

Casen 2022

Principales cambios metodológicos

1. Antecedentes

El Ministerio de Desarrollo Social y Familia (en adelante, el Ministerio), junto al Panel de Expertos y Expertas de Casen (en adelante, el Panel), definieron introducir un conjunto de mejoras metodológicas a la Encuesta de Caracterización Socioeconómica Nacional 2022 (Casen 2022), con el objetivo de asegurar la calidad de sus estimaciones, aprovechando la disponibilidad de nuevos insumos basados en el Censo 2017, las nuevas tecnologías y conocimiento asociado a encuestas.

La Encuesta Casen 2022 considera las siguientes mejoras metodológicas con respecto a versiones anteriores del estudio: actualización del Marco Muestral de Viviendas a partir del Censo de Población y Vivienda 2017, ajustes en el diseño muestral de la encuesta con base en la actualización del marco muestral, cambios en la modalidad de recolección y el *software* utilizado para la aplicación de la encuesta, y una nueva metodología de calibración de los factores de expansión conocida como *raking*. Cada uno de los cambios metodológicos introducidos fueron debidamente estudiados por la División Observatorio Social, a cargo de la Encuesta, en conjunto con el Instituto Nacional de Estadísticas (INE) y con la asesoría permanente de la Comisión Económica para Latinoamérica y el Caribe (Cepal).

El objetivo de este documento es sintetizar los principales cambios metodológicos implementados en Casen 2022 a fin de que sean tomados en cuenta al momento de utilizar los resultados de la Encuesta.

El documento se estructura de la siguiente forma. En el capítulo 2, se describe la actualización del marco muestral de viviendas y el estudio realizado para determinar su efecto. En el capítulo 3, se presentan los cambios asociados al diseño muestral de la encuesta. En el capítulo 4, se explican los ajustes vinculados a la modalidad de recolección y al *software* utilizado para la aplicación de la encuesta. En el capítulo 5, se resume la nueva metodología de calibración de los factores de expansión. Finalmente, el capítulo 6, presenta una síntesis de los principales cambios metodológicos de la encuesta.

2. Actualización del Marco Muestral de Viviendas

2.1. Antecedentes

Hasta el año 2020, la Encuesta Casen utilizó el Marco Muestral 2016 (MM 2016), compuesto por un marco de manzanas y otro de secciones, ambos elaborados a partir de información del Censo 2002 y actualizados al Precenso del año 2016. Luego, en el año 2020 el INE puso a disposición el nuevo Marco Muestral de Viviendas construido a partir del Censo 2017 (MMV 2020).

Debido a que el MM 2016 se encontraba desactualizado, tanto en su cobertura como en la conformación de las Unidades Primarias de Muestreo (UPM)¹, el Ministerio estimó necesaria la migración al nuevo marco muestral de viviendas (MMV 2020), considerando que un marco muestral completo y preciso es requisito fundamental para asegurar la calidad de los procesos de diseño, selección muestral, recolección y análisis.

En la Cuadro 1, se presentan los principales atributos del nuevo marco muestral MMV 2020 de forma comparativa con el MM 2016:

¹ Las Unidades Primarias de Muestreo (UPM) son áreas geográficas conformadas por viviendas particulares a partir de las cuales se seleccionan las unidades secundarias y finales de muestreo, que son las viviendas.

Cuadro 1. Comparación principales atributos MMV 2020 versus MM2016

	MM2016	MMV 2020
Organización general	Dos marcos excluyentes: <ol style="list-style-type: none"> 1. Marco de manzanas para el área urbana 2. Marco de secciones para el área rural Estratificación y conformación de UPM es distinta para cada marco.	1 solo marco para el territorio nacional
Conformación UPM	<u>Zona urbana</u> : manzanas (predios urbanos delimitados por calles). Gran variabilidad en el número de viviendas. <u>Zona rural</u> : secciones. Agrupaciones homogéneas de viviendas.	Unidades homogéneas en cuanto al número de viviendas: Promedio de 200 viviendas en el área urbana (rango [160, 240]) y 90 viviendas en el área rural (rango [70, 110]).
Estratificación del Marco	Estratificación geográfica: comuna-área (3 áreas). En el área urbana se considera una tercera variable de estratificación que corresponde al grupo de tamaño de la manzana (5 grupos de tamaño).	Se mantiene estratificación geográfica (2 áreas), pero se actualiza el límite urbano censal (LUC), que es la línea imaginaria que separa el área urbana de la rural. Se incorpora estratificación socioeconómica a nivel de UPM.
Cobertura del Marco	Exclusión de Áreas de Difícil Acceso (ADAs), que no están en el MM (22 comunas). Exclusiones al MM para marco de selección Casen: <ul style="list-style-type: none"> - Manzanas de 7 viviendas o menos. - Manzanas seleccionadas en la Encuesta Nacional de Empleo (ENE) del INE. 	Cobertura de todo el territorio nacional. Se definen Áreas Especiales (AE), que determinan la dificultad del trabajo de recolección: 1) Tiempo trayecto; 2) Clima; 3) Altitud; 4) Transporte especial; 5) Insularidad; 6) Acceso pedestre; 7) Estado de los caminos; 8) Ingreso restringido. Exclusiones al MMV para marco de selección Casen: <ul style="list-style-type: none"> - Áreas Especiales: 10 comunas y partes de otras 11 comunas.

Fuente: elaboración propia a partir del Informe de Diseño Muestral elaborado por el INE².

2.2. Piloto Cambio de Marco Casen 2022

Debido a las posibles implicancias de la actualización del marco muestral en los principales indicadores de la Encuesta Casen, 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.

El objetivo general del estudio fue estimar el efecto de utilizar el MMV 2020 sobre los principales indicadores de la encuesta Casen.

Los objetivos específicos fueron los siguientes:

- Estimar el efecto del cambio de marco muestral en las siguientes mediciones: tasa de pobreza por ingresos, tasa de pobreza multidimensional e ingreso promedio de los hogares a nivel nacional.
- Identificar si las características de la población difieren en las estimaciones obtenidas a partir de los diferentes marcos y las estrategias de diseño muestral asociadas.

El Cuadro 2 resume las principales características del estudio:

² El Informe de Diseño Muestral del INE de Casen 2022 será publicado durante agosto de 2023 en la página web del observatorio: <https://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2022>

Cuadro 2. Ficha técnica del estudio

Instituciones participantes	<ul style="list-style-type: none"> - INE: diseño muestral, recolección, procesamiento y análisis de los datos - Ministerio: coordinación del proyecto, diseño y construcción del cuestionario, y del análisis de los resultados
Características del diseño muestral	<ul style="list-style-type: none"> - Representatividad: nacional - Muestra objetivo: 11.995 viviendas distribuidas en dos submuestras de igual magnitud - Efecto mínimo detectable establecido por diseño: 2,08% - Comunas incluidas: 149
Características del cuestionario	<ul style="list-style-type: none"> - Cuestionario reducido, que permite medir: pobreza por ingresos, pobreza multidimensional, set corto <i>Washington Group</i>, identidad de género y orientación sexual. - Duración promedio: 45 minutos.
Características de la recolección	<ul style="list-style-type: none"> - Inicio: 10 de enero – Término: 04 de marzo de 2022 - Modalidad de aplicación: entrevista cara a cara asistida mediante dispositivo móvil (CAPI).

Fuente: Elaboración propia a partir de Piloto Cambio de Marco.

2.2.1. Metodología

2.2.1.1. Diseño muestral

Para alcanzar una potencia estadística del 80%, se definió una muestra total de alrededor de 12.000 viviendas, la cual fue dividida en dos submuestras independientes distribuidas a lo largo del territorio nacional. La primera muestra consideró 5.997 viviendas seleccionadas a partir del MM 2016; mientras que la segunda consideró 5.998 viviendas seleccionadas desde el MMV 2020.

El diseño muestral correspondiente al MM 2016 es probabilístico, estratificado y bietápico, donde los estratos muestrales están conformados por la dupla comuna-área. Adicionalmente, en el área urbana, las UPM se clasifican según su tamaño, conformándose 5 grupos de tamaño. Así, los estratos del área urbana quedan compuestos por la terna comuna-área-grupo de tamaño, en los cuales la muestra es seleccionada de forma independiente.

Respecto a la estrategia muestral utilizada para la muestra obtenida desde el MMV 2020 el diseño también es probabilístico, estratificado y bietápico. Sin embargo, dado que el MMV 2020 está conformado por UPM homogéneas en tamaño, la estratificación se realiza directamente a nivel de comuna-área, sin diferenciar por grupo de tamaño.

El diseño del estudio permite detectar diferencias para la tasa de pobreza por ingresos (efecto mínimo detectable) del 2% con un nivel de significancia del 10%.

2.2.1.2. Definición de criterios de inclusión y exclusión de comunas

El diseño muestral de la encuesta Casen, requiere un mínimo de UPM por estrato y, a su vez, que el número de UPM sea proporcional al tamaño de las comunas; lo que, para efectos del estudio, no era factible si todas las comunas del marco muestral eran incluidas, dadas las restricciones presupuestarias del proyecto. Por ello, y buscando maximizar la posibilidad de encontrar un efecto en la prueba, el INE en conjunto con el Ministerio, definieron un flujo de criterios de inclusión y exclusión de comunas que considera lo siguiente:

- Se excluyen las ADA.
- Se incluyen: capitales regionales y provinciales con 10.000 o más hab., comunas donde hubo

dinamismo del cambio el marco³, comunas pertenecientes al 70% con más habs. de la región y comunas con más del 10% de pobres de la región.

Esto significó la inclusión de 149 comunas de las 345 que tiene el MMV 2020.

2.3. Principales resultados

El análisis de resultados se efectuó, principalmente, por medio de *test* de hipótesis, donde la hipótesis nula es que los principales indicadores de la encuesta estimados a partir del MM 2016 con el diseño muestral original son iguales a aquellos estimada a partir del MMV 2020 con el diseño muestra actualizado.

Adicionalmente, se compararon las tasas de respuesta para distintos dominios de interés; se revisó la presencia de subpoblaciones entre ambas muestras; y, finalmente, se compararon los principales indicadores de la encuesta.

Para determinar la magnitud del efecto del cambio de marco en las tasas de pobreza por ingresos y multidimensional, se estimaron ambas tasas y sus distintas especificaciones, y se realizaron test de medias para determinar si las diferencias obtenidas eran estadísticamente significativas. El Cuadro 3 presenta las diferencias en las estimaciones, junto al estadístico obtenido al realizar un *test* de diferencia de medias.

Cuadro 3: Estadísticos a nivel de hogar e individuos de indicadores de pobreza por ingreso y pobreza multidimensional.

Tipo de pobreza	Indicador	Estadístico t a nivel de hogar	Diferencia absoluta estimación a nivel de hogar (porcentaje)	Estadístico t a nivel de individuo	Diferencia absoluta estimación a nivel de individuo (porcentaje)	Diferencia
Pobreza por ingresos	Pobreza	1,26	0,49	1,65	0,78	No significativo
Pobreza multidimensional	5 dimensiones	0,61	0,74	0,65	0,99	No significativo
	4 dimensiones	0,16	0,16	0,01	0,01	No significativo

Fuente: Elaboración propia a partir de Piloto Cambio de Marco.

Del Cuadro 3, es posible ver que ante las distintas tasas de pobreza y sus descomposiciones no presentan diferencias significativas.

De la misma manera, se realizaron pruebas de diferencia de medias para las principales corrientes de ingreso, encontrándose que tanto para el ingreso total del hogar como para el ingreso autónomo del hogar no había diferencias significativas a un 95% de confianza (se obtuvo un estadístico de 1,3 y 1,7, respectivamente).

Respecto a las carencias que conforman la medida de 4 y 5 dimensiones de pobreza multidimensional,

³ Las comunas donde hubo dinamismo en el cambio de marco son aquellas donde los siguientes indicadores son relevantes: i) Proporción de viviendas fuera del Límite Urbano Censal (LUC) del MM 2016 y dentro del nuevo LUC del MMV 2020, ii) Proporción de viviendas pertenecientes a áreas RAU (Resto de Área Urbana) excluidas: en el MM 2016 se excluyó el área RAU de las comunas con viviendas en el marco de manzanas y viviendas urbanas en el marco de secciones para evitar duplicidades en la selección y iii) Proporción de viviendas pertenecientes a manzanas con 7 viviendas o menos.

Para categorizar a las comunas como dinámicas o no, se definió un indicador agregado de dinamismo como una combinación lineal de las tres variables anteriores. El resultado fue un grupo dinámico de 50 comunas y otro poco o no dinámico de 274 comunas.

se realizaron pruebas de diferencias de medias y no se encontró ninguna diferencia significativa.

Posteriormente, con el objetivo de determinar si existen o no variables sociodemográficas que inciden en la probabilidad de que la vivienda pertenezca al MMV 2020 en lugar del MM2016, se modeló la probabilidad de que la vivienda pertenezca al MMV 2020, dado un vector de covariables definido, el cual se estimó a través de una función de probabilidad logística que toma valor 1 cuando la variable latente es mayor a 0. El modelo establece una relación no lineal entre la probabilidad de que la variable dependiente tome valor 1 (dado el vector de covariables) y los coeficientes, y permite garantizar que el valor estimado de la variable dependiente se encontrará entre 0 y 1.

En el Cuadro 4, se presentan los efectos marginales de los modelos estimados. Si bien en los cuatro modelos se utilizan las mismas variables, los modelos varían según si son estimados a nivel de hogar o individuo y si se utiliza el factor de expansión respectivo. En los modelos de hogar, las variables que reflejan características de las personas se refieren a la jefatura del hogar.

Cuadro 4: Modelos Logit

Variable	Estimación a nivel de hogar				Estimación a nivel de individuo			
	Sin factor		Con factor		Sin factor		Con factor	
Sexo (1: Mujer; 0: Hombre)	0.0136	(0.171)	-0.00381	(0.819)	0.00794	(0.228)	-0.000125	(0.988)
Tramo edad (1: 30 a 44 años)	0.00495	(0.818)	-0.00871	(0.829)	0.00229	(0.809)	0.00237	(0.903)
Tramo edad (1: 45 a 59 años)	-0.0138	(0.515)	-0.00164	(0.968)	-0.0116	(0.228)	0.00207	(0.908)
Tramo edad (1: 60 o más años)	-0.0355	(0.112)	0.0233	(0.584)	-0.0333***	(0.001)	0.0185	(0.363)
Área (1: Urbano; 0: Rural)	-0.0478***	(0.000)	0.00672	(0.823)	-0.0488***	(0.000)	0.00737	(0.808)
Años de escolaridad	0.000982	(0.423)	0.00381	(0.101)	0.0000401	(0.963)	0.00189	(0.372)
Inmigrante	-0.00495	(0.816)	-0.0123	(0.750)	0.0327**	(0.019)	-0.0178	(0.657)
Pueblo originario	0.0264*	(0.068)	0.00382	(0.885)	0.0174*	(0.067)	-0.0293	(0.232)
Tamaño del hogar	-0.00212	(0.518)	0.00867*	(0.097)	-0.00219	(0.274)	0.00848	(0.143)
Desocupado	0.0193	(0.454)	-0.00922	(0.821)	0.0146	(0.322)	-0.0182	(0.476)
Inactivo	-0.0206	(0.106)	-0.00855	(0.663)	-0.0125	(0.105)	-0.0131	(0.336)
N	10994		10994		24710		24710	

Efectos marginales; *p*-values en paréntesis

* *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

Fuente: Elaboración propia a partir de Piloto Cambio de Marco.

A nivel de hogar y sin factor de expansión, los resultados muestran que únicamente pertenecer a una zona urbana y la pertenencia a un pueblo originario por parte de la jefatura de hogar tienen un efecto en la probabilidad de pertenecer al MMV 2020. El resto de las variables no son significativas. A nivel de hogar y con factor de expansión, la única variable significativa es el tamaño del hogar. Luego, a nivel de persona y sin factor de expansión, los resultados muestran que tener 60 años, vivir en una zona urbana, ser inmigrante o pertenecer a un pueblo originario tiene efecto en la probabilidad de pertenecer al MMV 2020. Finalmente, a nivel de persona y con factor de expansión, ninguna variable es significativa.

Para interpretar los resultados es relevante considerar que los factores de expansión se encuentran calibrados por sexo, edad y zona; por lo que dichos resultados no deberían considerarse para estas variables. Por lo tanto, y según los modelos a nivel de personas, tramo de edad y zona sí tienen un efecto en la probabilidad. Además, en este estudio es esperable que esto suceda, dado que la actualización del marco muestral por definición podría implicar diferencias en ciertas características de las personas y hogares que lo componen.

Finalmente, se realizó una caracterización de los grupos en pobreza según indicadores a nivel de hogar entre marcos y se evaluó la existencia de diferencias estadísticamente significativas. El Cuadro 5 presenta la estadística descriptiva de lo anterior.

Cuadro 5: Caracterización de grupos en pobreza

Indicador	Pobreza por ingresos		Pobreza multidimensional (4D)		Pobreza multidimensional (5D)	
	Marco 2016	Marco 2017	Marco 2016	Marco 2017	Marco 2016	Marco 2017
Escolaridad promedio jefatura hogar	11,0	10,9	8,7	9,1	8,9	9,2
Edad promedio jefatura hogar	42,2	46,4	52,0	52,0	52,7	52,7
Proporción mujeres jefatura de hogar	56,4	49,4	47,9	53,0	46,6	53,1
Proporción Urbano	90,6	87,1	79,2	78,0	76,9	76,7

Fuente: Elaboración propia a partir de Piloto Cambio de Marco.

El cuadro anterior, muestra que no existen diferencias significativas entre las estimaciones realizadas a partir del MM 2016 y aquellas obtenidas desde el MMV 2020.

En conclusión, los resultados del estudio muestran que no existen diferencias significativas entre las estimaciones realizadas a partir del MM 2016 y aquellas obtenidas desde el MMV 2020.

3. Diseño muestral Casen 2022

Como consecuencia de la actualización del marco muestral descrita en el apartado anterior, el diseño muestral de Casen experimentó algunos cambios relevantes, que se resumen a continuación.

3.1. Cobertura

A diferencia de las versiones anteriores de la encuesta, en el marco de selección de Casen 2022 se incluyen las UPM seleccionadas para la Encuesta Nacional de Empleo (ENE) del INE y no hay exclusión de las manzanas con 7 viviendas o menos. Además, de las 22 comunas que anteriormente se excluían por constituir Áreas de Difícil Acceso (ADA), en el marco de selección Casen 2022 solo se excluyen 10 comunas identificadas como Áreas Especiales (AE) y solo algunas UPM de las 11 comunas restantes.

3.2. Diseño muestral y estratificación

El diseño muestral sigue siendo probabilístico, estratificado y bietápico. No obstante, ya no se clasifica a las manzanas según su tamaño, debido a que, como se señala en el Cuadro 1, en el nuevo marco muestral estas son homogéneas en cuanto al número de viviendas. Esto implica que en el área urbana la muestra ya no se estratifica según el tamaño de las manzanas.

Por otro lado, dada la disponibilidad en el nuevo marco muestral de una variable de clasificación por nivel socioeconómico de las UPM, en la muestra de Casen 2022 se agregó el nivel socioeconómico⁴ de la UPM como variable de estratificación. Esto significa que los estratos ahora están compuestos por la terna comuna, área y nivel socioeconómico. Los estratos de diseño son 767.

3.3. Número de viviendas por UPM

Dada la homogeneidad de las UPM en cuanto al número de viviendas, se definió seleccionar un número fijo según área: 8 viviendas en el área urbana y 10 viviendas en el área rural⁵. Anteriormente, el número de viviendas seleccionadas variaba según el tamaño de la unidad primaria de muestreo.

⁴ El Nivel Socioeconómico de la UPM fue determinado por el INE utilizando las siguientes variables: escolaridad, ocupados/as, hijos/as nacidos/as y materialidad de la vivienda.

⁵ En algunas regiones se seleccionaron 10 viviendas en el área urbana y 12 en el área rural para dar cuenta de las tasas de no respuesta de versiones anteriores de la Encuesta.

3.4. Conformación estratos y conglomerados de varianza

Desde Casen 2011 en adelante se han creado estratos y conglomerados para efectos de la estimación de la varianza, que reagrupan los estratos y conglomerados utilizados en el diseño muestral, según el número de UPM y viviendas efectivamente logradas. Una UPM se define como lograda cuando se logró aplicar la encuesta en al menos una de las viviendas seleccionadas dentro de dicha UPM.

Las reglas de agrupación utilizadas hasta Casen en Pandemia 2020 buscaban que cada conglomerado tuviera al menos 30 viviendas en su interior, para lo cual primero se colapsaban UPM, y si ello no era suficiente, estratos. Esto generaba una gran disparidad entre los estratos y conglomerados de diseño, y aquellos utilizados para la estimación de la varianza, lo que incide en la estimación de la varianza.

Para Casen 2022, se definió agrupar las UPM con otras de características similares únicamente en dos casos: cuando solo se logró aplicar la encuesta en solo una UPM dentro del estrato de muestreo, o bien, cuando solo se logró aplicar la encuesta en una vivienda dentro de una UPM. El primer caso implica el colapso del estrato con una UPM, con otro estrato de similares características y el segundo, significa el colapso de la UPM con una vivienda, con otra UPM perteneciente al mismo estrato.

4. Modalidad y *software* de recolección

4.1. Antecedentes

La Encuesta Casen 2022 fue aplicada presencialmente por personas encuestadoras por medio del uso de *Tablets*. El cuestionario fue programado en el *software* desarrollado por el grupo de datos del Banco Mundial para la aplicación de encuestas llamado *Survey Solutions* (SUSO)⁶, y, a diferencia de Casen en Pandemia 2020, corresponde al cuestionario en su versión completa.

SUSO es un sistema de softwares gratuito desarrollado por el Banco Mundial en colaboración con la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, por sus siglas en inglés). El desarrollo de esta herramienta comenzó el año 2012 y ha tenido distintas versiones que se han ido actualizado periódicamente por el equipo de desarrolladores del Banco⁷. La plataforma ofrece distintas funcionalidades para la puesta en marcha de encuestas complejas y de gran escala, habiendo sido utilizada incluso en algunos Censos de Población y Vivienda.

La versión 2022 de la encuesta Casen es la primera aplicada íntegramente en dispositivo móvil, así como también en utilizar como *software* de recolección *Survey Solutions*. Sin embargo, ello es el resultado de un conjunto de estudios y pruebas a cargo del Ministerio, llevados a cabo a partir del año 2016, que permitieron conocer con antelación a la aplicación de Casen 2022 los efectos del cambio de modalidad desde el papel a dispositivo móvil, así como el desempeño del *software* del Banco Mundial en encuestas a hogares. A continuación, se resume la evidencia recopilada.

4.2. Estudios asociados al cambio de la modalidad de recolección⁸

El Ministerio ha desarrollado dos estudios para evaluar el efecto del cambio de modalidad de aplicación de la Encuesta Casen desde PAPI (*Paper and Pencil Interview*- PAPI) a CAPI (*Computer Assisted Personal*

⁶ Para más información acerca de *Survey Solutions* visitar: [Introducción a Survey Solutions](#)

⁷ El *software* cuenta con actualizaciones frecuentes que añaden nuevas funcionalidades, mejoras y corrección de errores. La última versión disponible a la fecha de elaboración de esta nota es la 23.06 y fue liberada en junio de 2023.

⁸ Para mayores antecedentes sobre el uso de dispositivos móviles en la Encuesta Casen ver: [Nota técnica N°4: CAPI y Aseguramiento de Calidad \(QA\)](#).

Interviewing- CAPI). El primero fue realizado por el Centro de Microdatos de la Universidad de Chile (CMD) entre diciembre de 2016 y agosto de 2017, en las áreas urbana y rural de la Región Metropolitana. Y consistió en la aplicación de alrededor de 1.000 entrevistas en papel y 1.000 entrevistas en *Tablet* para detectar diferencias en la tasa de pobreza por ingresos por medio del Efecto del Tratamiento Promedio (ATE) entre papel y *Tablet*.

El segundo estudio fue efectuado por el Centro de Encuestas y Estudios Longitudinales (CEEL) de la Pontificia Universidad Católica de Chile entre agosto de y octubre de 2017, como parte de las actividades asociadas a la Encuesta Casen 2017. Esta prueba fue aplicada en las áreas urbana y rural de las regiones de Valparaíso, Metropolitana y Biobío. Al igual que el primer estudio, consideró una muestra objetivo de alrededor de 1.000 viviendas en papel y 1.000 en *Tablet*, y el Efecto del Tratamiento Promedio (ATE) entre papel y *Tablet*, como método de análisis.

El estudio del CMD, mostró que las tasas de pobreza por ingresos y multidimensional no presentan diferencias estadísticamente significativas por modalidad de aplicación. Mientras que la prueba desarrollada por el CEEL tuvo como resultado que únicamente la tasa de pobreza multidimensional a nivel de personas es significativamente mayor utilizando *Tablet*. En cuanto a los ingresos promedios, los resultados del estudio realizado por el CMD muestran que no existen diferencias significativas según modalidad. En tanto, el estudio del CEEL encontró diferencias significativas solo en el ingreso autónomo del hogar y los subsidios del hogar.

No obstante, los resultados obtenidos, ambas instituciones recomendaron aplicar la Encuesta Casen a través de dispositivo móvil, debido, principalmente, a que permite mejorar los procesos, la calidad de los datos y la oportunidad de la entrega de información.

Adicionalmente, el Ministerio revisó la experiencia internacional asociada a la modalidad de recolección de encuestas, que sugiere fuertemente el uso de dispositivos electrónicos en reemplazo del papel, debido a las ventajas que ello conlleva en cuanto a la calidad de los datos y la eficiencia de los procesos. Esta práctica ha sido adoptada por las principales encuestas de los países desarrollados desde fines de los años ochenta y más fuertemente a partir de la década de los noventa.

Asimismo, el CEEL en el contexto de los preparativos para Casen en Pandemia 2020 desarrolló distintas pruebas de usabilidad y aseguramiento de la calidad o QA⁹ para aplicaciones presenciales por medio de *Tablet*¹⁰, con el objetivo de garantizar el correcto funcionamiento del aplicativo para una encuesta de las características de Casen.

4.3. Elección de *Survey Solutions* como software de recolección

La elección de SUSO como *software* de recolección de la Encuesta Casen, se realizó a partir de su desempeño positivo en encuestas CAPI tanto del Ministerio como del INE, principalmente, en la Encuesta Social COVID-19, Piloto Cambio de Marco Casen 2022 y la IX Encuesta de Presupuestos Familiares (EPF). En particular, se consideró la evaluación realizada por el Ministerio y el INE acerca del desempeño del *software* en el Piloto de Cambio de Marco con miras a su uso en Casen 2022, que incluyó una comparación con respecto a la alternativa de utilizar un *software* de desarrollo propio, como se hizo en Casen en Pandemia 2020.

⁹ *Quality Assurance*, por sus siglas en inglés.

¹⁰ El paso desde modalidad presencial a modalidad mixta secuencial fue definido como respuesta a la crisis sanitaria a raíz del COVID-19. Para más detalles, revisar: [Nota técnica N°1 Casen en Pandemia 2020: Posibles efectos en las estimaciones de ingresos debido a la modalidad de aplicación en respuesta a crisis sanitaria COVID-19](#)

En dicha evaluación se consideraron distintas dimensiones relevantes para la elección de un *software* CAPI, establecidas a partir de un estudio desarrollado por la Universidad de Maryland por encargo del Banco Mundial¹¹. Estas dimensiones fueron las siguientes: herramientas para el diseño de instrumentos, desempeño durante la aplicación de la encuesta, herramientas para el manejo de casos y su asignación, necesidades de recursos de infraestructura y personal capacitado, herramientas para extender las funcionalidades del *software*.

La evaluación tuvo como resultado la recomendación de utilizar *Survey Solutions* frente a un desarrollo propio del contratista Casen, principalmente, por los motivos descritos a continuación:

- Permite estandarizar el diseño del cuestionario a lo largo de los distintos levantamientos, independiente del proveedor de la encuesta. Adicionalmente, el producto puede verse beneficiado por las actualizaciones y mejoras que se vayan realizando al software, el que se encuentra en constante desarrollo.
- Facilita la acumulación de conocimiento para el diseño de cuestionarios dentro de los equipos encargados de su diseño. Esto redundará en un mayor control sobre cómo se implementa el cuestionario.
- Dispone de la documentación necesaria para el correcto desarrollo de cuestionarios en su plataforma. En la página web de *Survey Solutions* se encuentra disponible la documentación de cada componente del software, un foro para consultas y un banco de preguntas que puede ser usado como referencia para la construcción de otros cuestionarios.
- Permite la navegación no lineal para las personas encuestadoras, emulando la flexibilidad de la aplicación en papel.
- La infraestructura necesaria para su implementación es compatible con los requerimientos ya definidos por el Ministerio. Además, el *software* disminuye las necesidades de equipos especializados en la programación de aplicaciones para la implementación del cuestionario, permitiendo que los equipos a cargo del diseño del cuestionario puedan también participar de su programación.

5. Metodología de calibración del factor de expansión

5.1. Antecedentes

El factor de expansión de la encuesta Casen considera cuatro ponderadores: 1) Ponderador de diseño o selección y ajuste por falta de cobertura, 2) Ponderador ajustado por elegibilidad, 3) Ponderador ajustado por no respuesta y 4) Ponderador de calibración.

Hasta Casen 2017, se utilizaron las proyecciones de población asociadas al Censo 2002 para cada estrato comuna-área. Sin embargo, a partir de la revisión de las nuevas proyecciones de población basadas en el Censo 2017, el Ministerio, junto al Panel, definieron la necesidad de analizar posibles cambios en la metodología de los factores de expansión. Para ello, a través de un convenio de colaboración entre el Ministerio y el INE, se determinó realizar un diagnóstico sobre la metodología actual de los factores de expansión y, en consecuencia, una propuesta de actualización de la metodología.

El resultado fue la definición de dos pasos a seguir: primero, el ajuste de los factores de expansión, considerando las nuevas proyecciones de población basadas en el Censo 2017; y segundo, cambios en

¹¹ Comparative Assessment of Software Programs for the Development of Computer-Assisted Personal Interview (CAPI) Applications. University of Maryland College Park, IRIS Center, 2011. Disponible en: [Evaluación comparativa de softwares para la aplicación de encuestas CAPI](#).

la metodología de calibración del factor de expansión que permitieran dar cuenta de las principales características demográficas de la población. El ajuste a los totales poblacionales, según las nuevas proyecciones de población, fue implementado en Casen en Pandemia 2020. Mientras que el ajuste vinculado a la metodología de calibración se incluye por primera vez en Casen 2022 y se describe a continuación.

5.2. Nueva metodología de calibración: *raking*

La calibración es el último ajuste que se aplica al factor de expansión de la Encuesta Casen. En el caso de Casen 2022, se utiliza el método *raking*, que permite incorporar información auxiliar relacionada a la variable de interés de la Encuesta y tiene tres objetivos: 1) que la encuesta pueda reproducir los totales poblacionales conocidos, 2) mejorar la precisión de las variables de interés y 3) mejorar la precisión de las estimaciones.

La calibración emplea información proveniente de las proyecciones de población del Censo 2017 al 30 de noviembre de 2022¹², según región, área, sexo y grupos de edad. Esta información se utiliza como un vector de totales poblacionales, llamados marginales de calibración, que para Casen 2022 son los siguientes:

- 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+].

Los nuevos marginales de calibración permiten subsanar tanto la subrepresentación como la sobrerrepresentación de determinadas subpoblaciones encuestadas relacionadas con la tasa de pobreza. Por ejemplo, en versiones anteriores de la encuesta se observó que los hombres entre 25 y 49 años estaban subrepresentados con respecto a las proyecciones de población del INE, lo que, dado que en general corresponden a personas con ingresos, podría implicar una subestimación de la tasa de pobreza. Por el contrario, en versiones anteriores de la encuesta, se observó que las mujeres del área rural mayores de 50 años estaban sobre representadas, lo que, considerando que corresponden a población sin o con bajos ingresos, podría significar una sobrestimación de la tasa de pobreza.

La nueva metodología de calibración permite mejorar el ajuste entre la muestra y los totales poblacionales, permitiendo reducir los sesgos por falta de respuesta y subcobertura.

¹² Fecha definida a partir del inicio y término del periodo de recolección de la Encuesta Casen 2022.

6. Síntesis

Las mejoras metodológicas introducidas la Encuesta Casen 2022 derivan de la actualización de los insumos basados en el Censo 2017, así como de nuevos desarrollos asociados a las metodologías de recolección de datos y a la calibración de los factores de expansión.

Los estudios asociados a los distintos ajustes implementados concluyeron que, independiente de si tienen o no efectos en las estimaciones asociadas a la Encuesta, son fundamentales para mejorar los procesos asociados Casen, así como para asegurar la calidad de los datos recopilados por la Encuesta y la oportunidad de la entrega de información.

7. Referencias

Banco Mundial (2018). Advancing CAPI/CAWI technology with Survey Solutions. Disponible en: [Advancing CAPI/CAWI technology with Survey Solutions](#)

IRIS Center (2011). Comparative Assessment of Software Programs for the Development of Computer-Assisted Personal Interview (CAPI) Applications. University of Maryland College Park. Disponible en: [Comparative Assessment of Software Programs for the Development of Computer-Assisted Personal Interview \(CAPI\) Applications](#)

Ministerio de Desarrollo Social y Familia (2020). Casen en Pandemia 2020, “Nota técnica N°1: Modalidad de aplicación Casen en Pandemia 2020”. Disponible en: [Nota técnica N°1: Modalidad de aplicación Casen en Pandemia 2020](#)

Ministerio de Desarrollo Social y Familia (2021). Casen en Pandemia 2020, “Nota técnica N°4: CAPI y Aseguramiento de Calidad (QA)”. Disponible en: [Nota técnica N°4: CAPI y Aseguramiento de Calidad \(QA\)](#)

Ministerio de Desarrollo Social y Familia (2022). Casen en Pandemia 2020, “Nota técnica N°8: 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](#)