direcciones georreferenciadas provenientes del MINVU, que contienen localizaciones de 3,2 millones de direcciones a nivel nacional, contenidas prácticamente en su totalidad en áreas urbanas del país. A partir de este insumo, se establecen seis actividades secuenciales de trabajo que contemplan la estandarización, mejora e integración de estas direcciones al Marco Geoestadístico Institucional.

Estas actividades parten con el proceso de ajuste cartográfico que corresponde a la integración de las direcciones de MINVU a la cartografía INE, este ajuste permite asociar cada dirección a las manzanas censales y corregir su posición en base a los distintos parámetros cartográficos y geodésicos, con lo que ambas cartografías fueron construidas. Luego se define el estándar de direcciones, el que se basa en la identificación geográfica, componentes principales y componentes complementarios, los que permiten definir la localización de la dirección en una estructura u organización única y segmentada que facilita su construcción, mantención y administración. Posteriormente, cada dirección requiere de la individualización de cada elemento que componga el MGE. Para esta individualización se garantiza que cada objeto que lo componga sea único y específico, a través de un identificador generado a partir de la coordenada geográfica y vinculada al Maestro de direcciones.

Para la construcción de este identificador, denominado ID_ADDR, se utiliza un número único generado mediante un algoritmo de código abierto. Este identificador garantiza la personalización única e inequívoca de cada dirección georreferenciada, el que es integrado a cada dirección del MGE. Luego, a partir de los nombres de calle y de los ejes de la ACON, se realiza una normalización de estos, identificando las distintas formas en que se puede escribir un nombre al interior de una comuna y normalizándolo a uno en común dentro de esta. Con este proceso se busca que una calle, de la cual se tiene conocimiento que es la misma, no tenga diferencias en cómo se escribe su nombre, asegurando con ello, el proceso de vinculación. El MGE requiere de nombres de calles normalizados con el fin de identificar cada calle con un nombre válido y único al interior de una comuna. Para este proceso se utiliza como insumo el maestro de calles normalizado, que se actualiza en paralelo al marco de edificaciones, utilizando códigos de calles únicos que permiten traspasar cada uno de estos nombres de calles normalizados a su respectiva edificación. Esta asociación se realiza automáticamente para las direcciones georreferenciadas, con el apoyo de desarrollos de programación en Python; del resultado, aproximadamente entre el 10% y el 20% se debe asociar manualmente en aquellas direcciones que no lograron ser asociadas en Python.

De esta forma, cada dirección de una edificación se asocia e integra al Marco Geoestadístico Institucional de manera directa y codificada, manteniendo atributos tales como el tipo de vía y nombres de calle del eje vial.

Finalmente, el maestro de direcciones, al tener como información de origen los listados de viviendas y otros usos de Precenso 2016, cuenta no solo con listados de direcciones (nombre de calle y altura domiciliaria), sino que también con códigos únicos de manzana (MANZENT) que permiten asociar direcciones en áreas geográficas al interior de cada comuna.

A fin de obtener un mínimo error, en la vinculación del maestro de direcciones al MGE, se construye una nomenclatura en base a los componentes básicos de dirección como lo son el nombre de la vía, la altura domiciliaria y la unidad geográfica que la contiene: la manzana. De esta forma, se obtiene una vinculación de los listados de viviendas y otros usos del maestro de direcciones al MGE, otorgándoles una posición precisa en el frente de manzana que le corresponde. El éxito de esta vinculación depende de la calidad y de la disponibilidad de insumos como direcciones validas en el maestro de direcciones y de direcciones georreferenciadas.

Priorización de UPM a enumerar

La priorización permite identificar aquellas UPM que requieren la actualización de sus directorios, definido por aquellos que demandan generar nuevamente un listado de direcciones de viviendas y aquellos donde se puede utilizar el directorio institucional disponible. Los porcentajes se encuentran previamente definidos, debiendo seleccionar de la muestra Casen 2022 un 40% de las UPM para ser actualizadas en terreno a través de un proceso de enumeración y el 60% restante, obtener y validar sus directorios a través de un proceso de revisión en gabinete. Este método de priorización utiliza varias componentes, dentro de las que se indican:

- **Priorización Cartográfica:** Esta priorización comienza ejecutando una metodología de priorización en base a una serie de variables que se aplican a todas las manzanas de la ACON que presenten una variación importante en los totales de viviendas, cambios de estrato y/o cambios de estructura, producto del análisis comparativo entre la información levantada en el Censo 2017 e información levantada en las distintas muestras enumeradas por la institución, información de registros administrativos recopilados post Censo (CRF⁸¹), así como otras fuentes de información (planos de loteo, campamentos, tomas, imágenes de satélite, entre otros).

El resultado de la metodología de priorización es un producto cartográfico que se genera sobre la manzana ACON el que identifica, a través del método binario, presencia o ausencia de cada variable para cada manzana (ACON)⁸², permitiendo establecer por sumatoria en cada manzana ACON, aquellas que presentan un mayor número de variables presentes en ella. Posteriormente, al aplicar el método de priorización, las manzanas ACON se cruzan geográficamente con las áreas

⁸¹ Para Casen 2022, la información de CRF incluye datos de viviendas hasta el primer semestre del 2021.

⁸² Las Manzanas ACON corresponden a la base de datos geométrica y tabular de la estructura urbana de las ciudades, generadas a partir del proceso de la actualización continua.

de UPM seleccionadas en la muestra Casen 2022, considerando que las UPM son conglomerados de viviendas y en algunos casos, su área corresponde a una parte de la manzana ACON. Para determinar si la UPM está afectada por la variable que indica el resultado de la priorización, se realiza una revisión geográfica de la selección en un post proceso para depurar y validar que el resultado de la priorización aplica a la UPM seleccionada.

La metodología utilizada para la priorización de la muestra Casen 2022 incorpora distintas variables, dependiendo de la información disponible a la fecha de ejecución del proceso y del área a priorizar (urbana o rural), cada una de ellas, alerta sobre la presencia de ciertas características en cada UPM, las que indican posibles cambios directamente basados en la estructura de las manzanas, cantidad de viviendas, calidad de las direcciones y, en definitiva, en la calidad del directorio disponible para la selección.

Uno de los factores más preponderantes dentro de las variables, que se consideran estructurales, son todas aquellas que se ven afectadas por la incorporación de nuevas construcciones con fines habitacionales y que son capturadas por el INE, a través de los Certificados de Recepción Final (CRF) los que para esta priorización corresponden a capturas hasta junio del año 2021.

El Cuadro IX.1 presenta las variables incluidas en el modelo de priorización para áreas urbanas que incluye 14 variables, que se aplican a las manzanas de la versión cartográfica (ACON) disponible:

Cuadro IX.1. Variables de la metodología de priorización cartográfica en áreas urbanas

Variables	Descripción	Ámbito de actualización
Cambio Codificación Censo vs Precenso 2016 (PC16)	Manzana creada Post Precenso 2016 (no tienen trazabilidad de directorios de viviendas)	Directorio
Cambios Estructurales en Manzanas	Actualización de estructuras y códigos de manzanas (fusión-división-creación)	Directorio
Incorporación Formal de Viviendas	Manzana que tiene actualización de viviendas por CRF (crecimiento / decrecimiento)	Direcciones / Directorio
Incorporación por Loteo	Manzana que tiene ingreso de viviendas por Loteo confirmadas por imágenes satelitales	Direcciones / Directorio
Incorporación por Imagen	Manzana que tiene actualización por imagen satelital, confirmado por estructuras homogéneas cercanas.	Direcciones / Directorio
Informalidad-Campamentos	Manzana identificada en un área de dinamismo informal (base de campamentos)	Viviendas / Estructura / Direcciones / Directorio
Áreas Pendientes	Manzana identificada en un área marcada como Pendiente en proceso de actualización (inconsistencia y/o incerteza de viviendas y/o estructuras)	Viviendas / Estructura / Direcciones / Directorio
Precenso 2016 vs Censo 2017	Considera diferencias de ± 5 viviendas entre viviendas Precenso 2016 versus Censo 2017	Viviendas / Directorio
Cambio Estrato	Identifica Manzana con cambios de estrato post Ajuste Geográfico (rural a urbano) y Manzanas fuera del Limite Urbano Censal (LUC) 2017	Directorio
Actualización Marco Maestro - División UPM	Manzana que dado un crecimiento formal (CRF) requiere de un directorio para facilitar la División y/o reconfiguración de UPM	Directorio
Direcciones vs RRAA	Diferencias entre el directorio mejorado Precenso 2016 vs los RRAA de Direcciones (MINVU, SISS y ADDR_CRF)	Direcciones
Hacinamiento	Identifica manzanas con 15% o más de viviendas con hogares en situación de hacinamiento crítico.	Viviendas / Directorio
Allegamiento	Identifica Manzanas con 15% o más de viviendas con hogares en situación de allegamiento externo.	Viviendas / Directorio
Revisión Directorio Precenso 2016	Identifica áreas marcadas para terreno en revisión del directorio de Precenso 2016, donde se analizan inconsistencias de la información en relación con las actualizaciones disponibles.	Viviendas / Directorio

Elaborado por el Instituto Nacional de Estadísticas.

El Cuadro IX.2, presenta las variables incluidas en el modelo de priorización para áreas rurales, que incluye 6 variables, que se aplican a cada UPM del MMV 2020 de la selección muestra Casen:

Cuadro IX.2. Variables de la metodología de priorización cartográfica de áreas rurales

Variables	Descripción de la variable	Ámbito de incidencia
Actualización Formal	Identifica CRF en el ámbito rural	Direcciones / Directorio
Cambios Estrato	Identifica UPM con cambios en el límite urbano censal (LUC).	Directorio
Informalidad-Campamentos	Manzana identificada en un área de dinamismo informal (base de campamentos)	Viviendas / Estructura / Direcciones / Directorio
Aldeas	Identificación de UPM de Aldeas con características urbanas (potencial de dinamismo)	Viviendas / Directorio
Precenso 2016 – Censo 2017	Diferencias de +5 viviendas Censo 2017 versus Precenso 2016	Viviendas
Áreas pendientes	Identificación de áreas de crecimiento no reportado en cartografía (ACON) 2020	Viviendas / Estructura / Directorio

Elaborado por el Instituto Nacional de Estadísticas.

En términos de selección de las UPM elegidas por el concepto de priorización cartográfica tienen mayor preponderancia aquellas que poseen un mayor número de variables estructurales y que incidan en el cambio del directorio institucional disponible.

- **Priorización Metodológica:** Corresponde al análisis que se realiza a las viviendas del directorio contenidas en la UPM. La finalidad de esta priorización es corroborar en los directorios la existencia de viviendas suficientes para la selección de la muestra. Las UPM que quedan marcadas por esta condición son todas incorporadas al proceso de enumeración en terreno. Para ello se evalúa lo siguiente:
 - El rango de viviendas es $\leq 0,6$ o $\geq 1,4$ entre las viviendas marco del directorio de UPM versus las viviendas particulares ocupadas (VPO) presentes en el directorio de la UPM del Marco de viviendas 2020. Las UPM fuera de rango, indican que no existen viviendas suficientes para la selección en base a la cantidad de viviendas particulares ocupadas. Estos casos se presentan cuando se tiene viviendas que estaban construidas al momento del Censo, pero aún estaban en condición de desocupadas. Por ello, en estos casos, se prioriza para actualizar su ocupación, como también muchas veces completar direcciones formales para disponer de un mejor directorio.
 - No existe directorio. Estos casos corresponden a cuando la UPM ha sido creada solo con viviendas por CRF o parte de ellas y no se dispone del listado de direcciones que complementen el directorio para su posterior selección de viviendas. Esto ocurre en el MMV 2020 por su proceso de actualización, el que incluye incorporación de viviendas por CRF, pero no todas las capturas post Censo disponen de un listado de direcciones formales o definitivas

para actualizar sus directorios, por lo tanto, la única fuente de captura para estos casos es la enumeración en terreno y por ello deben priorizarse de manera obligatoria.

- **Priorización Regional:** Con base en este análisis, también se incorpora el conocimiento regional, donde los equipos que permanentemente salen a terreno disponen de información actualizada del territorio, que puede incluir conocimiento de cambios estructurales dentro de sus áreas geográficas, como, por ejemplo, la instalación de campamentos, y erradicaciones o cambios en estructuras viales, como también de situaciones contingentes que puedan incidir en la priorización.
- **Otras consideraciones en la priorización:** Entre otros aspectos considerados para el análisis de la priorización están las observaciones derivadas de la construcción del Marco Maestro, donde se integran a la priorización todas aquellas UPM que puedan contener viviendas con ubicaciones tentativas derivadas del Censo 2017.

Por último, que dentro de la priorización se descartan aquellas UPM que poseen directorios actualizados en otros levantamientos de terreno realizados en el INE entre los años 2020 y 2022.

Enumeración

Los requerimientos para el proceso de enumeración de la muestra Casen, están relacionados principalmente con los productos cartográficos, herramientas informáticas y de sistemas de información geográfico requeridos para la planificación, procesos operativos y ejecución de las actividades de enumeración, considerando que este proceso para Casen 2022 es 100% digital para áreas urbanas y rurales. Dentro de los requerimientos se encuentran:

- **Mapas Teselares:** Corresponde a un mapa geográfico actualizado, que utiliza para su construcción la base cartográfica ACON. Este corresponde a paquete de teselas de mapas denominadas TPK y que es utilizado para visualizar las unidades primarias de Muestreo (UPM) sobre dispositivos móviles de captura (DMC) durante el proceso de enumeración en terreno. Se requiere un TPK por cada UPM priorizada para terreno.
- **Visualizador Web:** Plataforma informática donde se realiza la supervisión y control de carga operativa y análisis del levantamiento en terreno, tanto para áreas urbanas y rurales.
- **Aplicativo de Enumeración Casen:** Plataforma informática que se incorpora en el DMC y que utiliza archivos de transferencia de información geográfica WKT⁸³ utilizados para representar dentro de los servicios web, las unidades de levantamiento definidas para el proceso de

⁸³ Well Know as Text, corresponde a un estándar OGC de transferencia de información geográfica.

enumeración. Este aplicativo requiere de un archivo Spatialite⁸⁴, donde se almacenan los atributos y geometrías de las unidades de levantamiento, dentro de la aplicación, estas áreas se interpretan como unidades de asignación y actúan como restrictores de levantamiento, es decir, no puede recolectarse información fuera de estas áreas.

- **Visualizador Geográfico GEOCAsEN 2022 Enumeración:** Visualizador geográfico que contiene la información cartográfica institucional, de la muestra Casen 2022 y su selección, requerido para la ejecución y planificación del proceso operativo utilizado por supervisores(as) y enumeradores (as) para reconocer las UPM previamente a la visita y recorrido en terreno. Su principal función es disponer de manera geoespacial información oficial y actualizada, que sea útil para la gestión del proyecto. Con esta herramienta institucional podrán explorar, consultar y analizar elementos de forma visual interactuando con las capas de información publicadas de cada área seleccionada en la muestral.
- **Formulario de Enumeración:** Instrumento de recolección de información, que contiene una serie de ítems que buscan la caracterización de la dirección y del sitio y/o edificación al interior de una UPM. El formulario de enumeración se encuentra inserto dentro del aplicativo del DMC.
- **Formulario de supervisión directa:** Instrumento mediante el cual se evalúa el desempeño del recolector, según protocolos, a través de una pauta de observación, la que se aplica durante el acompañamiento al recolector en el proceso de enumeración.
- **Formulario de Actualización Cartográfica:** instrumento en el cual se registran cambios en la conformación de manzanas que componen la UPM.

Directorios institucionales

Uno de los componentes principales de la conformación del directorio de la selección de viviendas de la muestra Casen es disponer de información de los directorios institucionales disponibles. Estos corresponden a directorios de Precenso 2016, como también todos aquellos directorios que estén disponibles en procesos de enumeración provenientes de otras muestras, correspondientes a UPM que han sido seleccionadas para muestras de encuestas sociales o económicas, enumeradas por el INE y que son coincidentes con la selección de la muestra Casen 2022. Estas muestras corresponden a enumeraciones en terreno ejecutadas desde el año 2020 a la fecha, las que han sido priorizadas previamente por el método de priorización definidos en los puntos IV.2.3 mencionados anteriormente. Este proceso de validación de directorios en gabinete se genera a través de la plataforma informática

⁸⁴ Base de datos espacial en formato sqlite 3.

denominada SGOH⁸⁵, correspondiente al sistema informático de gestión utilizado para la revisión y validación de dichos directorios.

Anexo N°3.2. Descripción de procesos de enumeración

El desarrollo del proceso de enumeración integró varias actividades que se describen a continuación.

Dotación

El equipo encargado del proceso de enumeración está definido por una serie de perfiles de cargo con distintas responsabilidades y objetivos de trabajo, en este proceso se integran equipos permanentes del Departamento de infraestructura estadística de INE quienes interactúan en el flujo de generación del proceso de generación de directorios actualizados en terreno. A continuación, se describe cada una de sus funciones dentro del proyecto Casen:

- **Coordinador de Insumos Geográficos:** realiza la planificación general de las diferentes etapas de la enumeración, la definición de las estrategias, la coordinación con los diferentes equipos y regiones, el diseño, proyección y monitoreo de las actividades de terreno.
- **Coordinador SIG:** coordina, en conjunto con sus analistas, los procesos de extracción y generación de directorios de los marcos y es responsable de la generación de los directorios finales que se envían a selección. Además, es el responsable de la gestión de las plataformas informáticas y geográficas que dan soporte al proceso de enumeración.
- **Coordinador Cartografía:** Coordina, en conjunto con sus analistas, la priorización cartográfica y los productos digitales, requeridos para el proceso de enumeración.
- **Coordinador de Infraestructura y RM:** Efectúa el control de las actividades asociadas al operativo regional y la coordinación con los equipos a nivel central.
- **Coordinador Regional:** Efectúa la programación de actividades del levantamiento, el diseño e implementación de las estrategias operativas, además, de la coordinación y control del equipo de recolección (supervisores – enumeradores), responsable del estado de avance regional según plan operativo.
- **Gestor Técnico:** Las actividades del rol guardan relación con el enlace de información para priorización de áreas de recolección y gestionar técnicamente los recursos

⁸⁵ Sistema de Gestión Operativa de Hogares.

destinados al proyecto. También se encarga de desarrollar el análisis espacial de las áreas de selección con otras muestras para la generación de directorios.

- **Supervisor Operativo:** Organiza, coordina y monitorea las actividades del equipo de recolectores, junto con las actividades de supervisión y retroalimentación en terreno.
- **Enumerador:** Persona a cargo de recolectar la información asociada al proceso de enumeración en terreno. Para ello, debe recorrer el territorio asignado, recolectar la información base en cada edificación y el uso de estas, comunicar hallazgos o dificultades que se presentaran en el proceso a su supervisor. A su vez, es el encargado de registrar las actualizaciones cartográficas de las áreas enumeradas.
- **Supervisor de Análisis:** El objetivo del rol, es orientar a los analistas, resolver consultas, asignar cargas y monitorear constantemente al equipo de trabajo, revisar así los rechazos identificados por los analistas y canalizarlos a los encargados de grupo, verificar y revisar actualizaciones que presentaron modificaciones en terreno y realizar supervisiones aleatorias a formularios revisados por cada uno de los analistas asignados a su equipo.
- **Analistas de datos:** Este rol ejerce el control de calidad del 100% de los formularios recolectados en terreno, así como la revisión en gabinete de las unidades no priorizadas previo a la conformación del directorio de viviendas seleccionadas, con el fin de identificar que los registros efectivamente correspondan a lo requerido para el estudio. En el caso de la Encuesta Casen, viviendas particulares ocupadas, y que las direcciones, tipos de vías y accesos sean acordes a la unidad analizada.
- **Analista Geoespacial:** La función del rol es la de disponer de los insumos cartográficos, desde el modelo de priorización cartográfica, los insumos para la ejecución de la recolección de información en terreno y la generación de mapas móviles.
- **Analista BBDD:** Su función es disponer de los directorios de terreno y gabinete en el sistema institucional para que puedan estar disponibles para la selección de primera y segunda etapa.

El equipo que se integró al proceso de la enumeración de las UPM priorizadas consta de un total de 340 funcionarios, indicados en Tabla IX.10, distribuidos según cargas de trabajo derivadas de la selección de la muestra para cada una de las regiones, apoyados por los equipos internos del área de Geografía del INE en cada una de ellas. El análisis de los directorios generados en terreno como los de gabinete se realiza centralizadamente en la Región Metropolitana.

Tabla IX.10. Dotación equipo operativo para la enumeración a nivel de región

Región	Coordinador Regional	Supervisor Operativo	Recolector	Supervisor de Análisis	Analistas de Datos	Analistas Geoespaciales
Nivel País	16	42	239	8	33	2
Arica y Parinacota	1	2	10	-	-	-
Tarapacá	1	2	13	-	-	-
Antofagasta	1	2	10	-	-	-
Atacama	1	2	11	-	-	-
Coquimbo	1	2	9	-	-	-
Valparaíso	1	4	24	-	-	-
Metropolitana	1	8	51	8	33	2
O'Higgins	1	2	13	-	-	-
Maule	1	3	15	-	-	-
Ñuble	1	2	9	-	-	-
Biobío	1	4	21	-	-	-
Los Ríos	1	2	12	-	-	-
La Araucanía	1	3	15	-	-	-
Los Lagos	1	2	11	-	-	-
Aysén	1	1	7	-	-	-
Magallanes	1	1	8	-	-	-

Elaborado por el Instituto Nacional de Estadísticas.

Capacitación

La capacitación para el equipo de enumeración tuvo la modalidad telemática vía Teams⁸⁶, fundamentado en los aforos permitidos en el plan “Paso a Paso nos cuidamos”⁸⁷. Esta capacitación indica los criterios y metodologías utilizadas en el levantamiento de información, además de la usabilidad del dispositivo móvil de captura (DMC). Dicha capacitación contempló cuatro etapas:

- Preparación de la capacitación e insumos.
- Capacitación de Coordinadores Regionales y Supervisores Operativos.
- Capacitación de Enumeradores.
- Capacitación de Analistas.

Como herramienta de apoyo, se dispuso de la plataforma interna “Academia Virtual INE”, donde a través de acceso remoto, cada integrante del equipo de trabajo, según su rol, refuerza de manera interactiva

⁸⁶ Corresponde a un chat empresarial que permite trabajar juntas a personas de un mismo equipo, aunque no pertenezcan a la misma organización. Funciona como un escritorio virtual, a través del cual se puede acceder a todas las actividades necesarias para trabajar, desde cualquier dispositivo con conexión a Internet.

⁸⁷ El Plan Paso a Paso nos cuidamos se implementó en julio de 2020 como una estrategia para enfrentar la pandemia según la situación sanitaria de cada zona. Desde su implementación, la situación sanitaria, económica, social del país ha cambiado, por lo que ha sido necesario evaluar y actualizar los alcances de este plan, el que se ha desarrollado en un estado de alerta sanitaria. Este plan tiene 3 fases: bajo impacto sanitario, medio impacto sanitario y alto impacto sanitario, que verían de acuerdo a la situación sanitaria de cada territorio. <https://saludresponde.minsal.cl/nuevo-plan-paso-a-paso/>.

cada uno de los conceptos revisados en la capacitación online. Dentro de esta actividad, también se dispusieron de instructivos y manuales a cada uno de los integrantes, según el rol abordado dentro del proyecto.

El material de apoyo generado consistió en 3 manuales para enumeradores, supervisores y coordinadores, 8 infografías, 2 pautas de supervisión y 4 Presentaciones.

El proceso de capacitación comenzó de manera formal para todo el equipo que se integró al proyecto Casen, partiendo del lunes 14 hasta el 25 de marzo de 2022. Como el proceso de contratación de personal fue variando en los primeros meses de inicio del proyecto, esta actividad fue permanente para aquellos nuevos ingresos a nivel nacional.

Proceso operativo

Las actividades de recolección, supervisión y análisis de la enumeración se desarrollaron a partir del miércoles 28 de marzo de 2022, posterior a la etapa de capacitación. La distribución de las áreas a enumerar quedó supeditada a la definición de la priorización para cada una de las UPM y estratos, iniciando por UPM urbanas, por contener mayor carga de UPM (85,8%), frente a las UPM rurales (14,2%). La distribución se puede observar a nivel regional en la Tabla IX.11 donde, los porcentajes obtenidos, tanto en el área urbana como rural, son calculados con respecto al total nacional de UPM priorizadas. El término del proceso de enumeración finalizó el 31 de Julio en regiones con mayor carga y dificultades en el proceso de enumeración (Regiones Metropolitana, Valparaíso, La Araucanía).

Tabla IX.11. Distribución regional de UPM priorizadas según área urbano/rural. Porcentajes según total nacional.

Región	UPM Muestra Casen 2022		Urbano			Rural		
	Total UPM	UPM Terreno	Total UPM	UPM Terreno	% Terreno	Total UPM	UPM Rural	% Terreno
Nivel País	12.545	4.987	10.602	4.280	85,8	1.943	707	14,2
Arica y Parinacota	364	180	335	167	3,3	29	13	0,3
Tarapacá	512	229	482	208	4,2	30	21	0,4
Antofagasta	563	224	525	199	4,0	38	25	0,5
Atacama	506	201	451	181	3,6	55	20	0,4
Coquimbo	500	201	412	152	3,0	88	49	1,0
Valparaíso	1.326	441	1.138	389	7,8	188	52	1,0
Metropolitana	2.814	1.088	2.668	1.048	21,0	146	40	0,8
O'Higgins	785	315	589	221	4,4	196	94	1,9
Maule	834	355	635	282	5,7	199	73	1,5
Ñuble	544	218	380	162	3,3	164	56	1,1
Biobío	1.180	469	1.004	413	8,3	176	56	1,1
La Araucanía	821	335	604	291	5,8	217	44	0,9
Los Ríos	634	254	470	176	3,5	164	78	1,6
Los Lagos	642	253	485	192	3,9	157	61	1,2
Aysén	207	95	143	81	1,6	64	14	0,3
Magallanes	313	129	281	118	2,4	32	11	0,2

Elaborado por el Instituto Nacional de Estadísticas

Proceso Supervisión

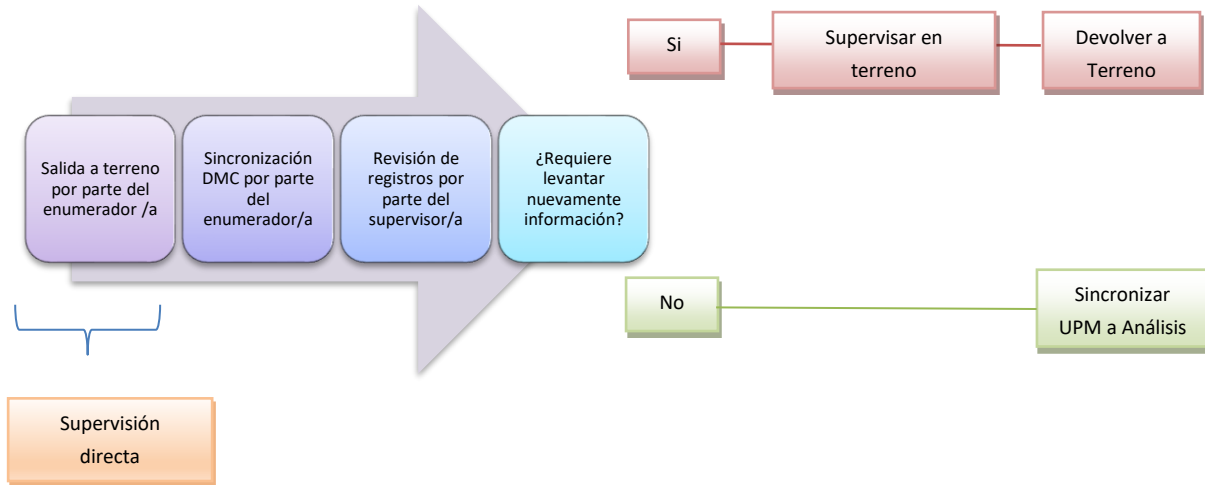
La etapa de supervisión se desarrolla de forma paralela a la recolección, dado que contempló tanto la coordinación y asignación de las cargas de trabajo al equipo de enumeradores, como las actividades de supervisión en terreno que permiten alcanzar los objetivos asociados al control de calidad de los datos colectados.

De esta forma, el supervisor en primer lugar debe mantener un control de las unidades que han sido levantadas en terreno, para ello dispuso del sistema informático Visualizador web, donde debe asignar las áreas de trabajo a cada uno de los enumeradores de su equipo, además de verificar y validar la información levantada en terreno, una vez que es sincronizada por el enumerador.

En segundo lugar, este operativo contempla una supervisión directa para cada uno de los enumeradores, es decir, al menos una vez a la semana el supervisor debe acompañar en terreno a cada uno de los enumeradores de su equipo, con la finalidad de evaluar la aplicación y los procedimientos en terreno que deben ser aplicados por cada enumerador y retroalimentar in situ, la labor de ellos. En muchos casos y debido a la contingencia del proceso de enumeración, el supervisor hacía el acompañamiento constante al equipo de enumeración. Esta supervisión se efectuó para el 20% de la muestra regional.

En el cumplimiento de este propósito, el supervisor contó con una pauta de evaluación, tanto para el área urbana como rural, que se completa por cada unidad supervisada, la que fue digitada en un formulario de Microsoft Forms con el fin de obtener un reporte del estado de avance de la supervisión y de las labores de cada recolector. El proceso de supervisión se indica en el Diagrama IX.1.

Diagrama IX.1. Proceso de supervisión



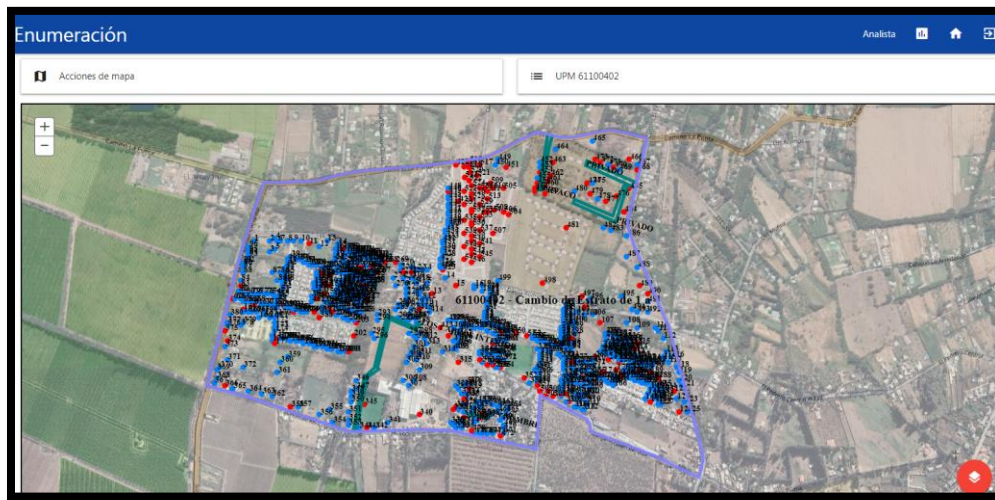
Elaborado por el Instituto Nacional de Estadísticas.

Proceso Análisis

La etapa de análisis correspondió a la revisión del 100% de las unidades seleccionadas para el operativo, una vez realizado y revisado el trabajo de campo, tanto por el enumerador como por el supervisor. Esta actividad se ejecutó centralizadamente por la sección de Insumos geográficos desde el 4 de abril hasta el 5 de agosto del 2022, realizada en forma paralela al levantamiento en terreno, y tiene como propósito controlar la calidad y completitud de la información.

Esta actividad se realiza a través del sistema informático asociado al proyecto denominado Visualizador Web, para todas las UPM urbanas y rurales a nivel nacional, realizando análisis de consistencia y concordancia, tanto a nivel de registro como de modificaciones que haya sufrido la cartografía (fusiones, divisiones de manzanas). La visualización del proceso de enumeración para el análisis se puede observar en la Imagen IX.1.

Imagen IX.1. Visualizador WEB con viviendas enumeradas



Elaborado por el Instituto Nacional de Estadísticas.

Otras consideraciones en el proceso de enumeración en terreno

Durante el proceso de enumeración generado entre los meses de marzo a julio del presente año, sucedieron una serie de situaciones que, de alguna manera afectaron la planificación original y el desarrollo satisfactorio del proceso de enumeración. Dentro de estos aspectos se consideran:

- **Contagios COVID:** Durante la enumeración en terreno desarrollada en las 16 regiones del país, el plan “Paso a Paso nos cuidamos”, las comunas integradas en la selección de la muestra Casen 2022 se encontraron en su mayoría en fases bajo y medio impacto sanitario. Este escenario, a pesar de ser menos restrictivo en términos de aforo, no estuvo exento de que el contacto entre los funcionarios y la ciudadanía aumentara las probabilidades de contagio, generando en algunos equipos la paralización total o parcial del levantamiento en terreno. Durante el desarrollo de la actividad se presentaron licencias médicas por COVID en todas las regiones del país, debiendo aislar los equipos para evitar los contagios masivos y evitar así, la paralización total del equipo Casen u otros funcionarios del INE. De esta forma, fue necesario prestar mayor atención a la calendarización de las actividades, considerando los tiempos requeridos para las solicitudes de los vehículos y kits sanitarios utilizados por el equipo de enumeración.
- **Rechazo por parte del informante:** Durante el periodo de levantamiento sucedieron una serie de hechos que afectaron directamente a los equipos en terreno específicamente al acceso a la información. A través de redes sociales (WhatsApp, TikTok, Instagram, Twitter entre otros), se generó una serie de información falsa (videos, textos y audios) indicando que personal del INE pasaría preguntando por información para otros fines distintos al proceso de enumeración, información que fue “viral” y masificada a nivel nacional. Esto generó un rechazo masivo por

parte del informante, en especial durante los meses de abril y mayo, donde la ciudadanía negaba directamente el acceso a la información al enumerador INE.

Esto impactó en varios ámbitos dentro del desarrollo normal del proceso de enumeración; el más relevante, el rechazo con violencia, donde algunos equipos tuvieron que poner denuncias en carabineros y PDI por ser consideradas situaciones que afectaban la integridad física del enumerador; y por otra parte un rechazo pasivo, donde el informante negaba, en palabras simples, dar información al enumerador o simplemente, estando en la vivienda, se negaba a abrir la puerta. Esto implicó implementar una serie de acciones, como generar infografías para RRSS y para que los equipos gestionaran previamente las visitas a las áreas de trabajo o intervenciones por parte de la Directora Nacional y Directoras(es) Regionales en programas de televisión y radio locales para informar a la ciudadanía del proceso de enumeración Casen. Esto trajo como consecuencia que en ciertos sectores donde se dieron condiciones de rechazo, se tuvo que fortalecer los equipos de terreno con más personal e integrar a funcionarios INE en procesos de gestión y acompañamiento, además de realizar muchas revisitas (5 o 6) hasta lograr conseguir la información, involucrando mayor tiempo del planificado. Paralelamente, como consecuencia adicional, existió mucha rotación de los equipos de enumeradores, ya que en ciertos casos el trabajo se tornó hostil y peligroso, por lo que desistían seguir en el proceso de enumeración.

- **Contingencia Nacional:** Se enfoca directamente a los estados de excepción, hechos violentos sucedidos en algunas comunas del país y sectores peligrosos donde, en ciertas UPM no se pudo acceder o en algunas de ellas no se podía generar el trabajo en jornadas de trabajo normal.

Anexo N°3.3. Descripción de validación y mejora de Directorios de Gabinete

A continuación, se describen los principales procesos de validación y mejoras de los directorios.

Directorios de Precenso 2016

Los procesos de actualización de los directorios de Precenso 2016, están asociados a poder validar direcciones de viviendas y otros usos en base al proceso de actualización permanente ACON. Para ello, se dispuso de una plataforma geográfica denominada GEODIRECTORIO, cuya función principal era poner a disposición de manera interactiva y a diversos usuarios del proceso, información geográfica que posibilite hacer un análisis territorial de las manzanas, ejes de calles e información complementaria de direcciones, que permita finalmente la mejora del directorio. Esta herramienta permite explorar, consultar, analizar y editar elementos de la base cartográfica, interactuando con las capas de información disponibles. Este proceso utilizó insumos base para su ejecución, como son la vinculación de la base de directorio de Precenso a la manzana censal, RRAA con direcciones georreferenciadas, CRF, bases cartográficas del ACON, áreas de campamentos y tomas entre otros.

Paralelamente cada analista dispuso de los directorios de viviendas y otros usos en la plataforma SGOH, correspondiente al sistema informático utilizado para revisión de gabinete de directorios generados en procesos de enumeración. Esta actividad permite revisar manzana a manzana y a través del orden de recorrido cada una de las direcciones en el GEODIRECTORIO, sus caras de manzanas y coincidencia con el SGOH. En caso de detectar diferencias, la información fuente de corrección corresponde a la que deriva del proceso ACON. Cuando una manzana presenta CRF o áreas de campamentos o diferencias de estructura con la manzana censal, esta se marcaba en el GEODIRECTORIO indicando su validación, la cual se considerada una validación con reparos, en cuyo caso la manzana ACON quedando marcada para “TERRENO”. Esto implica que el directorio ha sido revisado, pero sus direcciones están incompletas en cuando a sus componentes principales y complementarios, es decir, pueden faltar datos o especificaciones en su listado de direcciones, como también el uso y ocupación de esos registros. Esta etapa de revisión posteriormente se transforma en una variable de priorización para marcar las UPM e integrarlas al proceso de enumeración en terreno para actualizar dichos directorios. Cuando las direcciones del listado no disponen de información, estos son obtenidos desde fuentes de información disponibles en el GEODIRECTORIO (cuando existe la información) a través del eje de calle para tipos de vías y nombres de calle, como también numeraciones desde RRAA de la base georreferenciada de MINVU.

Esta actividad se realizó primariamente a nivel nacional con cada equipo de infraestructura regional, quienes conocen a cabalidad su territorio, ya que son quienes realizan de manera constante procesos de enumeración en terreno y poseen mayores referencias de cambios o situaciones particulares del territorio.

En el proceso de revisión de gabinete para Casen 2022, esta revisión se realiza en base a la selección de la muestra para el proyecto, donde los directorios se validan a partir de la información más actualizada del ACON (versión de actualización hasta diciembre del 2021) y con el MGE disponible a la fecha, de manera de poder incorporar actualizaciones posteriores a la revisión del directorio.

Directorios Otras Muestras

El procedimiento de integración de directorios provenientes de otras muestras deriva específicamente de UPM que han sido seleccionadas para muestras de encuestas sociales o económicas, enumeradas por el INE y que son coincidentes con la selección de la muestra Casen 2022. Estas muestras corresponden a enumeraciones en terreno ejecutadas desde el año 2020 a la fecha, las que han sido priorizadas previamente. Esta priorización corresponderá al análisis realizado sobre las áreas de selección de la muestra, basados en la necesidad de identificar a todas las unidades que presenten cambios o dinamismo de la información, que están asociados a las variaciones o cambios de estructura de la(s) manzana(s), cantidad de viviendas contenidas en ellas, uso u ocupación de cada edificación, como también disponer de las direcciones asociadas a las viviendas, que permiten llegar al informante y con

ello levantar la encuesta. Las UPM que son coincidentes con la muestra, pasan por un análisis de su levantamiento y si este no presenta cambios estructurales es reutilizado, en caso contrario pasa a formar parte de las UPM priorizadas para terreno. Estos directorios son cargados en la plataforma SGOH para su revisión y validación, al igual que se hace con los directorios provenientes de Precenso 2016. En ellos, se validan los componentes principales de la dirección y sus observaciones. Se da especial énfasis en direcciones sin nombre, sin número y sin observaciones, considerando que el objetivo de esta validación sea una dirección que pueda ser ubicada y reconocida en terreno al momento del levantamiento de la encuesta. A modo general, es importante indicar que formalmente no todas las viviendas cuentan con Nombre de Calle y N° domiciliario, principalmente en áreas rurales y sectores de informalidad en áreas urbanas. Por ello, no se puede contar con esa información para la totalidad del directorio de la muestra, siendo el campo observaciones y las coordenadas del registro las que complementan esa falta de información.

Dentro de las 625 UPM de otras muestras disponibles integradas al directorio de gabinete se encuentran las de la Muestra del Empleo mensual que se vienen actualizando desde el año 2020 hasta la fecha, la muestra piloto PIAAC en el año 2020, la muestra de Presupuestos Familiares (EPF) del año 2021 (primer semestre) y Estudio Piloto Cambio de Marco Casen 2022 (segundo semestre). El número de UPM por muestras utilizadas para el directorio Casen se indican en la Tabla IX.12.

Tabla IX.12. UPM de otras muestras por año

Año	Total UPM	Muestras
2020	205	PIAAC / ENE
2021	382	ENE / EPF / Estudio Piloto Cambio Marco Casen
2022	38	ENE

Elaborado por el Instituto Nacional de Estadísticas.

ANEXO N°4. STOCK DE VIVIENDAS SEGÚN FACTOR DE SELECCIÓN

Región	Tamaño muestral	Stock viviendas	Mínimo	Máximo
Total país	106.856	6.306.378	2	1.241
Arica y Parinacota	3.698	73.019	2	87
Tarapacá	4.156	104.778	3	181
Antofagasta	4.580	189.748	2	179
Atacama	5.060	105.715	7	129
Coquimbo	4.180	282.783	7	219
Valparaíso	10.984	665.422	4	257
Metropolitana	22.804	2.454.676	3	1.241
O'Higgins	6.672	337.515	10	170
Maule	7.090	395.955	11	308
Ñuble	4.696	191.780	12	137
Biobío	9.800	570.471	2	411
La Araucanía	7.006	364.059	4	306
Los Ríos	5.400	149.895	11	110
Los Lagos	5.466	315.173	12	275
Aysén	2.070	40.484	5	105
Magallanes	3.194	64.904	3	96

Elaborado por el Instituto Nacional de Estadísticas.

ANEXO N°5. COVARIABLES CONSIDERADAS EN EL MODELO PROPENSITY SCORE

La tabla presenta las variables utilizadas para la creación del modelo propensity score, su nivel de desagregación y la fuente desde donde fueron obtenidas

Variable	Glosa	Nivel	Fuente	Año
TR	Tasa de respuesta en la UPM (variable dependiente)	UPM	Casen	2022
Viv_bas_upm	Porcentaje de viviendas con basura en el sector (con basura en varios lugares o basura en 3 o 4 lugares).	UPM	Hoja de ruta Casen	2022
Viv_acc_upm	Porcentaje de viviendas con acceso controlado (casas en condominio y departamentos con acceso controlado)	UPM	Hoja de ruta Casen	2022
Viv_est_upm	Porcentaje de viviendas de estado bueno o muy bueno	UPM	Hoja de ruta Casen	2022
Viv_van_upm	Porcentaje de viviendas con presencia de vandalismo, grafitis o daño deliberado a la propiedad en el sector (con 3 o más áreas con grafiti o daño deliberado)	UPM	Hoja de ruta Casen	2022
Prom_visita	Promedio de visitas a las viviendas de la UPM por parte del encuestador	UPM	Hoja de ruta Casen	2022
G_edad1	Porcentaje de personas del grupo 0 – 14 años sobre el total de la población	UPM	Censo	2017
P_mujeres	Proporción de mujeres sobre el total de la población	UPM	Censo	2017
Pct_ext_n	Proporción de población extranjera sobre el total de la población.	UPM	Censo	2017
Prop_esc_baja	Proporción de la población de 15 o más años que poseen escolaridad baja (educación escolar incompleta)	UPM	Censo	2017
Prop_esc_media	Proporción de la población de 15 o más años que poseen escolaridad media (enseñanza media completa o equivalente)	UPM	Censo	2017
TOC_2021	Tasa de ocupación formal dependiente anual sobre el total de población de 15 años o más.	Comuna	MDSF-DataSocial (RRAA)	2021
RSH_0_40	Porcentaje de personas en la comuna inscritas en el Registro Social de Hogares en el tramo de calificación socioeconómica 0% – 40%	Comuna	MDSF-DataSocial (RRAA)	2022
Rend_medio_bajo	Porcentaje de establecimientos educacionales en categoría de desempeño medio-bajo en el Sistema Nacional Evaluación del Desempeño (SNED)	Comuna	Ministerio de Educación	2019
Prest_salud_priv	Porcentaje de prestadores de salud privados	Comuna	DEIS - Ministerio de Salud (RRAA)	2023
Def_enf_sist_circ	Porcentaje de defunciones por enfermedades al sistema circulatorio	Comuna	DEIS - Ministerio de Salud (RRAA)	2020

Elaborado por el Instituto Nacional de Estadísticas.

ANEXO N°6. STOCK DE VIVIENDAS SEGÚN FACTOR DE NO RESPUESTA

Región	Tamaño muestral (elegibles logrados)	Stock viviendas	Mínimo	Máximo
Total país	70.751	6.110.928	2	3.069
Arica y Parinacota	2.573	69.626	212	663
Tarapacá	2.735	101.479	4	256
Antofagasta	2.943	181.570	3	331
Atacama	3.081	101.730	9	225
Coquimbo	2.909	277.844	9	454
Valparaíso	7.288	625.897	6	494
Metropolitana	12.831	2.411.313	13	13.663
O'Higgins	4.957	329.384	12	275
Maule	5.104	378.161	14	575
Ñuble	3.290	181.943	93	1.706
Biobío	7.014	557.702	2	651
La Araucanía	4.880	344.151	6	413
Los Ríos	3.842	142.141	42	809
Los Lagos	3.906	303.546	15	466
Aysén	1.443	39.484	6	136
Magallanes	1.955	64.956	4	178

Elaborado por el Instituto Nacional de Estadísticas.

ANEXO N°7. STOCK POBLACIONAL CALIBRADO A PROYECCIONES DE POBLACIÓN EN VIVIENDAS PARTICULARES OCUPADAS REGIÓN - ÁREA

Región	Área	N° personas	Stock personas	Factor Regional persona	Factor Provincial	Factor Comunal
Total país		202.231	19.878.573	19.878.573	19.850.493	19.842.685
Arica y Parinacota	Urbana	7.333	237.133	237.133	231.984	232.683
	Rural	819	20.929	20.929	25.984	25.288
Tarapacá	Urbana	8.021	379.008	379.008	382.227	381.476
	Rural	670	19.691	19.691	16.474	17.226
Antofagasta	Urbana	8.211	687.535	687.535	686.032	686.475
	Rural	818	23.425	23.425	24.948	24.228
Atacama	Urbana	7.947	291.323	291.323	285.605	285.973
	Rural	1.064	27.036	27.036	32.756	32.389
Coquimbo	Urbana	6.276	709.119	709.119	684.114	682.746
	Rural	1.960	153.726	153.726	178.733	180.101
Valparaíso	Urbana	16.697	1.835.838	1.835.838	1.827.578	1.820.541
	Rural	3.855	165.447	165.447	165.056	171.013
Metropolitana	Urbana	35.551	8.025.026	8.025.026	8.004.061	8.003.261
	Rural	3.123	308.047	308.047	328.936	329.679
O'Higgins	Urbana	9.606	765.502	765.502	750.614	745.485
	Rural	4.258	247.312	247.312	262.201	267.329
Maule	Urbana	9.670	864.170	864.170	848.464	844.919
	Rural	4.069	292.660	292.660	308.367	311.912
Ñuble	Urbana	5.550	367.956	367.956	361.113	363.477
	Rural	3.156	150.046	150.046	156.889	154.525
Biobío	Urbana	15.833	1.493.749	1.493.749	1.470.539	1.472.958
	Rural	4.081	184.357	184.357	207.570	205.150
La Araucanía	Urbana	8.734	737.917	737.917	703.284	714.005
	Rural	4.843	287.583	287.583	322.217	311.496
Los Ríos	Urbana	7.368	298.582	298.582	290.995	291.427
	Rural	3.235	111.541	111.541	119.131	118.699
Los Lagos	Urbana	7.374	676.006	676.006	645.955	646.184
	Rural	3.333	228.312	228.312	239.064	234.864
Aysén	Urbana	2.873	88.536	88.536	85.855	83.536
	Rural	874	19.525	19.525	22.208	22.249
Magallanes	Urbana	4.512	167.392	167.392	169.571	168.205
	Rural	517	14.144	14.144	11.968	13.186

Elaborado por el Instituto Nacional de Estadísticas.

ANEXO N°8: STOCK POBLACIONAL CALIBRADO A LAS PROYECCIONES DE POBLACIÓN DE 18 AÑOS O MÁS POR REGIÓN-ÁREA

Región	Área	N° personas	Stock personas	Factor OSIG Regional
Total país		92.461	15.408.912	15.408.912
Arica y Parinacota	Urbana	3.102	179.247	179.247
	Rural	309	16.097	16.097
Tarapacá	Urbana	3.379	276.540	276.540
	Rural	286	16.608	16.608
Antofagasta	Urbana	3.481	516.270	516.270
	Rural	366	20.612	20.612
Atacama	Urbana	3.394	214.983	214.983
	Rural	558	21.079	21.079
Coquimbo	Urbana	2.767	536.198	536.198
	Rural	933	121.566	121.566
Valparaíso	Urbana	7.633	1.439.657	1.439.657
	Rural	1.692	129.038	129.038
Metropolitana	Urbana	16.014	6.279.955	6.279.955
	Rural	1.310	233.698	233.698
O'Higgins	Urbana	4.869	584.033	584.033
	Rural	2.211	196.941	196.941
Maule	Urbana	4.595	659.141	659.141
	Rural	1.921	233.084	233.084
Ñuble	Urbana	2.844	283.968	283.968
	Rural	1.601	120.387	120.387
Biobío	Urbana	7.286	1.157.839	1.157.839
	Rural	1.908	144.795	144.795
La Araucanía	Urbana	3.996	563.155	563.155
	Rural	2.124	222.418	222.418
Los Ríos	Urbana	3.305	231.569	231.569
	Rural	1.416	87.387	87.387
Los Lagos	Urbana	3.330	515.612	515.612
	Rural	1.588	183.022	183.022
Aysén	Urbana	1.315	65.099	65.099
	Rural	478	15.905	15.905
Magallanes	Urbana	2.138	131.106	131.106
	Rural	312	11.903	11.903

Elaborado por el Instituto Nacional de Estadísticas.

X. BIBLIOGRAFÍA

- Alvarado, M., Pizarro, M. & Guarda, N. (2020). Documentos de trabajo. La nueva metodología de calibración de la Encuesta Nacional de Empleo: método de calibración Raking. Obtenido de: https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/documentos-de-trabajo/documentos-de-trabajo/calibracion-factores-ene-metodologia.pdf?sfvrsn=cc1f6a74_8
- Cochran, W. (1998). *Técnicas de Muestreo*. México D.F: Compañía Editorial Continental, S.A. de C.V.
- Groves, R. (2004). *Survey Errors and Survey Cost*. Nueva York: John Wiley & Sons.
- Groves, R. M. (1998). *Nonresponse in Household Interview Surveys*. Nueva York: John Wiley.
- Guerrero, J., Guerrero, M., & Marín, A. (2020). <https://www.ine.cl/inicio/documentos-de-trabajo/documento/>. Obtenido de <https://www.ine.cl/inicio/documentos-de-trabajo/documento/estratificaci%C3%B3n-socioecon%C3%B3mica-del-marco-muestral-de-viviendas-2017>
- Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). *Applied Survey Data Analysis*. USA: Chapman & Hall / CRC Press.
- INE Chile. (2015). *Actualización de límites urbanos censales (LUC)*. Santiago: INE Chile.
- INE Chile. (2019b). *Nuevo Marco Muestral de Viviendas para encuestas de Hogares (MMV 2017)*. Santiago: INE Chile.
- INE Chile. (2020). *Documento Metodológico Encuesta Nacional de Empleo (ENE)*. Santiago: INE Chile.
- Instituto Nacional de Estadísticas. (2017). <https://www.ine.cl/docs/default-source/geodatos-abiertos>. Obtenido de https://www.ine.cl/docs/default-source/geodatos-abiertos/manuales/censo/alcances-base-cartografica-censo-2017.pdf?sfvrsn=d513cb51_6
- Instituto Nacional de Estadísticas. (2017). *Portal de mapas/geodatos abiertos*. Obtenido de <https://www.ine.cl/herramientas/portal-de-mapas/geodatos-abiertos>
- Instituto Nacional de Estadísticas. (2020). <https://www.ine.cl/docs/default-source/documentos-de-trabajo/>. Obtenido de https://www.ine.cl/docs/default-source/documentos-de-trabajo/20200318-lineamientos-medidas-de-precisi%C3%B3n.pdf?sfvrsn=f1ab2dbe_4.
- Instituto Nacional de Estadísticas. (2020). <https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion>. Obtenido de <https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion>

desocupacion/metodologia/espanol/informe-de-simulaciones-de-tamanos-muestrales-2020.pdf?sfvrsn=63d3ca6c_3

Kish, L. (1965). *Survey Sampling*. Nueva York: John Wiley & Sons.

MDSF. (2017). *Factores de expansión para el análisis de preguntas de orientación sexual e identidad de género*.

MDSF. (2020). *Metodología de Diseño Muestral Casen en Pandemia 2020*. Santiago: Ministerio de Desarrollo Social y Familia.

OCDE. (2007). *Glossary of Statistical Terms*. Organisation for Economic Cooperation and Development, Nueva York.

Organización de Naciones Unidas. (2009). *Encuestas de hogares en los países en desarrollo y transición*. Nueva York: Departamento de Asuntos Económicos, División Estadística.

Rao, P. (2000). *Sampling Methodologies with Applications*. Estados Unidos: Chapman & Hall/CRC.

Sarndal, C., Swensson, B. & Wretman, J. (1992). *Model Assisted Survey Sampling*. Nueva York: Springer

Wolter, K. (1985). *Introduction to Variance Estimation*. Nueva York: Springer

Yates, F., & Grundy, P. M. (1953). Selection without replacement from within strata with probability proportional to size. *Journal of the Royal Statistical Society: Series B (Methodological)*, 15(2), 253-261.