

METODOLOGÍA



INSTITUTO NACIONAL DE ESTADÍSTICAS

Subdirección Técnica

Departamento de Presupuestos Familiares

Departamento de Investigación y Desarrollo

VIII ENCUESTA DE PRESUPUESTOS FAMILIARES

Período de Información: julio 2016 - junio 2017

Metodología – VIII Encuesta de Presupuestos Familiares

Junio / 2018

El uso de un lenguaje que no discrimine ni marque diferencias entre hombres y mujeres ha sido una preocupación en la elaboración de este documento. Sin embargo, y con el fin de evitar la sobrecarga gráfica que supondría utilizar en castellano "o/a" para marcar la existencia de ambos sexos, se ha optado por utilizar -en la mayor parte de los casos- el masculino genérico, en el entendido de que todas las menciones en tal género representan siempre a hombres y mujeres, abarcando claramente ambos sexos.

Dirección: Av. Presidente Bulnes 418, Santiago de Chile

Casilla de correo: 498 – Correo 3

Teléfono: 562 2892 4000

Sitio web: www.ine.cl

Correo electrónico: ine@ine.cl

Facebook: @ChileINE

Twitter: @INE_Chile

ÍNDICE

1. PRESENTACIÓN.....	8
2. DICCIONARIO DE ACRÓNIMOS	10
3. INTRODUCCIÓN	11
4. ANTECEDENTES Y OBJETIVOS.....	13
4.1 Antecedentes del Estudio	13
4.2 Experiencia de encuestas de ingresos y gastos aplicada a hogares de otros países.....	14
4.3 Antecedentes metodológicos de la VIII EPF en una perspectiva histórica	17
4.4 Objetivos y usos de la encuesta	19
5. MARCO TEÓRICO-CONCEPTUAL E INSTRUMENTOS DE CAPTURA DE INFORMACIÓN.....	21
5.1 Marco teórico para la captura de información de hogares, gastos e ingresos	21
5.2 Prueba piloto de la VIII Encuesta de Presupuestos Familiares	36
5.3 Instrumentos de captura de información.....	43
6.DISEÑO ESTADÍSTICO DE LA ENCUESTA	50
6.1 Diseño muestral.....	50
6.2 Desarrollo de factores de expansión	69
6.3 Estimadores y medidas de precisión de la encuesta	83
6.4 Resultados de los principales indicadores de ingreso y gasto de la encuesta.....	93
7. ESTRUCTURA ORGANIZACIONAL Y PROCESO GENERAL DEL PROYECTO	97
7.1 Estructura organizacional del Departamento de Presupuestos Familiares	97
7.2 Proceso general de producción de la VIII EPF	100
8. METODOLOGÍA DE RECOLECCIÓN DE DATOS.....	106
8.1 Metodología de trabajo de campo	106
8.2 Metodología de supervisión.....	128
9. METODOLOGÍA DE PROCESAMIENTO DE DATOS	143
9.1 Procesamiento informático	143
9.2 Metodología de digitación.....	146
9.3 Metodología de codificación	151
9.4 Metodología de preparación de bases de datos	162

10. METODOLOGÍA DE ANÁLISIS.....	186
11. GLOSARIO	189
12. BIBLIOGRAFÍA	195
13. ANEXOS.....	200
Anexo N°1. Perfiles de cargo equipo de trabajo de campo.....	200
Anexo N°2. Afiche VIII EPF	202
Anexo N°3. Ejemplo de carta de apertura de viviendas VIII EPF	203
Anexo N°4. Díptico VIII EPF	205
Anexo N°5. Hoja de Ruta VIII EPF	207
Anexo N°6. Planilla de recepción de material.....	211
Anexo N°7. Formulario de supervisión directa en fase de apertura.....	213
Anexo N°8. Formulario de supervisión directa en fase de aplicación	215
Anexo N°9. Formulario de supervisión indirecta.....	218
Anexo N°10. Formulario de supervisión de hoja de ruta.....	230
Anexo N°11. Formulario de registro de folios supervisados	232
Anexo N°12. Perfiles y módulos asociados, sistema informático VIII EPF	233
Anexo N°13. Clasificación CIUO-88 según Gran grupo y Subgrupo principal	234
Anexo N°14. Módulos evaluados por la grilla técnica de mínimos de calidad VIII EPF	235
Anexo N°15. Matriz de exigencia para la elección del vecino cercano para la imputación de días faltantes de Libretas de Gastos Individuales	236
Anexo N°16. Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de trabajadores asalariados	238
Anexo N°17. Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de trabajadores honorarios	240
Anexo N°18. Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de trabajadores de negocios por cuenta propia.....	241
Anexo N°19. Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de profesionales independientes	243
Anexo N°20. Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de jubilaciones o pensiones de vejez	245
Anexo N°21. Algoritmos de ajuste al mes de referencia VIII EPF.....	246

ÍNDICE DE TABLAS

Tabla 1: Comparación entre EPF realizadas por el INE.	18
Tabla 2: Grupos de tratamiento y control	37
Tabla 3: Asignación de prueba piloto según submuestra operativa.....	38
Tabla 4: Indicadores de desempeño operativo entre formas	39
Tabla 5: Cantidad de libretas revisadas y porcentaje de libretas con al menos un error, por forma	40
Tabla 6: Pruebas de medias sobre el gasto total por libreta de procedencia y forma.....	40
Tabla 7: Pruebas de medias sobre el gasto total por división y forma (con arriendo imputado)	41
Tabla 8: Pruebas de medias sobre el ingreso disponible total por hogar.....	41
Tabla 9: Cobertura geográfica de la encuesta según capital regional o zona conurbada	51
Tabla 10: Distribución de manzanas y viviendas en el marco muestral (total y excluidas) según región.....	54
Tabla 11: Número de manzanas y viviendas en el marco muestral de selección de la VIII EPF, según grupo de tamaño de manzanas del marco muestral de la VIII EPF	56
Tabla 12: Número de manzanas y viviendas seleccionadas en la muestra de la VIII EPF, según grupo de tamaño de manzanas del marco muestral de la VIII EPF	57
Tabla 13: Tamaño objetivo, errores calculados de acuerdo con el gasto <i>per cápita</i> promedio mensual según VII EPF y tamaño con sobremuestreo	60
Tabla 14: Distribución del total de viviendas por macrozona, capital regional, comuna y clasificación socioeconómica. Muestra con sobremuestreo.....	62
Tabla 15: Distribución del total de viviendas a encuestar por submuestra mensual y clasificación socioeconómica con sobremuestra	64
Tabla 16: Manzanas reemplazadas en el proceso de selección de la muestra de la VIII EPF	66
Tabla 17: Manzanas reemplazadas según macrozona.....	66
Tabla 18: Total de unidades muestrales seleccionadas por nivel de estimación.....	69
Tabla 19: Total conglomerados y viviendas en el marco muestral según grupo tamaño	70
Tabla 20: Cantidad y porcentaje de estratos, conglomerados y viviendas de los estratos del marco de selección de la EPF, según completitud de la muestra y condición de estrato con muestra	70
Tabla 21: Cantidad y porcentaje de viviendas del marco de la EPF, según completitud de la muestra y grupo de tamaño por condición de estrato con muestra	71
Tabla 22: Total de viviendas contenidas en el marco muestral, estimación de viviendas expandidas en la encuesta y factor de ajuste por omisión de conglomerados según macrozona	72
Tabla 23: Distribución de viviendas seleccionadas según elegibilidad en el trabajo de campo	74
Tabla 24: Clasificación de viviendas según códigos de disposición final de la VIII EPF.....	76
Tabla 25: Indicadores de desempeño operativo a nivel de viviendas VIII EPF.....	77
Tabla 26: Ponderador de no respuesta por celda de ajuste	80
Tabla 27: Proyecciones de población al 31 diciembre 2016 por sexo y tramo etario según región de las comunas urbanas consideradas en la VIII EPF	82
Tabla 28: Total de conglomerados y <i>varunit</i> (pesudoconglomerados) según macrozona	88

Tabla 29: Frecuencia de pseudoconglomerados de varianza (<i>varunit</i>) según fase de agrupación.....	89
Tabla 30: Total de estratos y pseudoestratos de varianza (<i>varstrat</i>) según macrozona	90
Tabla 31: Frecuencia del número de pseudoconglomerados de varianza (<i>varunit</i>) que componen los <i>varstrat</i>	90
Tabla 32: Frecuencia del número de conglomerados que componen los pseudoconglomerados de varianzas (<i>varunit</i>).....	91
Tabla 33: Indicadores de precisión y confiabilidad del gasto promedio mensual de los hogares según área de estimación (excluye arriendo imputado)	93
Tabla 34: Indicadores de precisión y confiabilidad del ingreso disponible promedio mensual de los hogares según área de estimación (excluye arriendo imputado).....	94
Tabla 35: Indicadores de precisión y confiabilidad del gasto <i>per cápita</i> promedio mensual de los hogares según área de estimación (excluye arriendo imputado)	94
Tabla 36: Indicadores de precisión y confiabilidad del ingreso disponible <i>per cápita</i> promedio mensual <i>per cápita</i> de los hogares según área de estimación (excluye arriendo imputado).....	94
Tabla 37: Indicadores de precisión y confiabilidad del gasto promedio mensual de los hogares según división, total de capitales regionales (excluye arriendo imputado)	95
Tabla 38: Indicadores de precisión y confiabilidad del ingreso disponible promedio mensual de los hogares, total de capitales regionales (excluye arriendo imputado).....	96
Tabla 39: Distribución del personal de trabajo de campo VIII EPF	109
Tabla 40: Esquema de visitas VIII EPF	117
Tabla 41: Períodos y motivos de planes especiales	121
Tabla 42: Tipos de medidas adoptadas en planes especiales.....	122
Tabla 43: Clasificación de códigos de Hoja de Ruta según tipo de resultado de visita y momento de clasificación	126
Tabla 44: Indicadores de desempeño operativo empleados en la VIII EPF.....	127
Tabla 45: Visitas de supervisión por región	132
Tabla 46: Tipos de errores de digitación	150
Tabla 47: Niveles CIUO88 y CIUO-08.CL.....	156
Tabla 48: Clasificación CIUO-08.CL según Gran grupo y Subgrupo principal	157
Tabla 49: Divisiones CCIF-VIII EPF	159
Tabla 50: Códigos Residuales División 15	161
Tabla 51: Clasificador de Establecimientos	162
Tabla 52: Estadística de la no respuesta parcial en gastos individuales a nivel de personas que debían contestar la libreta	171
Tabla 53: Estadística de la no respuesta parcial en gastos individuales a nivel de días, considerando aquellas personas que debían contestar la libreta	171
Tabla 54: Comparación de estadística descriptiva entre datos imputados y sin imputar, a nivel muestral	174
Tabla 55: Estructura por división del clasificador de CCIF del gasto de consumo final de los hogares pre y post imputación de LGI, a nivel expandido (excluye arriendo imputado)	174

Tabla 56: Estadística de la no respuesta parcial de ingresos, a nivel personas	175
Tabla 57: Resultados muestrales del ingreso bruto promedio por persona según fuentes de ingreso imputadas por método <i>Hot Deck</i> aleatorio	177
Tabla 58: Resultados de la imputación de ingresos en la estructura del ingreso promedio del hogar por tipo de fuentes.....	178
Tabla 59: Estadística de la no respuesta parcial del gasto financiero desagregado, a nivel créditos	179
Tabla 60: Resultados de la imputación de la desagregación del gasto financiero.....	180

ÍNDICE DE DIAGRAMAS

Diagrama 1: Esquema de gasto de consumo (utilización de los ingresos)	26
Diagrama 2: Esquema del ingreso total de los hogares	32
Diagrama 3: Organigrama general del Departamento de Presupuestos Familiares en fase de trabajo de campo de la VIII EPF.....	99
Diagrama 4: Niveles 1 y 2 del GSBPM.....	100
Diagrama 5: Simplificación de la carta Gantt del GSBPM a nivel de fases aplicada al proceso de producción de la VIII EPF.....	101
Diagrama 6: Organigrama equipo de trabajo de campo VIII EPF	107
Diagrama 7: Fases de trabajo de campo en cada submuestra VIII EPF	114
Diagrama 8: Organigrama del equipo de supervisión y capacitación	130
Diagrama 9: Proceso informático para la generación de bases de datos para procesamiento.....	145
Diagrama 10: Organigrama del equipo de digitación VIII EPF.....	147
Diagrama 11: Flujograma del proceso de digitación VIII EPF.....	148
Diagrama 12: Organigrama del equipo de codificación VIII EPF	152
Diagrama 13: Niveles del clasificador CCIF	159
Diagrama 14: Períodos de referencia de los meses de julio y agosto	181

1. PRESENTACIÓN

El Instituto Nacional de Estadísticas (INE) pone a disposición de la ciudadanía la Metodología de la VIII Encuesta de Presupuestos Familiares (EPF), cuyo trabajo de campo se realizó entre julio de 2016 y junio de 2017. Este proceso de producción estadística es uno de los que cuenta con mayor trayectoria en la institución, representando la continuación del trabajo que el INE realiza en materia de gastos e ingresos de los hogares desde el año 1956.

Con el ingreso de Chile a la Organización para la Cooperación y el Desarrollo Económico (OCDE) en 2010, se estableció el compromiso país de actualizar el Índice de Precios de al Consumidor con una periodicidad máxima de cinco años. Por ello la EPF, como principal fuente de información para el IPC, se comenzó a realizar con la misma periodicidad quinquenal.

A partir de la séptima versión de la encuesta, se incorporan recomendaciones conceptuales y metodológicas de diversos organismos internacionales y oficinas estadísticas. Por ejemplo, se adoptan normas Internacionales indicadas por la Oficina Estadística de la Unión Europea (Eurostat) en su manual “Household Budget Surveys in the EU”, además de guiarse constantemente por la información contenida en el Programa para el Mejoramiento de las Encuestas y la Medición de las Condiciones de Vida en América Latina y el Caribe (Mecovi), que fue ejecutado desde 1996 hasta 2004 por el Banco Interamericano de Desarrollo (BID), el Banco Mundial y la Comisión Económica para América Latina y el Caribe (Cepal), conjuntamente con las instituciones y agencias especializadas de los países participantes. Asimismo, las definiciones utilizadas en la encuesta han sido armonizadas con el Sistema de Cuentas Nacionales, en cuanto es el principal marco conceptual de la EPF.

Este estudio posee diversos usos de alto interés público. Además de ser principal insumo para la próxima actualización de los bienes y servicios que componen la canasta del Índice de Precios al Consumidor (IPC) y sus ponderaciones, la encuesta es utilizada por diversos organismos, públicos y privados, siendo una información valiosa para las instituciones públicas y los organismos dedicados a la investigación económica, dando origen a investigaciones y publicaciones de diversos ámbitos, tales como estudios sobre calidad de vida, nutrición, gasto en salud entre otros.

Por estas razones, dando cumplimiento al compromiso permanente del INE en cuanto a transparentar su trabajo ante la opinión pública, en las siguientes páginas ponemos a disposición de quien desee consultar, las decisiones metodológicas tomadas en las distintas etapas de la encuesta: desde la organización del proyecto, el trabajo en terreno, el procesamiento de los datos y su análisis final.

La presente versión de la encuesta incorpora importantes mejoras metodológicas respecto a sus ediciones anteriores, entre las cuales destacan mejoras de los métodos de imputación, la incorporación de criterios más exigentes para la evaluación de la calidad de la información entregada por los hogares participantes, ajustes a los factores de expansión

de la encuesta, entre otras. Todas ellas han sido integradas a la metodología de la encuesta, a fin de mejorar la calidad de sus estimaciones.

Por último, y en nombre del Instituto Nacional de Estadísticas, deseo expresar nuestros sinceros agradecimientos a todos quienes colaboraron a lo largo de los cuatro años que llevó la realización de la VIII EPF, especialmente, a los ciudadanos que, con alto sentido cívico, dispusieron de su tiempo para brindar la información solicitada, a las más de 280 personas que trabajaron directamente en este proyecto, al resto de los equipos del INE que brindaron soporte a las distintas áreas del proceso productivo y a quienes integran las Direcciones Regionales del INE. Su trabajo, esfuerzo y colaboración resultaron fundamentales para el logro de los objetivos propuestos.

Guillermo Pattillo Álvarez
Director Nacional
Instituto Nacional de Estadísticas

2. DICCIONARIO DE ACRÓNIMOS

BID: Banco Interamericano de Desarrollo

CCIF: Clasificación del Consumo Individual por finalidades (CCIF). Classification of Individual Consumption According to Purpose (COICOP)

CEPAL: Comisión Económica para América Latina y el Caribe

CIUO: Clasificación Internacional Uniforme de Ocupaciones

CISE: Clasificación Internacional de la Situación en el Empleo

ENE: Encuesta Nacional de Empleo

ENUT: Encuesta Nacional sobre Uso del Tiempo

ENUSC: Encuesta Nacional Urbana de Victimización

EPF: Encuesta de Presupuestos Familiares

ESI: Encuesta Suplementaria de Ingresos

EUROSTAT: Oficina Estadística de la Unión Europea. Statistical Office of the European Union.

FMI: Fondo Monetario Internacional

GSBPM: (sigla en inglés) Modelo Genérico del Proceso de Producción Estadística. Generic Statistical Business Process Model.

INE: Instituto Nacional de Estadísticas

INEGI: Instituto Nacional de Estadísticas y Geografía de México

IPC: Índice de Precios al Consumidor

ISFLSH: Instituciones Sin Fines de Lucro que Sirven a los Hogares

OCDE: Organización para la Cooperación y el Desarrollo Económico (OECD en inglés - Organisation for Economic Co-operation and Development)

OIT: Organización Internacional de Trabajo

ONU: Organización de las Naciones Unidas

UNECE: (sigla en inglés) Comisión Económica de las Naciones Unidas para Europa. United Nations Economic Commission for Europe.

SCN: Sistema de Cuentas Nacionales

SEN: Sistema Estadístico Nacional

3. INTRODUCCIÓN

El Instituto Nacional de Estadísticas (INE) es el organismo responsable de producir y difundir las estadísticas oficiales de Chile, proporcionando información confiable y accesible a los usuarios para la toma de decisiones. De igual modo, es el responsable de articular el Sistema Estadístico Nacional (SEN), de manera que este provea al país información pertinente, precisa, fiable, oportuna, accesible, clara, coherente y comparable a nivel nacional e internacional.

La Encuesta de Presupuestos Familiares (EPF) producida por el INE, constituye un estudio fundamental para la actualización del Índice de Precios al Consumidor (IPC). La primera EPF se realizó en los años 1956–1957, para luego aplicarse con una periodicidad de diez años, hasta la VI versión 2006-2007. Desde la VII EPF, llevada a cabo entre noviembre de 2011 y octubre de 2012, la encuesta se realiza con una periodicidad quinquenal, conforme al compromiso país adquirido con la Organización para la Cooperación y el Desarrollo Económico (OCDE), en cuanto a contar con una estructura del gasto de los hogares que permita la actualización del IPC al menos cada cinco años.

La EPF es una encuesta económica aplicada a hogares, cuyo objetivo es producir información sobre los gastos e ingresos en que estos incurren en un período de tiempo determinado. Es una encuesta que requiere de un alto nivel de participación de los hogares que son encuestados por ella, razón por la cual es uno de los procesos de producción estadística más complejos que existen entre aquellos asociados a las estadísticas de hogares del país.

Desde su séptima versión, la encuesta ha recogido recomendaciones conceptuales y metodológicas de distintos organismos internacionales y buenas prácticas de diversas oficinas estadísticas. En su actual versión, la EPF introduce mejoras al proceso de producción, relacionadas, por ejemplo, con perfeccionamientos en los cuestionarios, la introducción de nuevos ítems para mejorar la captura de gastos y la optimización de los procesos de digitación y codificación de datos, además de la estandarización de la organización de su proceso productivo, adaptándolo al estándar internacional Modelo Genérico del Proceso de Producción Estadística (GSBPM, por su sigla en inglés).

La Metodología de la VIII EPF forma parte de una serie de documentos de referencia¹ y consulta que el INE pone a disposición del público para quienes deseen conocer instrucciones, procedimientos y decisiones metodológicas del estudio, dando cumplimiento, de este modo, a las orientaciones generales para la documentación de metodologías de la institución.

1 Junto a la presente metodología de la VIII EPF, en la página web institucional del INE se pone a disposición del público la siguiente documentación: "Informe de calidad de la VIII EPF", "Documento de principales resultados de la VIII EPF", "Manual de usuario de la base de datos de la VIII EPF" y los manuales para la capacitación de los distintos equipos de la encuesta ("Manual de trabajo de campo para encuestadores" Tomos I y II, "Manual de trabajo de campo para encargados de grupo", "Manual de trabajo de campo para supervisores" y "Manual de codificación").

La Metodología de la VIII EPF está enfocada a quienes deseen conocer los métodos y decisiones asociadas a la producción de la última versión de la encuesta. Su estructura muestra el siguiente orden:

- En primer lugar, se presentan los antecedentes y objetivos del estudio. Se revisan experiencias internacionales respecto a este tipo de encuesta, para luego realizar una presentación de los antecedentes nacionales que alimentan la presente versión de la EPF. Aquel primer capítulo finaliza exponiendo los objetivos y usos del estudio.
- En el siguiente capítulo se aborda el marco teórico conceptual en el cual se basa la encuesta, abordando los principales conceptos empleados por el estudio, los criterios técnicos utilizados para capturar información de gasto e ingreso de los hogares, la prueba piloto de la encuesta y los instrumentos utilizados para la recolección de datos.
- A continuación, se presenta el diseño estadístico de la encuesta, revisando las características del diseño muestral, el desarrollo de los factores de expansión con los correspondientes ajustes en sus ponderaciones, los estimadores y medidas de precisión y una breve síntesis de los resultados de los indicadores principales de ingreso y gasto de la encuesta.
- En el siguiente capítulo se expone la estructura organizacional de la encuesta y el flujo general asociado al proceso de producción estadística de la VIII EPF, basándose en el estándar internacional del Modelo Genérico del Proceso de Producción Estadística (GSBPM por sus siglas en inglés).
- Posteriormente, se explica la metodología de recolección de datos (equipos de trabajo de campo y supervisión), haciendo una revisión del modo de organización, capacitación y metodología de trabajo de los equipos involucrados en esta fase.
- En el capítulo siguiente, se aborda la metodología de procesamiento de datos, revisando los subprocesos de procesamiento Informático, digitación, codificación y preparación de base de datos.
- El penúltimo capítulo reporta la metodología de análisis seguida por la encuesta para la presentación de resultados.
- Finalmente, se agrega un glosario asociado a las palabras y expresiones relevantes del documento, además de la bibliografía y anexos del presente documento.

4. ANTECEDENTES Y OBJETIVOS

4.1 Antecedentes del Estudio

En términos generales, las encuestas de presupuestos familiares (EPF) o de ingresos y gastos² son estudios realizados a hogares, cuyo objetivo es conocer las pautas de consumo y la estructura de gasto de la población de estudio, así como información referente a los ingresos de esta.

Históricamente, el principal objetivo para llevar a cabo este tipo de encuestas ha sido recolectar información sobre los gastos de consumo de los hogares para su uso en la actualización de los índices del IPC³. Sin embargo, la riqueza de la información que capturan ha permitido ampliar su uso a otras áreas, tales como conocer algunos aspectos del nivel de vida, estudiar tendencias de consumo, indagar respecto a los patrones alimenticios de población, obtener información para cuentas nacionales y utilizar su información para el cálculo de las líneas de pobreza, entre otras.

La evolución de este tipo de encuestas ha sido incentivada a través de los requerimientos propios de los organismos encargados de realizarlas, en función de las necesidades de actualización de información⁴, así como de recomendaciones de diversos organismos internacionales, que buscan que la información producida cumpla con ciertos principios de calidad asociados a la producción estadística. Así, se han transformado en encuestas multipropósito que, dada la relevancia y la diversidad de sus usos, requieren de la realización de un proceso de producción estadística con los más altos estándares de calidad.

En línea con lo anterior, el desarrollo de la producción estadística de la VIII EPF procuró incorporar las mejores prácticas internacionales, además de seguir lineamientos y recomendaciones internacionales de la Oficina Estadística de la Unión Europea (Eurostat, 2003), de la Organización Internacional del Trabajo, de Naciones Unidas (ONU *et al.*, 2016; UNECE, 2016) y del Programa para el Mejoramiento de las Encuestas y la Medición de las Condiciones de Vida en América Latina y el Caribe (MECOVI, 2004). Si bien la EPF es un instrumento aplicado a hogares, se debe tener presente que es una encuesta de carácter económico y como tal, el marco conceptual en el cual se basa es el Sistema de Cuentas Nacionales (SCN).

Aunque más adelante se desarrollarán en detalle los elementos teóricos que sustentan la encuesta, es importante mencionar que el concepto de consumo ocupa una relevancia central en la encuesta. Al respecto, la Organización Internacional de Trabajo (OIT, 2003c) sostiene:

2 En otros países de la región, las encuestas de presupuestos familiares se conocen por este denominativo, y es el caso por ejemplo en Colombia, Ecuador, México, entre otros.

3 Para conocer objetivos de las encuestas de presupuestos familiares en Unión Europea ver "Household Budget Surveys in the EU - Methodology and recommendations for harmonisation – 2003", Capítulo 3: Objetivos, Usos y Usuarios, página 9.

4 El propósito principal para la realización de esta encuesta es la actualización del insumo para la elaboración de la canasta del IPC. Para el caso chileno, dado el ingreso a la OCDE, se sugiere la actualización del IPC al menos cada cinco años, en consecuencia, la periodicidad de realización de la EPF cambia de diez a cinco años.

Los hogares tienen necesidades y deseos personales que se satisfacen directamente mediante el consumo de bienes y servicios resultantes de actividades que son productivas en sentido económico. Estos bienes y servicios se conocen como bienes y servicios de consumo y su valor individual se define como gastos de consumo en ese bien o servicio (pág. 30).

La satisfacción de las necesidades de los hogares se puede realizar a través de distintos canales:

- Mediante transacciones en el mercado, utilizando instrumentos monetarios o no.
- Mediante la producción de bienes y servicios al interior de los hogares.
- Mediante transferencias sociales recibidas desde el Estado o instituciones sin fines de lucro que sirven a los hogares (ISFLSH).

La suma de los gastos realizados a través de estos tres canales se conoce como **gasto de consumo final efectivo** de los hogares, que, para propósitos de medición de condición de vida o bienestar económico, sería la medida más adecuada a utilizar, ya que considera todos aquellos bienes y servicios que tiene a disposición el hogar para la satisfacción de sus necesidades y no solo aquellos que los hogares adquirieron en el mercado (Eurostat, 2003). Sin embargo, existen dificultades prácticas asociadas a la medición del **gasto de consumo final efectivo** de los hogares, ya que el cálculo de algunas transferencias sociales (valorización de educación, salud y otros bienes recibidos por los hogares) complejizarían la aplicación de esta encuesta.

Por lo tanto, se recomienda que la EPF busque medir el **gasto en consumo final** de los hogares, que contabiliza los gastos realizados a través de transacciones en el mercado, utilizando o no instrumentos monetarios como medio de pago (Eurostat, 2003, pág. 25)⁵. Además, dado que el principal uso de las EPF es la actualización del IPC, este agregado es más adecuado que el gasto en consumo efectivo, ya que los gastos son contabilizados a precios de mercado.

En los próximos apartados se describen antecedentes de este tipo de encuestas en el ámbito internacional, luego se presentan antecedentes históricos de las encuestas de presupuestos familiares en Chile, para finalizar con los objetivos y usos de la VIII EPF.

4.2 Experiencia de encuestas de ingresos y gastos aplicada a hogares de otros países

Parte fundamental del trabajo de diseño metodológico de la VIII EPF fue la revisión de experiencias internacionales. Dada la complejidad y heterogeneidad de realidades sociales, se hace necesario comparar la implementación de los conceptos utilizados por las distintas EPF, de modo de extraer buenas prácticas.

Los manuales elaborados por la Oficina Estadística de la Unión Europea sobre este tipo de encuestas, además de entregar directrices conceptuales y recomendaciones sobre la

5 Considerado las dificultades prácticas asociadas a la medición del Gasto de Consumo Final Efectivo, Eurostat recomienda la medición del Gasto de Consumo Final para los Estados miembros de la Unión Europea.

recopilación de información sobre los gastos en los hogares, provee comparaciones entre las EPF aplicadas en los distintos países de la Unión Europea (Eurostat, 2003; Eurostat, 2015). En ese sentido, estos documentos sirven como un importante instrumento para la elaboración del marco metodológico y conceptual de la VIII EPF.

4.2.1 Periodicidad de la encuesta

Mientras que países como Alemania, Canadá, España, Estados Unidos, Inglaterra e Italia cuentan con EPF continuas, Chile, al igual que otros países, realiza un levantamiento cada cierto período de tiempo, debido básicamente a los elevados costos asociados. A partir del año 2010, con el ingreso de Chile a la OCDE, Chile realiza la EPF cada cinco años (al igual que Austria, Francia, Colombia, Costa Rica, República de Irlanda y Portugal, entre otros).

Mientras que México realiza esta encuesta de forma bienal, Australia la realiza cada seis años y otros países como Argentina y Uruguay la realizan cada diez años. Se observa que otros países, como Bolivia, no cuentan con una periodicidad definida.

4.2.2 Período de presencia en el hogar

Las encuestas de presupuestos familiares poseen diferentes períodos de referencia y, en general, utilizan diversos formularios que son aplicados en más de una visita a los hogares. Respecto a la captura de los gastos del día a día de los hogares encuestados, muchos países suministran libretas para que los miembros de los hogares encuestados autorreporten sus gastos durante un período de tiempo determinado.

Respecto al período de colaboración de cada hogar, en Chile la encuesta tiene una duración de una quincena, lapso en que los miembros del hogar deben entregar información sobre sus gastos e ingresos en seis libretas. Durante este período, se entrega una libreta especial a cada miembro del hogar de 15 años o más para que registre sus gastos individuales o del día a día⁶.

En un contexto internacional, los países utilizan distintas metodologías de recolección de la información. Mientras países como Argentina, Brasil, Costa Rica, Ecuador, Italia, México, Perú y Uruguay utilizan una semana para medir el gasto de consumo individual, Australia, Austria, Canadá, Colombia, EEUU, España, Francia, Irlanda, Portugal, Sudáfrica y Reino Unido utilizan dos semanas.

No existe consenso a nivel internacional sobre el período óptimo para registrar los gastos diarios de los hogares. Entre los argumentos a favor de hacerlo durante dos semanas, se explica que así se mitigaría el efecto de una semana de gasto atípica en el hogar. Por otro lado, un argumento a favor de utilizar solo una semana apunta a que el desgaste que se produce en los informantes en la segunda semana podría aumentar el porcentaje de días sin registros.

Aunque no existe acuerdo respecto a este tema, cierta experiencia indica que no es adecuado extender el tiempo de colaboración de los hogares más allá de dos semanas. En ese sentido, es concluyente el caso de Sudáfrica, país en el que se implementó en 2005-2006 un período de referencia de un mes, siendo reducido a dos semanas en la

6 En el apartado 5.3.2 se entregan más antecedentes respecto a esta libreta de recolección de información.

siguiente versión de 2011-2012. El motivo fue el acusado agotamiento experimentado por los hogares, producto del extenso período de registro (Statistics South Africa, 2012).

Teniendo en consideración los antecedentes internacionales y los resultados de la prueba piloto realizada el año 2015, donde se probó reducir el tiempo de colaboración de los hogares a una semana⁷, en la VIII EPF se optó por mantener el período de registro quincenal (aproximadamente dos semanas).

4.2.3 Períodos de recuerdo

Existen ciertos tipos de gastos que por su naturaleza no quedan bien registrados en las libretas de gastos diarios. Aquellos cumplen con ciertas características, por ejemplo, no ser estacionales (en la mayoría de los casos), tener una baja frecuencia de adquisición y, además, representar un gasto importante, en términos del presupuesto total del hogar. En este sentido, los denominados bienes o productos semidurables y durables que pertenecen al consumo final de los hogares (OIT, 2003c) se definen como aquellos que no reducen su capacidad de satisfacer necesidades y que son utilizados repetidamente durante un largo intervalo de tiempo⁸.

Para medir de mejor forma estos gastos, las encuestas de presupuestos familiares han introducido distintos períodos de referencia para consultar respecto a ellos. En muchos casos se apela al recuerdo de los informantes, consultando sobre los gastos en bienes semidurables o durables de manera retrospectiva, a través de distintos espacios de tiempo. Luego, estos gastos son ajustados al período de referencia informado mediante diversos algoritmos de ajuste⁹.

Cabe mencionar que no existe acuerdo internacional sobre qué bienes y servicios preguntar en cada período de recuerdo ni tampoco cuáles son los períodos óptimos. En consecuencia, los países utilizan diferentes períodos de recuerdo, siendo los más recurrentes 3, 6 y 12 meses. La experiencia sugiere que mientras más importante sea el gasto desde el punto de vista del presupuesto familiar, más sencillo es de recordar; por lo tanto, los gastos de mayor cuantía pueden ser consultados utilizando períodos de referencia más amplios (U.S. BLS, 2004; U.S. BLS, 2006; Bradburn, 2010).

En la VIII EPF se mantuvieron los tres períodos de recuerdo utilizados en la VII versión de la encuesta: 3, 6 y 12 meses. Los productos más significativos desde el punto de vista del gasto se preguntaron utilizando períodos más amplios, debido a la capacidad de recuerdo ya señalada, mientras que los productos menos significativos fueron consultados en lapsos más reducidos respecto al momento de la entrevista.

7 Para más información sobre la prueba piloto de la encuesta, consultar el apartado 5.2.

8 En cierta medida, los criterios que hacen que un tipo de gasto se consulte con un período de recuerdo están correlacionados con la definición de productos semidurables y durables. Se sugiere revisar los apartados 5.1.2 y 5.3.4.

9 En Chile, la información se entrega mensualizada, mientras que otros países como España, Inglaterra, Italia o Portugal anualizan la información entregada sobre los gastos. Más adelante, en el apartado 9.4.4.2, se detalla la metodología de cada uno de los factores de ajuste temporal utilizados en la VIII EPF.

4.3 Antecedentes metodológicos de la VIII EPF en una perspectiva histórica

La Encuesta de Presupuestos Familiares es uno de los productos estadísticos del INE que cuenta con mayor tradición histórica y relevancia, ya que es el principal insumo¹⁰ para la actualización de la canasta de bienes y servicios que componen el Índice de Precios al Consumidor (IPC), por lo que sus desarrollos en el seno de la institución están íntimamente ligados.

A lo largo del tiempo, las encuestas han variado en sus objetivos, en su metodología de recopilación de los datos y en el tratamiento de estos. El primer estudio sobre el consumo en el Gran Santiago se realizó en el año 1928, con el fin de construir un índice de precios al consumidor. Casi tres décadas después, entre noviembre de 1956 y octubre de 1957 se realizó la primera Encuesta de Presupuestos Familiares en el Gran Santiago, cuyos resultados permitieron elaborar la canasta del IPC de 1957.

Las distintas versiones realizadas de la encuesta han desarrollado un período de trabajo de campo de un año calendario, con el fin de incorporar las distintas variaciones estacionales en la estructura del gasto de los hogares chilenos.

Las primeras seis encuestas se realizaron con una periodicidad de aproximadamente diez años:

- I EPF: noviembre 1956 - octubre 1957
- II EPF: septiembre 1968 - agosto 1969
- III EPF: diciembre 1977 - noviembre 1978
- IV EPF: diciembre 1987 - noviembre 1988
- V EPF: agosto 1996 - julio 1997
- VI EPF: noviembre 2006 - octubre 2007

Tras el ingreso de Chile a la OCDE en 2010, el país adquiere el compromiso de actualizar cada cinco años la canasta utilizada para construir el Índice de Precios al Consumidor. De este modo, a contar de la VII EPF la encuesta pasa a ser realizada con una periodicidad quinquenal (noviembre 2011 - octubre 2012), para luego de cinco años, entre julio de 2016 y junio de 2017, realizarse el trabajo de campo de la VIII EPF.

La encuesta ha evolucionado en diversos aspectos, de acuerdo con los requerimientos y exigencias de los usuarios de la encuesta y también a partir de las recomendaciones metodológicas de organismos internacionales.

Existen diversos aspectos en los que esta encuesta ha ido evolucionando a lo largo del tiempo. En la tabla 1 es posible observar una síntesis respecto a algunos de los cambios más relevantes abordados por las EPF.

¹⁰ Se indica que es el principal, mas no el único insumo el IPC, puesto que aquel producto también emplea información de Cuentas Nacionales y registros administrativos para ajustar las ponderaciones de los gastos en algunos productos (INE Chile, 2013b).

Tabla 1: Comparación entre EPF realizadas por el INE

Aspecto	VIII EPF	VII EPF	VI EPF	V EPF	IV EPF	III EPF	II EPF	I EPF
Ámbito geográfico	Capitales regionales y zonas conurbadas (61 comunas de la VII EPF, además de Chillán y Chillan Viejo, de la nueva Región de Ñuble).	Capitales regionales y zonas conurbadas (60 comunas de la VI EPF, además de Alto Hospicio).	Gran Santiago y ciudades capitales regionales con alguna de sus zonas conurbadas (60 comunas).	Gran Santiago (desde 1992, incluye Padre Hurtado), Iquique y Punta Arenas.	Gran Santiago	Santiago	Hogares particulares de todo el país, procesándose solo Santiago	Santiago, Providencia, Ñuñoa, San Miguel, Quinta Normal, Renca, Conchalí, Barrancas, La Cisterna, La Granja y Las Condes.
Período de trabajo de campo	julio 2016 - junio 2017	noviembre 2011 - octubre 2012	noviembre 2006 - octubre 2007	agosto 1996 - julio 1997	diciembre 1987 - noviembre 1988	diciembre 1977 - noviembre 1978	septiembre 1968 - agosto 1969	noviembre 1956 - octubre 1957
Período de registro	Una quincena	Una quincena	Una quincena	Una quincena	3 meses	3 meses	3 meses	3 meses
Período de recuerdo	3, 6 y 12 meses.	3, 6 y 12 meses.	6 y 12 meses	1 mes	No hubo	No hubo	No hubo	No hubo
Criterio de medición gasto	Consumo adquirido (para gastos periódicos, consumo pagado)	Consumo adquirido (para gastos periódicos, consumo pagado)	Consumo adquirido (excepcionalmente gastos de consumo pagado)	Consumo adquirido	Consumo adquirido	Consumo pagado	Consumo pagado	Consumo pagado
Concepto de gasto	Gasto de consumo final	Gasto de consumo final	Gasto de consumo final efectivo					
Concepto de ingreso	Ingreso total disponible	Ingreso total disponible	Ingreso corriente disponible					

Fuente: Instituto Nacional de Estadísticas (INE).

4.3.1 Período de trabajo de campo

En la VIII EPF, al igual que en versiones anteriores de la encuesta, el período de trabajo de campo es de doce meses consecutivos, con el fin de incorporar en la estructura del gasto de los hogares, las variaciones estacionales de dicha variable. En esta versión, el trabajo de campo oficial de la encuesta se realizó entre julio de 2016 y junio de 2017.

Durante el mes de junio de 2016 se realizó una marcha blanca del proceso, que consistió en un ensayo del trabajo de campo durante un mes completo. Esta actividad tuvo como objetivo poner en marcha la ejecución del proceso en todo el país, con el fin de detectar eventuales desviaciones logísticas respecto al proceso de producción y mejorarlas antes del inicio oficial de la encuesta.

La VII EPF realizó su trabajo de campo entre noviembre de 2011 y octubre de 2012, entregando sus resultados en septiembre de 2013. La VIII EPF, en cambio, adelantó su inicio al mes de julio, respondiendo a la solicitud realizada por los equipos del IPC respecto a poder contar con los resultados finales de la encuesta antes de septiembre de 2018. Por

esta razón, el proceso productivo completo se adelanta para poder entregar los resultados a los equipos de precios antes del segundo semestre de 2018.

4.3.2 Ámbito de representatividad

El ámbito de representatividad de las primeras cinco versiones de la encuesta fue la zona urbana del Gran Santiago. El área de estimación se extendió a partir de la VI EPF, pasando a ser una encuesta de alcance nacional, y como consecuencia el ámbito geográfico se amplió al conjunto de capitales regionales del país, más algunas de sus zonas conurbadas. De este modo, junto a una correcta captura de los gastos de consumo final de los hogares, la ampliación del área de estimación permite elaborar de forma robusta una canasta de bienes y servicios más representativa de la realidad chilena urbana, con ponderadores reales de los gastos dentro del presupuesto de los hogares urbanos.

En la VIII EPF, se agregaron a la muestra las comunas de Chillán y Chillán Viejo. Si bien ninguna de ellas correspondía a una capital regional en el momento del diseño de la encuesta, se decidió incluirlas tomando en consideración la tramitación de la ley que desde 2018 establecería la creación de la nueva Región de Ñuble, cuya capital regional es Chillán¹¹. La variación de la nueva canasta del IPC base 2018, calculado a partir de la encuesta, se publicará en febrero de 2019, por lo tanto, fue relevante incorporar las comunas asociadas a la nueva capital regional para que pudiesen ser incorporadas en la nueva canasta del IPC¹².

4.4 Objetivos y usos de la encuesta

Si bien a nivel institucional, la Encuesta de Presupuestos Familiares es la principal fuente de información para la actualización de la canasta de bienes y servicios que componen el Índice de Precios al Consumidor (IPC) y sus respectivas ponderaciones, también posee otros usos relevantes a nivel nacional.

La EPF es la principal fuente de información para actualizar las líneas de indigencia y pobreza, que se utilizan en la medición oficial de pobreza por ingresos del país, a cargo del Ministerio de Desarrollo Social (MDS)¹³. Además, los resultados de la encuesta constituyen un insumo para la estimación del gasto de los hogares en la compilación de cuentas nacionales realizada por el Banco Central. También son usados por el Ministerio de Salud para la estimación del gasto de bolsillo en salud de los hogares y como base para diversos estudios de calidad de vida, nutrición, estudios de patrones de consumo, entre otros¹⁴.

11 La Región de Ñuble es la nueva región del país que entrará en vigencia en septiembre de 2018, según la Ley N° 21.033 que crea la Región de Ñuble y las provincias de Diguillín, Punilla e Itata. Para más información visitar: <http://bcn.cl/21vsb>.

12 Una situación similar tuvo lugar durante la realización de la VI EPF (2006-2007), ya que mientras se estaba realizando la encuesta se crearon las dos regiones de Arica y Parinacota y de Los Ríos. Las capitales regionales de esas nuevas regiones no estuvieron consideradas en el diseño original, por lo que se tuvieron que introducir de forma posterior al inicio del trabajo de campo oficial de la encuesta.

13 Respecto a la actualización de esta información, si bien no existe una definición metodológica explícita, existe una recomendación emanada desde la Comisión para la Medición de la Pobreza (CMP) publicada el año 2014 en que recomienda realizar esta actualización cada diez años (CMP, 2014, pág. 13).

14 A modo de ejemplo, para revisar usos relacionados con el estudio de patrones de consumo alimentario y nutrición, consultar Crovetto y Uauy (2012; 2014), Crovetto *et.al.* (2014) y Araneda *et.al.* (2016).

A continuación, se presentan los objetivos del estudio para luego revisar los principales usos de la encuesta junto a las consideraciones generales de uso.

4.4.1 Objetivos del estudio

Objetivo principal:

Identificar la estructura y características del gasto en consumo final de los hogares urbanos, en las capitales regionales de Chile y algunas de sus zonas conurbadas, con un período de referencia de un año.

Objetivo secundario:

Identificar la estructura del ingreso total disponible de los hogares urbanos, en las capitales regionales de Chile y algunas de sus zonas conurbadas, con un período de referencia de un año.

4.4.2 Principales usos

A continuación, se describen los principales usos de la encuesta con algunas consideraciones asociadas a ellos, además de los objetivos del estudio y los clasificadores utilizados.

- Ser la principal fuente de información para la actualización de la canasta de bienes y servicios que componen el IPC y sus respectivas ponderaciones.
- Recopilar información de los tipos de establecimiento donde realizan sus compras los hogares urbanos.
- Ser la principal fuente de información para la actualización de las líneas de indigencia y pobreza para la medición oficial de pobreza por ingresos del país.
- Servir de insumo para la estimación en la compilación de cuentas nacionales (SCN) de diferentes componentes del sector institucional de los hogares.
- Servir de insumo para la estimación del gasto de bolsillo en salud realizada por el Ministerio de Salud.
- Recopilar información relativa a la condición de tenencia de las viviendas urbanas.
- Disponer de información socioeconómica de los hogares urbanos.
- Servir de fuente de información para el desarrollo de estudios de calidad de vida, nutrición, estudios de patrones de consumo, entre otros.

4.4.3 Consideraciones generales de uso

El diseño muestral de la VII EPF permite realizar estimaciones consistentes y estadísticamente significativas con frecuencia temporal anual y en términos geográficos para las áreas de estimación de Gran Santiago, Resto de capitales regionales y Total de capitales regionales (Gran Santiago + Resto de capitales regionales).

Si bien el trabajo de campo de la encuesta se realiza a un nivel más desagregado en términos geográficos (manzana, comuna, región), los estimadores no son estadísticamente significativos a estos niveles.

5. MARCO TEÓRICO-CONCEPTUAL E INSTRUMENTOS DE CAPTURA DE INFORMACIÓN

El presente capítulo presenta las definiciones usadas en la VIII EPF, las cuales integran recomendaciones planteadas por Eurostat (2003), la Organización Internacional del Trabajo (OIT, 2003a; OIT, 2003b; OIT, 2003c) y el Sistema de Cuentas Nacionales 2008 (ONU *et al.*, 2016)¹⁵. Junto con ello, al igual que su versión anterior, la VIII EPF incorpora recomendaciones y buenas prácticas internacionales vinculadas con este tipo de encuestas, las que son descritas en el presente documento¹⁶.

La adopción de criterios y buenas prácticas internacionales de medición ha permitido una mayor estandarización de conceptos incorporados en el proceso de creación de los instrumentos que capturan la información y en las distintas fases asociadas al proceso de producción estadística.

Este apartado se divide en tres partes: en primer lugar, se dan a conocer los principales conceptos asociados a la definición de hogar, junto con los criterios que permiten capturar la información sobre gastos e ingresos, los que son transversales a los distintos instrumentos de registro. En segundo lugar, se presenta la prueba piloto de la VIII EPF, un insumo fundamental para actualizar los cuestionarios y la metodología de trabajo de campo. En tercer lugar, se describen las libretas utilizadas para capturar dicha información y realizar una caracterización sociodemográfica y económica del hogar, para así dar cumplimiento a los objetivos de la encuesta.

5.1 Marco teórico para la captura de información de hogares, gastos e ingresos

5.1.1 Concepto de hogar

La definición de hogar tiene una relevancia fundamental en la encuesta, debido a que corresponde a la principal unidad de análisis del estudio (INE Chile, 2016a). La encuesta justamente busca identificar la estructura de gastos de consumo e ingresos disponibles de los hogares. Una vez definido el concepto de hogar, se requiere describir las características sociodemográficas de los mismos con el fin de analizar la información de la principal unidad de análisis de la EPF.

Diversos organismos internacionales definen el hogar como una o más personas, con o sin lazos de parentesco, que comparten domicilio y satisfacen sus necesidades de alimentación y otros bienes esenciales a través de un presupuesto común (ONU, 2008; CEPAL, 2011). En tanto, Eurostat (2003) e investigadores como Hoffmeyer-Zlotnik & Warner (2009) identifican cuatro dimensiones que sostienen el concepto de un hogar:

15 El Sistema de Cuentas Nacionales 2008 fue traducido de manera oficial por CEPAL con el objetivo de proveer precisiones técnicas y lingüísticas propias de la actualización.

16 Por ejemplo, la incorporación de sobremuestreo en el diseño muestral de la encuesta, seguimiento a los equipos de campo a través de los indicadores de desempeño operativo, incorporación del estándar de procesos del GS-BPM, mejoras al tratamiento de la no respuesta parcial, ajuste a las calibraciones de los factores de expansión, entre otras.

- a. **Corresidencia:** implica que las personas viven juntas, es decir, comparten una vivienda particular con una dirección registrada.
- b. **Financiamiento doméstico:** involucra compartir gastos e ingresos¹⁷, lo que implica mantener un presupuesto común, mayoritariamente de alimentación y alojamiento.
- c. **Organización doméstica:** supone la mantención y limpieza conjunta de los espacios del hogar, y el compartir comidas.
- d. **Familia:** asume la existencia de una relación legal o de sangre, a través del matrimonio, adopción o tutela afectiva.

Por otro lado, el Sistema de Cuentas Nacionales 2008 define el hogar como “un grupo de personas que comparten la misma vivienda, que juntan, total o parcialmente, su ingreso y su riqueza y que consumen colectivamente ciertos tipos de bienes y servicios, sobre todo los relativos a la alimentación y el alojamiento” (ONU *et al.*, 2016, pág. 723). Esta definición especifica que algunas decisiones relativas al consumo del hogar se deben tomar en conjunto y que cada miembro debe tener algún grado de derecho sobre los recursos colectivos.

Durante el 2014-2015 el INE propuso una definición estandarizada del concepto de hogar, en el marco de los trabajos desarrollados por el Grupo de Trabajo de Hogar¹⁸ compuesto por representantes de distintos departamentos de la institución liderados por el Sistema Integrado de Estadísticas Sociales (SIES). El objetivo de este grupo fue diagnosticar, discutir e investigar sobre las experiencias de las distintas encuestas que utilizan como unidad de análisis el hogar, para la creación de un marco conceptual común que posibilitara la complementariedad y comparabilidad entre encuestas realizadas por el Sistema Estadístico Nacional.

La estandarización de la identificación de hogar llevó al grupo a definirlo como “una o más personas que, unidas o no por vínculo de parentesco, residen en una misma vivienda o parte de ella y se benefician de un mismo presupuesto, principalmente para alimentación” (INE Chile, 2015a, pág. 1).

La VIII EPF adopta esta definición de hogar que integra las recomendaciones internacionales y estandariza a nivel institucional y para el Sistema Estadístico Nacional el concepto de hogar para su medición (INE Chile, 2015a), tomando en cuenta las siguientes consideraciones:

- a. El hogar está constituido por “*una o más personas...*”: un hogar puede estar formado por una sola persona (que se beneficia de su propio presupuesto) o por varias que se benefician de un mismo presupuesto, principalmente para alimentación.
- b. “*...que unidas o no por vínculos de parentesco...*”: para constituir un hogar no es necesario que las personas sean parientes. Si bien en muchos casos será posible

17 Eurostat (2003) identifica la agrupación de ingresos como una dimensión distinta a la del gasto. Para efectos de esta metodología, se agrupó en el ítem de financiamiento.

18 Este grupo surgió en respuesta a las recomendaciones internacionales que sugieren la coordinación y estandarización de la producción estadística. Esto implica generar productos estadísticos que sigan lineamientos, criterios y metodologías comunes.

encontrar hogares compuestos por familiares, las personas sin parentesco, como amigos o estudiantes, podrían también constituir un hogar si mantienen algún grado de vínculo económico entre sí, incorporando así los cambios sociales y culturales que pueden ver transformadas las estructuras de los hogares en Chile.

- c. *“...residen en una misma vivienda o parte de ella...”*: para que una persona forme parte de un hogar en la vivienda, debe residir habitualmente en ella. Por ello es necesario corroborar que las personas no sean residentes habituales de otra vivienda, ya sea en la dirección que se realiza la entrevista u otro lugar de residencia.
- d. *“...se benefician de un mismo presupuesto...”*: todas las personas que componen un hogar, sin importar si aportan o no con ingresos al presupuesto, como es el caso de los bebés, niños o cualquier persona que no genere ingresos, se benefician de un presupuesto común.
- e. *“...principalmente para alimentación”*: considerando la diversidad de gastos que puede realizar un hogar con cargo a un mismo presupuesto, se ha establecido que la alimentación es el gasto más generalizado, y permite distinguir con mayor claridad entre grupos con presupuestos diferenciados (pág. 1).

En consecuencia, de acuerdo a la definición antes detallada, aquellas personas que residan en una misma vivienda, pero mantengan presupuestos separados de alimentación, son considerados como hogares diferentes.

El hogar puede estar constituido por una o más personas, por lo cual es necesario especificar a quiénes se considera como miembro del hogar. La importancia de las categorías radica en que incluir o excluir personas afecta las características sociodemográficas, tales como el tamaño y composición de los hogares que responden la encuesta.

No existe consenso a nivel internacional respecto a la definición de miembro del hogar. No obstante, Eurostat (2003) recomienda aplicar el concepto desarrollado por las Estadísticas de la Unión Europea sobre Ingreso y Condiciones de Vida (EU-SILC), donde se consideran como miembros del hogar a las siguientes personas:

- Residentes habituales, relacionados o no con otros miembros del hogar.
- Pensionistas o inquilinos (durante al menos 6 meses).
- Visitantes (por al menos 6 meses).
- Servicio doméstico (durante al menos 6 meses).
- Residentes habituales, pero temporalmente ausentes de la vivienda.
- Estudiantes del hogar educados lejos de la vivienda.
- Personas ausentes por largos períodos, pero que mantienen un vínculo con el hogar (Eurostat, 2003, págs. 19-20).

Por otro lado, la OIT identifica que la pertenencia habitual a un hogar se conforma de tres maneras: 1) personas que están presentes en la vivienda al momento la entrevista, 2) individuos que legalmente residen en la vivienda o 3) personas que son residentes habituales (OIT, 2003a).

La VIII EPF establece que una persona es miembro del hogar si reside habitualmente en la vivienda, es decir, que haya habitado o dormido ahí regularmente al menos cuatro días de la semana en los últimos seis meses. A la vez, aquel residente debe compartir presupuesto de alimentación, situación que excluye a pensionistas e inquilinos, así como al personal de

servicio doméstico, en tanto se entiende que, en estos casos, la provisión de alimentos y/o alojamiento es parte de un intercambio económico¹⁹ (ONU *et al.*, 2016).

En el caso de las personas que se incorporan de forma reciente a un hogar, el criterio de residencia durante los últimos seis meses se sustituye por la intención de permanecer durante un período igual o mayor en el nuevo lugar.

Además, siguiendo recomendaciones internacionales, se consideran miembros del hogar a aquellas personas que, al momento de realizar la entrevista, no se encontraban en la vivienda, pero que cumplen alguno de los siguientes requisitos: 1) personas que durante el período de aplicación de la encuesta se encontraban temporalmente ausentes por motivos laborales y no tienen residencia habitual en otra vivienda particular²⁰, 2) personas que por algún motivo residen en viviendas colectivas²¹ y 3) estudiantes ausentes que no tienen otra residencia habitual.

Finalmente, dentro de un hogar se clasifica a los integrantes según las siguientes categorías: jefe/a de hogar, esposo/a o cónyuge, conviviente civil (AUC), conviviente de hecho o pareja, hijo/a, hijo/a del cónyuge o pareja, hermano/a, padre o madre, cuñado/a, suegro/a, yerno/nuera, nieto/a, abuelo/a, otro pariente y otro no pariente²².

En cuanto al concepto de jefatura de hogar, en los censos de población y encuestas de hogares se aplican, en general, los siguientes tres criterios²³ para su identificación (INE Chile, 2016a; Budlender, 2003; DIESA, 1984):

- Autoasignación o propia definición: clasificación como jefe/a de hogar que el integrante del hogar entrevistado hace de sí mismo o de otra persona miembro de este.
- Criterio de autoridad: identificación del integrante del hogar como una autoridad, es decir, como la persona que ejerce autoridad para determinar reglas u órdenes, tomando las decisiones más importantes del hogar.
- Criterio económico: expresado en:
 - o Provisión principal de recursos al hogar: corresponde al integrante que hace el mayor aporte en términos económicos a la mantención del hogar, es decir, la persona *proveedora principal*.
 - o Tenencia de la vivienda (sea como propietaria o como arrendataria): es decir, se aplica el concepto anglosajón *householder* (INE Chile, 2016a).

19 Dado el proceso de estandarización llevado a cabo por el INE, este concepto de hogar adoptado por la VIII EPF es comparable al utilizado en la Encuesta Nacional de Uso del Tiempo 2015 (Enut) y Encuesta Nacional Urbana de Seguridad Ciudadana 2016 (Enusc).

20 Se podrían encontrar ausentes por motivos de vacaciones, trabajo, hospitalización u otro motivo. Ejemplos de estos casos son miembros del hogar en vacaciones, choferes de buses o vendedores viajeros, entre otros.

21 Las viviendas colectivas corresponden a hoteles, residenciales, hospitales, cuarteles y cárceles, entre otros. Para más detalles respecto a la implementación operativa de este concepto, revisar el manual de trabajo de campo, tomo I de la encuesta (INE Chile, 2018b).

22 Estas categorías también incluyen a pensionistas y servicio doméstico y parientes de este, que no son considerados como integrantes del hogar.

23 Los que no son mutuamente excluyentes entre sí.

Por su parte, la VIII EPF determinó como jefe de hogar a **quien sea reconocido como tal por los miembros del hogar, sin necesidad de cumplir con otra característica**, por lo que puede o no coincidir con quien ejerce las funciones de sustentador principal²⁴ o de administrador de gastos del hogar²⁵. En la encuesta, esta figura se utiliza para organizar el registro del resto de los integrantes del hogar, ordenando las relaciones de parentesco de sus miembros en relación a la jefatura reportada.

5.1.2 Criterios de captura del gasto

La VIII EPF es una encuesta económica enmarcada dentro de los lineamientos conceptuales planteados por el SCN para estandarizar la medición de la actividad económica. Para lograr esto, el SCN recopila información de cinco unidades institucionales (mutuamente excluyentes entre sí), correspondientes a hogares, sociedades no financieras, sociedades financieras, unidades de Gobierno e instituciones sin fines de lucro que sirven a los hogares (ISFLSH). Los conceptos utilizados en esta metodología se sustentan en el **sector hogares**.

Un tema fundamental en la economía de los países corresponde a las decisiones respecto al modo en que los hogares distribuyen su ingreso entre consumo y ahorro²⁶. Por lo tanto, es necesario identificar el gasto en bienes y servicios de consumo al que pueden acceder con ese ingreso (ONU *et al.*, 2016). En este sentido, existen dos definiciones relevantes asociadas al consumo de los hogares:

- i. **Consumo final:** es el monto del gasto en bienes y servicios de consumo.
- ii. **Consumo final efectivo:** corresponde al gasto en consumo final de los hogares, más el valor de las transferencias sociales²⁷ a disposición del hogar.

Según las definiciones del SCN, un bien o servicio de consumo corresponde a aquellos utilizados por los hogares para la satisfacción directa de sus necesidades, en donde el gasto se define como *el monto que los compradores pagan, o conviene pagar, a los vendedores a cambio de bienes o servicios que estos suministran a los compradores* (ONU *et al.*, 2016).

Siguiendo las recomendaciones internacionales, la VIII EPF utilizó el concepto de **gasto en consumo final** de los hogares, ya que contabiliza los gastos realizados a través de transacciones en el mercado.

De acuerdo con la definición de gasto en consumo final de los hogares (ver diagrama 1), se excluyen los gastos en activos fijos en forma de viviendas o de objetos valiosos²⁸, así como también los gastos en consumo intermedio. Estos últimos corresponden a bienes

24 Sustentador o proveedor principal es quien más aporta al presupuesto del hogar, considerando los últimos seis meses.

25 Administrador/es de gastos del hogar, corresponde a quien/es regularmente realiza/n las compras del hogar, considerando los últimos seis meses.

26 Para efectos de la VIII versión de la encuesta, nos enfocaremos en el concepto de consumo final del hogar.

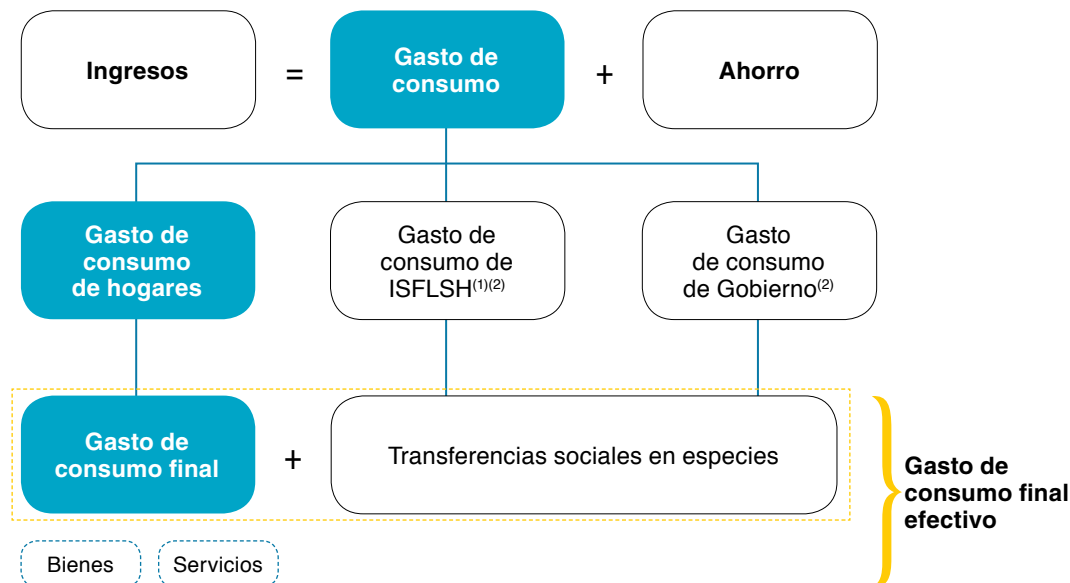
27 Realizadas por unidades gubernamentales e instituciones sin fines de lucro que sirven a los hogares, las que se definen como una transacción mediante la cual una unidad institucional suministra a otra un bien, un servicio o un activo sin recibir a cambio ningún bien, servicio o activo como contrapartida (ONU *et al.*, 2016).

28 La adquisición de una vivienda corresponde a la formación bruta de capital fijo, en tanto los objetos valiosos conservan o aumentan su valor y no son utilizados de la misma manera que los bienes de consumo (ONU *et al.*, 2016).

y servicios que los hogares adquieren para el desarrollo de sus actividades productivas. Solo aquellos bienes y servicios adquiridos para actividades no productivas forman parte de su gasto de consumo final.

El grupo más importante de actividades no productivas desarrolladas por los hogares, y para las cuales adquieren bienes y servicios en el mercado, corresponde al desarrollo de tareas para los mismos hogares, tales como limpieza de la vivienda, preparación de alimentos o cuidado de menores, entre otros.

Diagrama 1: Esquema de gasto de consumo (utilización de los ingresos)



(1) ISFLSH: Instituciones sin Fines de Lucro que sirven a los hogares.

(2) El esquema considera el gasto de consumo que realizan las ISFLSH y Gobierno destinado a los hogares.

Fuente: Elaboración propia en base a Sistema de Cuentas Nacionales 2008 (2016) y Eurostat (2003).

Para el correcto registro del gasto en consumo final de bienes y servicios, es importante considerar dos aristas fundamentales. La primera consiste en el **momento en que se registra el gasto** y la segunda se relaciona **con el tipo de bien o servicio que se está adquiriendo**.

Respecto a la primera arista, existen tres criterios para el registro de los gastos de consumo, es decir, para los gastos en bienes y servicios adquiridos, pagados o utilizados por los hogares para la satisfacción de sus necesidades. En la 17° CIET, estas consideraciones respecto al momento de registrar el gasto en bienes o servicios se sintetizan de la siguiente forma (OIT, 2003c):

- **Criterio de adquisición del gasto:** si se adquieren durante el período de referencia, independientemente de la fecha en que se utilizaron o pagaron en su totalidad.
- **Criterio de uso del bien o servicio:** si se utilizaron durante el período de referencia, independientemente de la fecha en que se adquirieron o pagaron en su totalidad.
- **Criterio de pago del bien o servicio:** si se pagaron durante el período de referencia, independientemente de la fecha en que se adquirieron o utilizaron.

Cabe señalar que los hogares adquieren un bien cuando se convierten en propietarios de estos, es decir, la adquisición de bienes se produce al mismo tiempo que se contrae una obligación de pago con el vendedor, independientemente de si se realiza o no la transacción monetaria en el mismo momento²⁹. Por su parte, la utilización del bien o servicio está relacionada con el consumo o uso de este en un período de referencia³⁰. Por último, el pago corresponde a la transacción efectiva que el hogar se compromete a realizar por un bien o servicio. Es decir, adquirir o pagar un bien o servicio no implica necesariamente consumirlo de forma inmediata.

Para la medición del gasto, la VIII EPF utilizó principalmente el criterio de **gasto de consumo adquirido**³¹, un concepto que resulta de mejor comprensión para los informantes, y que además es recomendado y usado internacionalmente (OIT, 2003c; Eurostat, 2003; OCDE, 2013; ONU *et al.*, 2016). Por lo tanto, el momento de registro del gasto en la VIII EPF corresponde al momento en que el hogar adquiere el derecho de propiedad sobre un bien o el derecho a la prestación de un servicio, independientemente de si el consumo o pago de ese bien se efectúa durante el período de referencia de la encuesta.

Un ejemplo del criterio de gasto en consumo adquirido son las compras realizadas a crédito. En estos casos se debe registrar el valor total del bien o servicio, tal como si se hubiera comprado al contado, ya que al momento de adquirir el bien se adquiere el compromiso legal de adquisición. En otras palabras, el gasto en consumo adquirido corresponde al valor precio contado del gasto que se realiza en bienes y servicios, sin considerar si se han comprado en cuotas. Ello es así debido a que, por lo general, la información sobre intereses y otros tipos de cargos asociados al pago a crédito son desconocidos por el informante³².

No obstante, para la medición de aquellos gastos que los hogares realizan con frecuencia periódica (mensual, bimestral, trimestral, semestral o anual), se utiliza el concepto de **gasto de consumo pagado**, debido a que existen gastos cuyo pago es posterior a la entrega del servicio, tales como el consumo de electricidad y agua del hogar. Además, y como criterio excepcional, se utiliza el criterio de **uso de los gastos** para registrar los bienes y servicios asociados a la realización de viajes³³.

La segunda arista del gasto en consumo final de bienes y servicios se relaciona con el tipo de consumo. Esta definición se fundamenta en la durabilidad física del bien, es decir, si se puede utilizar solamente una vez para fines de consumo o si puede emplearse de forma repetida durante un plazo determinado (ONU *et al.*, 2016). Según la OIT (2003c), los tipos de bienes³⁴ utilizados se pueden clasificar en:

29 En compras a plazo, el comprador ejerce un completo control sobre el bien, aun cuando el pago no se haya realizado en su totalidad.

30 Por ejemplo, la compra de un pasaje en bus por anticipado, se adquiere el derecho de propiedad de ese servicio, pero su consumo o uso será en el futuro.

31 OIT (2003b), Párrafo 142, página 36: "Para los IPC utilizados como índice para el seguimiento de la inflación o como índice de compensación y para las cuentas nacionales, probablemente sea más indicado el método basado en las adquisiciones".

32 Técnicamente, esos gastos deben ser declarados en la Libreta de Gastos en Servicios Financieros y de Seguros (LSF), la cual se describirá en el apartado 5.3.5.

33 Esto se implementa de esta forma considerando las condiciones especiales asociadas a los gastos relacionados a viajes de los hogares en los que este criterio de medición del gasto facilita el registro de estos.

34 Se considera (convencionalmente) que todos los servicios se consumen inmediatamente en el momento de la prestación de este (OIT, 2003c).

- **Bienes durables/semidurables:** son aquellos que pueden utilizarse numerosas veces en un período superior a un año, sin reducir su capacidad de satisfacer necesidades al hogar.
- **Bienes no durables:** son los que se consumen total e inmediatamente (y desaparecen) en el plazo de un año.

En el marco de la VIII EPF, se especifican los criterios de decisión para incluir o no un gasto en consumo final de los hogares en la encuesta. Cabe destacar que algunos de los gastos mencionados son excluidos de la definición por razones operativas³⁵, aunque estén comprendidos en sus definiciones conceptuales.

5.1.2.1 Criterios de decisión del gasto en consumo final

Adicionalmente, se describen algunas reglas de decisión para gastos específicos, con el objetivo de disminuir el número de casos que necesitan arbitraje por parte del encuestador y así lograr estimaciones lo menos sesgadas posible. Los criterios son los siguientes:

- En el caso de que una vivienda cambie de contrato de algún servicio que paga de forma periódica (electricidad, agua, entre otros), dicho gasto será incluido o excluido de las estimaciones dependiendo si durante el período de referencia está consumiendo o no dicho servicio.
- Para un servicio que esencialmente se presta en forma continua, donde puede existir la opción de pago mensual, trimestral o anual (primas de seguro, mensualidades universitarias entre otros), el gasto se registra en el momento del pago, buscando capturar, para los casos en los que sea posible, información referida a la periodicidad³⁶.
- Si el valor de un servicio no se puede determinar hasta concluir la prestación, el gasto se registra en el momento en que servicio fue cobrado (por ejemplo, el consumo de luz, el gasto se realiza después de su consumo).

En los siguientes apartados se revisan gastos considerados como parte del consumo final de los hogares, gastos que se excluyen de las estimaciones por definiciones técnicas u operativas y, por último, el registro de gastos que se consideran casos especiales.

5.1.2.2 Gastos considerados como parte del consumo final de los hogares

Existen algunas definiciones técnicas que deben ser adaptadas según criterios operativos. Conforme a ello, en la VIII EPF, los gastos efectuados por los hogares considerados como parte del consumo final corresponden a:

- Las contrataciones de servicios y compras de bienes durables y semidurables, que se recogen mediante el criterio de las adquisiciones (gasto en consumo adquirido), y que se registran cuando el hogar adquiere un compromiso de pago y el derecho de propiedad. Estos bienes y servicios deben ser de consumo final y

35 Los gastos excluidos serán revisados en el apartado 5.1.2.4 Casos especiales.

36 En los formularios de gasto se registra la periodicidad de pago de ciertos bienes y servicios con el fin de realizar el ajuste temporal de los gastos de la mejor forma posible. Más información respecto al ajuste temporal aplicado a los gastos en el apartado 9.4.4.2.

no para actividades productivas. Se incluyen en los bienes durables automóviles, refrigeradores y televisores, entre otros, considerando que son adquiridos en el momento en que son puestos a disposición del hogar, independiente de su vida útil.

- Los pagos realizados a la administración pública por servicios prestados, tales como pagos por licencias de conducir, certificados y pasaportes, entre otros.
- Los impuestos incluidos en el precio final de los bienes y servicios.
- El arriendo pagado por los hogares.
- El gasto en servicio doméstico, en cuyo caso el hogar cumple el rol de empleador del personal doméstico (asalariado).
- Los materiales y mano de obra para pequeñas reparaciones en viviendas. Aquellos gastos que son destinados a grandes reparaciones no se incluyen como gastos de consumo de los hogares, ya que son considerados como inversión. Las reparaciones y mantenimiento de los bienes durables también son considerados como gastos de consumo final³⁷.
- Las inscripciones a clubes o sociedades, debido a que generan un derecho a utilizar sus instalaciones.
- Los regalos comprados por el hogar, con el objetivo de ser donados o llevados como presente a otros hogares.
- Finalmente, se incluye como gasto en consumo final ciertas imputaciones, tales como: 1) el valor de los bienes procedentes del auto suministro, 2) el valor de los bienes y servicios recibidos como retribución por el trabajo realizado por miembros del hogar (pago en especies) y 3) los arriendos imputados a los propietarios de viviendas o a los que las utilizan gratuitamente.

5.1.2.3 Gastos no considerados como parte del consumo final de los hogares

Los registros que no se consideran como gasto en consumo final del hogar y que, por ende, son excluidos de las estimaciones de gasto de la encuesta, corresponden a los siguientes casos:

- Inversiones realizadas por los hogares, como compras de viviendas y grandes reparaciones de viviendas³⁸. También se excluyen la adquisición de objetos valiosos³⁹ y de activos financieros (bonos, acciones, etc.).
- Valorización de regalos recibidos por otro hogar.
- Gastos en primas de seguros de vida, por considerarse inversión.

37 Las grandes reparaciones se consideran grandes mejoras, tales como reconstrucciones, renovaciones o ampliaciones. Estos gastos no se incluyen en el gasto en consumo final, debido a que corresponde a la formación bruta de capital fijo.

38 Las grandes reparaciones de viviendas en arriendo, cuando son realizadas por los propietarios, son consideradas como un bien de consumo intermedio.

39 Los bienes durables de alto valor monetario que no se deterioran con el paso del tiempo se excluyen del IPC, porque no se utilizan con fines de consumo y se adquieren fundamentalmente como depósito de valor. Son principalmente obras de arte, piedras y metales preciosos y las joyas elaboradas en estos últimos materiales.

- Algunos impuestos y tasas pagadas a la administración pública, tales como pagos por licencias de propiedad o uso de vehículos, embarcaciones o aeronaves, así como las licencias de caza, tiro al blanco o pesca, consideradas como impuestos.
- Intereses y cargos administrativos asociados al retraso de un pago, ya sea por compras realizadas a crédito o por el pago atrasado de las cuentas del hogar (moras).
- Servicios provistos por el aparato estatal: administración pública, defensa, justicia y ordenamiento jurídico, entre otros.
- Servicios domésticos y personales producidos y consumidos dentro del mismo hogar o entre hogares, en forma no remunerada. Sin embargo, los bienes y servicios adquiridos por los hogares para la realización de las actividades que desarrolla el servicio doméstico sí se incluyen dentro de su gasto de consumo final del hogar.
- Bienes de consumo intermedio de los hogares. Tal como se mencionó anteriormente, estos bienes se utilizan para actividades productivas, por lo tanto, no son considerados en el gasto en consumo final del hogar.

5.1.2.4 Casos Especiales

En el siguiente apartado, se presentan los criterios para el registro de gastos especiales.

5.1.2.4.1 Gastos en servicio doméstico

La remuneración al servicio doméstico puede incluir pagos en dinero y/o en especies. Estos últimos normalmente comprenden alimentación y alojamiento del empleado⁴⁰. La inclusión de alimentos y vivienda pagados al servicio doméstico implicaría una reclasificación de estos gastos del hogar según la finalidad. Por convención, los servicios domésticos se valoran solamente por la remuneración en su condición de asalariados que se paga en dinero y/o en especies, distintos a la alimentación y vivienda (ONU *et al.*, 2016).

5.1.2.4.2 Gastos en primas de seguros

Estos se clasifican en dos tipos: seguros de vida y otros tipos de seguros (seguros relacionados con la vivienda, la salud, el transporte y otros riesgos, como seguros frente a terceros o el seguro de accidentes personales).

Los organismos internacionales recomiendan la exclusión de los seguros de vida de las EPF, pues son considerados una inversión. Consecuentemente, en la VIII EPF, los gastos declarados por el hogar relacionados con el pago de primas de seguros de vida fueron excluidos de la estimación del gasto del hogar. En ese sentido, también se excluyeron las primas o seguros de ahorro, los que se vinculan directamente a desembolsos por concepto de seguros de vida.

5.1.2.4.3 Donaciones recibidas en especie

Las transferencias sociales provenientes desde el sector institucional gubernamental o de las instituciones sin fines de lucro que sirven a los hogares son de difícil medición, debido a que los hogares desconocen los montos asociados a dichas transferencias. Por lo tanto, en

⁴⁰ Los que en algunas ocasiones puede incluir el alojamiento de su propia familia.

la práctica, se considera que el gasto en consumo final es realizado por quien adquiere los bienes para donarlos⁴¹. Los hogares reciben estos bienes y servicios como transferencias sociales, las que permiten medir el consumo efectivo del hogar, tal como se revisó en el diagrama 1.

En el caso de los regalos entre hogares, si respetáramos el criterio de gasto en consumo adquirido, el hogar que recibe el regalo (bien) debería registrar el gasto, debido a que tiene el derecho de propiedad sobre este. No obstante, por razones operativas se adopta el criterio de que el gasto lo registra el hogar que lo realiza⁴² (OIT, 2003c).

5.1.2.4.4 Gastos de los hogares propietarios de empresas no constituidas como sociedad

Los gastos que los hogares realizan con fines empresariales o productivos se excluyen del gasto de consumo de los hogares, ya que se considera consumo intermedio (ONU *et al.*, 2016). Teóricamente esta distinción no resulta difícil, sin embargo, operativamente presenta diversas dificultades cuando el mismo bien o servicio puede utilizarse con fines empresariales o de consumo final en el hogar.

Los gastos con fines productivos no pueden identificarse simplemente sobre la base del tipo de bien o servicio comprado. Por lo tanto, el énfasis se realiza en la captura de la información, capacitando al encuestador para que en casos ambiguos solicite al informante la desagregación del gasto realizado con fines productivos y del realizado como parte del consumo final del hogar, siendo solo este último el que se registra en la encuesta.

5.1.3 Criterio de captura de los ingresos

Los ingresos de un hogar representan los recursos disponibles que poseen para el consumo y el ahorro en un período determinado de tiempo. La medición del ingreso es relevante, porque permite aproximarse al bienestar económico de un hogar, siendo considerado uno de los indicadores más importantes y accesibles del bienestar, representado a través de los recursos económicos disponibles para satisfacer sus necesidades (OIT, 2003c). Adicionalmente, la información de ingresos proporcionada por el hogar permite dar cumplimiento al objetivo secundario de la encuesta, que consiste en identificar la estructura del ingreso total disponible de los hogares.

La definición del ingreso que percibe el hogar dependerá de las partidas que se consideren en su construcción. En primer lugar, el ingreso total del hogar se define conceptualmente como todos los ingresos, monetarios o en especie (bienes y servicios) que son percibidos por el hogar o sus integrantes en un período determinado (OCDE, 2013). Es decir, tal como refleja el diagrama 2, el ingreso total de un hogar corresponde a la suma de los ingresos del trabajo⁴³, ingresos de la propiedad, ingresos derivados de la autoprovisión de bienes producidos por el hogar⁴⁴ y las transferencias corrientes recibidas.

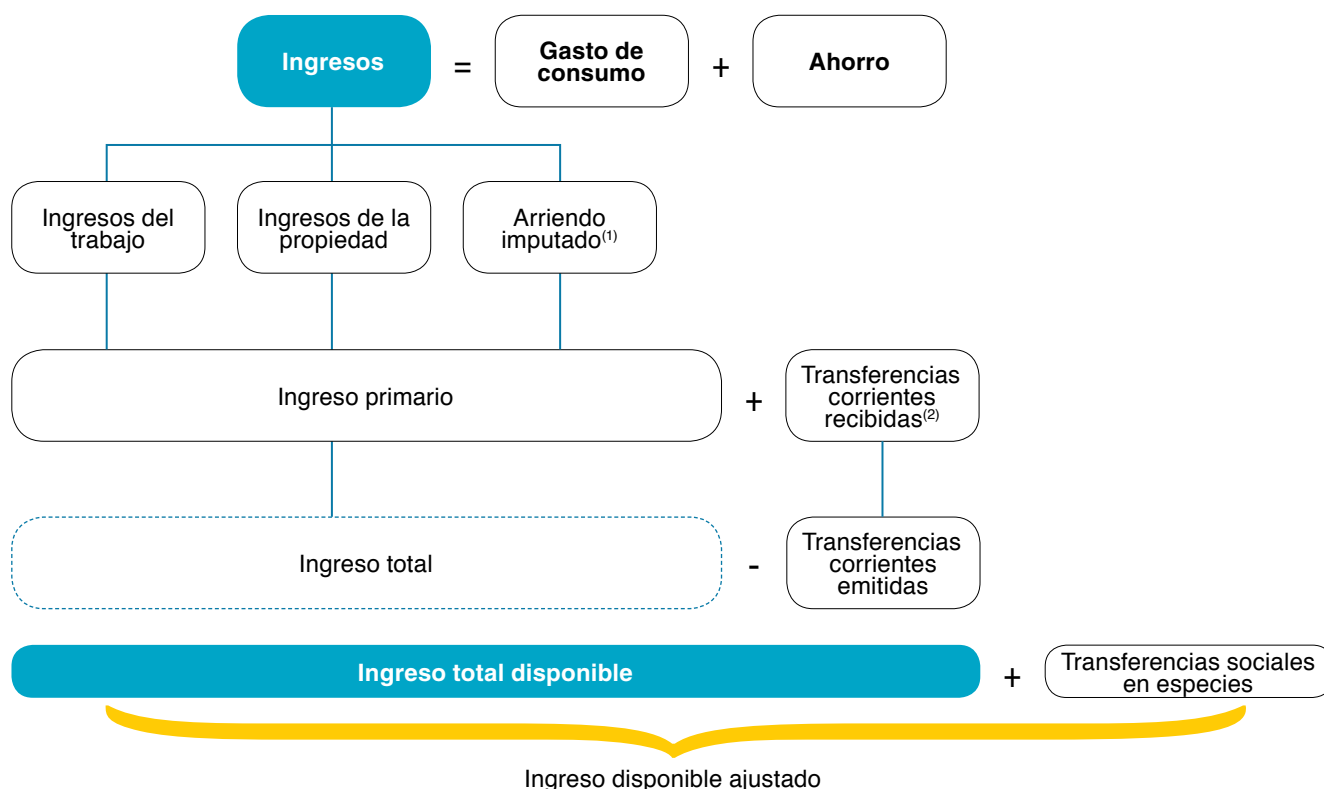
41 En el caso de las transferencias, serían el Gobierno y las instituciones sin fines de lucro que sirven a los hogares los que deberían registrar los bienes donados como gasto en consumo final.

42 Esto se realiza así, porque el hogar que realiza el gasto conoce en detalle el costo asociado a este.

43 Se realiza la distinción de los trabajadores dependientes (ingreso bruto) e independientes (ingreso neto = ingreso bruto – costos operacionales).

44 Corresponde al valor neto de los bienes y servicios producidos para el trueque o para el propio consumo, tales como la agricultura de subsistencia. Además, considera el valor del arriendo imputado de los hogares y la estimación del valor de la mano de obra del servicio doméstico.

Diagrama 2: Esquema del ingreso total de los hogares



(1) El esquema de Cuentas Nacionales para los ingresos incorporan en esta partida los ingresos provenientes de los servicios domésticos no remunerados y del consumo de bienes durables los cuales son excluidos de la medición de esta encuesta considerándose solo el arriendo imputado.

(2) En el caso chileno, las jubilaciones poseen componentes asociados a la seguridad social, componentes asociados a capitalización individual o una combinación de ambos componentes. En la VIII EPF, no es posible diferenciar entre estos componentes, por lo que se incluyen todos en las transferencias corrientes recibidas.

Fuente: Elaboración propia en base a Eurostat (2003) y OECD (2013).

El ingreso corresponde a una variable de flujo que impacta en la variación de la riqueza o patrimonio neto de los hogares, que contribuye al gasto en consumo final. El SCN define el ingreso disponible como “el monto máximo que un hogar puede permitirse gastar en bienes o servicios de consumo durante el período, sin reducir su patrimonio neto real, es decir, financiar los gastos sin disminuir sus activos financieros o no financieros, ni aumentando su deuda” ⁴⁵ (ONU *et al.*, 2016, pág. 185).

La Encuesta Suplementaria de Ingresos (ESI) del INE busca medir el ingreso disponible que perciben las personas y los hogares residentes del país por medio del ingreso corriente, es decir, los ingresos nominales en el período de referencia de la encuesta. Su objetivo es caracterizar los ingresos laborales de las personas que son clasificadas como ocupadas y los ingresos de los hogares provenientes de otras fuentes, tanto a nivel nacional como regional. Dentro de los ingresos de la ocupación se incluyen los de las personas ocupadas dependientes, los de las personas ocupadas independientes y aquellos de otros trabajos. Respecto a los ingresos de otras fuentes, se incorporan los de la propiedad, por arriendo

⁴⁵ El período de referencia del Sistema de Cuentas Nacionales es el período contable, que es al cual se refieren las Cuentas Nacionales. En la encuesta, se utiliza como período de referencia para la captura de ingresos el mes anterior, ya que, en general, dicho lapso refleja los pagos de sueldos y salarios, cuentas del hogar y créditos, entre otros aspectos.

imputado, por ahorro u otro tipo de activos financieros, además de los ingresos por transferencias⁴⁶ (INE Chile, 2017).

La Encuesta de Caracterización Socioeconómica Nacional (Casen), del Ministerio de Desarrollo Social, utiliza el mismo marco conceptual que la EPF y la ESI, a saber, el SCN. De acuerdo con sus objetivos, esta busca disponer de información que permita medir las diferentes categorías de ingreso que reciben las personas y los hogares, con el fin de caracterizarlos socioeconómicamente y estimar pobreza (Observatorio Social, 2015)⁴⁷.

En este contexto y según los objetivos de la encuesta, siguiendo las recomendaciones internacionales respecto a este tipo de estudios (OCDE, 2013; OIT, 2003c; ONU *et al.*, 2016), la VIII EPF midió el ingreso disponible del hogar, correspondiente a la suma de todos los ingresos primarios y las transferencias corrientes recibidas⁴⁸ menos las contribuciones sociales y las transferencias corrientes emitidas.

El ingreso disponible del hogar (diagrama 2) se compone de los ingresos que se describen a continuación:

1. **Ingresos del trabajo:** se construyen en base a los ingresos generados por una actividad económica laboral. Estos incluyen:
 - a. **Ingresos del trabajo asalariado:** asociados al trabajo dependiente, incluyen los sueldos y salarios, comisiones, propinas, bonificaciones o bonos en efectivos, indemnización por despido y contribuciones de seguridad social.
 - b. **Ingresos del trabajo independiente:** son los que reciben las personas que trabajan como empleadores o cuenta propia y corresponden a la ganancia de la unidad económica menos los costos de los insumos.
 - c. **Otros ingresos del trabajo:** captura los ingresos asociados a segundas actividades económicas, así como los ingresos de personas que se declararon desocupadas o inactivas durante la semana de referencia, pero tuvieron ingresos del trabajo durante el mes anterior⁴⁹.
2. **Ingresos de la propiedad:** corresponden al flujo de ingresos netos que se obtienen de tenencia de activos financieros. Incluye el retorno de dichos activos (interés ganado de los depósitos), recibos de ingresos por inversiones, anualidades y otros

46 Las transferencias incluidas en esa encuesta incluyen jubilaciones/pensiones, seguro de desempleo, subsidios del Estado, transferencias entre hogares, ingresos del exterior y otras transferencias monetarias recibidas.

47 En el caso de la encuesta Casen, se utiliza el concepto de **ingreso disponible** para la medición de la pobreza oficial, el cual se construye restando al ingreso total (ingreso primario o autónomo + transferencias corrientes percibidas) los impuestos directos, las contribuciones a la seguridad social y las transferencias corrientes a otros hogares (CEPAL, 2015).

48 Las transferencias se definen como la entrega de recursos (monetarios o en especies) que no implican un intercambio. Las transferencias corrientes corresponden a transferencias monetarias (bienes o servicios), las que no incluyen las transferencias sociales en especies.

49 En la VIII EPF, la clasificación de Otros Ingresos del Trabajo considera a las personas que cambiaron de trabajo durante el mes de referencia, es decir, declaran un trabajo en el Registro de Personas en el Hogar que es distinto al declarado en la Libreta de Ingresos. Además, incluye a las personas que se declararon desocupadas o inactivas durante la **semana** de referencia al momento de aplicar la encuesta y que, sin embargo, recibieron ingresos asociados al trabajo durante el **mes** de referencia.

pagos regulares de seguros de vida y fondos privados de pensiones (excluidos de la seguridad social), además de las rentas netas recibidas por el arriendo de activos no financieros, tales como recursos naturales, tierras, edificios, equipos, patentes o derechos de autor.

3. Ingresos por arriendo imputado: corresponden al valor estimado de la prestación del servicio de alojamiento en las viviendas ocupadas por sus propietarios⁵⁰.

4. Transferencias corrientes recibidas: son flujos monetarios recibidos por los hogares, que no generan relaciones de reciprocidad respecto a quienes las emiten. Estas incluyen los ingresos asociados a las jubilaciones⁵¹, pensiones, otros beneficios en efectivo asociados a la seguridad social o al empleo, beneficios de asistencia social en efectivo del Gobierno, transferencias desde otros hogares (dentro y fuera del país) y transferencias en efectivo recibidas desde instituciones sin fines de lucro que sirven a los hogares. Este tipo de transferencias excluye las transferencias sociales en especies que los Gobiernos y las instituciones sin fines de lucro que sirven a los hogares emiten a estos.

5. Transferencias corrientes emitidas: son flujos monetarios emitidos por los hogares que incluyen impuesto a la renta, contribuciones a la seguridad social, pago por pensiones alimenticias, transferencias a hogares (dentro y fuera del país) y transferencias a instituciones sin fines de lucro que sirven a los hogares.

Los ingresos del trabajo (1), de la propiedad (2) y por arriendo imputado (3) corresponden al ingreso primario del hogar. En tanto, el ingreso total del hogar se construye con la suma del ingreso primario más las transferencias corrientes recibidas.

Finalmente, la construcción del ingreso disponible del hogar corresponde al ingreso total (definido previamente) menos los impuestos, los descuentos legales (cotizaciones previsionales y de salud) y las transferencias corrientes emitidas.

El ingreso disponible ajustado del hogar se construye sumando el ingreso disponible del hogar y las transferencias sociales en especie. Estas últimas corresponden a las transferencias emitidas por el Gobierno y las instituciones sin fines de lucro que sirven a los hogares en forma gratuita o a precios menores a los de mercado (por ejemplo, copago en educación) tales como alimentos, viviendas y servicios de salud, enseñanza, cultura y recreación (ONU, 2001). En la VIII EPF, se miden algunas de las transferencias en especies mencionadas, sin embargo, no es posible construir el componente en su totalidad.

Una vez definido el concepto de ingreso a medir, la VIII EPF, estableció que el criterio de captura del ingreso fuese el de **ingreso devengado**, lo que significa que se registran los

50 El ingreso por arriendo imputado corresponde a un concepto que mide el valor de la prestación del servicio de alojamiento en las viviendas ocupadas por sus propietarios (arriendo imputado). El arriendo imputado es parte de un concepto más amplio conocido como **ingresos de autoprovisión de servicios producidos por el hogar** el cual además, incluye el valor de los servicios domésticos no remunerados y el valor de los servicios aportados por el consumo de bienes durables. No obstante, los ingresos provenientes de los servicios domésticos no remunerados y del consumo de bienes durables son excluidos de esta encuesta y han sido tradicionalmente excluidos de la medición de encuestas y de Cuentas Nacionales (ONU, 2001).

51 Las jubilaciones poseen componentes asociados a la seguridad social y a componentes asociados a capitalización individual (cuentas de ahorro individuales) o una combinación de ambos componentes. En la VIII EPF, no es posible diferenciar entre estos componentes, por lo que se incluyen todos en las transferencias corrientes recibidas.

ingresos generados por la actividad económica en el período de referencia (mes anterior), sin importar si estos realmente fueron recibidos durante ese período. En otras palabras, corresponde a los ingresos que el informante recibió o debió haber recibido durante el mes anterior.

En el diagrama 2 se aprecia también la partida del **ingreso disponible ajustado**, correspondiente al ingreso disponible, definido anteriormente, más las transferencias sociales en especie realizadas por el Gobierno y las instituciones sin fines de lucro que sirven a los hogares. Estos componentes corresponden a transferencias no monetarias que percibe un hogar, por ejemplo, la educación gratuita o subvencionada o los servicios de atención médica financiados con fondos públicos.

El concepto de ingreso disponible ajustado corresponde a un indicador más cercano al concepto de bienestar de los hogares (OCDE, 2013) y se utiliza como contrapartida al concepto de gasto de consumo final medido por la encuesta.

Si bien la VIII EPF, al igual que la VII EPF, posee una pregunta respecto a transferencias sociales en especie recibidas por los hogares, no realiza una medición detallada de estas⁵², debido a que su objetivo secundario corresponde a identificar la estructura del **ingreso total disponible**. Si el objetivo de la encuesta migrara a la medición del bienestar económico de los hogares a través del ingreso disponible ajustado, también se debería modificar la forma de medir el consumo, considerando el gasto en consumo final efectivo (diagrama 2), el cual incluye las transferencias sociales recibidas por los hogares (ONU *et al.*, 2016).

5.1.4 Clasificadores utilizados

El equipo de codificación EPF es el encargado de asignar códigos a los registros recopilados en los distintos cuestionarios de la encuesta. Específicamente, se asignaron códigos a los registros de bienes y servicios, ocupaciones y establecimientos registrados en los cuestionarios aplicados en las viviendas seleccionadas en el estudio.

Para esta tarea, el equipo de codificación utilizó cuatro clasificadores⁵³:

- **Clasificación Internacional Uniforme de Ocupaciones (CIUO):** se empleó a un nivel de dos dígitos, los que describen los subgrupos principales. Se utiliza el clasificador de 1988 y una adaptación nacional de su versión 2008.
- **Clasificación del Consumo Individual por Finalidades (CCIF):** se empleó para clasificar funcionalmente el SCN y así estandarizar los bienes y servicios capturados por la encuesta. Se utiliza este clasificador hasta el quinto nivel de desagregación, que describe el producto⁵⁴. Los primeros tres niveles del clasificador (que describen las clases de producto) se adoptaron según el clasificador internacional, mientras que los siguientes dos niveles (que respectivamente describen las subclases y los

52 La VIII EPF valoriza las transferencias sociales recibidas en el módulo de transferencias recibidas de la Libreta de Ingresos pero no de forma exhaustiva.

53 Estos clasificadores se encuentran desarrolladas de forma más extensa en el apartado 9.3.4 de Clasificadores utilizados en el subproceso de codificación.

54 La clasificación CCIF cuenta con cinco niveles de desagregación: división (primer nivel), grupo (segundo nivel), clase (tercer nivel), subclase (cuarto nivel) y producto (quinto nivel).

productos) fueron adaptados a la realidad nacional. La adaptación fue realizada por el INE con el apoyo de instituciones externas⁵⁵.

- **Clasificador de Establecimientos INE:** fue empleado para estandarizar los establecimientos que proveen bienes y servicios. Cuenta con 17 categorías y es utilizado como fuente de información para la recolección de precios por parte de los equipos de IPC.
- **Clasificadores de Establecimientos Educativos:** el uso dado a este clasificador fue de estandarizar los establecimientos de educación básica, media y superior. Corresponde al conjunto de listados de establecimientos reconocidos por el Ministerio de Educación (Mineduc). Para los establecimientos de nivel de enseñanza básica y media se utiliza el Rol de Base de Datos (RBD) del Ministerio de Educación.

5.2 Prueba piloto de la VIII Encuesta de Presupuestos Familiares

Antes de la realización de la VIII versión de la EPF, se propusieron algunos cambios a los instrumentos, además de la posibilidad de ajustar la metodología de trabajo de campo. Estos ajustes fueron probados a través de una prueba piloto, realizada en el segundo semestre del año 2015.

5.2.1 Objetivos de la prueba piloto

El objetivo general de la prueba piloto fue conocer el efecto de los ajustes a los cuestionarios y metodología de trabajo de campo de la VII EPF en la captura y procesamiento de información de esta encuesta.

Por su parte, los objetivos específicos se definieron en los siguientes términos:

- Medir y comparar el efecto de los ajustes en las variables objetivo de la encuesta.
- Indagar en la comprensión del encuestador y los informantes respecto de los conceptos utilizados en cada cuestionario.
- Examinar la naturaleza de las fuentes de error observadas en las preguntas de cada cuestionario.
- Medir y comparar el efecto de los ajustes a los cuestionarios y metodología de trabajo de campo en la duración de las entrevistas de cada instrumento.
- Medir el efecto de los ajustes a los cuestionarios y la metodología de trabajo de campo en las tasas de respuesta total.

La prueba piloto constituyó un proceso fundamental para la decisión de los cuestionarios y metodología de trabajo de campo que se implementarían en la VIII EPF, a partir de la comparación de los resultados de tres estrategias metodológicas alternativas, entre las cuales también se incluyeron los cuestionarios y metodología de trabajo de campo de la VII

⁵⁵ La adaptación de la CCIF ha contado con el apoyo de Cepal, del Ministerio de Desarrollo Social y del Ministerio de Salud.

EPF. Este análisis permitió contar con instrumentos óptimos para la captura de información requerida por la encuesta, a la vez que permitió ajustar el diseño de distintas fases del proceso de producción estadística.

5.2.2 Diseño muestral de la prueba piloto

La prueba piloto contó con un diseño experimental asociado a una muestra dividida (*split panel test*⁵⁶). En este tipo de diseño se comparan los resultados de versiones diferentes de cuestionarios, aplicados a muestras comparables de la población en un mismo tiempo. Según U.S. Census Bureau (2003), el método de muestras divididas resulta suficiente para la evaluación de encuestas que aportan datos sensibles para la política pública.

En la literatura, los grupos de tratamiento corresponden al subconjunto de unidades estadísticas, en este caso hogares de la población objetivo, que se ven expuestos al tratamiento definido por el diseño experimental –cambios propuestos a los instrumentos y a la metodología de trabajo de campo, en el contexto de la prueba–. La comparación con un grupo de control (donde se mantiene la antigua versión del cuestionario y no se ajusta la metodología de campo), permite cuantificar el efecto de las modificaciones a los instrumentos y a la metodología de trabajo de campo, en cuanto a la validez y fiabilidad de los resultados de la encuesta⁵⁷. La denominación que se utilizó para cada uno de estos grupos en la prueba piloto se detalla en la tabla 2.

Tabla 2: Grupos de tratamiento y control Prueba Piloto

Forma	Versión cuestionarios EPF	Versión metodológica trabajo de campo
A	VII EPF	15 días sin ajustes
B	Cuestionarios EPF modificados	15 días con ajustes
C	VII EPF	7 días con ajustes

Fuente: Instituto Nacional de Estadísticas (INE), Prueba piloto VIII Encuesta de Presupuestos Familiares (EPF).

El diseño muestral de la prueba piloto no buscó tener representatividad estadística en los mismos niveles de representatividad que la VII EPF. Más bien, solo se advirtió la necesidad de contar con desagregación a nivel de estratos socioeconómicos⁵⁸, considerando las diferencias en las tasas de no respuesta detectadas en este nivel en la VII EPF⁵⁹. Debido a factores presupuestarios, el trabajo de campo de la prueba se realizó únicamente en la Región Metropolitana. Las comunas seleccionadas fueron Santiago, Macul y Ñuñoa, desde donde se obtuvieron las manzanas pertenecientes a los estratos socioeconómicos bajo, medio y alto, respectivamente. Además de dicho criterio, su selección respondió a que en aquellas comunas se contaba con una enumeración de viviendas reciente y suficiente para seleccionar las unidades muestrales de la prueba.

56 Para más información respecto a este tipo de diseños experimentales, revisar U.S. Census Bureau (2013).

57 El grupo de control representa un contrafactual que busca dar respuesta a la interrogante: ¿qué hubiese pasado si no se hubiese incorporado el o los tratamientos?

58 Las manzanas del Marco Muestral de hogares son estratificadas socioeconómicamente según información proveniente del Censo de Población y Vivienda del año 2002 a través del método Princals. Para mayor información respecto a este método, consultar Guerrero (2003a; 2003b).

59 Parte de esta revisión está disponible en INE Chile (2014a).

Tal como se observa en la tabla 2, el grupo de control correspondió a la forma A, donde se aplicaron los mismos cuestionarios y metodología de trabajo de campo de la VII EPF, con un período de colaboración de una quincena (período para la recolección de gastos individuales). En la forma B se introdujeron cambios a los cuestionarios, manteniendo el tiempo de colaboración quincenal empleado en la VII EPF, aunque con ajustes al esquema de visitas de aquella versión⁶⁰. A diferencia de lo anterior, en la forma C se mantuvieron los cuestionarios de la VII EPF, introduciéndose ajustes a la metodología de trabajo de campo, reduciéndose el período de colaboración de los hogares de quince a siete días (u ocho, dependiendo de la semana) y ajustándose la aplicación de los cuestionarios⁶¹.

Se planificaron tres meses de trabajo de campo (junio, julio y agosto de 2015) para el operativo de la prueba piloto, durante los cuales se buscó encuestar a los hogares residentes en las 603 viviendas que componían la muestra total de la prueba piloto. Las viviendas se distribuyeron en los tres meses de trabajo, considerando la existencia de un equilibrio entre estratos socioeconómicos en cada mes y la optimización del traslado de los encuestadores. Los resultados de la asignación final de las viviendas a las distintas formas de la prueba piloto se presentan a continuación.

Tabla 3: Asignación de prueba piloto según submuestra operativa

Tipo de metodología	N = 603	Junio 2015				Julio 2015				Agosto 2015			
Forma A		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
CSE bajo	65	11	12	12	11	10	9						
CSE medio	68	12	11	11	11	11	12						
CSE alto	68	10	12	11	12	12	11						
Total	201	33	35	34	34	33	32						
Forma B		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
CSE bajo	65	12	11	11	11	9	11						
CSE medio	68	11	11	11	12	12	11						
CSE alto	68	11	12	12	11	11	11						
Total	201	34	34	34	34	32	33						
Forma C		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
CSE bajo	65	4	6	5	7	6	5	6	6	4	6	6	4
CSE medio	68	6	5	6	6	6	6	6	5	6	5	5	6
CSE alto	68	6	6	6	5	5	6	5	6	6	5	6	6
Total	201	16	17	17	18	17	17	17	17	16	16	17	16

Fuente: Instituto Nacional de Estadísticas (INE), Prueba Piloto VIII EPF.

60 Los ajustes al esquema de visitas consisten en 1) la realización de la entrevista de la Libreta de Gastos del Recuerdo solo en la segunda visita (antes se entrevistaba en dos oportunidades), 2) la aplicación de la Libreta de Gastos en Servicios Financieros y Seguros en la tercera visita (antes se aplicaba en la cuarta visita) y 3) el retiro de la Libreta de Gastos Individuales solo desde el día siguiente al término del período de colaboración (antes se realizaba desde el último día de la quincena).

61 Dado que el período de recolección de la encuesta en esta forma de la prueba piloto se redujo a siete días, se ajustó el régimen de visitas para aplicar el resto de los cuestionarios en un menor número de visitas.

5.2.3 Resultados principales de la prueba piloto

Se revisarán a continuación los criterios que se utilizaron para comparar las distintas formas (A, B y C) de la prueba, seguido por las principales conclusiones y recomendaciones surgidas de dicha comparación.

5.2.3.1 Especificación de criterios de comparación entre formas

Para definir los ajustes a incorporar al cuestionario de la VIII EPF, se tomó información de tres fuentes cuantitativas en la comparación de las distintas formas (A, B y C) de la prueba.

Uno de los componentes del análisis fue la comparación de los principales indicadores de desempeño operativo, que dan cuenta de la efectividad del trabajo de campo. Los indicadores seleccionados fueron las tasas de contacto, cooperación, rechazo y respuesta. La comparación de estas tasas se realizó mediante pruebas z de proporciones, bajo supuesto de normalidad. Con ellas se determinó la existencia de diferencias estadísticamente significativas entre las tasas obtenidas para cada forma.

La reducción del período de colaboración de quince a siete días concentró buena parte de los análisis de la prueba piloto, ya que se evaluó la ventaja potencial asociada a la posibilidad de aumentar las tasas de respuesta de la encuesta, disminuir el riesgo de subdeclaración en la Libreta de Gastos Individuales y aminorar el agotamiento del informante durante este período. Tras la aplicación de la prueba, los resultados sugieren una disminución en la tasa de rechazo y un aumento en la tasa de respuesta (tabla 4). No obstante, solo la diferencia en la tasa de rechazo fue significativa, observándose que los niveles de rechazo en la forma A fueron mayores en 7,89 pp a los presentados en la Forma B (el valor p es de 0,04).

Tabla 4: Indicadores de desempeño operativo entre formas⁽¹⁾

Forma	Tasa de contacto	Tasa de cooperación	Tasa de rechazo	Tasa de respuesta	Desactualización del marco
Forma A (15d)	91,37%	61,11%	30,46%	55,84%	6,19%
Forma C (7d)	89,23%	66,67%	22,56%	59,49%	6,25%
Valor p (H_0 : iguales)	0,4753	0,2781 ($<$) 0,1390	0,0771 ($>$) 0,0386	0,4659	0,98

(1) Los valores p reportados son de dos colas, en la segunda línea se copiaron los valores p de una cola para los casos más bajos.

Fuente: Instituto Nacional de Estadísticas (INE), Prueba piloto VIII Encuesta de Presupuestos Familiares (EPF).

Una segunda fuente analizada fueron los errores detectados en cada forma durante la aplicación de la encuesta. La caracterización de estos errores resultaba importante, pues era parte del registro que se tenía sobre el control de la calidad de los datos capturados. Cabe señalar que para la revisión se hizo mayor énfasis en las libretas cuya aplicación fue modificada (Libreta de Gastos Individuales, Libreta de Gastos del Recuerdo y Libreta de Gastos en Servicios Financieros y Seguros)⁶².

Para el análisis de esta fuente no se realizaron pruebas de significancia y, en su defecto, se realizó un análisis descriptivo, comparando la distribución de errores según libretas,

62 La explicación respecto a los instrumentos de captura de información se realiza en el apartado 5.3, donde son descritos con mayor detalle.

módulos y tipos de errores. Un resumen de la distribución de errores de revisión de libretas está disponible en la tabla 5.

Tabla 5: Cantidad de libretas revisadas y porcentaje de libretas con al menos un error, por forma

Libretas	Forma piloto							
	A		B		C		Total	
	N°	% con error	N°	% con error	N°	% con error	N°	% con error
Hoja de Ruta	108	39,80	101	42,60	109	53,20	318	45,30
Registro de Personas en el Hogar	110	73,60	102	65,70	115	72,20	327	70,60
Libreta de Gastos Individuales	109	61,47	102	52,94	114	60,53	325	58,46
Libreta de Gastos del Hogar	109	66,10	102	71,60	115	73,00	326	70,20
Libreta de Gastos del Recuerdo	107	58,90	103	80,60	113	64,60	323	67,80
Libreta de Ingresos	108	43,50	102	54,90	113	64,60	323	54,50
Libreta de Gastos en Servicios Financieros y Seguros	108	58,30	101	32,70	113	62,80	322	51,90

Fuente: Instituto Nacional de Estadísticas (INE), Prueba piloto VIII Encuesta de Presupuestos Familiares (EPF).

Además de la información presentada en la tabla 5, se realizaron análisis de errores detectados en revisión según libretas. En general, se observó que la forma C elevaba la cantidad de errores registrados en los instrumentos.

Finalmente, se consideró relevante para la prueba piloto el análisis del comportamiento de las variables objetivo de la encuesta, el gasto y los ingresos, frente a los cambios propuestos: los cambios planteados fueron evaluados considerando el gasto total promedio y el ingreso disponible promedio para las distintas formas, a fin de establecer si existían diferencias de medias entre ellas. A nivel del gasto total promedio, las comparaciones según libreta de procedencia y forma se detallan en la tabla 6.

Tabla 6: Pruebas de medias sobre el gasto total por libreta de procedencia y forma

Libretas	Efecto forma B	Valor P	Efecto forma C	Valor P
	(Diff=A-B)	(H0: Diff=0)	(Diff=A-C)	(H0: Diff=0)
Libreta de Gastos Individuales	95.311	0,129	75.671	0,222
Libreta de Gastos del Hogar	-127.168	0,063 ⁽²⁾	42.629	0,416
Libreta de Gastos del Recuerdo	-156.935	0,039 ⁽¹⁾	32.038	0,611
Libreta de Ingresos	19.176	0,312	-21.343	0,514
Libreta de Gastos en Servicios Financieros y Seguros	-16.375	0,354	19.476	0,274

(1) Significancia al 5%.

(2) Significancia al 10%.

Fuente: Instituto Nacional de Estadísticas (INE), Prueba piloto VIII Encuesta de Presupuestos Familiares (EPF).

Al realizar análisis de diferencias de gasto por libreta entre formas, se evidenció un aumento significativo del gasto registrado en la forma B en el caso de las libretas de Gasto del Recuerdo y del Hogar. Dichos cambios eran esperables, ya que aquellas sufrieron los mayores cambios, derivados de la incorporación de nuevos módulos que buscaban mejorar la captura de gastos.

Además, debido a que algunos cambios fueron introducidos con el objetivo de mejorar la captura de ciertos ítems⁶³, también se realizaron pruebas a nivel de divisiones⁶⁴, para determinar si aquel objetivo era logrado con los ajustes. La tabla 7 presenta los resultados de estas pruebas.

Tabla 7: Pruebas de medias sobre el gasto total por división y forma (con arriendo imputado)

Divisiones CCIF	Efecto forma B (Diff=A-B)	Valor P (H0: Diff=0)	Efecto forma C (Diff=A-C)	Valor P (H0: Diff=0)
1	29.905	0,074 ⁽¹⁾	7.678	0,66
2	-1.094	0,896	-10.143	0,312
3	-15.545	0,603	-2.190	0,939
4	-56.742	0,079 ⁽¹⁾	1.380	0,962
5	-14.867	0,439	9.351	0,572
6	-15.784	0,611	39.752	0,028 ⁽¹⁾
7	-74.235	0,267	47.826	0,381
8	8.138	0,294	-1.075	0,911
9	-8.218	0,786	-5.514	0,858
10	-35.777	0,416	42.613	0,247
11	-27.911	0,324	15.379	0,457
12	-5.570	0,823	45.036	0,047

(1) Significancia al 10%.

Fuente: Instituto Nacional de Estadísticas (INE), Prueba piloto VIII Encuesta de Presupuestos Familiares (EPF).

Respecto a los análisis realizados de los ingresos disponibles de los hogares, en términos globales, no se observaron diferencias estadísticas significativas entre las formas tal como se aprecia en la tabla 8.

Tabla 8: Pruebas de medias sobre el ingreso disponible total por hogar

Efecto forma B (Diff=A-B)	Valor P (H0: Diff=0)	Efecto forma C (Diff=A-C)	Valor P (H0: Diff=0)
-72.523,66	0,7585	264.282,50	0,2147

Fuente: Instituto Nacional de Estadísticas (INE), Prueba piloto VIII Encuesta de Presupuestos Familiares (EPF).

63 Considerando comparaciones con registros externos, como Cuentas Nacionales, se identificó que algunos ítems, como algunos gastos asociados a combustible o vestuario, no quedaban bien registrados en la encuesta. Al respecto, se incorporaron ajustes en la prueba piloto, para buscar capturar de mejor forma estos ítems.

64 Las divisiones corresponden al primer nivel de desagregación de la Clasificación del Consumo Individual Según Finalidades (CCIF). Para mayor detalle consultar el apartado 5.1.4.

Se debe considerar que uno de los principales ajustes que se evaluaron en esta prueba fue la reducción del período de autorreporte de gastos diarios, de quince días a una semana de registro. Este cambio no solo afecta la duración del período de colaboración en un hogar, sino que también el régimen de visitas del resto de los cuestionarios, en una encuesta que contempla en su metodología más de un contacto con las unidades seleccionadas.

De este modo, la evaluación de los ajustes no solo se enfocó en conocer el desempeño de la encuesta en los hogares en particular, sino que también en observar la interacción entre el diseño de los cuestionarios, el régimen de visitas y la ejecución y organización de las distintas fases contempladas en el trabajo de campo.

Con el fin de complementar los análisis cuantitativos de la prueba, también se efectuaron análisis cualitativos. Para ello, se realizaron grupos focales con los encuestadores luego de finalizar el trabajo de campo de la prueba. Los grupos focales se orientaron principalmente a la evaluación de los encuestadores respecto de dos grandes áreas: 1) las modificaciones a la metodología de trabajo de campo y 2) el diseño de los cuestionarios en sus diferentes formas. Se buscó que la información reportada por los encuestadores permitiera clarificar y/o confirmar algunos resultados de la encuesta y, al mismo tiempo, mejorar tanto el diseño de cuestionarios como la metodología utilizada para el trabajo de campo.

Respecto al primer punto, los grupos focales reforzaron la elección de la metodología de recopilación de información durante quince días, ya que, de acuerdo al relato de los encuestadores, en la metodología de siete días se genera un mayor agotamiento por parte de los encuestadores y encuestados, al tener solo dos días entre visitas, lo que reducía el tiempo disponible para la revisión de las encuestas aplicadas. Por otra parte, respecto al diseño de instrumentos, durante los grupos focales en general la forma B del cuestionario es mejor evaluada por los encuestadores en comparación con la forma A, lo que refuerza también la conveniencia de las modificaciones a los instrumentos probadas en la prueba piloto.

5.2.3.2 Principales conclusiones y recomendaciones

La principal recomendación emanada de los resultados de la prueba piloto de la VIII EPF fue mantener la recolección del gasto en la modalidad de una quincena de colaboración, considerando principalmente el resguardo respecto a la obtención del gasto (variable objetivo), la mejor calidad del dato obtenido (basado en la menor cantidad de errores de registro) y la mejor valoración de los encuestadores en la evaluación cualitativa (debido a mayor disponibilidad de tiempo para revisar encuestas y generar vínculos con el hogar en un período quincenal).

Al mismo tiempo, se recomendó implementar los cambios adicionales propuestos a la metodología de trabajo de campo, aplicados en la forma B, tales como el retiro de libretas de registro diario un día después de la fase de aplicación de encuestas, la reducción a una sola entrevista asociada a los gastos de recuerdo y el adelantamiento de la entrevista asociada a la captura de los gastos en servicios financieros a la tercera visita de aplicación⁶⁵. Este conjunto de modificaciones se sustenta en base a un conjunto de beneficios identificados, tales como la mejora en la obtención del gasto (variable objetivo),

⁶⁵ Durante la VII EPF, los gastos del día a día se retiraban hasta el último día de aplicación de la encuesta y la entrevista asociada a los gastos en servicios financieros se realizaba la última visita en el hogar. Para más información consultar respecto a la Metodología de la VII EPF (INE Chile, 2013).

la mantención de la duración global de la aplicación de la encuesta en los hogares (la suma del tiempo entre todas las visitas) y la disposición de mayor tiempo de recuperación para los gastos asociados a servicios financieros, que son datos delicados y sensibles.

Respecto a los cambios propuestos para los cuestionarios, la recomendación principal fue aceptar las modificaciones relativas a flujos y formas de preguntar, incluso sobre temáticas nuevas. Los cambios se fundaron en hallazgos tales como la mejora en la obtención del gasto (variable objetivo), la mayor calidad del dato obtenido (basado en el número de errores a depurar por forma, que mostró ser mayor en la forma A que en la forma B), la ausencia de diferencias significativas en la duración promedio de las entrevistas a los hogares entre distintos tipos de cuestionarios y una mejor valoración de los encuestadores respecto de los cuestionarios modificados, desde una mirada cualitativa.

Sobre las recomendaciones referidas al aspecto y diagramación de los cuestionarios, se planteó una modificación en las portadas de las libretas, alternando colores para resaltar información relevante y utilizando logos institucionales y del producto para identificar la encuesta. Adicionalmente, se propuso utilizar una pestaña en la primera página de la libreta de captura de información sociodemográfica del hogar, que permitiera registrar por orden el nombre de los miembros del hogar, facilitando el registro en los módulos localizados en las siguientes páginas de ese cuestionario. Del mismo modo, se sugirió una redistribución de preguntas en algunas páginas, con el objetivo de evitar que existieran celdas de respuesta demasiado pequeñas en los cuestionarios.

Finalmente, cabe mencionar que en la VII EPF, la información contenida en las libretas de información era codificada y luego digitada en el sistema informático de la encuesta. Para la realización de la prueba piloto de la VIII EPF, se invirtió este proceso. Si bien no era un objetivo de la prueba, se comprobó que primero digitar y luego codificar ayudaba a mejorar el rendimiento y calidad del proceso de codificación, acortando sus tiempos, a la vez que reducía parte de los errores no muestrales asociados a la digitación de los datos. Teniendo en cuenta estos resultados, aquel cambio en el flujo de producción de la encuesta también fue incorporado en la VIII EPF.

5.3 Instrumentos de captura de información

Siguiendo los lineamientos conceptuales revisados en el apartado 5.1, a continuación se describen los seis cuestionarios o libretas utilizadas en la VIII EPF para capturar la información de los hogares. Las libretas se pueden clasificar según las siguientes tres temáticas: 1) una libreta que consulta sobre características sociodemográficas del hogar, 2) cuatro libretas que capturan gasto y 3) una que registra los ingresos del hogar⁶⁶. Las libretas que tienen el objetivo de medir el gasto se diferencian entre sí, fundamentalmente por el tipo de bienes y servicios que incluye cada una y por el período de referencia respecto al que el gasto es consultado.

En los apartados siguientes se revisan los objetivos y características de cada libreta, junto con los períodos de referencia en los que se captura la información.

66 Cabe señalar que la Libreta de Ingresos también registra una porción pequeña del gasto de los hogares. No aparece mencionada dentro de los instrumentos que miden gasto debido a que su objetivo fundamental es medir ingresos y no gasto.

5.3.1 Registro de Personas en el Hogar (RPH)

El primer instrumento de captura de información es el Registro de Personas en el Hogar (RPH). Esta libreta busca caracterizar socio-demográficamente a los miembros de los hogares urbanos, residentes en las capitales regionales y algunas de sus zonas conurbadas. Estas características permiten no solo analizar los patrones de gasto de distintos tipos de hogares, sino que también posibilitan los ejercicios de imputación y de cálculo de factores de expansión.

La información de cada hogar se captura a través de los módulos que se describen a continuación, estableciendo enlaces con el resto de los instrumentos:

- **Características de los miembros del hogar:** captura información demográfica de los integrantes del hogar y establece los lazos de parentesco que estos tienen con quien ejerce el rol de jefe de hogar. Una de las características importantes a capturar en este módulo es la edad de los distintos miembros del hogar, ya que esta información permite definir quién debe responder la Libreta de Gastos Individuales (5.3.2), instrumento base para la captura de gastos.
- **Educación:** además de capturar información básica sobre curso y nivel más alto alcanzado por cada miembro del hogar, recoge información sobre la existencia de estudios actuales, formales o complementarios. Asimismo, para quienes estudian en educación formal, caracteriza la dependencia del establecimiento. Esta información permite la validación y depuración de los gastos por concepto de educación, registrados en la Libreta de Gastos del Hogar (5.3.3).
- **Condición de actividad económica:** registra información respecto a la situación laboral de una persona, dentro del marco de medición de la fuerza de trabajo⁶⁷, para los miembros del hogar de 15 años y más. Adicionalmente, caracteriza la ocupación principal de las personas que se declaran ocupadas e identifica los tipos de ingreso que perciben las personas de 15 años o más. La caracterización de la actividad económica permite determinar qué módulos de la Libreta de Ingresos (5.3.6) se deben consultar para cada integrante del hogar, disminuyendo el número de preguntas a realizar y la probabilidad de rechazo de la Libreta de Ingresos.
- **Salud y previsión:** recopila información sobre la afiliación y cotización de los miembros del hogar de 18 años y más a sistemas de salud y previsión, identificando el sistema al que pertenecen. En el caso de cotizar a un sistema previsional, la caracterización de este permite la estimación del descuento obligatorio a los asalariados. Contar con esta información permite la construcción del ingreso disponible del hogar.
- **Vivienda principal⁶⁸:** captura información sobre la tenencia de la vivienda principal e identifica los gastos asociados al uso de esta. Además, registra información necesaria para el cálculo del arriendo imputado⁶⁹.

67 La secuencia de preguntas se realiza siguiendo los criterios de la OIT, implementados también en la actual Encuesta Nacional de Empleo (ENE) del INE.

68 Se refiere a la vivienda utilizada como residencia habitual, es decir, aquella donde la persona habita o duerme al menos cuatro días a la semana, durante los últimos seis meses.

69 El arriendo imputado corresponde a los beneficios del hogar por habitar en la vivienda (propia, arrendada a un valor inferior al de mercado, cedida, gratuitamente u otras condiciones). Como los beneficios no son monetarios, se realiza una estimación, por parte del propietario, arrendatario o beneficiario, para acercarnos a las ganancias que ha percibido el hogar por el hecho de habitar la vivienda.

- **Vivienda secundaria:** indaga sobre la tenencia de vivienda secundaria y los gastos asociados a esta. La tenencia de vivienda secundaria se define por cuatro características: 1) es una vivienda propia (pagada, pagándose, herencia o sucesión compartida), 2) el hogar la tuvo a disposición el mes anterior, 3) el hogar la tuvo a disposición al menos tres meses durante el último año y 4) su uso principal es la recreación o el descanso. Esta información se utiliza para determinar si debe aplicarse o no una serie de preguntas de la Libreta de Gastos del Hogar, asociadas a los gastos generados por dicha vivienda⁷⁰.

5.3.2 Libreta de Gastos Individuales (LGI)

La Libreta de Gastos Individuales (LGI) captura el gasto en bienes y servicios que realiza cada integrante del hogar de 15 años o más. En este instrumento se registran los gastos personales realizados con frecuencia diaria. Estos deben registrarse el mismo día en el cual se efectúan, debido a que la capacidad de recuerdo del informante para este tipo de gastos es baja⁷¹.

Este es el único instrumento autoaplicado de la encuesta y se utiliza por los informantes como un cuadernillo en el que día a día registran los gastos realizados durante la quincena de aplicación de la VIII EPF. Los gastos deben ser registrados sin importar si estos se realizan para beneficio propio, de otro miembro del hogar o como regalo a otro hogar⁷².

Otro objetivo de esta libreta es obtener información sobre el autosuministro de los hogares que posean algún tipo de negocio perteneciente a los integrantes del hogar. De esta forma, si un bien fue adquirido a través del retiro de productos desde un negocio, este se debe registrar en la LGI y el monto debe ser contabilizado al valor venta o precio de mercado.

5.2.3 Libreta de Gastos del Hogar (LGH)

En la Libreta de Gastos del Hogar (LGH) se registra el gasto en bienes y servicios considerando dos objetivos. El primero consiste en pesquisar todos los gastos que realiza el hogar en consumo final y que presentan algún tipo de regularidad (mensual, trimestral, anual, etc.). El ejemplo más claro son los gastos en servicios básicos de una vivienda, tales como agua y electricidad, entre otros. El segundo objetivo es recopilar todos los gastos relacionados con la educación de los integrantes del hogar que se encuentran estudiando actualmente⁷³, tanto en educación formal como complementaria.

Para la captura de los gastos periódicos se utiliza el criterio de consumo pagado, revisado en el punto 5.1.2, debido a que se consulta por lo que el hogar pagó o debería haber pagado por el consumo del bien y servicio durante el mes anterior (período de referencia). De este modo, no se contabiliza como gasto el cobro de consumos anteriores a dicho período o los gastos por intereses en mora asociados a esa deuda.

70 Para definir la vivienda secundaria, se establecen todas estas condiciones, ya que se busca diferenciarla de viviendas que el hogar arrienda con el fin de generar ingreso y en dichos casos los gastos asociados a esa vivienda corresponderían a gastos de consumo intermedio.

71 Este registro es apoyado por los encuestadores, quienes realizan seguimiento de las anotaciones realizadas y recuperan en base al recuerdo los gastos de aquellos días que no presentan registro.

72 En el siguiente link: <https://youtu.be/S0PyTW7SlrA> se puede consultar un video generado para difusión de la encuesta donde se explica a los informantes la forma de llenar esta libreta.

73 Información declarada previamente en RPH.

La LGH se divide en los siguientes módulos:

- **Gastos periódicos relativos a la vivienda principal:** registra los gastos periódicos relacionados con el funcionamiento de la vivienda, que normalmente presentan una frecuencia mensual de pago. Debido a lo anterior, se consulta al hogar por el valor del gasto asociado al consumo del mes anterior a la aplicación de la encuesta.
- **Gastos periódicos relativos a la vivienda secundaria:** captura los gastos periódicos relativos a la/s vivienda/s secundaria/s de aquellos hogares que declaren poseerla/s.
- **Gastos periódicos en telecomunicaciones del hogar:** captura los gastos periódicos relativos al consumo de servicios de telecomunicaciones, tales como televisión, telefonía e internet, entre otros.
- **Gastos periódicos en aseo municipal:** captura los gastos periódicos relativos al pago de aseo municipal.
- **Gastos relativos al uso de vehículos con fines domésticos:** captura gastos periódicos vinculados al uso de vehículos, tales como combustible, estacionamiento y peaje, entre otros.
- **Gastos en servicio doméstico contratados por el hogar:** captura gastos periódicos vinculados a la contratación de servicios domésticos, tales como limpieza de la vivienda, cuidado de niños y mantenimiento de jardín, entre otros.
- **Otros gastos periódicos del hogar:** recopila información de otros gastos pagados de manera periódica por el hogar, tales como suscripciones a revistas y gimnasio, entre otros.
- **Gastos de educación formal:** permite registrar gastos de los integrantes del hogar que asisten a educación formal⁷⁴ dentro de un año académico, los que se relacionan con el pago de matrícula, mensualidades, cuota del centro de padres, cuota de incorporación, exámenes de admisión u otros derechos.
- **Gasto en otros cursos y talleres:** este módulo tiene por objetivo registrar dos tipos de gasto: el relacionado con educación complementaria a la formal (por ejemplo, cursos de capacitación laboral y clases particulares) y el que corresponde a gastos en cursos, talleres o clases contratadas con fines recreativos.
- **Otros gastos relativos a educación:** captura gastos relacionados con la educación, que no hayan sido descritos con anterioridad en la libreta.

5.3.4 Libreta de Gastos del Recuerdo (LGR)

La Libreta Gastos del Recuerdo (LGR) captura aquellos gastos en bienes y servicios que no son frecuentes en el hogar, pero que pueden representar un monto importante desde el punto de vista del presupuesto del hogar. Los gastos en bienes y servicios se capturan en tres períodos acotados de recuerdo: desde el momento de la entrevista hasta 3, 6 y 12 meses previos al mes en que se aplica la encuesta en el hogar⁷⁵.

74 Este módulo debe ser respondido por los miembros del hogar que asisten a establecimientos de educación formal, reconocidos y regulados por el Ministerio de Educación, integrantes que son identificados previamente en RPH.

75 Cabe aclarar que los períodos de recuerdo más amplios contienen a los más acotados.

Para cada período de recuerdo se capturan gastos en productos específicos, los que son definidos en base a la capacidad de recuerdo del hogar. En primera instancia, los períodos establecidos en el cuestionario se corroboraron o modificaron de acuerdo con los resultados de las pruebas del cuestionario, el análisis de encuestas internacionales y a través del estudio de la frecuencia de compra en base a los datos de la VII EPF.

Los bienes y servicios que se consultan al hogar se agrupan en tablas de productos afines, los que para cada período corresponden a:

Período de recuerdo 3 meses.

- Tabla 1: Mantenimiento y reparación del hogar (servicios y materiales).
- Tabla 2: Consultas médicas, cirugías ambulatorias, dentistas y otros profesionales de la salud.
- Tabla 3: Exámenes.
- Tabla 4: Artículos de corrección médica.
- Tabla 5: Libros y espectáculos.
- Tabla 6: Vestuario.
- Tabla 7: Calzados y zapatillas.
- Tabla 8: Repuestos, accesorios y mantenimiento para vehículos.
- Tabla 9: Artículos de telefonía y fotografía.
- Tabla 10: Maletas, bolsos, carteras, joyas y relojes.
- Tabla 11: Equipos de deporte, campamento e instrumentos musicales.
- Tabla 12: Artefactos eléctricos pequeños para el hogar y herramientas (manuales y eléctricas).

Período de recuerdo 6 meses.

- Tabla 13: Muebles.
- Tabla 14: Computación.
- Tabla 15: Artefactos electrónicos.
- Tabla 16: Viajes realizados con fines vacacionales o de descanso.
- Tabla 17: Viajes comprados (aún no realizados)⁷⁶.
- Tabla 18: Grandes artefactos para el hogar (eléctricos o no).

Período de recuerdo 12 meses

- Tabla 19: Vehículos nuevos.
- Tabla 20: Vehículos usados.
- Tabla 21: Hospitalizaciones.
- Tabla 22: Otros gastos en período de recuerdo no considerados con anterioridad⁷⁷.

76 Si bien se consulta por los gastos realizados en viajes aún no realizados, según lo explicado en el apartado 5.1.2, estos son registrados como gastos según el criterio de uso (tabla 16), por lo que los gastos consultados en la tabla 17 se utilizan como control y no son imputados al gasto de los hogares.

77 Se registran otros gastos de gran incidencia en el presupuesto del hogar y con baja frecuencia de compra, tales como gastos en compra de leña, gastos en fiestas y ceremonias como matrimonios, bautizos, etc.

5.3.5 Libreta de Gastos en Servicios Financieros y Seguros (LSF)

La Libreta de Gastos en Servicios Financieros y Seguros (LSF) busca capturar el gasto asociado al uso de instrumentos financieros. Los gastos financieros corresponden a intereses, gastos en administración, comisiones y seguros asociados.

La libreta se divide en los siguientes módulos:

- Créditos hipotecarios.
- Tarjetas de crédito de establecimientos comerciales.
- Crédito de consumo y crédito automotriz.
- Cuentas corrientes y líneas de crédito.
- Cuentas vistas o cuentas RUT.
- Tarjetas de crédito bancarias.
- Créditos universitarios.
- Seguros.

Cada módulo, excepto el de cuenta vista o cuenta RUT y el de seguros, se consulta en dos grupos de preguntas. El primer grupo (zona blanca del módulo) busca caracterizar el crédito en cuanto al monto solicitado, plazo y valor de la cuota. El segundo grupo (zona ennegrecida del módulo) busca capturar, por medio de comprobantes o cupones de pago, los cobros por servicios financieros, es decir, intereses, cobros de administración, comisiones y seguros asociados. Cabe señalar que el segundo grupo de preguntas solo se completa si el informante pone a disposición del encuestador los comprobantes de pago.

El objetivo del segundo grupo de preguntas es contar con información detallada del crédito, lo cual permite calcular la Carga Anual Equivalente (CAE) del crédito solicitado, la que se desagrega para obtener los distintos componentes del gasto financiero.

5.3.6 Libreta de Ingresos (LI)

La Libreta de Ingresos (LI) responde al objetivo secundario de la EPF, es decir, permite identificar la estructura del ingreso total disponible de los hogares urbanos del país.

En la LI se registra información respecto a los ingresos monetarios y no monetarios de todos los integrantes del hogar de 15 años o más⁷⁸. El período de referencia para la LI corresponde al mes anterior al de la entrevista y se registran los ingresos generados en dicho período, sin importar si estos realmente se recibieron durante ese mes (criterio de ingreso devengado).

Los ingresos se capturan en los siguientes módulos:

- **Ingresos del trabajo asalariado:** captura los ingresos provenientes del trabajo dependiente⁷⁹. Estos datos incluyen información desde el salario bruto hasta los descuentos legales o voluntarios. Adicionalmente, se capturan remuneraciones en especies y/o servicios que puedan ser recibidas.

78 No se capturan ingresos derivados del trabajo infantil.

79 Entendido como aquel que realiza sus actividades laborales bajo subordinación y dependencia, es decir, tiene jefe, cumple un horario, tiene continuidad en el trabajo, periodicidad de la remuneración, entre otros.

- **Ingresos del trabajo independiente:** recopila información de los ingresos generados a partir de negocios por cuenta propia o por profesionales independientes que dirigen y organizan su propia actividad.
- **Otros ingresos del trabajo:** captura ingresos derivados por dos tipos de fuentes: ingresos provenientes de segundas actividades e ingresos derivados de una actividad laboral, realizada el mes anterior, distinta a la actividad principal declarada en el RPH para la semana de referencia, utilizada para determinar la fuerza de trabajo.
- **Ingresos por jubilaciones y/o pensiones de vejez:** registra ingresos de jubilaciones y/o pensiones de vejez provenientes de ahorros previsionales y, en algunas situaciones, aportes fiscales (pilar básico solidario y pensiones provenientes de cajas de ahorro fiscales, entre otras).
- **Ingresos por otras pensiones y transferencias recibidas:** captura ingresos que no son generados por una actividad económica. Incluye pensiones alimenticias, pensiones de invalidez o discapacidad, montepíos, pensiones asistenciales, entre otras. Además, registra transferencias (monetarias y en especies) recibidas desde otros hogares, dentro o fuera del país.
- **Otros ingresos de carácter no habitual:** recopila información sobre ingresos que son significativos en el presupuesto del hogar, pero que no son ni periódicos ni habituales, tales como herencias, loterías e indemnizaciones, entre otros. El objetivo es identificar una fuente de ingresos no corriente, que explique una estructura de gasto anómala durante el período de referencia. Estos ingresos no son considerados para el cálculo del ingreso disponible del hogar.
- **Ingresos de la propiedad y ganancias por tenencia de instrumentos financieros:** captura ingresos derivados de la tenencia de propiedades (muebles o inmuebles) y de instrumentos financieros, como acciones y cuentas de ahorro, entre otros.
- **Transferencias emitidas:** registra información sobre el total de transferencias periódicas a otros hogares e instituciones sin fines de lucro, en dinero y/o en especies.

6. DISEÑO ESTADÍSTICO DE LA ENCUESTA

El diseño estadístico de la encuesta se describe a lo largo de este capítulo. En primer lugar, se presentan los elementos que conforman el diseño muestral, que dan origen a la muestra de viviendas seleccionadas para la VIII EPF. Luego, se aborda el desarrollo del cálculo del factor de expansión, revisando cada una de sus etapas de ajustes. A continuación, se presentan los estimadores y medidas de precisión de la encuesta. Finalmente, se expone la metodología de estimación de los indicadores principales de ingreso y gasto obtenidos de la encuesta.

La VIII EPF, al igual que en su versión previa, cuenta con representatividad estadística para las áreas Gran Santiago, Resto de capitales regionales y Total de capitales regionales. Entre las mejoras metodológicas-estadísticas realizadas a la VIII EPF, tal como se ha comentado anteriormente, se destaca el aumento en la cobertura, incorporando a la muestra las comunas de Chillán y Chillán Viejo, ya que ambas constituirán el área urbana de la capital de la Región de Ñuble.

6.1 Diseño muestral

6.1.1 Características generales del diseño muestral

El diseño muestral de la encuesta corresponde a un muestreo estratificado por comuna, clasificación socioeconómica y clasificación de las manzanas según grupo de tamaño (considerando el total de viviendas que las componen).

Dentro de cada estrato, la selección de la muestra es aleatoria y se realiza en dos etapas:

- Selección de unidades de primera etapa (o unidades primarias de muestreo), que corresponden a manzanas, definidas por delimitaciones geográficas fijas.
- Selección de unidades de segunda etapa (o unidades secundarias de muestreo), que corresponden a viviendas particulares ocupadas, desde las cuales se obtiene la información respecto a los hogares (unidad de análisis) que habitan en dichas viviendas⁸⁰.

6.1.1.1 Objetivo diseño muestral

El diseño muestral de la VIII EPF busca determinar y distribuir el número de viviendas, que permita recopilar información de los hogares, para cada nivel de estimación, con el objeto de obtener resultados fidedignos y consistentes con los estándares nacionales e internacionales.

El diseño muestral de la VIII EPF determina una muestra representativa a nivel de Capitales regionales del país y para las zonas de Gran Santiago y Resto de capitales regionales, las que en conjunto concentran el 72,11% de la población de viviendas del área urbana del país⁸¹.

80 El apartado 5.3. da cuenta de los instrumentos utilizados para la recolección de información de los hogares que son encuestados.

81 Esta información proviene del Marco Muestral de Manzanas del INE actualizado al año 2015 según se explica en el apartado 6.1.2.

6.1.1.2 Población objetivo

La población objetivo la constituyen los hogares y las personas que residen habitualmente en las viviendas particulares ocupadas de las capitales regionales del país (incluyendo sus zonas conurbadas más importantes). Se excluyen del estudio las viviendas con más de cinco hogares junto con las personas que las residen⁸².

6.1.1.3 Cobertura geográfica y niveles de estimación

La cobertura geográfica está conformada por Gran Santiago y Resto de capitales regionales del país, incluyendo sus zonas conurbadas más importantes. A continuación, la tabla 9 presenta información respecto a las comunas presentes en la encuesta en cada región del país.

Tabla 9: Cobertura geográfica de la encuesta según capital regional o zona conurbada

Región	Capital Regional o Zona Conurbada
Arica y Parinacota	Arica
Tarapacá	Gran Iquique: Iquique, Alto Hospicio
Antofagasta	Antofagasta
Atacama	Copiapó
Coquimbo	Gran La Serena: La Serena, Coquimbo
Valparaíso	Gran Valparaíso: Valparaíso, Viña del Mar, Quilpué, Concón, Villa Alemana
Metropolitana	Gran Santiago: Cerrillos, Cerro Navia, Conchalí, El Bosque, Estación Central, Huechuraba, Independencia, La Cisterna, La Florida, La Granja, La Pintana, La Reina, Las Condes, Lo Barnechea, Lo Espejo, Lo Prado, Macul, Maipú, Padre Hurtado, Pedro Aguirre Cerda, Peñalolén, Providencia, Pudahuel, Puente Alto, Quilicura, Quinta Normal, Recoleta, Renca, San Bernardo, San Joaquín, San Miguel, San Ramón, Santiago, Vitacura y Ñuñoa.
O'Higgins	Rancagua
Maule	Talca
Ñuble ⁽¹⁾	Gran Chillán: Chillán, Chillán Viejo
Biobío	Gran Concepción: Concepción, Chiguayante, Penco, San Pedro de la Paz, Talcahuano, Hualpén
La Araucanía	Gran Temuco: Temuco, Padre Las Casas
Los Ríos	Valdivia
Los Lagos	Puerto Montt
Aysén	Coihaique
Magallanes	Punta Arenas

(1) Estas comunas no estaban presentes en la muestra de la VII EPF y corresponden al área conurbada de la capital de la Región de Ñuble.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

82 Al establecer la metodología de trabajo de campo, se decidió excluir del proceso de entrevistas aquellas viviendas con más de cinco hogares residentes. El sustento de esta decisión alude a razones operativas, ya que una vivienda con esas características multiplica el número de visitas e instrumentos a aplicar en esas unidades, generando un importante desequilibrio en las cargas de trabajo de encuestadores. Como procedimiento de control ante esta medida, se cuenta con el registro de contactos (e intentos) a cada unidad seleccionada en la muestra, incluyendo una o más observaciones del encuestador o encargado de grupo. En la revisión y depuración de los códigos de disposición final se identificaron ocho viviendas con estas características, es decir, el 0,03% de las unidades registradas en Hoja de Ruta. Dada la cantidad de casos, se considera que la cobertura de la población objetivo no se vio dañada por esta decisión en específico.

La encuesta presenta representatividad estadística para las siguientes áreas de estimación:

- Gran Santiago.
- Resto de capitales regionales.
- Total de capitales regionales (Gran Santiago + Resto de capitales regionales).

6.1.1.4 Método de muestreo

El diseño muestral obedece a un muestreo probabilístico, estratificado y bietápico. Las unidades muestrales de primera etapa (UPE o unidades primarias de muestreo) son las manzanas y presentan delimitaciones geográficas fijas. Cada manzana es clasificada dentro de la comuna a la que pertenece, según clasificación socioeconómica⁸³ y el número de viviendas que la componen.

Las unidades muestrales de segunda etapa (USE o unidades secundarias de muestreo) son las viviendas particulares que se encuentran al interior de cada manzana.

Para realizar la distribución de la muestra de manzanas, se consideró la distribución porcentual de viviendas según grupo de tamaño, en la que se truncó la proporción correspondiente a los grupos de tamaño más pequeños (grupos de tamaño 1 y 2) debido a los altos costos asociados al trabajo de campo en estas manzanas.

Las manzanas se seleccionaron con igual probabilidad dentro de cada estrato. Mientras que la selección de viviendas al interior de cada manzana, se realizó también con igual probabilidad.

Al interior de las viviendas seleccionadas, se entrevistaron a todos los hogares residentes. La encuesta buscó capturar información de gastos e ingresos de todas las personas miembros de los hogares residentes de cada vivienda seleccionada⁸⁴.

6.1.2 Características del marco de muestreo

El marco muestral de la VIII EPF es el Marco Muestral de Manzanas que utiliza el INE para las zonas geográficas urbanas⁸⁵, conformado por unidades geográficas con límites fijos denominados manzanas (o conglomerados). El Instituto posee información cartográfica y del número de viviendas que contiene cada manzana según los datos del Censo de Población y Vivienda del año 2002 parcialmente actualizada al año 2015.

El Marco Muestral de Manzanas del INE es utilizado para seleccionar las muestras de las principales encuestas de hogares que se realizan en el país, tales como la Encuesta de Caracterización Socioeconómica Nacional (Casen), la Encuesta Nacional Urbana de Seguridad Nacional (Enusc) y la Encuesta Nacional sobre Uso del Tiempo (ENUT), entre otras.

83 Según lo explicado en el apartado 5.2.2, las manzanas del Marco Muestral de hogares son estratificadas socioeconómicamente según información proveniente del Censo de Población y Vivienda del año 2002 a través del método Princals (INE Chile, 2014; Guerrero, 2003b).

84 Salvo en el caso de ingresos donde se excluyen los ingresos de menores de 15 años (para mayor información, revisar el apartado 5.1.3.)

85 El Marco Muestral utilizado para zonas rurales se denomina Marco Muestral de Secciones y dada la cobertura de la VIII EPF no es utilizado en esta encuesta.

Hasta el año 2009, la actualización del marco se realizaba mediante registros administrativos, provenientes de los certificados de recepción final de las nuevas construcciones, otorgados mensualmente por las direcciones de obras de las municipalidades, que captura el Formulario Único de Edificación del INE, lo que permite mantener actualizado el número total de viviendas en la manzana e incorporar aquellas creadas con posterioridad al Censo de Población y Vivienda 2002.

Adicionalmente, considerando el importante crecimiento demográfico en el área urbana del país, en el año 2014 el Ministerio de Desarrollo Social firmó un convenio con el INE para actualizar una parte del marco desde donde se extraería la muestra de la Encuesta de Caracterización Socioeconómica Nacional (Casen) 2015. Esta actualización se realizó a 2.351 manzanas, que equivale a 1,96% de un total de 119.970 manzanas que posee el marco, y a 9,7% de un total de 24.182 manzanas que poseen las 61 comunas elegidas para la actualización. La elección de estas 61 comunas se debió a que eran las que presentaban mayor cantidad de manzanas no elegibles, ya fuera por motivos de cambio de estructura, o por crecimiento o decrecimiento en cuanto al número de viviendas que las componían.

El Marco Muestral de Manzanas, respecto al cual se selecciona la muestra de la EPF, se encuentra organizado en forma jerárquica, de acuerdo a la división político administrativa del territorio. Las unidades se ordenan en forma descendente según los niveles jerárquicos: región, provincia y comuna. Al interior de cada comuna, se conforma la división censal que da origen a las áreas geográficas denominadas urbano y rural. En el caso de la EPF solo se seleccionan áreas urbanas de las comunas que son capitales regionales, o bien que forman conurbaciones en centros urbanos importantes.

6.1.2.1 Cobertura del marco de muestreo

Según el glosario de términos estadísticos de la OCDE, la cobertura corresponde a “la población desde la que se pueden extraer las observaciones para un tema en particular” (OCDE, 2007, pág. 156). Es importante cubrir adecuadamente la población respecto a la que se está realizando el estudio ya que, si no se abarcan adecuadamente todos los componentes de la población estudiada, surgen los errores de cobertura, los que corresponden a divergencias entre la población objetivo y el marco (OCDE, 2007).

La cobertura geográfica de la VIII EPF incluye las ciudades del Gran Santiago, más el Resto de capitales regionales y algunas de sus zonas conurbadas. No obstante, se excluyeron del muestreo manzanas debido a los siguientes motivos operativos:

- Manzanas de tamaño pequeño, es decir, aquellas manzanas que tienen siete viviendas o menos de acuerdo al marco muestral.
- Manzanas de difícil acceso o alta peligrosidad.

Las manzanas excluidas del marco muestral de la EPF por razones de alto costo (manzanas de siete viviendas o menos) o de seguridad corresponden a 7.265 y representan 0,9% del total de viviendas del marco muestral de la encuesta.

Además, cuando se realizó la afijación proporcional de la muestra al interior de las comunas, ocurrió que, para algunos grupos de tamaño (específicamente grupos grandes), la cantidad de viviendas asignada era menor a la cantidad de viviendas predeterminadas a seleccionar por manzana en estos grupos, lo cual implicaba seleccionar solo un conglomerado o manzana por estrato. No obstante, en un diseño estratificado por conglomerados, se requiere seleccionar al menos dos unidades (manzanas) por cada estrato.

Debido a lo anterior, se definió no seleccionar muestra en aquellos estratos donde la cantidad de viviendas asignada era menor a la cantidad de viviendas predeterminadas a seleccionar por manzana, redistribuyéndose el total de viviendas asignado a estos grupos de tamaño en otros grupos, buscando, en la medida de lo posible, asignarlos a la misma comuna. Esto se traduce en la exclusión de 225 estratos (comuna, condición socioeconómica y grupos de tamaño) asociados a 4.861 manzanas que en total concentran 8,3% de las viviendas del marco (mayor información en el apartado 6.2.1.3).

En total, realizando las consideraciones de tamaño y seguridad anteriormente mencionadas, se excluyeron 12.126 manzanas (13,8%) del marco muestral, lo que representa un 9,2% del total de viviendas presentes en las comunas de selección de la EPF.

Si bien las manzanas excluidas por razones de tamaño son contempladas en la corrección de los factores de expansión, según un ajuste de omisión de conglomerados⁸⁶, y la población presente en la encuesta es ajustada a las proyecciones de población⁸⁷, de todas formas persisten viviendas excluidas de la cobertura. Al utilizar los datos de la encuesta para fenómenos que puedan estar asociados al tamaño de las manzanas, en particular a manzanas de tamaño pequeño, se debe ser cauteloso respecto de la inferencia que se realice sobre la población a partir de los datos de la encuesta.

Tabla 10: Distribución de manzanas y viviendas en el marco muestral (total y excluidas) según región

Marco muestral de viviendas de las comunas estudiadas en la VIII EPF					Exclusiones			
Región	Total de manzanas	Manzanas (%)	Total de viviendas	Viviendas (%)	Total de manzanas	Manzanas (%) respecto total marco	Total de viviendas	Viviendas (%) respecto total marco
Total	87.690	100	2.931.070	100	12.126	13,8	269.657	9,2
Arica y Parinacota	2.424	2,8	58.057	2	280	0,3	1.050	0,0
Tarapacá	2.163	2,5	76.265	2,6	183	0,2	3.342	0,1
Antofagasta	3.199	3,6	85.395	2,9	431	0,5	6.309	0,2
Atacama	1.691	1,9	39.826	1,4	202	0,2	2.089	0,1
Coquimbo	4.246	4,8	102.723	3,5	448	0,5	4.936	0,2
Valparaíso	10.176	11,6	296.174	10,1	1.674	1,9	14.222	0,5
Metropolitana	42.145	48,1	1.657.248	56,5	6.706	7,6	201.230	6,9
O'Higgins	2.193	2,5	66.120	2,3	171	0,2	1.748	0,1
Maule	2.369	2,7	63.995	2,2	212	0,2	2.341	0,1
Ñuble	1.877	2,1	54.048	1,8	195	0,2	3.401	0,1
Biobío	6.376	7,3	204.191	7	671	0,8	14.761	0,5
Araucanía	3.260	3,7	83.973	2,9	353	0,4	3.283	0,1
Los Ríos	1.340	1,5	38.160	1,3	145	0,2	4.576	0,2
Los Lagos	1.770	2	48.340	1,6	222	0,3	3.215	0,1
Aysén	723	0,8	17.055	0,6	70	0,1	1.009	0,0
Magallanes	1.738	2	39.500	1,3	163	0,2	2.145	0,1

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

86 Al respecto ver el apartado 6.2.1.3 de ajuste por omisión de manzanas.

87 Al respecto ver el apartado 6.2.4 de ajuste de calibración donde se ajusta la proyección de la encuesta a las proyecciones poblacionales según región, sexo y tramo etario.

6.1.2.2 Estratificación del marco de muestreo

Dado el tipo de encuesta y el diseño muestral, previo a la selección de la muestra se aplicó un proceso de estratificación, agrupando elementos homogéneos entre sí. Posteriormente, la muestra fue seleccionada de modo que los distintos grupos o estratos estuviesen representados en ella. Se buscó que los estratos fuesen homogéneos en su interior, pero heterogéneos entre sí, con el fin estadístico de reducir la varianza de las estimaciones.

Considerando las características de la VIII EPF, se utilizó un marco estratificado geográfica y socioeconómicamente.

6.1.2.3 Estratificación geográfica

La estratificación geográfica para este estudio está dada por la división político-administrativa y, en orden jerárquico, corresponde a región, provincia y comuna.

6.1.2.4 Estratificación socioeconómica

Para la clasificación socioeconómica de las manzanas se utiliza la información recabada en el Censo de Población y Vivienda 2002. Este es el insumo para crear la clasificación de las comunas, es decir, la estratificación del marco muestral⁸⁸.

Se identificaron las manzanas del marco muestral que existían en los registros originales del Censo de Población y Vivienda 2002 y a cada una de ellas se les asignó el puntaje construido en esa oportunidad para su estratificación. Con este procedimiento, se logró segmentar aproximadamente el 82,7% de las manzanas del marco muestral.

El 17,3% de las manzanas restantes no lograron ser identificadas debido a cambios estructurales, como, por ejemplo, manzanas con viviendas construidas con posterioridad al Censo 2002. A estas manzanas, se les imputó el puntaje promedio de las manzanas más cercanas que estuvieran dentro de la misma comuna, distrito y zona censal, y que además pertenecieran al mismo grupo de tamaño.

Al combinar la información del marco muestral y de Censo 2002, se ordenaron las manzanas según el puntaje asignado y se formaron los deciles, para construir los siguientes estratos socioeconómicos:

- Alto: decil 10.
- Medio: decil 4-9.
- Bajo: decil 1-3.

⁸⁸ Para mayor información respecto al método de estratificación utilizado para la generación de los estratos socioeconómicos revisar Guerrero (INE Chile, 2014).

6.1.2.5 Estratificación del grupo de tamaño

La estratificación de las manzanas en grupos de tamaño corresponde al agrupamiento de las manzanas en seis grupos, de acuerdo a la cantidad de viviendas que las conforman. Estas son:

- Grupo 0: menos de 8 viviendas.
- Grupo 1: 8-23 viviendas.
- Grupo 2: 24-44 viviendas.
- Grupo 3: 45-81 viviendas.
- Grupo 4: 82-154 viviendas.
- Grupo 5: más de 154 viviendas.

Es importante recordar que se excluyeron las manzanas conformadas por menos de ocho viviendas.

Para la VIII EPF, considerando el objetivo de la encuesta y con el fin de tener presentes a grupos de población de distintas características en la selección de las unidades, se definió el estrato de muestreo en función de la combinación de las variables comuna (63 comunas), clasificación socioeconómica (alto, medio y bajo) y grupo de tamaño (cinco grupos). A partir de esta combinación se generaron 810 estratos de muestreo, todos correspondientes a zonas urbanas de las comunas incluidas en el marco de selección de la encuesta. No obstante, fue necesario excluir 225 estratos que no contaban con suficientes viviendas para hacer la selección adecuadamente. De esta forma, finalmente la muestra quedó constituida por 585 estratos.

La tabla 11 muestra la distribución de los grupos de tamaño en el marco muestral respecto al cual fue seleccionada la muestra de la VIII EPF, mientras que la tabla 12 presenta la distribución de muestra seleccionada en la VIII EPF.

Tabla 11: Número de manzanas y viviendas en el marco muestral de selección de la VIII EPF, según grupo de tamaño de manzanas del marco muestral de la VIII EPF

Grupo de tamaño	Rango de viviendas	Total de manzanas	% de manzanas	Total de viviendas	% de viviendas
Total		75.564	100	2.661.413	100
Grupo 1	8 a 23	36.827	48,7	588.780	22,1
Grupo 2	24 a 44	26.995	35,7	840.775	31,6
Grupo 3	45 a 81	7.176	9,5	411.876	15,5
Grupo 4	82 a 154	2.772	3,7	306.792	11,5
Grupo 5	155 y más	1.794	2,4	513.190	19,3

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Tabla 12: Número de manzanas y viviendas seleccionadas en la muestra de la VIII EPF, según grupo de tamaño de manzanas del marco muestral de la VIII EPF

Grupo de tamaño	Rango de viviendas	Total de manzanas	% de manzanas	Total de viviendas	% de viviendas
Total		2.124	100	25.361	100
Grupo 1	8 a 23	834	39,3	3374	13,3
Grupo 2	24 a 44	498	23,5	4882	19,3
Grupo 3	45 a 81	410	19,3	6558	25,9
Grupo 4	82 a 154	189	8,9	4792	18,9
Grupo 5	155 y más	193	9,1	5775	22,7

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

6.1.3 Cálculo y distribución del tamaño muestral

6.1.3.1 Cálculo del tamaño muestral

La Encuesta de Presupuestos Familiares es una de las encuestas económicas aplicadas a hogares más complejas que existen. En la etapa de cálculo del tamaño muestral se generaron diversas simulaciones con la muestra objetivo considerando mayor representatividad estadística (se probaron tamaños muestrales para una muestra con representatividad regional, para cada una de las capitales regionales, además de incorporar las áreas urbano-rural).

Considerando la extensión de la encuesta y los altos costos asociados al proceso de producción estadística, se decidió mantener las áreas de estimación de la encuesta, mejorando los niveles de precisión de los datos, buscando reducir el error asociado a las estimaciones para los indicadores principales.

Una de las mejoras metodológicas importantes en el proceso de selección y definición del tamaño de la muestra de la VIII EPF fue la incorporación de una sobremuestra a la muestra objetivo, con el fin de alcanzar los tamaños muestrales estimados. Mientras que la VII EPF utilizó un proceso de selección de muestra con reemplazo, la VIII EPF introdujo la estrategia de sobremuestreo. La diferencia radica en que los reemplazos son utilizados a medida que las unidades de segunda etapa (viviendas) rechazan participar en el estudio, mientras que la muestra con sobremuestreo considera *a priori* (en base a información recabada en anteriores versiones de la encuesta) la cantidad teórica de viviendas requeridas para cumplir con los tamaños objetivos.

La estructura temporal de la VIII EPF fue rígida (los hogares fueron asignados a un período fuera del cual no podían ser encuestados) y el grado de participación solicitada a los hogares fue elevado (cada hogar participó del estudio durante quince días), por lo que resultó aún más conveniente y recomendado emplear un proceso de selección con sobremuestreo, buscando encuestar la totalidad de las unidades seleccionadas. Esta estrategia permitió contrarrestar en parte la no respuesta total.

Para el cálculo del tamaño con sobremuestreo, se utilizó la tasa de no respuesta y la tasa de desactualización del marco muestral a nivel comunal obtenida de la VII EPF. Con ello se calcularon los tamaños regionales con sobremuestreo, a través de la agregación de las comunas y, finalmente, los tamaños por nivel de estimación, como una agregación de la o las regiones que lo componen.

El cálculo del tamaño muestral utilizado correspondió a un diseño probabilístico, bietápico y estratificado. Los estimadores necesarios para el diseño de la VIII EPF fueron obtenidos a partir de las estimaciones realizadas con la VII versión de la encuesta, incluyendo el gasto *per cápita* promedio mensual del hogar, el cual se definió como la variable de interés del diseño. Así también, partiendo de la VII EPF, se estableció la variable de interés, el número de viviendas logradas y el error absoluto y relativo asociado a la variable de interés, además de la tasa de no respuesta y de desactualización del marco.

Dado que la VII EPF y la VIII EPF comparten el mismo objetivo, se empleó el siguiente esquema para el cálculo del tamaño muestral según nivel de estimación:

1. Obtención del tamaño de muestra

El procedimiento se realizó mediante la siguiente fórmula para cada nivel de estimación:

$$m_1 = m_0 \cdot \frac{(er_0)^2}{(er_1)^2} \quad (1)$$

Donde,

m_0 es el tamaño muestral en cada nivel de estimación para la VII EPF.

m_1 es el tamaño muestral en cada nivel de estimación para la VIII EPF.

er_0 es el error relativo obtenido en la VII EPF.

er_1 es el error relativo deseado en la VIII EPF.

A continuación, se describe el desarrollo de la fórmula (1):

Dado que el tamaño muestral bajo muestreo aleatorio simple es:

$$m_d = \frac{Z_{1-\alpha/2}^2 \cdot S_d^2(\hat{\theta})}{E_d^2} \quad (2)$$

Donde

$Z_{1-\alpha/2}$ es el percentil $1-\alpha/2$ de la distribución normal estándar

$S_d^2(\hat{\theta})$ es la cuasi varianza del estimador

E_d es el error prefijado (en este caso es el error absoluto o relativo)

Se consideran tamaños muestrales en dos momentos en el tiempo ("0" y "1") para estimar el mismo parámetro poblacional.

El tamaño de la muestra inicial en el período "0", denotado por m_0 , viene dado por:

$$m_0 = \frac{Z_{1-\alpha/2}^2 \cdot S_0^2(\hat{\theta})}{e_0^2} \quad (3)$$

Donde

- m_0 es el tamaño muestral en el período “0”
- $Z_{1-\alpha/2}$ es el percentil $1-\alpha/2$ de la distribución Normal estándar
- $S_0^2(\hat{\theta})$ es la cuasi varianza del estimador de la muestra disponible en el período “0”
- e_0 es el error absoluto o relativo efectivo en el período “0”

Análogamente, el tamaño de la muestra en el período “1”, denotado por m_1 , sujeto a un error absoluto o relativo establecido *a priori*, e_1 , viene dado por:

$$m_1 = \frac{Z_{1-\alpha/2}^2 \cdot S_1^2(\hat{\theta})}{e_1^2} \quad (4)$$

Donde

- m_1 es el tamaño muestral en el período “1”.
- $Z_{1-\alpha/2}$ es el percentil $1-\alpha/2$ de la distribución normal estándar.
- $S_1^2(\hat{\theta})$ es la cuasivarianza del estimador de la muestra disponible en el período “1”.
- e_1 es el error absoluto o relativo efectivo en el período “1”.

Al examinar el cociente entre ambos tamaños (tiempo 0 y 1), de las ecuaciones (3) y (4), se obtiene:

$$\frac{m_0}{m_1} = \left(\frac{e_1}{e_0} \right)^2 \frac{S_0^2(\hat{\theta})}{S_1^2(\hat{\theta})} \quad (5)$$

Como se está estimando el mismo parámetro poblacional en dos momentos del tiempo, se asume que la dispersión se mantiene constante, esto es, que las cuasi varianzas $S_0^2(\hat{\theta})$ y $S_1^2(\hat{\theta})$ son iguales, lo que deriva en la siguiente relación:

$$m_1 = m_0 \left(\frac{e_0}{e_1} \right)^2 \quad (6)$$

Luego, relacionando los errores con los tamaños muestrales en dos períodos de tiempo “0” y “1”, permite tener una fórmula para disminuir los errores absolutos o relativos ante el aumento o disminución entre m_0 y m_1 , relación frecuentemente empleada en el cálculo de tamaños muestrales, en el contexto de diseños complejos, como es el caso de la VIII EPF.

2. Distribución de la muestra a nivel regional y comunal

Tras establecer los tamaños muestrales según nivel de estimación, se realizó la distribución de la muestra a nivel regional de forma proporcional según el total de viviendas en el marco. Luego de haber obtenido los tamaños de muestra a nivel regional, se repitió este proceso de distribución a nivel comunal.

3. Corrección por no respuesta

Se corrigió el tamaño muestral m_1 de cada comuna utilizando la tasa de no respuesta ajustada por no elegibilidad (tnr_{ne})⁸⁹ obtenida en la VII EPF a nivel comunal. Esa tasa se utilizó como una medida de desempeño de las viviendas logradas en trabajos de campo anteriores, ya que refleja los posibles cambios de estado de las viviendas, las dificultades de contacto y el rechazo a responder de los informantes. Así, el tamaño muestral final se corrigió con la tasa de no respuesta, de la siguiente forma:

$$m_2 = \frac{m_1}{1 - tnr_{ne}} \quad (7)$$

Donde el tamaño, m_2 , es el tamaño muestral sobredimensionado a ser utilizado en el programa de trabajo de campo para poder lograr el tamaño m_1 en la muestra.

Finalmente, los tamaños m_2 obtenidos en cada comuna se agregaron por región y luego, por los niveles de estimación, para finalmente obtener el tamaño nacional sobredimensionado.

La tabla 13 presenta los tamaños objetivos, los errores calculados según el gasto *per cápita* promedio mensual de los hogares y la muestra con sobremuestreo.

Tabla 13: Tamaño objetivo, errores calculados de acuerdo con el gasto *per cápita* promedio mensual según VII EPF y tamaño con sobremuestreo

Nivel	Muestra objetivo	Gasto <i>per cápita</i> promedio mensual	Total viviendas VII EPF	Coeficiente de variación	Error absoluto	Error relativo	Muestra con sobremuestreo
Total de capitales regionales⁽¹⁾	13.503	268.067	3.009.720	2,96%	15.537	5,80%	25.361
Gran Santiago	6.800	269.330	1.741.414	4,23%	22.279	8,30%	14.705
Resto de capitales regionales ⁽¹⁾	6.703	266.398	1.268.306	3,01%	15.607	5,90%	10.656

(1) Se considera la incorporación de las comunas de Chillán y Chillán Viejo, comunas que compondrán la capital de la futura Región de Nuble.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

A modo de síntesis, el cálculo del total de viviendas a encuestar se realizó considerando una muestra (con sobremuestreo) de 25.361 viviendas a encuestar, cuyo error relativo asociado es de 5,8% para la estimación del gasto *per cápita* promedio mensual por hogar del total de capitales regionales. En el caso del Gran Santiago se consideró una muestra (con sobremuestreo) de 14.705 viviendas, mientras que para el Resto de capitales regionales esta fue de 10.656, con errores relativos esperados de 8,3% y 5,9%, respectivamente, para el gasto *per cápita* promedio mensual por hogar.

89 La tasa de no respuesta ajustada por no elegibilidad se obtiene considerando las viviendas elegibles, viviendas de elegibilidad desconocida y viviendas no elegibles según la siguiente fórmula: $1 - (\text{entrevistadas} / (\text{viviendas elegibles} + \text{viviendas de elegibilidad desconocida} + \text{viviendas no elegibles}))$.

A diferencia de la VII EPF, en el cálculo del tamaño muestral se incorporaron las comunas de Chillán y Chillán Viejo, pertenecientes a la nueva Región de Ñuble. Al no contar con información proveniente de la VII EPF sobre estas comunas, se utilizaron los resultados de Casen 2013 para determinar qué comunas podrían representarlas. Para esto, se analizó el comportamiento de la población de dichas comunas, respecto a otras similares en términos del ingreso disponible *per cápita* promedio mensual y del total de viviendas según el Marco Muestral 2015, dentro de la macrozona sur⁹⁰ y según el estrato socioeconómico⁹¹.

El criterio utilizado para definir cuáles comunas serían las mejores representantes de Chillán y Chillán Viejo consistió en elegir aquellas que presentaran menor distancia euclidiana con respecto a estas, siendo Talcahuano y Hualpén las que resultaron más parecidas a Chillán y Chillán Viejo, respectivamente.

6.1.3.2 Distribución geográfica de la muestra

Una vez determinado el tamaño muestral objetivo por nivel de estimación y nivel socioeconómico, se procedió a distribuir estos tamaños por región y luego por comuna, en forma proporcional según la distribución del marco muestral.

Teniendo los tamaños comunales, se obtuvo el tamaño con sobremuestreo, según las tasas de no respuesta a nivel comunal provenientes de la VII EPF. Luego, se agregaron estos tamaños a nivel regional para, finalmente, obtener los tamaños con sobremuestreo para cada nivel de estimación definido.

El procedimiento consideró varios aspectos que se señalan a continuación:

- La distribución de la muestra objetivo a nivel regional se realizó en forma proporcional según la distribución de viviendas del marco muestral, con respecto al total de viviendas de cada área de estimación (Gran Santiago y Resto de capitales regionales).
- Luego, a nivel de comuna se realizó el mismo procedimiento, según el total de viviendas de cada región.
- Con las tasas de no respuesta obtenidas en la VII EPF, se determinó el tamaño comunal con sobremuestreo y se distribuyó este tamaño entre todos los grupos de tamaño que conforman cada comuna, de igual forma que en la etapa anterior.
- Cabe destacar que, tal como se describió en el apartado 6.1.2.1, dado que el trabajo de campo en manzanas de menor tamaño (grupos 1 y 2) es más costoso en términos operativos, se truncó la proporcionalidad obtenida según la distribución del marco, dando como resultado la asignación de un menor tamaño muestral en estos grupos.

En la tabla 14, se presenta la distribución geográfica del total de viviendas por clasificación socioeconómica y capital regional o zona conurbada.

90 Al respecto, ver tabla 14. La macrozona sur está formada por las regiones de: La Araucanía, Los Ríos, Los Lagos, Aysén y Magallanes.

91 Dado que no se contaba con información de gastos para revisar comunas de similar comportamiento con el objetivo de imputar las tasas de no respuestas con el fin de calcular el tamaño a encuestar con sobremuestreo, se consideró el comportamiento del ingreso disponible *per cápita* promedio mensual dada su alta correlación con el gasto *per cápita* promedio mensual.

Tabla 14: Distribución del total de viviendas por macrozona, capital regional, comuna y clasificación socioeconómica. Muestra con sobremuestreo

Macrozona	Capital regional y/o zona conurbada	Comuna	Total	Clasificación socioeconómica		
				Bajo	Medio	Alto
Total capitales regionales			25.361	6.720	14.233	4.408
Total macrozona norte			3.007	726	1.606	675
Norte	Arica	Arica	436	101	282	53
	Gran Iquique	Iquique	397	70	170	157
		Alto Hospicio	189	133	56	0
	Antofagasta	Antofagasta	784	115	419	250
	Copiapó	Copiapó	394	106	214	74
	Gran La Serena	La Serena	336	76	169	91
		Coquimbo	471	125	296	50
Total macrozona centro			5.765	1364	3.664	737
Centro	Gran Valparaíso	Valparaíso	648	126	522	0
		Concón	146	0	82	64
		Viña del Mar	975	125	569	281
		Quilpué	373	21	295	57
		Villa Alemana	267	81	186	0
	Rancagua	Rancagua	479	98	303	78
	Talca	Talca	455	152	237	66
	Gran Concepción	Concepción	548	122	312	114
		Chiguayante	200	75	93	32
		Penco	95	53	42	0
		San Pedro de La Paz	262	81	136	45
		Talcahuano	346	117	229	0
		Hualpén	209	47	162	0
	Gran Chillán	Chillán	293	102	191	0
		Chillán Viejo	469	164	305	0
Total macrozona sur			1.884	743	831	310
Sur	Gran Temuco	Temuco	530	269	190	71
		Padre Las Casas	116	97	19	0
	Valdivia	Valdivia	306	103	152	51
	Puerto Montt	Puerto Montt	366	116	176	74
	Coyhaique	Coyhaique	253	107	134	12
	Punta Arenas	Punta Arenas	313	51	160	102
Total macrozona Gran Santiago			14.705	3887	8.132	2.686
Región Metropolitana	Gran Santiago	Santiago	856	315	541	0
		Cerrillos	199	110	89	0
		Cerro Navia	293	194	99	0
		Conchalí	316	93	223	0

Macrozona	Capital regional y/o zona conurbada	Comuna	Total	Clasificación socioeconómica		
				Bajo	Medio	Alto
Región Metropolitana	Gran Santiago	El Bosque	434	185	249	0
		Estación Central	343	210	133	0
		Huechuraba	154	0	113	41
		Independencia	206	114	92	0
		La Cisterna	247	0	247	0
		La Florida	959	176	596	187
		La Granja	292	143	149	0
		La Pintana	351	339	12	0
		La Reina	325	0	0	325
		Las Condes	844	0	69	775
		Lo Barnechea	139	0	42	97
		Lo Espejo	242	81	161	0
		Lo Prado	287	211	76	0
		Macul	303	0	303	0
		Maipú	1249	0	1249	0
		Ñuñoa	559	0	183	376
		Pedro Aguirre Cerde	253	90	163	0
		Peñalolén	523	98	315	110
		Providencia	532	0	114	418
		Pudahuel	511	157	354	0
		Quilicura	364	16	348	0
		Quinta Normal	292	131	161	0
		Recoleta	366	139	227	0
		Renca	316	129	187	0
		San Joaquín	201	98	103	0
		San Miguel	255	0	193	62
		San Ramón	266	199	67	0
		Vitacura	295	0	0	295
		Puente Alto	1198	401	797	0
		San Bernardo	563	155	408	0
		Padre Hurtado	172	103	69	0

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

6.1.3.3 Distribución de la muestra en submuestras mensuales

El período de captura de información de la VIII EPF es de 12 meses consecutivos.

Esto significa que las 25.361 viviendas seleccionadas (13.503 viviendas objetivo) son entrevistadas a lo largo de un año calendario, de forma gradual, con la finalidad de capturar estacionalidades de diversos tipos e índices, que pueden impactar directamente sobre la distribución de los gastos del hogar.

Con respecto a la captura de información, tal como se especifica con detalle en el apartado 8.1.2, es necesario señalar que los hogares reportan información durante quince días consecutivos, por lo tanto, el total de viviendas a encuestar fue distribuido en 24 submuestras operativas, equivalentes a 12 meses con 2 quincenas por mes⁹².

Debido a la forma de expandir los datos, todas las viviendas a encuestar en una manzana son entrevistadas en un solo mes. Por lo tanto, previo a la distribución de cada vivienda en una de las 24 submuestras operativas, se asignó cada manzana a un mes en particular.

Así, el conjunto de viviendas a entrevistar en cada manzana cumplía con los siguientes criterios:

- A nivel nacional, el total de viviendas por submuestra mensual debía ser similar. Cada mes de trabajo de campo se consideró como una réplica a menor escala de la muestra total de la encuesta. El total de viviendas a encuestar de 25.361 (esperando lograr respuesta de 13.503 viviendas) se repartió en cada mes, por lo que cada submuestra mensual tenía, en promedio, aproximadamente 2.113 viviendas (esperando respuesta objetivo de 1.125 viviendas).
- Se definieron cuatro macrozonas entre las cuales se distribuyeron las submuestras mensuales. El criterio para distribuir las viviendas en las macrozonas consideró que cada macrozona debía tener, en cada mes, presencia de los tres estratos socioeconómicos de la muestra⁹³.
- Dado que se buscó que cada submuestra mensual sea una réplica a menor escala de la muestra completa, en cada submuestra mensual debía estar representada cada capital regional o zona conurbada en su conjunto. Por lo tanto, cada capital regional o zona conurbada, según correspondía, debía tener distribuidas, de la forma más homogénea posible, el total de viviendas en las 12 submuestras mensuales.

La tabla 15 presenta la distribución del total de viviendas a encuestar según submuestra mensual y clasificación socioeconómica.

Tabla 15: Distribución del total de viviendas a encuestar por submuestra mensual y clasificación socioeconómica con sobremuestra

Clasificación socio-económica	Total viviendas	Submuestra mensual											
		1	2	3	4	5	6	7	8	9	10	11	12
Total	25.361	2.146	2.131	2.135	2.128	2.140	2.107	2.127	2.116	2.107	2.071	2.071	2.082
CSE Bajo	6.720	559	525	551	563	546	564	592	559	561	590	579	531
CSE Medio	14.233	1.198	1.229	1.197	1.188	1.210	1.174	1.170	1.190	1.225	1.137	1.125	1.190
CSE Alto	4.408	389	377	387	377	384	369	365	367	321	344	367	361

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

92 Dado que los meses son divididos en submuestras operativas denominadas quincenas, existen quincenas de 14, 15 o 16 días dependiendo del mes de trabajo de campo.

93 Cabe señalar que, debido al tamaño de la muestra, la distribución por submuestra mensual a nivel CSE se realizó a nivel de macrozona y no de región.

6.1.4 Selección de unidades muestrales

Tal como se ha mencionado, la selección de unidades muestrales se realizó en dos etapas. En una primera etapa se seleccionaron manzanas (unidades de primera etapa), mientras que en la segunda etapa se seleccionaron las viviendas particulares ocupadas (unidades de segunda etapa).

6.1.4.1 Selección de unidades de primera etapa (UPE - manzanas)

La selección de la muestra de unidades de primera etapa se realizó desde el marco muestral considerando la implementación del diseño complejo, utilizando como estratos de diseño la combinación de las variables *comuna*, *clasificación socioeconómica* y *grupo de tamaño*.

En cada estrato de la muestra, se realizó la selección de forma independiente mediante un método aleatorio sistemático con igual probabilidad, ordenando los registros por las variables *comuna*, *distrito*, *zona* y *manzana*, estrategia que busca garantizar una buena cobertura de las diferentes áreas a considerar.

La tabla 16 muestra el número de manzanas seleccionadas. Partiendo de una muestra inicial de 2.120 manzanas, se realizó una revisión que originó reemplazos en las siguientes tres fases:

Fase 1: corresponde a reemplazos de aquellas manzanas que han participado en la Encuesta Nacional de Empleo. En esta fase fueron reemplazadas 71 manzanas, en forma biunívoca.

Fase 2: corresponde a la verificación de manzanas realizada en la unidad de cartografía. En esta fase fueron reemplazadas 374 manzanas. En algunos casos, por falta de manzanas disponibles, no fue posible reemplazar una manzana en forma biunívoca dentro del estrato de muestreo (dentro de su mismo grupo de tamaño), por lo que se tuvieron que reemplazar por dos o más manzanas de grupos de tamaños más pequeños. Esta situación da origen a que se utilizaran 386 manzanas de reemplazo.

Fase 3: corresponde a la verificación de viviendas en terreno que da origen al reemplazo de manzanas debido a ausencia de viviendas suficientes para selección. Por este motivo fue necesario reemplazar 29 manzanas dando origen a 35 manzanas reemplazantes.

Adicionalmente, fue necesario eliminar 14 manzanas debido a que no contaban con suficientes viviendas elegibles para la selección y, por tanto, las viviendas asociadas fueron redistribuidas en otras manzanas, en la medida de lo posible, dentro del mismo estrato de muestreo.

En resumen,

- En la fase de reemplazos realizados durante el proceso de selección de la muestra, de las 2.120 manzanas correspondientes a la muestra original, se reemplazaron 304, por un total de 322, pudiendo especificar que se reemplazaron dos manzanas fuera de su clasificación socioeconómica (CSE) y tres manzanas fuera de su comuna.
- En la fase de reemplazos solicitados por la Unidad de Cartografía, 33 manzanas se debieron reemplazar fuera de su CSE y 41, por una de otro tamaño. Finalmente, de los reemplazos realizados en terreno, cuatro manzanas se reemplazaron en otro grupo de clasificación socioeconómica y 11 manzanas se reemplazaron en otro grupo de tamaño.

Tabla 16: Manzanas reemplazadas en el proceso de selección de la muestra de la VIII EPF

Condición	Manzanas seleccionadas inicialmente	Manzanas reemplazadas				Conglomerados reemplazados (%)			
		ENE	Cartografía	Terreno	Total ⁽¹⁾	ENE	Cartografía	Terreno	Total
Reemplazadas	2.120	71	374	29	304	3,3	17,6	1,4	14,3
Reemplazantes		71	386	35	322	3,3	18,2	1,7	15,2

(1) La columna "total" no coincide con la suma de las columnas anteriores, ya que algunas de las manzanas de reemplazo fueron a su vez reemplazadas por otras manzanas.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

- En la fase de reemplazo de manzanas, se reemplazaron dos manzanas fuera de su clasificación socioeconómica y tres manzanas fuera de su comuna.
- En la fase de reemplazos solicitados por la Unidad de Cartografía, 33 manzanas se debieron reemplazar fuera de su clasificación socioeconómica y 41 por una de otro grupo de tamaño. Finalmente, de los reemplazos realizados en terreno, cuatro manzanas se reemplazaron en otra clasificación socioeconómica y 11 manzanas se reemplazaron en otro grupo de tamaño.

Cabe mencionar que, en cada uno de estos casos, el reemplazo en otra categoría fue forzado por la falta de disponibilidad de manzanas, ya que siempre se buscó una manzana sustituta lo más parecida posible a la original en términos de grupo tamaño y clasificación socioeconómica.

La distribución de estos casos se detalla en la siguiente tabla, considerando únicamente las manzanas reemplazadas finales (es decir, contando solo la manzana final para los casos en donde hubo reemplazos sucesivos).

Tabla 17: Manzanas reemplazadas según macrozona

Macrozona	Manzanas en el marco	Manzanas seleccionadas	Manzanas reemplazadas ⁽¹⁾	Manzanas reemplazadas en otro estrato
Norte	12.179	257	38	4
Centro	20.068	485	113	9
Sur	7.878	173	36	13
Gran Santiago	35.439	1.191	118	39

(1) Considera únicamente manzanas que originalmente se encontraban en la muestra, es decir, no se cuentan manzanas que entraron como reemplazo y luego fueron a su vez reemplazadas. Incluye manzanas reemplazadas en otro estrato.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

En términos formales, la probabilidad de incluir la i -ésima manzana perteneciente a la comuna h , al grupo de clasificación socioeconómica k y al grupo de tamaño g viene dada por:

$$P_{hkg}(i) = \frac{n_{hkg}}{N_{hkg}}, \forall i \in s_1 \quad (8)$$

Donde,

- n_{hkg} Es el número de manzanas seleccionadas en el estrato hkg .
- N_{hkg} Es el número de manzanas en el estrato hkg según el marco muestral actualizado.
- s_1 Es el conjunto de manzanas seleccionadas.

6.1.4.2 Proceso de enumeración de manzanas (UPE)

Una vez que se han seleccionado las manzanas a encuestar, se realizó un proceso de enumeración (o empadronamiento), que consistió en visitar las manzanas seleccionadas y recolectar información de todas las viviendas de estas manzanas con el fin de determinar el uso dado a las edificaciones. Este proceso permite identificar de forma actualizada⁹⁴ aquellas viviendas que son utilizadas para uso habitacional particular, que son finalmente las unidades seleccionadas en la segunda etapa (viviendas particulares ocupadas).

Durante el proceso de enumeración, algunas manzanas todavía podían estar sujetas a cambios (reemplazos) por diversos motivos:

- La estructura del conglomerado en terreno difiere de la enviada en el croquis, debido a que esta pudo haberse fusionado con otra o dividido debido al cambio de límites por cierre o construcción de nuevas calles. Este caso es el más frecuente.
- La manzana puede estar temporalmente destruida porque se van a construir nuevas edificaciones, ya sean casas o edificios.
- El conglomerado no tiene el número de viviendas mínimo requerido para la selección de las viviendas.
- Es imposible ingresar a las viviendas del conglomerado porque están dentro de un edificio que tiene un conserje que les impide el ingreso.
- La mayoría de las viviendas cambiaron de uso o destino a otro que no es de vivienda.
- El conglomerado se encuentra afectado por algún tipo de catástrofe natural como aluviones, terremotos u otros.

Para realizar los reemplazos de las unidades de primera etapa en el proceso de enumeración, al igual que en los procesos previos de reemplazo, se buscaron manzanas que compartían características similares en cuanto a tamaño (número de viviendas) y ubicación (la misma comuna, distrito y zona, o la misma comuna y distrito, o la misma comuna, dependiendo si al nivel más desagregado se encuentra una de tamaño similar).

6.1.4.3 Selección de unidades de segunda etapa (USE - viviendas particulares ocupadas)

Una vez finalizado el proceso de enumeración de manzanas, la selección de las unidades de segunda etapa (viviendas particulares ocupadas) se realizó al interior de las manzanas seleccionadas, en forma aleatoria y de manera sistemática. Lo anterior implica que la probabilidad de selección de unidades de segunda etapa es una probabilidad condicional a la selección de una unidad de primera etapa.

94 Si bien se cuenta con información de las viviendas particulares ocupadas proveniente del Marco Muestral, es necesario verificar esta información de forma previa al proceso de selección con el fin de actualizar la información desajustada debido al paso del tiempo.

Con el objetivo de no concentrar la muestra en manzanas grandes, se definió una cota mínima de cuatro viviendas y una máxima de 30 viviendas a encuestar por conglomerado. Esto permite que manzanas de todos los tamaños estén representadas en la muestra en forma proporcional, a través de un número fijo de viviendas, según el grupo de tamaño al que pertenezcan.

La probabilidad condicional de incluir la j -ésima vivienda de la manzana i que pertenece a la comuna h , al grupo de clasificación socioeconómica k y al grupo de tamaño g viene dada por:

$$P_{hkgi}(j|i) = \frac{m_{hkgi}}{M'_{hkgi}}, \forall j \in s_2 \quad (9)$$

Donde,

m_{hkgi} es el número de viviendas a encuestar en la manzana i , que pertenece al estrato hkg .

M'_{hkgi} es el número de viviendas particulares (actualizado según enumeración) en la manzana i que pertenece al estrato hkg .

s_2 es el conjunto de viviendas seleccionadas.

El algoritmo de selección sistemático de viviendas, que es aplicado también en otras encuestas, se estandariza para que, independientemente de quien aplique el algoritmo y bajo las mismas condiciones de ordenamiento previo, las mismas viviendas sean seleccionadas. Este procedimiento es análogo a introducir una semilla de aleatorización para selecciones aleatorias, las cuales arrojan también las mismas selecciones.

El procedimiento se describe a continuación:

Paso 1

Considerando solo las viviendas particulares identificadas en el proceso de enumeración, se ordenaron los elementos de forma ascendente según las variables región, comuna, grupo de tamaño, distrito censal, zona censal, manzana y número de orden de la vivienda.

Paso 2

Se calculó el período $Q = \frac{M'_{hkgi}}{m_{hkgi}}$, donde las variables siguen la nomenclatura de la ecuación (9).

Paso 3

Se calculó el arranque “A” o primera vivienda seleccionada, como un número aleatorio entre 1 y el período Q . Con el fin de asegurar la réplica de los cálculos, la semilla aleatoria que se estableció en esta selección correspondió a 33.333.333.

Paso 4

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

$$A, (A+Q), (A+2 \cdot Q), (A+3 \cdot Q) \cdots A + (m_i - 1) \cdot Q$$

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

El número de unidades seleccionadas (primera y segunda etapa) se muestra en la siguiente tabla:

Tabla 18: Total de unidades muestrales seleccionadas por nivel de estimación

Unidades Muestrales	Total de capitales regionales	Gran Santiago	Resto de capitales regionales
Manzanas	2.124	1.198	926
Viviendas	25.361	14.705	10.656

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

6.2 Desarrollo de factores de expansión

El factor de expansión se interpreta como la cantidad de unidades en la población que representa una unidad en la muestra.

De acuerdo con el diseño muestral, el factor de expansión se calcula en función del total de manzanas y viviendas seleccionadas en cada estrato, así como también del número de manzanas totales en el estrato (según información reportada por el marco muestral) y del número de viviendas particulares enumeradas.

Al tratarse de un muestreo bietápico, la fórmula de cálculo del factor de expansión depende de las probabilidades de selección de las unidades de primera y segunda etapa.

6.2.1. Ponderador de selección de unidades de primera etapa (manzanas)

6.2.1.1 Probabilidad de selección de manzanas

La probabilidad de incluir cada una de las manzanas (n_{hkg}) en la muestra del estrato hkg está dada por $P_{hkg}(i)$, presentado anteriormente en la ecuación (8).

6.2.1.2 Inverso de las probabilidades de selección

Posterior al cálculo de las probabilidades de selección de manzanas, se calcula el ponderador de selección de unidades de primera etapa, tal como se indica a continuación:

$$w_{hkgi} = \frac{1}{P_{hkg}(i)} = \frac{N_{hkg}}{n_{hkg}}, \forall i \in s_1 \quad (10)$$

Donde,

w_{hkgi} es el ponderador de selección de la i -ésima unidad primaria de muestreo (UPM).

6.2.1.3 Ajuste por omisión de manzanas

El ajuste por omisión de manzanas tiene como finalidad corregir los ponderadores de selección de manzanas según la cantidad de viviendas que se excluyeron de la muestra de la VIII EPF por motivos operativos y presupuestarios.

Tal como se menciona en el apartado 6.1.2.1, previa selección de la muestra de la encuesta, se descartaron *a priori* del marco muestral las unidades primarias de muestreo conformadas por siete viviendas o menos, además de aquellas manzanas pertenecientes a estratos que no contaban con suficientes viviendas para extraer una muestra adecuada.

La tabla 19 muestra la cantidad de manzanas que, según el marco muestral, tienen menos de ocho viviendas o pertenecen al grupo de tamaño "0", lo que representa 8,3% de los conglomerados y 0,9% del total de viviendas presentes en el marco.

Tabla 19: Total conglomerados y viviendas en el marco muestral según grupo tamaño

Grupo tamaño del conglomerado	Total		Porcentaje	
	Conglomerados	Viviendas	Conglomerados	Viviendas
Total	87.690	2.931.070	100	100
Menos de 8	7.265	27.492	8,3	0,9
De 8 a 23	39.102	625.670	44,6	21,3
De 24 a 44	28.468	886.243	32,5	30,2
De 45 a 81	7.664	438.897	8,7	15
De 82 a 154	3.025	334.764	3,4	11,4
Mayor a 154	2.166	618.004	2,5	21,1

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

La siguiente tabla entrega información respecto a aquellas viviendas que fueron excluidas de la muestra dada la omisión de manzanas realizada considerando todas las razones antes mencionadas.

Tabla 20: Cantidad y porcentaje de estratos, conglomerados y viviendas de los estratos del marco de selección de la EPF, según completitud de la muestra y condición de estrato con muestra

Condición de estrato con muestra	Total			Porcentaje		
	Estratos	Conglomerados	Viviendas	Estratos	Conglomerados	Viviendas
Marco sin exclusión de unidades	991	87.690	2.931.070	100	100	100
Marco de selección de unidades VIII EPF	585	75.564	2.661.413	59,0	86,2	90,8
Unidades excluidas de la selección	406	12.126	269.657	41,0	13,8	9,2

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

A continuación, la tabla 21 presenta información respecto a la distribución de la omisión de estratos según los grupos de tamaño de las manzanas, considerando la cantidad de viviendas presentes en cada grupo.

Si se consideran las exclusiones, se observa un porcentaje de omisión de viviendas de 9,2%, siendo los conglomerados que tienen más de 154 viviendas los que presentan mayor porcentaje de omisión de vivienda, con 17,0% (ver tabla 21).

Tabla 21: Cantidad y porcentaje de viviendas del marco de la EPF, según completitud de la muestra y grupo de tamaño por condición de estrato con muestra

Muestra	Grupo tamaño del conglomerado	Estrato con muestra de viviendas			Condición de estrato con muestra (%)		
		Sí	No	Total	Sí (%)	No (%)	Total (%)
Muestra total	Total	2.661.413	269.657	2.931.070	90,8	9,2	100
	Menos de 8	-	27.492	27.492	-	100	100
	De 8 a 23	588.780	36.890	625.670	94,1	5,9	100
	De 24 a 44	840.775	45.468	886.243	94,9	5,1	100
	De 45 a 81	411.876	27.021	438.897	93,8	6,2	100
	De 82 a 154	306.792	27.972	334.764	91,6	8,4	100
	Mayor a 154	513.190	104.814	618.004	83,0	17,0	100

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

En este contexto, utilizar el ponderador de selección de manzanas sin ajustar por la omisión de los conglomerados resultaría en una subestimación del número total de viviendas por manzana. Para evitar lo anterior, se realiza un ajuste que consiste en aumentar los ponderadores de selección de las manzanas de forma proporcional a la cantidad de viviendas excluidas.

Dentro de los escenarios evaluados para definir el nivel al cual se iba a aplicar este ajuste, se consideran las combinaciones de las variables *macrozona*, *clasificación socioeconómica* y *grupo de tamaño del conglomerado*. Dado que la muestra no garantiza representatividad para cada una de las combinaciones posibles de estas variables y considerando que el grupo que contiene las manzanas con menos de ocho viviendas (grupo tamaño "0") fue excluido de la cobertura, se definió realizar el ajuste a nivel de macrozonas. Para determinar el ajuste \hat{R}_l a realizar por macrozona l , se calculó la razón entre M_l , el número total de viviendas en la macrozona l y el número total de viviendas estimadas con el ponderador de selección de la manzana a la cual pertenece:

$$\hat{R}_l = \frac{M_l}{\sum_{i \in \Omega_l} w_{hkgi} \cdot M_{hkgi}}, \forall l = 1, 2, 3, 4 \quad (11)$$

Donde,

M_l es el número de viviendas en la macrozona l .

Ω_l es el conjunto de las manzanas seleccionadas en la macrozona l .

w_{hkgi} es el ponderador de selección de la manzana i del estrato hkg .

M_{hkgi} es el número de viviendas de la manzana i del estrato hkg .

El ponderador de ajuste por omisión de manzanas es el mismo para todas las unidades muestrales pertenecientes a una misma macrozona.

Posteriormente, se calculó un ponderador de selección de manzanas corregido por omisión, w'_{hkgi} , en adelante *ponderador de selección de manzanas corregido*, como:

$$w'_{hkgi} = w_{hkgi} \cdot \hat{R}_i, \forall i \in \Omega_i \quad (12)$$

Los resultados obtenidos, con la muestra de la VIII EPF, para los diferentes niveles de agregación en los cuales se realiza el ajuste se presentan a continuación:

Tabla 22: Total de viviendas contenidas en el marco muestral, estimación de viviendas expandidas en la encuesta y factor de ajuste por omisión de conglomerados según macrozona

Macrozona	Total viviendas contenidas en el marco de selección de la EPF	Estimación del total de viviendas a partir de la muestra de la encuesta	Factor de ajuste por omisión de conglomerado
Total VIII EPF	2.931.070	2.196.319	1,335
Norte	362.266	292.155	1,24
Centro	684.528	553.207	1,237
Sur	227.028	165.075	1,375
Gran Santiago	1.657.248	1.185.882	1,398

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

En la tabla 22, se observa el total de viviendas contenidas en el marco muestral y el total de viviendas expandido por la muestra según macrozona, destacando que el mayor ajuste a realizar para compensar la falta de cobertura del marco se debió hacer en la macrozona Gran Santiago, con un valor de 1,398.

Cabe mencionar que, mientras más alejado esté el factor o ponderador de ajuste de omisión de conglomerados de la unidad, mayor será la distorsión del factor de expansión base o inverso de la probabilidad de selección del conglomerado, por lo que conviene que estos ajustes sean lo más homogéneos posible entre ellos.

6.2.2 Ponderador de selección de unidades de segunda etapa (viviendas particulares ocupadas)

Los pasos anteriormente descritos están vinculados a la probabilidad de selección de las manzanas, no obstante, la última unidad de selección de la muestra son las viviendas particulares ocupadas. En el presente apartado se describen las probabilidades de selección de las viviendas, así como los ajustes y ponderadores asociados a estas.

6.2.2.1 Probabilidad condicional de selección de viviendas

La probabilidad condicional de incluir la j -ésima vivienda de la i -ésima manzana seleccionada en el estrato hkg está dada por

$$P_{hkgi}(j|i) = \frac{m_{hkgi}}{M'_{hkgi}}, \forall j \in s_2 \quad (13)$$

Donde la notación es la misma que la utilizada en la ecuación (9). Al aplicar la definición de probabilidad condicional, se puede determinar la probabilidad $P_{hkgi}(j|i)$ de selección de la vivienda j ubicada en la manzana i del estrato hkg en la muestra:

$$P_{hkgi}(j) = P_{hkg}(i) \cdot P_{hkgi}(j|i) \quad (14)$$

Donde $P_{hkgi}(j)$ corresponde a la probabilidad de selección de la vivienda j .

6.2.2.2 Inverso de las probabilidades de selección

Dada la probabilidad condicional de selección de las viviendas, el inverso está dado por:

$$w_{hkgij} = \frac{1}{P_{hkg}(i) \cdot P_{hkgi}(j|i)} = \frac{1}{P_{hkg}(i)} \cdot \frac{1}{P_{hkgi}(j|i)} = w_{hkgi} \cdot \frac{1}{P_{hkgi}(j|i)} \quad (15)$$

Sin embargo, el ponderador de selección ya cuenta con un ajuste por omisión w'_{hkgi} (ver 6.2.1.3). Al reescribir el ponderador de selección de la vivienda junto con la incorporación del ajuste por omisión de manzanas, la expresión queda:

$$w_{hkgij} = w'_{hkgi} \cdot \frac{1}{P_{hkgi}(j|i)} = w_{hkgi} \cdot \hat{R}_i \cdot \frac{1}{P_{hkgi}(j|i)} = w_{hkgi} \cdot \hat{R}_i \cdot \frac{M'_{hkgi}}{m_{hkgi}} \quad (16)$$

Donde,

w_{hkgi} es el ponderador de selección de manzanas sin corregir.

\hat{R}_i es el ajuste por omisión de conglomerados.

M'_{hkgi} es el número de viviendas particulares (actualizado según enumeración) en la manzana i que pertenece al estrato hkg .

m_{hkgi} es el número de viviendas a encuestar en la manzana i , que pertenece al estrato hkg .

6.2.2.3 Ajuste por elegibilidad

La población objetivo de la VIII EPF está constituida por las personas que residen habitualmente en las viviendas particulares ocupadas. Sin embargo, pese a todos los esfuerzos realizados por mantener marcos muestrales válidos y actualizados para esta población, siempre existe la posibilidad de errores de inclusión (o exclusión) en el marco.

Por ejemplo, fallas en el proceso de enumeración pueden derivar en el registro de inmuebles no elegibles (casas de veraneo, negocios, etc.) en el marco, así como también pueden ocurrir cambios en el uso de la vivienda u otro tipo de situaciones entre el proceso de enumeración y el momento de la entrevista, la cuales afectan la elegibilidad de las viviendas.

Por esta razón, es importante ajustar las probabilidades de selección de las viviendas seleccionadas con el objetivo de ponderar el hecho que una proporción de estas no es elegible y para otras viviendas simplemente se desconoce el estado de su elegibilidad.

Al terminar el trabajo de campo, todas las viviendas seleccionadas inicialmente son clasificadas en tres grandes grupos:

- **Elegibles:** edificaciones identificadas como viviendas particulares ocupadas (incluye tanto viviendas entrevistadas como no entrevistadas).
- **No elegibles:** edificaciones identificadas como negocios, viviendas colectivas, viviendas deshabitadas, viviendas de veraneo, viviendas destruidas, entre otros.

- **De elegibilidad desconocida:** edificaciones cuyo estado no se pudo determinar. Este es el caso, por ejemplo, de unidades que nunca fueron enviadas a terreno, viviendas que no pudieron ser encontradas, entre otros.

La tabla 23 muestra la distribución de viviendas según los tres grandes grupos de clasificación asociados a su clasificación de trabajo de campo.

Tabla 23: Distribución de viviendas seleccionadas según elegibilidad en el trabajo de campo

Elegibilidad de viviendas	Frecuencia	Porcentaje
Total	25.361	100
No elegible	1.327	5,2
Elegible	23.901	94,2
Elegibilidad desconocida	133	0,5

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

El ponderador de selección de viviendas tiene valores válidos para las viviendas elegibles, no elegibles y de elegibilidad desconocida. A continuación, se describen los ajustes aplicados a estas categorías de viviendas.

6.2.2.3.1 Ajuste por elegibilidad desconocida

Inicialmente, todas las viviendas tenían un estatus conocido de elegibilidad: son o no viviendas particulares ocupadas. La muestra de viviendas seleccionadas incluye, en principio, solo viviendas elegibles, a saber, viviendas particulares ocupadas.

Sin embargo, durante el trabajo de campo este estado no se puede comprobar en algunas viviendas, las que terminan siendo clasificadas como de elegibilidad desconocida. En el caso de la VIII EPF, estas unidades correspondieron a 133 casos, 0,5% de las viviendas seleccionadas.

El peso muestral de las viviendas seleccionadas que en un principio fueron consideradas como elegibles y cuyo estado final cambia al de elegibilidad desconocida fue redistribuido proporcionalmente al interior de cada comuna entre las viviendas elegibles y no elegibles (estados de disposición conocidos según el trabajo de campo), con la finalidad de no subestimar el número total de viviendas de la población.

La proporción de los pesos a redistribuir se calculó como la razón entre dos estimaciones realizadas en base al ponderador de selección de viviendas (w_{hkgij}). 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. Ambas estimaciones se realizaron con el ponderador de selección de viviendas, esto es:

$$\hat{R}_{h, know} = \frac{\sum_{i \in \Omega_h} \sum_{j \in \Theta_i} w_{hkgij}}{\sum_{i \in \Omega_h} \sum_{j \in \Theta_{i, know}} w_{hkgij}} \quad (17)$$

Donde,

Ω_h es el conjunto de manzanas seleccionadas en la comuna h .

Θ_i es el conjunto de viviendas seleccionadas en la manzana i .

$\Theta_{i, know}$ es el conjunto de viviendas seleccionadas en la manzana i clasificadas como elegibles o no elegibles.

w_{hkgij} es el ponderador de selección de la vivienda j , de la manzana i y del estrato hkg .

La razón antes descrita ($\hat{R}_{h, know}$) es la misma para todas las unidades muestrales pertenecientes a una misma comuna.

Por último, el ponderador de selección de viviendas corregido por elegibilidad desconocida, w'_{hkgij} , está dado por:

$$w'_{hkgij} = w_{hkgij} \cdot \hat{R}_{h, know} \quad (18)$$

6.2.2.3.2 Ajuste por no elegibilidad

Según se aprecia en la tabla 23, del total de las 25.361 viviendas seleccionadas con sobremuestreo para la VIII EPF, 1.327 viviendas resultaron no elegibles (un 5,2% del total de viviendas), 133 viviendas presentaron elegibilidad desconocida (0,5% del total de viviendas) y 23.901 se declararon elegibles (94,2% del total de viviendas).

El objetivo analítico de la VIII EPF es producir inferencia sobre la población que reside en las 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.).

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

6.2.3 Ajuste por no respuesta

En esta etapa solo son consideradas las viviendas elegibles, pues son estas las que cumplen con el requisito de ser viviendas particulares ocupadas. Sin embargo, aun cuando dichas viviendas cumplan con los atributos necesarios, es posible que los residentes de algunas de ellas deseen no participar en el estudio.

Para efectos de la VIII EPF, la no respuesta total se puede definir como la pérdida total de la información de las unidades de análisis, en este caso, los hogares que residen en las viviendas particulares ocupadas seleccionadas en la muestra. Esta pérdida total de respuesta se denomina *no respuesta a la unidad*, que se diferencia de la *no respuesta al ítem*, (valores faltantes) en cuyo caso se obtiene una respuesta parcial de ciertas preguntas del cuestionario.

Una importante distinción utilizada en la EPF⁹⁵ se relaciona con la determinación de si la información entregada por cada hogar encuestado cumple con ciertos niveles mínimos de

95 Si bien esta distinción también es utilizada por otras encuestas, dado que la EPF es una encuesta con múltiples visitas y entrevistas, la categoría *break-off*, a diferencia de otras encuestas, representa parte importante de hogares que rechazan participar en el estudio.

calidad para ser incluido en los cálculos de los estimadores de la encuesta, permitiendo establecer la distinción entre hogares entrevistados con información completa o parcial (clasificados como entrevistados) y aquellos hogares que deben ser clasificados como rechazo al entregar información insuficiente respecto a los objetivos del estudio (*break-off*⁹⁶).

La tabla 24 muestra los códigos de disposición final obtenidos luego de desarrollar el trabajo de campo de la VIII EPF para las 25.361 viviendas seleccionadas con sobremuestreo. Del total de viviendas seleccionadas, 4.986 viviendas rechazaron participar en el estudio (R+BO).

Tabla 24: Clasificación de viviendas según códigos de disposición final de la VIII EPF

Código de disposición final de viviendas	Cantidad de viviendas
Elegibles	23.901
Entrevistada (E)	14.999
Interrumpida o <i>break-off</i> (BO) ⁽¹⁾	3.167
Rechazo (R)	1.819
Otros contactados (OC)	3.020
No contactados (NC)	896
Elegibilidad desconocida (ED)	133
No elegible (NE)	1.327
Total de viviendas con sobremuestra	25.361

(1) Si bien la categoría *Break-off* es una subcategoría del rechazo, para efectos de análisis de la encuesta, se presentan desagregados.
Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Considerando los indicadores de desempeño operativo, el 20,75% de las viviendas elegibles decidió rechazar participar en el estudio⁹⁷, ya sea entregando algo de información (BO) o bien sin entregar información (R).

96 (Aapor, 2015, pág. 23). Para más información respecto a la grilla técnica de mínimos de calidad de la VIII EPF, consultar el apartado 9.4.1.1.

97 Para efectos del cálculo de las viviendas elegibles, como criterio conservador, se tratan a las viviendas de elegibilidad desconocida como si hubiesen sido unidades elegibles, incorporándose a esta categoría.

Tabla 25: Indicadores de desempeño operativo a nivel de viviendas VIII EPF

Indicador	Cálculo del indicador	Valor VIII EPF (%)
Desactualización del marco muestra	$\frac{(ED + NE)}{(E + BO + R + OC + NC + ED + NE)}$	5,76
Tasa de contacto	$\frac{(E + BO + R + OC)}{(E + BO + R + OC + NC + ED)}$	95,72
Tasa de cooperación	$\frac{E}{(E + BO + R + OC)}$	65,2
Tasa de rechazo	$\frac{(BO + R)}{(E + BO + R + OC + NC + ED)}$	20,75
Tasa de respuesta	$\frac{E}{(E + BO + R + OC + NC)}$	62,41

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

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, por lo que una parte de la población no sería cubierta.

Si no hay relación entre el patrón de la no respuesta y la variable de interés, es posible que la no respuesta solo afecte la precisión (varianza) de las estimaciones. Sin embargo, dado que existe evidencia respecto a que la no respuesta se relaciona con características especiales de la población (como el nivel de ingresos), no realizar un ajuste por no respuesta podría sesgar las estimaciones obtenidas a partir de la muestra.

Si bien es imposible evitar la no respuesta total, en la práctica se recurre a estrategias para mitigarla. Una forma de asegurar la cantidad de respuestas necesarias para asegurar ciertos niveles de precisión de las estimaciones es aplicar un sobremuestreo de unidades con respecto a la tasa de respuesta de la misma encuesta, tal como se especificó en el apartado 6.1.3.1 de cálculo del tamaño de la muestra.

Existen, además, otras formas de corregir la no respuesta, que consisten en generar grupos de individuos homogéneos en cuanto a características disponibles, asociadas a los mismos sujetos o a unidades de muestreo superiores, tanto para los que responden como para los que no responden (por ejemplo, sexo, edad, lugar de residencia, nivel de ingresos de la comuna, etc.). De esta manera, se busca que, al interior de estos grupos, las unidades que no responden sean representados por las que sí responden, mediante un ajuste por no respuesta.

En el caso de la VIII EPF, para minimizar los problemas ocasionados por la no respuesta a la unidad, se realiza un ajuste para compensar la pérdida de unidades, denominado propensity score. Este método de ajuste consiste en estimar la probabilidad de respuesta de las viviendas que responden y de aquellas que no responden, y es un procedimiento común para el control de la no respuesta, propuesto en Little (1986) y más recientemente revisado en Brick (2013). Cabe destacar que al no contar con información de los residentes

de las viviendas donde no se realizan las entrevistas, ni tampoco de las características de estas, resulta imposible realizar el ajuste a nivel de viviendas.

Si se contara con información de las viviendas que no responden la encuesta (o de sus habitantes) sería posible modelar la probabilidad de responder de cada vivienda utilizando una regresión de variable dependiente dicotómica (por ejemplo, a través de modelos *probit* o *logit*)⁹⁸. Sin embargo, en la práctica solo es factible conocer cuántas viviendas no responden en cada manzana seleccionada para la encuesta, por lo tanto, los ajustes de no respuesta que se aplican a la encuesta se realizan a dicho nivel.

Para realizar el ajuste, el primer paso consistió en estimar la proporción de viviendas que respondieron en cada manzana. Luego, se utilizó un modelo de regresión lineal para predecir la tasa de respuesta de la encuesta. El modelo consideró un conjunto de variables auxiliares las cuales contenían información a nivel de manzanas, distrito censal, así como también información a nivel de comunas. La información utilizada proviene de otras encuestas o de registros administrativos disponibles, para así predecir la tasa de respuesta al interior de las manzanas.

La selección del modelo de estimación de la tasa de respuesta requirió comparar diferentes especificaciones lineales, dada la imposibilidad de obtener variables de respuesta a nivel de vivienda. De esta forma, se estimaron distintas regresiones mediante el método de Mínimos Cuadrados Ordinarios (MCO). Cabe destacar que en todas las especificaciones la variable dependiente correspondió a la tasa de respuesta a nivel de manzana. Los criterios de selección del modelo utilizados en esta ocasión son el R^2 -ajustado, el factor de inflación de la varianza (VIF) y los criterios de Akaike (AIC) y Schwarz (BIC)⁹⁹. Finalmente, el modelo seleccionado correspondió a:

$$TR_i = \beta_0 + \beta_1 \cdot \log(AF_i) + \beta_2 \cdot TALF_i + \beta_3 \cdot d_{ROBO_i} + \beta_4 \cdot REGION_i + u_i \quad (19)$$

Donde,

TR_i es la tasa de respuesta de la manzana i , definida como $(viviendas\ logradas_i / viviendas\ totales_i)$.

$\log(AF_i)$ es el logaritmo del avalúo fiscal de la manzana i .

$TALF_i$ es la tasa de alfabetización de la manzana i . Esta variable está a nivel comunal y se imputa su valor a todas las manzanas de la comuna.

d_{ROBO_i} es una variable dicotómica que toma valor 1 si la manzana i posee un nivel alto de robo con intimidación y 0 en otro caso.

d_{REGION_i} es la región a la cual pertenece la manzana, usando como categoría base a la Región Metropolitana.

u_i es el término de error aleatorio.

98 Otra forma de estimar probabilidades de responder, como alternativa de los modelos logit y probit, es la aplicación de modelos predictivos mediante árboles de decisiones, redes neuronales, entre otros.

99 Se consideran como criterios el R^2 -ajustado más alto, VIF menor a diez y menor en comparación a otros modelos, y un AIC y BIC menores.

Una vez que se determinó la probabilidad de contestar la encuesta de cada manzana, estas fueron ordenadas de forma creciente de acuerdo a sus tasas de respuesta predichas, agrupándose las unidades en veintiles ordenados (20 grupos de igual tamaño), en donde las viviendas pertenecientes a una misma manzana son asignadas al mismo veintil.

En resumen, la determinación de las *celdas de ajuste* se realizó en tres pasos:

1. Selección y estimación del modelo de predicción de la tasa de respuesta de manzanas.
2. Predicción de la tasa de respuesta para cada manzana.
3. Ordenamiento creciente de las tasas predichas y creación de veintiles o *celdas de ajuste*.

Si el modelo de predicción es adecuado, al interior de cada celda de ajuste las viviendas serán similares en cuanto a la variable 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 la encuesta deben ser redistribuidos entre las viviendas elegibles que sí lo hicieron. Para llevar a cabo esto, se calculó la razón de ajuste por no respuesta.

La determinación de la razón de ajuste se realizó al interior de cada veintil utilizando el ponderador de selección de viviendas corregido por elegibilidad (w'_{hkgij}). En el numerador se incluye la estimación del total de viviendas elegibles y en el denominador se incluye la estimación del total de viviendas elegibles que responden:

$$\hat{R}_{v,R} = \frac{\sum_{i \in \Phi_v} \sum_{j \in \Theta_{i,eleg}} w'_{hkgij}}{\sum_{i \in \Phi_v} \sum_{j \in \Theta_{i,eleg,R}} w'_{hkgij}} \quad (20)$$

Donde,

- v es el índice de la celda de ajuste o veintil.
- Φ_v es el conjunto de manzanas asignadas a la celda de ajuste v .
- $\Theta_{i,eleg}$ es el conjunto de viviendas pertenecientes a la manzana i y catalogadas como elegibles.
- $\Theta_{i,eleg,R}$ es el conjunto de viviendas pertenecientes a la manzana i , catalogadas como elegibles y que responden la encuesta.
- w'_{hkgij} es el ponderador de selección de viviendas ajustado por elegibilidad desconocida y no elegibilidad.

Tabla 26: Ponderador de no respuesta por celda de ajuste

Celda de ajuste de no respuesta o veintiles	Componentes del ponderador de corrección por no respuesta			
	Ponderador de selección de viviendas promedio	Viviendas elegibles	Viviendas que responden (R)	$\hat{R}_{v,R}$
Total	122,13	23.901	14.999	
1	145,05	1.454	638	2,34
2	148,37	1.745	916	1,79
3	110,67	1.364	773	1,73
4	99,55	1.242	697	1,78
5	117,5	1.232	705	1,82
6	114,3	1.327	760	1,70
7	105,58	944	543	1,66
8	117,15	1.122	634	1,80
9	126,14	1.175	697	1,65
10	127,03	1.044	645	1,66
11	75,92	1.162	713	1,68
12	122,51	1.127	696	1,56
13	138,77	1.156	753	1,55
14	118,27	1.304	840	1,54
15	106,48	1.113	792	1,36
16	113,57	1.045	757	1,35
17	122,39	1.055	785	1,29
18	138,69	1.114	858	1,28
19	135,87	1.062	817	1,32
20	143,74	1.114	980	1,10

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

La última columna de la tabla 26 muestra el valor del ajuste por cada veintil o celda. Cabe destacar que el valor del ajuste es decreciente con respecto a las celdas de ajuste construidas, es decir, en general (aunque se puede notar que no ocurre en todos los veintiles) los mayores ajustes se concentran en los veintiles más bajos (aquellos que tienen una menor tasa de respuesta predicha) y los menores ajustes se concentran en las celdas caracterizadas por una tasa de respuesta más alta.

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

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

Cabe destacar que este ponderador se asigna a las viviendas elegibles entrevistadas. A las viviendas elegibles no entrevistadas, se les asigna un valor en blanco (*missing value*).

6.2.4 Ajuste de calibración

Tal como se ha revisado, las unidades de selección de la encuesta son viviendas particulares ocupadas y hasta el momento, los factores de expansión de estas permiten obtener una estimación del total de viviendas particulares ocupadas. No obstante, si bien se seleccionan viviendas, las unidades de análisis son los hogares¹⁰⁰, que a su vez están compuestos por personas que residen en las viviendas seleccionadas. Por lo tanto, se requiere incorporar ajustes a los factores de expansión que permita ajustar la población presente en la encuesta a los *stocks* poblacionales.

De esta manera, si se aplica el factor de expansión ajustado por no respuesta y se asigna a cada persona dentro de la base de datos, se obtiene una estimación del total de personas. Sin embargo, este valor difiere de las estimaciones provenientes de otras fuentes, como las proyecciones de población generadas a partir de modelos demográficos.

Con el objetivo de ajustar la población de la encuesta a los *stocks* poblacionales calculados por la institución, para la VIII EPF, se realizó un ajuste o calibración al *stock* poblacional mediante la desagregación de las variables región, sexo y tramo etario (de 0 a 14 años, de 15 a 59 años y de 60 años o más).

Debido a la necesidad de obtener un solo factor de expansión a nivel vivienda, es decir, que todas las personas al interior de una vivienda compartan el factor de expansión, y al mismo tiempo obtener un factor calibrado a los *stocks* poblacionales según región, sexo y tramo etario, se utiliza el método de *raking*, que corresponde a un método iterativo que asegura que la expansión de la muestra alcance ciertos totales poblacionales, sujeto a minimizar la distancia entre el factor calibrado y el factor previo.

El método de calibración específico utilizado es uno de los propuestos en Deville y Särndal (1992), que entrega diversas funciones de distancia a minimizar, de las cuales se escoge la siguiente, denominada *método logístico*:

$$F(w_k, d_k) = w_k \log(w_k / d_k) - w_k + d_k \quad (22)$$

Donde,

w_k es el factor calibrado perteneciente a la unidad k ,

d_k es el factor previo perteneciente a la unidad k .

Si bien los autores proponen una serie de posibles formulaciones de la función objetivo, se selecciona aquella que garantiza que el algoritmo siempre arrojará una solución, y esta será positiva. Por otra parte, la implementación del algoritmo requiere utilizar las variables región, sexo y tramo etario como variables cuantitativas, es decir, se contabiliza a nivel de vivienda la cantidad de hombres y mujeres pertenecientes a cada tramo de edad, en cada región, lo cual origina 90 variables (seis para cada región), que son las que se busca asegurar que sumen el total poblacional tras la calibración.

Para poder realizar este ajuste al factor de expansión asociado a la muestra de la VIII EPF, la institución preparó proyecciones poblacionales estimadas al 31 de diciembre de

¹⁰⁰ Los hogares son unidades dinámicas y cambiantes y por razones operativas, en las encuestas de hogares del INE, se llega a ellos a través de las estructuras en las cuales residen, a saber, las viviendas.

2016 (en base al Censo 2002), que corresponde al mes central del período de trabajo de campo oficial de la VIII EPF; cabe mencionar que dichas proyecciones consideran solo las 63 comunas urbanas investigadas en este estudio. Es por ello que el stock poblacional al que se ajusta el factor corresponde a 11.095.466. Más detalles se exponen en la tabla 27.

Tabla 27: Proyecciones de población al 31 diciembre 2016 por sexo y tramo etario según región de las comunas urbanas consideradas en la VIII EPF

Región	Total	Hombres				Mujeres			
		Menor a 15	15-59	60 o más	Total	Menor a 15	15-59	60 o más	Total
Total	11.095.466	1.148.151	3.507.833	722.058	5.378.042	1.117.230	3.618.153	982.041	5.717.424
Arica y Parinacota	151.492	17.435	39.049	13.168	69.652	16.497	48.943	16.400	81.840
Tarapacá	298.070	37.091	85.526	16.836	139.453	36.514	101.019	21.084	158.617
Antofagasta	404.755	46.889	140.538	21.681	209.108	45.585	124.845	25.217	195.647
Atacama	171.317	20.411	54.718	10.395	85.524	19.844	54.547	11.402	85.793
Coquimbo	453.993	50.140	140.825	29.863	220.828	47.842	147.752	37.571	233.165
Valparaíso	949.540	88.279	299.894	73.152	461.325	85.079	301.378	101.758	488.215
Metropolitana	6.224.057	634.684	1.967.136	406.103	3.007.923	619.185	2.035.777	561.172	3.216.134
O'Higgins	252.532	25.434	81.962	18.534	125.930	24.450	78.772	23.380	126.602
Maule	252.895	26.692	79.392	16.132	122.216	25.799	82.544	22.336	130.679
Biobío	999.024	99.339	315.416	65.106	479.861	96.104	332.470	90.589	519.163
La Araucanía	357.950	37.871	113.511	18.682	170.064	37.556	121.734	28.596	187.886
Los Ríos	152.949	15.438	47.934	9.663	73.035	14.771	50.564	14.579	79.914
Los Lagos	244.860	28.506	83.516	11.005	123.027	28.329	79.913	13.591	121.833
Aysén	58.437	7.214	18.637	3.231	29.082	7.103	18.390	3.862	29.355
Magallanes	123.595	12.728	39.779	8.507	61.014	12.572	39.505	10.504	62.581

(1) Se consideran las proyecciones a 15 regiones, dado que a este nivel es cuando se incluye menos dispersión al realizar el ajuste por calibrado.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Tras haber evaluado diferentes escenarios según nivel al cual realizar el ajuste por calibrado al *stock* poblacional (a saber: regional a 16, regional a 15, mesozona¹⁰¹, macrozona y zona) se opta por elegir la opción regional a 15 debido a que fue el escenario en el cual el factor de expansión tuvo menor dispersión en cuanto a dicho ajuste.

Lo anterior permite generar un factor de expansión calibrado a nivel de viviendas y de personas que corresponde a:

$$w_{hkgij}^{cal} = w_{hkgij}^{NR} \cdot calib_{hkgij}^{rse} \quad (23)$$

Donde,

101 Las mesozonas constituyen un nivel de agregación intermedio entre región y macrozona, en donde se agrupan las regiones que tienen características similares, conformadas solo para fines de contrastar el impacto de calibrar a diferentes niveles de desagregación.

- w_{hkgij}^{cal} es el factor de expansión calibrado de la vivienda j , de la manzana i , perteneciente al estrato hkg .
- w_{hkgij}^{NR} es el factor de expansión ajustado por no respuesta.
- $calib_{hkgij}^{rse}$ es el ponderador de ajuste de calibración obtenido por el método de *raking*, asociado a cada persona de la vivienda j , de la manzana i , perteneciente al estrato hkg , realizado a nivel regional r , por sexo s (mujer u hombre) y tramo etario e .

6.3 Estimadores y medidas de precisión de la encuesta

Definidas las características del diseño muestral, se presentan a continuación las fórmulas utilizadas para la estimación de gastos e ingresos de los hogares con sus respectivas varianzas.

6.3.1 Estimador de totales de gastos e ingresos en hogares

La estimación del total del gasto (o ingreso) de los hogares se obtiene multiplicando el valor del gasto (o ingreso) en cada hogar por su factor de expansión, y luego sumando sobre todos los hogares de la muestra por nivel de estimación. Esto es:

Gasto total en la categoría de gasto tipo c :

$$\hat{G}_c^{n_e} = \sum_h \sum_k \sum_g \sum_i \sum_j \sum_l w_{hkgijl}^{cal} \cdot g_{c\ hkgijl}, \quad \forall l \in \{n_e\} \quad (24)$$

Donde,

$\hat{G}_c^{n_e}$ es el gasto total de los hogares del nivel de estimación n_e , en la categoría de gasto tipo c .

$g_{c\ hkgijl}$ es el valor del gasto tipo c medido en el hogar l de la vivienda j de la manzana i perteneciente al estrato hkg .

w_{hkgijl}^{cal} es el factor de expansión calibrado del hogar l en la vivienda j , de la manzana i , perteneciente al estrato hkg .

$\{n_e\}$ conjunto de todos los hogares que pertenecen al nivel de estimación n_e .

$\{c\}$ conjunto de hogares que presentan la característica c .

6.3.2 Estimador de gastos e ingresos promedio en hogares

La estimación del gasto (o ingreso) medio por hogar se obtiene dividiendo la estimación del gasto o ingreso total en la categoría c del área de estimación n_e , por la estimación de hogares del nivel de estimación n_e , incluso aquellos que no incurren en gastos (o ingreso) de tipo c . Esto es:

$$\hat{\bar{G}}_c^{n_e} = \frac{\hat{G}_c^{n_e}}{\hat{H}^{n_e}} \quad (25)$$

Siendo:

$$\hat{H}^{n_e} = \sum_h \sum_k \sum_g \sum_i \sum_j \sum_l w_{hkgijl}^{cal} \cdot hog_{hkgijl}^{n_e}, \quad \forall l \in \{n_e\} \quad (26)$$

Donde,

$$\hat{H}^{n_e}$$

es la estimación del total de hogares en el nivel de estimación n_e .

es una variable dicotómica (tipo *dummy*)¹⁰² asociada al hogar l de la vivienda j , de la manzana i , perteneciente al estrato hkg en el nivel de estimación n_e . Se define de la siguiente forma:

$$hog_{hkgijl}^{n_e} = \begin{cases} 1, & \forall l \in \{n_e\} \\ 0, & \forall l \notin \{n_e\} \end{cases}$$

Alternativamente, se puede calcular el promedio de gastos por hogar considerando solo a los hogares que realizan el gasto tipo c del nivel de estimación n_e , en cuyo caso, el denominador de la ecuación (25) se expresa como:

$$\hat{H}_c^{n_e} = \sum_h \sum_k \sum_g \sum_i \sum_j \sum_l w_{hkgijl}^{cal} \cdot hog_{c hkgijl}^{n_e}, \quad \forall l \in \{n_e\} \quad (27)$$

Donde,

$$\hat{H}_c^{n_e}$$

es la estimación del total de hogares que realizan gasto tipo c en el nivel de estimación n_e .

$$hog_{c hkgijl}^{n_e}$$

es una variable *dummy* asociada al hogar l , de la vivienda j , de la manzana i , perteneciente al estrato hkg en el nivel de estimación n_e que realiza gasto tipo c . Se define de la siguiente manera:

$$hog_{c hkgijl}^{n_e} = \begin{cases} 1, & \forall l \in \{n_e \cap c\} \\ 0, & \forall l \notin \{n_e \cap c\} \end{cases}$$

6.3.3 Estimador de proporciones

La estimación de la proporción del gasto por hogar en la categoría de gasto tipo c en relación al gasto total por hogar, se obtiene dividiendo la estimación del gasto (o ingreso) en la categoría de gasto tipo c por la estimación del gasto (o ingreso) total en el nivel de estimación. Esto es:

$$\hat{G}_c^{n_e} = \frac{\hat{G}_c^{n_e}}{\hat{H}^{n_e}} \quad (28)$$

Siendo:

102 Las variables *dummy* o ficticias corresponden a variables dicotómicas o binarias que toman los valores 0 y 1. Para definir las se debe establecer a qué acontecimiento se le asigna el valor 1, y a cuál el 0.

$$\hat{G}^{n_e} = \sum_h \sum_k \sum_g \sum_i \sum_j \sum_l w_{hkgijl}^{cal} \cdot g_{hkgijl}, \quad \forall l \in \{n_e\} \quad (29)$$

Donde,

\hat{G}^{n_e} la estimación del gasto (o ingreso) total en el nivel de estimación n_e .

g_{hkgijl} es el valor del gasto (o ingreso) medido en el hogar l de la vivienda j de la i manzana perteneciente al estrato hkg en el nivel de estimación n_e .

6.3.4 Estimador de gastos e ingresos *per cápita* en hogares

La estimación del gasto *per cápita* de los hogares se obtiene multiplicando el valor del gasto *per cápita* de cada hogar por su factor de expansión, y luego sumando sobre todos los hogares de la muestra por nivel de estimación. Esto es:

$$\hat{G}_c^{n_e PC} = \sum_h \sum_k \sum_g \sum_i \sum_j \sum_l w_{hkgijl}^{cal} \cdot g_c^{n_e PC}{}_{hkgijl}, \quad \forall l \in \{n_e\} \quad (30)$$

Donde,

$\hat{G}_c^{n_e PC}$ la estimación del gasto total (o ingreso) *per cápita* tipo c en el nivel de estimación n_e .

$g_c^{n_e PC}{}_{hkgijl}$ es el gasto tipo c *per cápita* medido en el hogar l de la vivienda j de la manzana i perteneciente al estrato hkg en el nivel de estimación n_e , que se calcula dividiendo el gasto tipo c medido en el hogar por el total de personas en el mismo.

6.3.5 Estimador de gastos e ingresos *per cápita* promedio de los hogares

La estimación del gasto medio *per cápita* por hogar se obtiene dividiendo la estimación del gasto o ingreso total *per cápita* en la categoría de gasto c del nivel de estimación n_e , por la estimación de hogares del nivel de estimación n_e , incluso aquellos que no incurren en gastos de tipo c . Esto es:

$$\hat{\bar{G}}_c^{n_e PC} = \frac{\hat{G}_c^{n_e PC}}{\hat{H}^{n_e}} \quad (31)$$

Al igual que en el caso anterior, en el numerador, se puede calcular el promedio de gastos por hogar considerando solo a los hogares que realizan el gasto tipo c del nivel de estimación n_e :

$$\hat{\bar{G}}_c^{n_e PC} = \frac{\hat{G}_c^{c PC}{}_{n_e}}{\hat{H}_c^{n_e}} \quad (32)$$

6.3.6 Medidas de precisión

Actualmente los objetivos analíticos de las encuestas van mucho más allá de la generación de cuadros de resumen básicos o totales de los parámetros de interés. Cada vez es más frecuente que, además de obtener estimaciones de los parámetros, se requiera además plantear hipótesis, comprobar o construir modelos y, hasta realizar evaluaciones de

impacto. Es por esto que cobra relevancia el hecho de que los resultados incluyan medidas adecuadas de precisión o exactitud de las estimaciones derivadas de los datos de las encuestas.

Una de las medidas claves de precisión en las encuestas por muestreo es la varianza de muestreo, un indicador de la variabilidad que se produce al elegir una muestra, en lugar de confeccionar una lista de toda la población. Otras formas de medir el error de muestreo son a través del error estándar, el coeficiente de variación y el efecto de diseño (ONU, 2009)¹⁰³; sin embargo, cabe mencionar que guardan entre sí relación algebraica, por lo que es posible deducir la expresión de cualquier de ellas a partir de las otras mediante la aplicación de operaciones matemáticas.

6.3.6.1 Varianza del estimador

Para el cálculo de la varianza de los estimadores, en la VIII EPF se utiliza el factor de expansión calibrado w_{hkgij}^{cal} , es decir, el factor que incluye todos los ajustes realizados a los ponderadores de selección (descritos en el apartado anterior). Los *softwares* estadísticos como SAS, SPSS, Stata y R actualmente poseen módulos para el cálculo de estimaciones, varianzas y coeficientes de variación, incluyendo los efectos de diseño cuando se trata de muestreo complejo. Sin embargo, para una correcta estimación de la varianza asociada a una estimación de algún parámetro de interés, es necesario cerciorarse que los estratos de muestreo y los conglomerados cumplan ciertas características. Este tema es abordado en apartado 6.3.6.1.2.

6.3.6.1.1 Estimación de la varianza

6.3.6.1.1.1 Varianza del total estimado del gasto o ingreso en la categoría de gasto tipo *c*

La varianza de un estimador de totales se obtiene mediante un estimador de razón $\left(\hat{R}_{c\ hkg}^{n_e} = \frac{\hat{G}_c^{n_e}}{\hat{H}_{hkg}^{n_e}} \right)$, sumando todos los hogares que realizan el gasto (o ingreso) tipo *c*, en el

estrato *hkg* y área de estimación n_e :

$$\hat{V}\left(\hat{G}_c^{n_e}\right) = \sum_i \frac{n_e}{(n_e - 1)} \cdot \left[w_{hkgij}^{cal} \cdot \left(g_{c\ hkgi}^{n_e} - \hat{R}_{c\ hkg}^{n_e} \cdot hog_{hkgi}^{n_e} \right) \right]^2, \quad \forall i \in \{n_e\} \quad (33)$$

Donde,

$\hat{G}_c^{n_e}$ es la estimación del gasto (o ingreso) tipo *c* en el nivel de estimación n_e .

$\hat{R}_{c\ hkg}^{n_e} = \frac{\hat{G}_c^{n_e}}{\hat{H}_{hkg}^{n_e}}$ es la razón entre la estimación del gasto tipo *c* y la estimación de hogares a nivel de estrato *hkg*; en el nivel de estimación n_e .

$hog_{hkgi}^{n_e}$ es el número de hogares en la manzana *i* en el estrato *hkg* en el nivel de estimación n_e .

103

Para más detalle consultar el capítulo 7 de Naciones Unidas (2009).

6.3.6.1.1.2 Varianza del promedio estimado del gasto o ingreso en la categoría de gasto tipo c

Considerando el promedio estimado del gasto como la razón entre el gasto total estimado en la categoría de gasto tipo c del estrato hkg en el nivel de estimación n_e dividido por el número total de hogares estimado en el estrato hkg en el nivel de estimación n_e , $\hat{R}_{c\ hkg}^{n_e} = \frac{\hat{G}_{c\ hkg}^{n_e}}{\hat{H}_{hkg}^{n_e}}$.

La varianza de la razón para cada nivel de estimación n_e se obtiene de acuerdo con la siguiente expresión:

$$\hat{V}(\hat{R}_{c\ hkg}^{n_e}) = \sum_i \frac{1}{(\hat{H}_{hkg}^{n_e})^2} \cdot \frac{n_e}{(n_e - 1)} \cdot \left[w_{hkgij}^{cal} \cdot \left(g_{c\ hkgi}^{n_e} - \hat{R}_{c\ hkg}^{n_e} \cdot hog_{hkgi}^{n_e} \right) \right]^2, \quad \forall i \in \{n_e\} \quad (34)$$

6.3.6.1.1.3 Varianza de la proporción estimada del gasto o ingreso en la categoría de gasto tipo c respecto del total estimado del gasto (o ingreso) total

Considerando la proporción estimada del gasto como la razón (o tasa) entre el gasto estimado en la categoría de gasto tipo c en el nivel de estimación n_e dividido por el gasto total estimado, $\hat{R}_{c\ hkg}^{n_e\ prop} = \frac{\hat{G}_{c\ hkg}^{n_e}}{\hat{G}_{hkg}^{n_e}}$, la varianza de la razón se obtiene según la siguiente expresión:

$$\hat{V}(\hat{R}_{c\ hkg}^{n_e\ prop}) = \sum_i \frac{1}{(\hat{H}_{hkg}^{n_e})^2} \cdot \frac{n_e}{(n_e - 1)} \cdot \left[w_{hkgij}^{cal} \cdot \left(g_{c\ hkgi}^{n_e} - \hat{R}_{c\ hkg}^{n_e\ prop} \cdot g_{hkgi}^{n_e} \right) \right]^2, \quad \forall i \in \{n_e\} \quad (35)$$

6.3.6.1.2 Variables que identifican el diseño muestral complejo

Al igual que versiones anteriores, la VIII EPF posee un diseño muestral probabilístico, estratificado y bietápico, es decir, un diseño muestral complejo. Bajo la estratificación, en un muestreo bietápico las unidades muestrales de primera etapa se suelen llamar conglomerados, que son grupos de unidades muestrales de segunda etapa (o viviendas) a partir de las cuales se obtiene información para estimaciones de la variable de interés. En un diseño muestral estratificado, la estimación de la varianza total se obtiene mediante la suma de las varianzas estimadas en cada estrato y la varianza de cada estrato se estima a partir de la variabilidad de la característica de interés entre los conglomerados.

En la VIII EPF, un estrato está compuesto por la combinación de las variables *comuna*, *clasificación socioeconómica* y *grupo de tamaño*. Por su parte, los conglomerados se identifican por medio de las *manzanas*. En total, para el trabajo de campo se consideraron 584 estratos y 2.094 conglomerados.

En un muestreo estratificado para que la estimación de la varianza sea factible, se deben seleccionar como mínimo dos conglomerados por estrato. Adicionalmente, para evitar estimaciones de varianza volátiles, se recomienda seleccionar un número mínimo de unidades muestrales (es decir, viviendas) por conglomerado.

En la práctica y por razones del diseño, y algunas veces por falta de respuesta, se dan casos de estratos compuestos por solo un conglomerado, situación recurrente en encuestas de hogares. Existen tres formas de subsanar este tipo de problema: 1) mantener esta información, lo que implica que no es posible estimar varianzas; 2) asumir que estos conglomerados fueron seleccionados “con certeza”, en cuyo caso los estratos no contribuyen a la estimación de la varianza total y se les asigna varianza 0, generando una subestimación de la varianza, y 3) crear unidades que agrupen estratos unitarios junto a otros estratos (denominadas pseudoestratos de varianza, *varstrat* o pseudoestratos) para efectos del cálculo de varianza.

De acuerdo con lo anterior, y en pos del objetivo de aplicar los más altos estándares de calidad estadística, se opta por la última opción, con la finalidad de poner a disposición las variables que permitan identificar las características del diseño complejo de la encuesta y producir estimaciones de varianza consistentes y estables (Wolter, 2007; Valliant, 2012)¹⁰⁴.

En ese sentido, también es frecuente que los conglomerados no contengan el mínimo de unidades requeridas para la estimación de la varianza, lo que hace necesaria la agrupación de varios conglomerados. Esto se traduce en la creación de una nueva variable como resultado del agrupamiento de dichos conglomerados, denominada pseudoconglomerados de varianza, *varunit* o pseudoconglomerados.

En el contexto de la VIII EPF, para la construcción de estas variables, se establece que cada estrato de varianza debe estar conformado por un mínimo de dos conglomerados (manzanas) y cada conglomerado de varianza debe contener entre 40 y 60 viviendas. Además, cada conglomerado de varianza debe estar contenido en solo un estrato de varianza.

6.3.6.1.2.1 Creación de conglomerados de varianza

En la práctica, pocas manzanas cumplen el criterio establecido de estar compuestas por al menos 40 viviendas, lo que hace necesario el agrupamiento de conglomerados con el fin de estabilizar el cálculo de la varianza. Esta nueva variable corresponde a los conglomerados de varianza.

A partir de un total de 2.094 manzanas de donde fue posible obtener información de las viviendas y hogares, se crearon 304 conglomerados de varianza. Es importante destacar que la distribución porcentual de los conglomerados de varianza por macrozona es similar a la presentada en las manzanas seleccionadas (ver tabla 28).

Tabla 28: Total de conglomerados y *varunit* (pseudoconglomerados) según macrozona.

Macrozona	Número de conglomerados con información	Número de pseudo-conglomerados (<i>varunit</i>)	Porcentaje de conglomerados	Porcentaje de pseudo-conglomerados (<i>varunit</i>)
Total	2.094	304	100,0	100,0
Norte	257	40	12,3	13,2
Centro	486	76	23,2	25,0
Sur	178	29	8,5	9,5
Región Metropolitana	1.173	159	56,0	52,3

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

El procedimiento aplicado para la creación de los conglomerados de varianza se puede separar en cinco fases, dependiendo de la forma de agrupamiento (ver tabla 29).

A continuación, se describen las cinco fases empleadas para la creación de los pseudoconglomerados de varianza (*varunit*):

104 Para mayor detalle consultar la sección 2.5. de Wolter (2007) y sección 15.5.3 de Valliant (2012).

Tabla 29: Frecuencia de pseudoconglomerados de varianza (*varunit*) según fase de agrupación

Fase	Frecuencia	Porcentaje	Frecuencia acumulada	Porcentaje acumulado
1	44	14,5	44	14,5
2	47	15,5	91	29,9
3	23	7,6	114	37,5
4	6	2,0	120	39,5
5	18	5,9	138	45,4
6	166	54,6	304	100

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

- **Fase 1:** al interior de cada estrato (combinación de *comuna*, *clasificación socioeconómica* y *grupo de tamaño*) se agruparon las viviendas hasta sumar entre 40 y 60 unidades. A todas las agrupaciones de este tamaño que se lograron formar, se les considera como *varunit* o pseudoconglomerado, mientras que aquellas combinaciones que no cumplieran con este criterio se abordaron en las siguientes fases. De esta manera se logró conformar el 14,5% de los *varunit*.
- **Fase 2:** las agrupaciones de la fase 1 que no lograron constituirse como pseudoconglomerados, es decir, aquellas con una cantidad menor de 40 unidades, en esta fase se agruparon por *comuna*, *clasificación socioeconómica* y por las distintas combinaciones posibles de las categorías de la variable *grupo de tamaño*, conformando un *varunit* o pseudoconglomerado con la primera combinación que contuviera entre 40 y 60 viviendas. Aquellas agrupaciones que no cumplieran con este criterio se trabajaron en las fases siguientes. En esta fase se logró conformar el 15,5% de los *varunit*.
- **Fase 3:** se trabajaron las agrupaciones de la fase anterior que tenían menos de 40 viviendas, y se agruparon por comuna y las diferentes categorías de la variable *clasificación socioeconómica*, conformándose los nuevos *varunit* o pseudoconglomerados con aquellas combinaciones que contenían entre 40 y 60 viviendas. Aquellos conglomerados que no cumplieran con este criterio quedaron pendientes para ser agrupados en las fases siguientes. En esta fase se logró conformar el 7,6% de los *varunit*.
- **Fase 4:** se trabajó con las agrupaciones con menos de 40 viviendas que no lograron constituir un pseudoconglomerado hasta la fase anterior con aquellas que sí pudieron agruparse, sumando el total de viviendas a nivel de la combinación *comuna* y *clasificación socioeconómica*, de manera que contuviesen entre 40 y 60 viviendas. Con este procedimiento se logró conformar un 2,0% de los *varunit*.
- **Fases 5:** durante esta fase se realizaron agrupaciones de manera manual bajo el criterio de que cada pseudoconglomerado debía contener entre 40 y 60 viviendas¹⁰⁵, cuidando que para los casos donde fuese necesario agrupar comunas diferentes (dentro de la misma región o macrozona), estas tuvieran características socioeconómicas similares. Bajo este criterio se logró obtener 60,5% de los *varunit*.

105 Si bien siempre se buscó que los pseudoestratos o *varunit* tuviesen entre 40 y 60 viviendas en su interior, por razones de agrupación, un estrato en particular quedó conformado por 39 viviendas.

6.3.6.1.2.2 Creación de estratos de varianza

La creación de estratos de varianza tiene como objetivo evitar la presencia de estratos unitarios, por lo que se realizan agrupaciones de los estratos del diseño muestral original. De esta manera se busca sortear posibles subestimaciones de la varianza poblacional asociada a las variables de interés.

En consecuencia, los estratos de varianza (pseudoestratos o *varstrat*) son una reagrupación de los 584 estratos considerados inicialmente en la muestra. Para su obtención se emplea la variable *comuna*, de manera que cada uno contenga al menos dos *varunit*. También fue necesario, debido a la cantidad de unidades muestrales, la unión de algunas comunas, lo cual se realizó en base a criterios de cercanía geográfica y características socioeconómicas similares. La aplicación de este procedimiento generó 56 estratos de varianza (ver tabla 30).

Tabla 30: Total de estratos y pseudoestratos de varianza (*varstrat*) según macrozona

Macrozona	Total de estratos muestrales	Total de pseudoestratos de varianza (<i>varstrat</i>)	Porcentaje de estratos	Porcentaje de pseudoestratos de varianza (<i>varstrat</i>)
Total	584	56	100	100
Norte	86	7	14,7	12,5
Centro	154	11	26,4	19,6
Sur	60	6	10,3	10,7
Región Metropolitana	284	32	48,6	57,1

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

En las tablas 23 y 24 se presentan las distribuciones del número de pseudoconglomerados de varianza (*varunit*) que componen los pseudoestratos de varianza (*varstrat*) y el número de conglomerados (manzanas) que componen los *varunit*.

Tabla 31: Frecuencia del número de pseudoconglomerados de varianza (*varunit*) que componen los *varstrat*

N° <i>varunit</i>	Frecuencia (N° de <i>varstrat</i>)	Porcentaje	Frecuencia acumulada	Porcentaje acumulado
2	8	14,3	8	14,3
3	9	16,1	17	30,4
4	8	14,3	25	44,6
5	8	14,3	33	58,9
6	6	10,7	39	69,6
7	7	12,5	46	82,1
8	2	3,6	48	85,7
9	3	5,4	51	91,1
10	1	1,8	52	92,9
11	1	1,8	53	94,6
13	2	3,6	55	98,2
14	1	1,8	56	100

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Tabla 32: Frecuencia del número de conglomerados que componen los pseudoconglomerados de varianzas (*varunit*)

N° conglomerados	Frecuencia (N° de <i>varunit</i>)	Porcentaje	Frecuencia acumulada	Porcentaje acumulado
2	14	4,6	14	4,6
3	46	15,1	60	19,7
4	37	12,2	97	31,9
5	38	12,5	135	44,4
6	34	11,2	169	55,6
7	24	7,9	193	63,5
8	25	8,2	218	71,7
9	20	6,6	238	78,3
10	15	4,9	253	83,2
11	16	5,3	269	88,5
12	6	2,0	275	90,5
13	10	3,3	285	93,8
14	5	1,6	290	95,4
15	5	1,6	295	97,0
16	3	1,0	298	98,0
17	1	0,3	299	98,4
19	3	1,0	302	99,3
20	1	0,3	303	99,7
24	1	0,3	304	100

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF)

6.3.6.1.3 Coeficiente de variación e intervalos de confianza

El coeficiente de variación es una medida de la variabilidad relativa que se expresa como el cociente entre el error estándar $\hat{\sigma}_{\hat{\theta}}$ y la estimación $\hat{\theta}$:

$$CV_{\hat{\theta}} = CV(\hat{\theta}) = \frac{\hat{\sigma}_{\hat{\theta}}}{\hat{\theta}} = \frac{\sqrt{V(\hat{\theta})}}{\hat{\theta}} \quad (36)$$

Donde,

$\hat{\theta}$ es la estimación de la variable de interés.

$\hat{\sigma}_{\hat{\theta}}$ es el error estándar de la estimación.

$V(\hat{\theta})$ es la varianza de la estimación.

En la inferencia estadística, se quiere estimar parámetros poblacionales a partir de datos de muestras observadas.

El **intervalo de confianza** es un rango de valores determinados a partir de los valores observados en la muestra. Sus límites, inferior y superior, se determinan de forma tal que la probabilidad de que el parámetro poblacional (de la variable de interés) se ubique entre ellos sea de $(1-\alpha)\%$. La probabilidad asociada al intervalo es también denominada como “nivel de confianza” del intervalo.

El intervalo de confianza se determina mediante la siguiente expresión:

$$\hat{\theta} \in \left[\hat{\theta} \mp t_{1-\alpha/2}^v \cdot \hat{\sigma}_{\hat{\theta}} \right]_{1-\alpha} \Leftrightarrow \left[\hat{\theta} - t_{1-\alpha/2}^v \cdot \hat{\sigma}_{\hat{\theta}} ; \hat{\theta} + t_{1-\alpha/2}^v \cdot \hat{\sigma}_{\hat{\theta}} \right]_{1-\alpha} \quad (37)$$

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

El error absoluto de la estimación corresponde a la máxima diferencia que se puede esperar entre el parámetro estimado y el verdadero parámetro poblacional con una probabilidad de $(1-\alpha) 100$, y se obtiene mediante la siguiente expresión:

$$E_{abs}(\hat{\theta}) = t_{1-\alpha/2}^v \cdot \hat{\sigma}_{\hat{\theta}} \quad (38)$$

El error relativo corresponde a la máxima variación porcentual que se puede esperar entre el parámetro estimado y el verdadero parámetro poblacional con una probabilidad del $(1-\alpha) \cdot 100\%$, y se obtiene por el cociente entre el error absoluto dividido por la misma estimación, esto es:

$$E_{rel}(\hat{\theta}) = t_{1-\alpha/2}^v \cdot \hat{\sigma}_{\hat{\theta}} / \hat{\theta} = t_{1-\alpha/2}^v \cdot CV_{\hat{\theta}} = E_{abs}(\hat{\theta}) / \hat{\theta} \quad (39)$$

Asintóticamente, cuando los tamaños muestrales son grandes, el percentil de la distribución t-Student converge al percentil de la distribución Normal, es decir, $t_{1-\alpha/2}^v \approx Z_{1-\alpha/2}$, lo que para un nivel de confianza del 95% corresponde al valor 1,96.

Dados los tamaños muestrales de la VIII EPF (considerando la unidad de análisis hogares), el que una estimación sea o no aceptable dependerá del coeficiente de variación obtenido y de la desagregación del nivel de estimación, ya que al analizar un subgrupo de la muestra¹⁰⁶ se debe considerar la menor cantidad de observaciones disponibles. Por tanto, es relevante considerar la cantidad de unidades con las que se busca realizar alguna inferencia estadística.

Para interpretar el error muestral, usualmente se selecciona el coeficiente de variación asociado a una estimación, para lo cual se recomienda considerar, a modo de referencia, los siguientes rangos:

- hasta un 5% se considera que la estimación es óptima en términos de precisión.
- sobre un 5% y hasta un 15% se considera una buena estimación en términos de precisión.
- sobre un 15% la estimación se considera deficiente en términos de precisión y por ende no es recomendable realizar análisis o inferencias sino solo para fines descriptivos.

Aun cuando el coeficiente de variación presente rangos adecuados en términos de precisión, las estimaciones podrían derivar en errores de interpretación por parte de los usuarios de las bases de datos, debido a la disparidad de los tamaños muestrales (número de hogares) que sustenta el cálculo de la estimación desagregada. Lo anterior podría

¹⁰⁶ Por ejemplo, niveles funcionales de la clasificación de productos o cruces de características de los integrantes del hogar.

implicar detectar estimaciones de gastos en productos con coeficientes de variación aceptables, pero calculadas en base a un número reducido de hogares (observaciones), lo que no permitiría dirimir si el valor obtenido en este estimador es o no confiable.

En cuanto al número de observaciones se recomienda que las estimaciones evaluadas consideren al menos seis hogares con información, en el contexto de la VIII EPF. En efecto, las variables de interés con una cantidad muy baja de observaciones presentan en principio una alta dispersión en la estimación de las medidas de error, lo que dificulta su interpretación como indicador de calidad.

Por consiguiente, análisis conjunto del número de conglomerados con información de la variable de interés y del coeficiente de variación efectivo, al nivel de desagregación evaluado, debe ser considerado para definir el uso que se le puede dar a las estimaciones:

- **Estimador confiable:** el estimador se puede utilizar según los objetivos propuestos.
- **Uso descriptivo:** se recomienda que el estimador se utilice para estadística descriptiva. Si se desea utilizar para otro tipo de análisis, se recomienda agruparlo a un nivel superior de la clasificación de consumo. En todos estos casos se recomienda incluir una nota técnica advirtiendo sobre las restricciones de la estimación que mencione el valor del coeficiente de variación.
- **Estimador poco confiable:** no se recomienda utilizar la estimación del gasto obtenida. Se debe reevaluar la estimación de consumo a un nivel superior de la clasificación de consumo, para allí evaluar su calidad y determinar su aptitud para un uso inferencial o solamente descriptivo.

6.4 Resultados de los principales indicadores de ingreso y gasto de la encuesta

A continuación, se presentan las estimaciones del gasto promedio y el ingreso disponible promedio de los hogares (con y sin arriendo imputado). Estos indicadores también se calculan a nivel *per cápita* para las áreas de estimación de Gran Santiago, Resto de capitales regionales (en su conjunto) y Total capitales regionales (Gran Santiago + Resto de capitales regionales).

Tabla 33: Indicadores de precisión y confiabilidad del gasto promedio mensual de los hogares según área de estimación (excluye arriendo imputado)⁽¹⁾

Área de estimación	Indicadores					CV
	Gasto promedio mensual	Error absoluto	Error relativo	Intervalo de confianza al 95%		
				Inferior	Superior	
Total capitales regionales	1.121.925	67.236	5,99%	1.054.689	1.189.161	3,04%
Gran Santiago	1.243.155	112.187	9,02%	1.130.968	1.355.342	4,56%
Resto de capitales regionales	969.048	57.767	5,96%	911.282	1.026.815	3,01%

(1) Expresado en pesos chilenos corrientes.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Tabla 34: Indicadores de precisión y confiabilidad del ingreso disponible promedio mensual de los hogares según área de estimación (excluye arriendo imputado)⁽¹⁾

Área de estimación	Indicadores					CV
	Ingreso disponible promedio mensual	Error absoluto	Error relativo	Intervalo de confianza al 95%		
				Inferior	Superior	
Total capitales regionales	1.193.456	79.937	6,70%	1.113.519	1.273.393	3,40%
Gran Santiago	1.344.840	131.654	9,79%	1.213.185	1.476.494	4,95%
Resto de capitales regionales	1.002.555	72.794	7,26%	929.761	1.075.349	3,67%

(1) Expresado en pesos chilenos corrientes.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Tabla 35: Indicadores de precisión y confiabilidad del gasto *per cápita* promedio mensual de los hogares según área de estimación (excluye arriendo imputado)⁽¹⁾

Área de estimación	Indicadores					CV
	Gasto <i>per cápita</i> promedio mensual	Error absoluto	Error relativo	Intervalo de confianza al 95%		
				Inferior	Superior	
Total capitales regionales	397.819	25.364	6,38%	372.455	423.182	3,24%
Gran Santiago	443.594	42.536	9,59%	401.057	486.130	4,85%
Resto de capitales regionales	340.095	23.284	6,85%	316.810	363.379	3,46%

(1) Expresado en pesos chilenos corrientes.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Tabla 36: Indicadores de precisión y confiabilidad del ingreso disponible *per cápita* promedio mensual de los hogares según área de estimación (excluye arriendo imputado)⁽¹⁾

Área de estimación	Indicadores					CV
	Ingreso disponible per cápita promedio mensual	Error absoluto	Error relativo	Intervalo de confianza al 95%		
				Inferior	Superior	
Total capitales regionales	425.608	30.367	7,13%	395.241	455.974	3,62%
Gran Santiago	481.944	50.744	10,53%	431.199	532.688	5,32%
Resto de capitales regionales	354.565	27.624	7,79%	326.942	382.189	3,94%

(1) Expresado en pesos chilenos corrientes.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

El objetivo principal de la encuesta indica que se debe identificar la estructura y características del gasto en consumo final, mientras que el objetivo secundario puntualiza sobre la estructura del ingreso total disponible. Por esta razón resulta importante evaluar la precisión y confiabilidad del gasto desagregado por división¹⁰⁷ del gasto y fuentes de ingreso.

La tabla 37 presenta información respecto a los indicadores de precisión y confiabilidad del gasto promedio mensual de los hogares a nivel de divisiones de gasto, mientras que en la tabla 38 se agrega información en relación al ingreso promedio mensual disponible de los hogares sin considerar arriendo imputado.

Tabla 37: Indicadores de precisión y confiabilidad del gasto promedio mensual de los hogares según división, total de capitales regionales (excluye arriendo imputado)⁽¹⁾

División	Glosa	Indicadores					
		Gasto promedio mensual	Error absoluto	Error relativo	Intervalo de confianza al 95%		CV
					Inferior	Superior	
Total		1.121.925	67.236	5,99%	1.054.689	1.189.161	3,04%
1 Alimentos y bebidas no alcohólicas		209.983	7.317	3,48%	202.666	217.300	1,77%
2 Bebidas alcohólicas, tabaco y estupefacientes		19.636	1.069	5,45%	18.567	20.705	2,77%
3 Prendas de vestir y calzado		38.690	2.380	6,15%	36.311	41.070	3,12%
4 Alojamiento, agua, electricidad, gas y otros combustibles		160.692	7.898	4,91%	152.794	168.589	2,50%
5 Muebles, artículos para el hogar y para la conservación ordinaria del hogar		70.778	8.328	11,77%	62.450	79.107	5,97%
6 Salud		84.791	10.454	12,33%	74.337	95.245	6,26%
7 Transporte		170.238	14.513	8,53%	155.725	184.751	4,33%
8 Comunicaciones		59.209	2.469	4,17%	56.740	61.678	2,12%
9 Recreación y cultura		76.167	6.034	7,92%	70.133	82.201	4,02%
10 Educación		72.596	8.588	11,83%	64.008	81.184	6,01%
11 Restaurantes y hoteles		73.096	5.725	7,83%	67.371	78.821	3,98%
12 Bienes y servicios diversos		86.050	6.968	8,10%	79.081	93.018	4,11%

(1) Expresado en pesos chilenos corrientes.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

107 El primer grado de división de la CCIF se denomina división y corresponde a 12 categorías funcionales que están desagregadas en la tabla 37. Tanto la EPF como el IPC comparten la utilización de las divisiones.

CONTINUA ➡

Tabla 38: Indicadores de precisión y confiabilidad del ingreso disponible promedio mensual de los hogares, total de capitales regionales (excluye arriendo imputado)⁽¹⁾

Desglose	Indicadores					
	Partidas del Ingreso promedio mensual del hogar	Error absoluto	Error relativo	Intervalo de confianza al 95%		CV
				Inferior	Superior	
Ingreso total del hogar	1.419.249	97.615	6,88%	1.321.634	1.516.863	3,49%
De la ocupación	1.192.803	89.164	7,48%	1.105.212	1.283.540	3,80%
Sueldos y salarios	930.962	70.550	7,58%	861.643	1.002.743	3,85%
Empleadores	50.534	11.585	22,92%	39.016	62.186	11,63%
Cuenta propia	163.920	17.851	10,89%	146.283	181.985	5,53%
Ingresos de otros trabajos	47.386	8.881	18,74%	38.566	56.329	9,51%
De otras fuentes	226.446	17.899	7,90%	208.810	244.607	4,00%
Jubilaciones o pensiones de vejez	89.041	6.804	7,64%	82.352	95.960	3,88%
Rentas de la propiedad ⁽²⁾	62.488	11.353	18,17%	51.213	73.919	9,21%
Transferencias recibidas ⁽³⁾	74.917	6.009	8,02%	68.978	80.996	4,06%
Transferencias corrientes emitidas	225.793	9.584	8,36%	207.057	244.827	4,24%
Impuestos a la renta	30.139	3.310	21,65%	23.653	36.705	10,98%
Contribuciones ⁽⁴⁾	157.428	5.548	6,94%	146.705	168.567	3,52%
Transferencias emitidas ⁽⁵⁾	38.226	2.253	11,61%	33.688	42.566	5,89%
Ingreso disponible del hogar	1.193.456	40.586	6,70%	1.115.184	1.275.103	3,40%

(1) Expresado en pesos chilenos corrientes.

(2) Incluye ingreso por arriendos de bienes inmuebles y muebles, así como ganancias por instrumentos financieros.

(3) Corresponde a la suma de las transferencias corrientes en efectivo recibidas desde otros hogares, el Gobierno y/o instituciones sin fines de lucro (ISFLSH), más las transferencias en especies recibidas únicamente desde otros hogares.

(4) Corresponde a las contribuciones totales a los sistemas de salud y previsión, ya sean obligatorias o voluntarias.

(5) Corresponde a la suma de las transferencias corrientes en efectivo pagadas a otros hogares y/o instituciones sin fines de lucro (ISFLSH).

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

7. ESTRUCTURA ORGANIZACIONAL Y PROCESO GENERAL DEL PROYECTO

El proceso de producción estadística de la VIII EPF, al igual que su versión anterior, siguió un modelo de organización en el que el área técnica y operativa se mantuvo en constante comunicación. Durante todo el proceso de preparación y posterior ejecución, los equipos buscaron continuos acuerdos respecto a cómo llevar a la práctica los distintos niveles de decisiones técnicas que deben considerarse en un proceso complejo como la EPF. El resultado fue una serie de decisiones consensuadas en un inicio, donde la puesta a prueba de dichas decisiones y la continua observación de los procesos permitió incorporar mejoras continuas durante la realización del trabajo de campo de la encuesta.

Este desarrollo también incorporó reuniones periódicas con las Subdirecciones Técnica y de Operaciones, en conjunto con los Departamentos de Estudio de Precios y Estadísticas de Precios, pertenecientes a dichos estamentos, como clientes institucionales principales de la información de la encuesta y quienes participaban de la toma de decisiones, con el fin de salvaguardar los usos internos de la encuesta.

Los equipos de trabajo de la VIII EPF implementaron el estándar internacional Generic Statistical Business Process Model (GSBPM)¹⁰⁸ para estructurar y sistematizar los procesos relacionados con la producción estadística de la encuesta. Para realizar este proceso de adaptación, el equipo de la encuesta consideró las distintas fases (o ámbitos de negocio en la adaptación institucional del estándar), para así delimitar responsabilidades asociadas a estas, además de definir responsabilidades asociadas a distintos niveles de desagregación (subprocesos, etapas y tareas).

En el contexto de la VIII EPF, la implementación del estándar GSBPM buscó generar un proceso de producción más eficiente y eficaz en sus subprocesos, perfeccionando la gestión del proyecto. Dentro del proceso de mejoras, se destacan algunos hitos, tales como 1) instaurar prácticas de trabajo orientada a procesos, 2) implementar protocolos estandarizados para los procesos, 3) monitorear y gestionar los resultados de los procesos e 4) incorporar tecnología para apoyar la gestión.

Por tanto, para comprender de mejor manera el proceder del proyecto, tanto en su etapa de diseño, construcción, recolección, procesamiento, análisis y posterior difusión de resultados, es que se describe, a modo de síntesis, su estructura organizacional y el proceso general de producción.

7.1 Estructura organizacional del Departamento de Presupuestos Familiares

El equipo de trabajo del Departamento de Presupuestos Familiares contó con una dotación de 280 personas para realizar el proceso de producción estadística de la VIII EPF. Dicho personal se distribuyó en los equipos de gestión central, de trabajo de campo, técnico, de codificación, de digitación, de supervisión y capacitación, y el equipo de informática (diagrama 3). Considerando los equipos que conforman el departamento, solo el equipo

¹⁰⁸ Más información disponible en <https://statswiki.unece.org/display/GSBPM>

de trabajo de campo tuvo una distribución a nivel nacional, para lo cual fue fundamental el apoyo otorgado por las direcciones regionales¹⁰⁹ del INE en temas administrativos, logísticos, de supervisión, coordinación y difusión de la encuesta. El resto de los equipos estuvo ubicado totalmente en la Región Metropolitana, donde se centralizaron procesos como la digitación de formularios y la codificación de datos.

Adicionalmente, se trabajó con el apoyo de la Subdirección Administrativa y se contó con el apoyo de otros departamentos y subdepartamentos del INE en diferentes tareas necesarias para la ejecución de la encuesta, tales como las habilitaciones de espacios de trabajo, la generación de perfiles de cargo, los procesos de selección, la contratación de personal, la compra de materiales, las actividades de difusión, los aspectos legales, las ayudas técnico-metodológicas y la gestión de viáticos, entre otras actividades.

Para mejor funcionamiento de la orgánica institucional del proyecto, el Departamento de Presupuestos Familiares generó una instancia de coordinación que sesionaba periódicamente, para facilitar la realización de trabajos conjuntos entre los distintos equipos que componen el departamento. Dicho comité de organización se denominó “equipo de gestión central” de la encuesta y era el encargado de coordinar los distintos procesos, ya sea entre los equipos que componen el departamento como aquellos que requieren la vinculación con equipos externos vinculados a la encuesta. Este equipo estuvo compuesto por el jefe del departamento, las jefaturas de los distintos equipos del departamento y por la persona encargada del apoyo de gestión del proceso.

¹⁰⁹ El INE dispone de oficinas regionales en cada una de las capitales regionales del país, las cuales participaron coordinando los equipos de campo involucrados en la producción de la VIII EPF. A la participación de las distintas Direcciones Regionales se sumó la participación de la Unidad Provincial de Ñuble.

```

graph TD
    DN[Dirección Nacional] --> SO[Subdirección de Operaciones]
    DN --> ST[Subdirección Técnica]
    DN --> SA[Subdirección Administrativa]
    
    SO --> DEH[Departamento de Estadísticas de Hogares]
    SO --> DEP[Departamento de Estadísticas de Precios]
    SO --> DEE[Departamento de Estudios de Precios]
    SO --> DID[Departamento de Investigación y Desarrollo]
    
    ST --> DEH
    ST --> DEP
    ST --> DEE
    ST --> DID
    
    SA --> DSGP[Departamento de Gestión de las Personas]
    SA --> DLS[Departamento de Logística y Servicios]
    SA --> DGF[Departamento de Gestión Financiera]
    SA --> DR[Direcciones Regionales]
    SA --> DJ[División Jurídica]
    SA --> DCI[Depto. de Comunicaciones e Imagen Corporativa]
    SA --> DTI[Departamento de Tecnologías de la Información]
    
    DEH --> EOC[Equipo de Codificación (15)]
    DEH --> REE[Revisores Equipo Técnico (2)]
    DEH --> ET[Equipo Técnico (1)]
    
    REE --> EOC
    REE --> REE
    
    ET --> ES[Equipo de Supervisión y Capacitación (19)]
    ET --> ED[Equipo de Digitación (24)]
    ET --> ETC[Equipo de Trabajo de Campo (8)]
    ET --> EI[Equipo de Informática (4)]
    
    ES --> EOC
    ES --> REE
    ES --> ET
    ES --> ED
    
    ED --> AM[Área Metropolitana (87)]
    ED --> DCC[Digitadores de trabajo de campo (2)]
    ED --> ARC[Área Resto de Capitales Regionales (103)]
    
    EOC --> EOC
    REE --> REE
    ET --> ET
    ED --> ED
    EI --> EI
    
    AGC[Apoyo Gestión Central (3)] --> ES
    AGC --> ED
    AGC --> ETC
    AGC --> EI
    
    S[Secretarías (2)] --> ES
    S --> ED
    S --> ETC
    S --> EI
    
    DPF[Depto. de Presupuestos Familiares (1)] --> ES
    DPF --> ED
    DPF --> ETC
    DPF --> EI
  
```

7.2 Proceso general de producción de la VIII EPF

Con el fin de ordenar el proceso productivo asociado a la generación de la VIII EPF, los equipos del Departamento de Presupuestos Familiares del INE implementaron una adaptación del modelo GSBPM (diagrama 4) propuesto por Naciones Unidas para ordenar los procesos estadísticos.

Diagrama 4: Niveles 1 y 2 del GSBPM

Administración de calidad y metadatos							
Especificación de necesidades	Diseño	Construcción	Recolección	Procesamiento	Análisis	Difusión	Evaluación
1.1 Identificación de necesidades.	2.1 Diseño conceptual.	3.1 Construcción de instrumentos de recolección.	4.1 Conformación del marco muestral y selección de la muestra.	5.1 Integración de datos.	6.1 Preparación de borradores de resultados.	7.1 Actualización de sistemas de salida.	8.1 Concentración de insumos para la evaluación.
1.2 Consulta y confirmación de necesidades.	2.2 Diseño de las descripciones de las variables.	3.2 Construcción o mejora de componentes.	4.2 Preparación de la recolección.	5.2 Clasificación y codificación.	6.2 Validación de los resultados.	7.2 Generación de productos de difusión.	8.2 Evaluación.
1.3 Definición de objetivos.	2.3 Diseño de la recolección.	3.3 Construcción o mejora de componentes de difusión.	4.3 Ejecución de la recolección.	5.3 Revisión y validación.	6.3 Interpretación y explicación de los resultados.	7.3 Gestión de la publicación de productos de difusión.	8.3 Determinación de un plan de acción.
1.4 Identificación de conceptos.	2.4 Determinación del marco muestral y tipo de muestreo.	3.4 Configuración de flujos de trabajo.	4.4 Cierre de la recolección.	5.4 Edición e imputación.	6.4 Aplicación del control a la divulgación.	7.4 Promoción de productos de difusión.	
1.5 Comprobación de la disponibilidad de datos.	2.5 Diseño del procesamiento y análisis.	3.5 Pruebas al sistema de producción.		5.5 Derivación de nuevas variables y unidades.	6.5 Finalización de resultados.	7.5 Gestión de soporte a usuarios.	
1.6 Elaboración del Plan General.	2.6 Diseño de los sistemas de producción y de los flujos de trabajo.	3.6 Prueba piloto del proceso estadístico.		5.6 Cálculo de ponderadores.			
		3.7 Finalización del sistema de producción.		5.7 Cálculo de agregaciones.			
				5.8 Finalización de los archivos de datos.			

Fuente: Modelo internacional GSBPM versión 5.0 publicado por UNECE (2016).

Es importante mencionar que en el esquema de organización por procesos GSBPM, que estructura la producción estadística, las fases y subprocesos pueden ser iterativos y paralelos (UNECE, 2016), lo que se debe considerar al revisar los distintos elementos que componen la estructura productiva de la VIII EPF, para comprender el funcionamiento de la encuesta con un enfoque sistémico.

En un proceso complejo de producción de datos como el de la VIII EPF, las fases de recolección de datos y procesamiento ocurren en paralelo, dado que la fase de procesamiento comienza a realizarse antes que la fase de recolección termine. Esto permite realizar compilaciones de resultados provisionales, ya que el tiempo de producción de la encuesta es una preocupación constante del Instituto. Es más, la fase de análisis del estudio pudo gatillar eventuales revisiones de la fase de procesamiento. Dicho lo anterior, el flujo planteado y la estructura que se revisa en la presente metodología no tienen necesariamente un orden lineal en la realización del proceso de producción estadística¹¹⁰. El diagrama 5 presenta una simplificación de la carta Gantt del proceso de producción de la VIII EPF, considerando las distintas fases asociadas al estándar GSBPM. Es necesario advertir que el diagrama es simplificado y que existen algunas fases que se extienden más allá de los plazos presentados¹¹¹.

Diagrama 5: Simplificación de la Carta Gantt del GSBPM a nivel de fases aplicada al proceso de producción de la VIII EPF

Fases GSBPM	2014	2015				2016				2017				2018			
	TR04	TR01	TR02	TR03	TR04	TR01	TR02	TR03	TR04	TR01	TR02	TR03	TR04	TR01	TR02	TR03	TR04
Especificación de Necesidades																	
Diseño																	
Construcción																	
Recolección																	
Procesamiento																	
Análisis																	
Difusión																	
Evaluación																	

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

A continuación se hace una revisión general de las ocho fases que componen el proceso de producción, examinando aquellos aspectos más relevantes para el funcionamiento de la VIII EPF.

110 El estándar del GSBPM reconoce que estas situaciones son habituales y si bien el estándar norma el proceso de producción estadística, este no debe asumirse como un flujo de procesos lineales en términos temporales (UNECE, 2016).

111 Por ejemplo, la fase de especificación de necesidades es un proceso amplio donde se evalúan distintos requerimientos de información a la encuesta y, dado que la EPF es un producto continuo del INE, puede estar presente durante las distintas fases del proceso productivo (con miras a las siguientes versiones de la encuesta). La fase de evaluación contempla la realización de encuestas de satisfacción de usuarios, las que se realizan entre 1 a 2 años de forma posterior a la publicación de los datos.

7.2.1 Especificación de necesidades

Respecto a la fase de especificación de necesidades, al ser la principal fuente de información para el IPC y por sus importantes usos a nivel de políticas públicas, la encuesta es parte de los productos de elaboración continua que lleva a cabo la institución, destacándose como uno de los productos estadísticos estratégicos del INE. A partir de la VII EPF, se genera un equipo de trabajo estable al interior del INE, para asegurar su realización y su proceso de continua mejora.

Al respecto, en la etapa de especificación de necesidades, la EPF tiene definidos los objetivos de alto nivel para los resultados estadísticos que genera el proceso de producción, por lo tanto, los objetivos del estudio se encuentran en línea con los requerimientos a los que la encuesta satisface y los principales usos para los cuales se genera.

7.2.2 Diseño

La etapa de diseño busca definir los productos que se generarán con el estudio, conceptos, metodologías, instrumentos de recolección y procesos de trabajo de campo. Al respecto, dado que la encuesta es un producto continuo al interior del INE, en la fase de diseño de la VIII EPF se planificaron aquellas acciones de mejora que fueron identificadas como necesarias en las etapas de evaluación de la VII EPF.

La fase de diseño de la encuesta fue un proceso amplio y participativo al interior del instituto, en el que se buscó continuar el perfeccionamiento del estudio, actualizando los procesos de acuerdo a las últimas versiones de estándares nacionales e internacionales, e incorporando buenas prácticas identificadas a nivel internacional, nacional e institucional.

Durante el año 2015, el equipo a cargo de la producción de la EPF generó trabajos de participación con distintos actores involucrados en este proceso, entre los cuales destaca el rol central que tuvieron las direcciones regionales, las que participaron tanto en la generación de protocolos administrativos del proyecto (arriendo y habilitación de oficinas, contratación de personal, provisión de materiales y vestuario, gestión de vehículos, despacho de encuestas, etc.), como en el desarrollo de la actualización de la metodología y protocolos de trabajo de campo de la encuesta (actividades de difusión en medios locales, apoyo en control de cumplimiento de tasas de logro, etc.).

Para la aplicación de la VIII EPF, las direcciones regionales tuvieron una participación activa en el proceso de diseño y planificación de la encuesta, así como también en su ejecución, especialmente durante el despliegue del trabajo de campo. El desarrollo de un trabajo participativo, sumado a la experiencia y al conocimiento del contexto regional de las direcciones regionales, permitió incorporar buenas prácticas institucionales al proceso.

La etapa de diseño de la encuesta también contempló la determinación del marco muestral y el tipo de muestreo que se utilizó¹¹². Respecto al trabajo realizado para la definición del tipo de diseño muestral, se trabajó con el Departamento de Investigación y Desarrollo de la Subdirección Técnica del INE. En el trabajo con dicho equipo se realizaron diversas

¹¹² Más detalles respecto al subproceso asociado a la determinación del marco muestral y tipo de muestreo de la VIII EPF se detallan en el capítulo 6, de diseño estadístico de la encuesta.

simulaciones y mesas de trabajo conjuntas para determinar el diseño que finalmente se incorporó en la encuesta.

Después de que las unidades de primera etapa fueron seleccionadas (manzanas), se realizó un proceso de **enumeración**, que consistió en visitar las unidades primarias para determinar la cantidad de viviendas particulares ocupadas que estas contenían, luego de lo cual se procedió a seleccionar las unidades de segunda etapa (viviendas particulares ocupadas).

7.2.3 Construcción

La etapa de construcción del proceso de producción de la encuesta abarcó desde las tareas asociadas a la actualización de los instrumentos hasta la ejecución de la prueba piloto, la que tuvo lugar en el segundo semestre de 2015.

En esta etapa se generaron los flujos de trabajo y se actualizaron los cuestionarios, considerando los resultados y recomendaciones de la versión previa de la encuesta, además de los aportes de contrapartes internas y externas a la institución.

Uno de los cambios testados en esta etapa se relacionó con el cambio de flujo en el proceso de producción estadística, invirtiendo los procesos de codificación y digitación. Mientras que en la VII EPF la información se codificaba en los cuestionarios, para la VIII EPF estos primero fueron digitados, para realizar el proceso de codificación de forma posterior y en plataformas informáticas. Esta modificación permitió optimizar el proceso en términos de calidad y de tiempo, además de introducir mejoras tales como la codificación automatizada, la que se describe en el apartado 9.3 del presente documento.

7.2.4 Recolección

La fase de recolección de información es uno de los procesos más extensos de la encuesta en términos de ejecución, ya que su trabajo de campo contempla un período de un año calendario.

La etapa de preparación de la recolección involucró el trabajo activo de las direcciones regionales y equipos administrativos del INE, ya que se requirió seleccionar personal, entrenarlo y asegurar los recursos para este pudiese realizar su trabajo (oficinas, equipos informáticos, materiales de oficina, encuestas, etc.).

La fase de recolección no solo abarcó la realización del trabajo de campo de la encuesta, sino que también incorporó el proceso de supervisión de la misma. La metodología de trabajo de campo se describe con detalle en el apartado 8.1, mientras que la metodología de la supervisión es abordada en el apartado 8.2 de este documento.

Si bien en el estándar internacional la etapa de digitación se encuentra asociada al subproceso de cierre de la recolección, en la adaptación al proceso de producción de la VIII EPF, considerando el volumen de información a digitar, el subproceso se dividió entre 1) la digitación de la Hoja de Ruta, inserta en la fase de recolección (descrito en el apartado 8.1.5.1) y 2) la digitación del resto de los instrumentos de recolección de información, inserta en la fase de procesamiento (apartado 9.1).

7.2.5 Procesamiento

Corresponde a la fase en que se digita la información de las libretas (cuestionarios)¹¹³ y se verifica la validez de los datos digitados, editándolos cuando fuese necesario y preparando la información para su publicación oficial.

En el contexto de la VIII EPF, el procesamiento contempló ocho subprocesos: 1) digitación de libretas, 2) codificación, 3) validación, 4) recuperación, 5) depuración, 6) imputación, ajuste temporal y creación de nuevas variables, 7) cálculo de ponderadores y 8) consolidación de las bases de datos y cálculo de agregaciones.

El subproceso de digitación consistió en ingresar al sistema los datos de los hogares recopilados en las libretas (cuestionarios)¹¹⁴, para luego realizar el subproceso de codificación, que consiste en asignar códigos, a partir de clasificadores establecidos, a distintas preguntas abiertas de la encuesta. Las etapas de validación, recuperación y depuración tuvieron por objetivo identificar y reducir los errores asociados al procesamiento de los datos. Por su parte, el subproceso de imputación, ajuste temporal y creación de nuevas variables buscó reducir en parte los sesgos asociados a la no respuesta parcial de la encuesta, llevar las distintas variables de ingresos y gasto a un período común de presentación de resultados (mes) y generar aquellas variables que se requerirían en los subprocesos siguientes.

El subproceso siguiente, de cálculo de ponderadores, comprende las etapas de ajustes al factor de expansión y cálculos de los pseudoestratos y pseudoconglomerados de varianza de la encuesta. Estas etapas son descritas con detalle en el capítulo de diseño estadístico, en los apartados 6.2 y 6.3.

Respecto al subproceso de consolidación de las bases de datos y cálculo de agregaciones, en él se generó la finalización de la revisión de bases de datos asociadas a los distintos instrumentos de recolección de información (libretas), para luego consolidarlas y construir las bases de datos a utilizar en la fase de análisis. Además, se calcularon los resultados de la encuesta y las medidas de precisión asociadas a estos (en esta etapa se aplican en los resultados las medidas de precisión descritas en el apartado 6.3).

7.2.6 Análisis

Esta fase tiene como propósito validar los resultados con fuentes externas y versiones anteriores de la encuesta, interpretar los resultados y prepararlos para su difusión.

En la VIII EPF, el análisis de la encuesta fue definido por cinco subprocesos: 1) validación de resultados con VII EPF y fuentes externas, 2) interpretación de los resultados, 3) preparación de tabulados, 4) preparación de contenido y controles de difusión y 5) cierre de resultados.

El primer subproceso de validación tuvo como propósito verificar que los datos fuesen consistentes entre las distintas versiones de la encuesta y otros datos afines, desarrollados

113 Tal como se explica en el apartado anterior, en la adaptación del estándar al proceso de producción de la VIII EPF, la digitación de libretas (salvo la Hoja de Ruta) se incorporó a la fase de procesamiento (subproceso de digitación de libretas).

114 Salvo de la Hoja de Ruta que se digita en la fase de recolección.

por la institución o por organismos externos. Adicionalmente, se definió como el único subproceso que podía volver a la fase de procesamiento si se detectaban inconsistencias al comparar resultados.

El segundo subproceso tuvo como propósito interpretar la información para tener insumos para los productos de difusión y para explicar cambios que no se esperaban. El tercer y cuarto subproceso tuvo como propósito preparar la información a publicar, ajustar el contenido de este material y aplicar los controles para su correcta difusión (por ejemplo, incorporando el análisis de anonimización de datos). El último subproceso buscó verificar la consistencia de los cuatro subprocesos anteriores y corroborar que la información estuviese en condiciones de ser difundida.

7.2.7 Difusión

Esta fase abarca la finalización y entrega de todos los productos asociados a la realización de la encuesta. La información fue publicada en la página web de la institución (www.inec.cl) y contempló la entrega de la presente metodología, tabulados oficiales, síntesis de resultados, manuales de usuario de la base de datos, metadatos de la encuesta, informe de calidad, documento de principales resultados y manuales de capacitación, entre otros.

7.2.8 Evaluación

La evaluación se inicia al momento de finalizar el proceso de producción de la encuesta y busca llegar recomendaciones para la mejora continua de la misma. Esta fase no debe confundirse con el proceso de mejora continua asociada a la gestión de la calidad, reconocida en el estándar GSBPM como un proceso transversal a la producción estadística.

En la EPF, esta fase se alimenta de diversos insumos generados a lo largo de todo el proceso de producción de la encuesta e incorpora información cualitativa y cuantitativa para la incorporación de mejoras a la próxima versión del estudio.

8. METODOLOGÍA DE RECOLECCIÓN DE DATOS

8.1 Metodología de trabajo de campo

El diseño de la metodología de trabajo de campo se basó fundamentalmente en la experiencia de la VII EPF y, por lo tanto, presenta leves diferencias con la anterior versión de la encuesta. Los ajustes realizados a aquella metodología provienen de la observación de buenas prácticas internacionales en los procesos de recolección de datos, del estudio de literatura especializada, del intercambio de experiencias con otras encuestas de gastos e ingresos¹¹⁵, del análisis de la no respuesta total de la VII EPF¹¹⁶ y de la retroalimentación aportada por personal del trabajo de campo de la anterior versión de la encuesta, a través de procesos de evaluación internos.

Este capítulo describe la metodología de trabajo de campo de la VIII EPF, explicando los diversos aspectos considerados a nivel organizacional, técnico y operativo, para producir datos de calidad y, de este modo, aportar al cumplimiento de los objetivos la encuesta.

En la sección Diseño organizacional del equipo de trabajo de campo (8.1.1), se describe la forma en que se organizó, constituyó y preparó al equipo de trabajo de campo, indicando los criterios considerados para definir la dotación de personal necesaria, el proceso de reclutamiento de personal y la capacitación necesaria para la encuesta.

Posteriormente, en la sección Características generales del proceso de trabajo de campo (8.1.2), se especifica el período de recolección, el modo en que se distribuyó operativamente la muestra y las técnicas de producción de información utilizadas. A continuación, en la sección Organización del proceso de trabajo de una submuestra (8.1.3), se exponen las distintas fases y actividades contempladas para el trabajo de campo de cada submuestra, detallando sus características más relevantes.

Luego de ello, se informan los Planes especiales (8.1.4) implementados durante el trabajo de campo, que introdujeron ajustes a la organización habitual del proceso, a fin de proteger la calidad y niveles de respuesta del estudio ante situaciones que constituían potenciales riesgos para el cumplimiento del proceso. Finalmente, se presentan las Herramientas de gestión y monitoreo del trabajo de campo (8.1.5), utilizadas para diagnosticar el avance del proceso, retroalimentar a los equipos y decidir acciones tendientes a mejorar los resultados.

8.1.1 Diseño organizacional del equipo de trabajo de campo

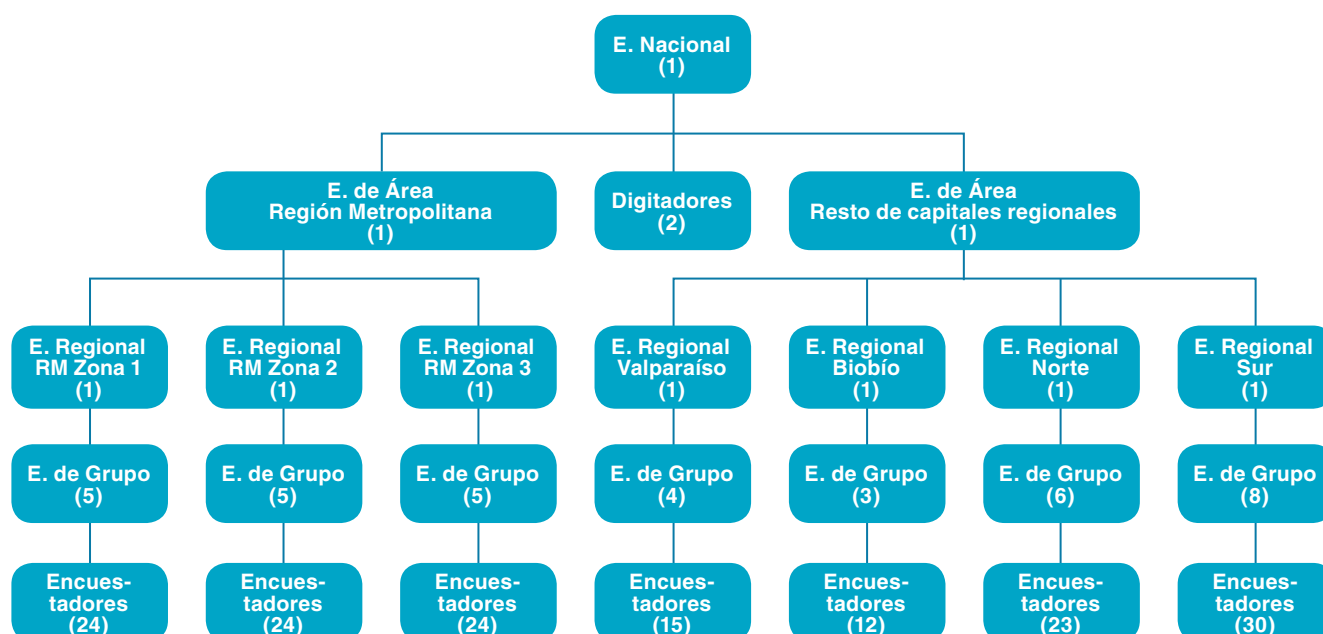
8.1.1.1 Organigrama del equipo de trabajo de campo

El equipo de trabajo de campo fue el único que tuvo un alcance nacional y estuvo compuesto por 200 personas, entre encuestadores de hogares, digitadores, encargados de grupo, encargados regionales, encargados de área y una encargada nacional. La estructura organizacional del equipo de trabajo de campo se presenta en el diagrama 6.

 115 Entre los intercambios realizados, se destaca la oportunidad de compartir experiencias con el Departamento Administrativo Nacional de Estadística (DANE) de Colombia.

116 Para mayor información respecto a análisis de la no respuesta total de la VII EPF revisar INE Chile (2014a).

Diagrama 6: Organigrama equipo de trabajo de campo VIII EPF



Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Desde la perspectiva de los ámbitos de responsabilidad, el equipo reconocía dos niveles de organización:

- **Coordinación nacional del trabajo de campo:** compuesta por los encargados regionales, de área y nacional, que eran responsables de la elaboración y capacitación de las directrices metodológicas asociadas al trabajo de campo, así como del monitoreo de los indicadores de desempeño operativo y la propuesta de acciones para la mejora de los mismos.

Esta coordinación estuvo subdividida en dos áreas¹¹⁷: Gran Santiago y Resto de capitales regionales. Cada una de ellas tuvo un encargado, quien trabajaba con el apoyo de encargados regionales (responsables de zonas específicas al interior de un área) en la organización, acompañamiento y monitoreo de los grupos de campo respectivos.

- **Grupos de campo:** eran coordinados por un encargado de grupo y estaban compuestos por entre tres y seis encuestadores de hogares, según la cantidad de viviendas a encuestar en cada territorio. Constituían la unidad básica del equipo nacional y fueron responsables directos de la difusión territorial del estudio y producción de datos en los hogares.

¹¹⁷ Se definen estas áreas en correspondencia a las áreas de estimación de la encuesta, para propiciar un seguimiento detallado de los resultados operativos y respuesta oportuna ante potenciales riesgos en sus niveles de respuesta.

8.1.1.2 Distribución del personal de trabajo de campo

La estimación de personal de trabajo de campo se realizó de acuerdo a la distribución de la muestra en cada región del país. Dado el efecto de la no respuesta¹¹⁸, la cantidad de encuestadores en cada región se estimó a partir de la muestra objetivo asignada, estableciendo una carga promedio de cinco viviendas objetivo por encuestador durante cada quincena. A la vez, a cada grupo de campo se asignó un encuestador adicional, como previsión ante la rotación de personal o ausencias laborales (licencias médicas, vacaciones, etc.).

Por su parte, la definición del número de encargados de grupo (y, por tanto, de grupos) se realizó en relación a la cantidad de encuestadores necesarios en cada región, considerando un mínimo de tres y un máximo de seis por encargado. Finalmente, se definió la existencia de encargados regionales, a cargo de la coordinación y seguimiento de los resultados de zonas específicas, delimitadas según las características sociodemográficas y geográficas comunes entre las regiones y/o comunas de cada área. La cantidad de personal de trabajo de campo en cada región del país se presenta en la tabla 39.

En la Región Metropolitana (RM), las zonas de coordinación fueron definidas territorialmente, de modo que estuviesen compuestas por comunas con características sociodemográficas similares y existiesen medios de transporte que facilitaran su conectividad interna. De este modo, se conformaron las zonas RM oriente, RM sur y RM norte. La asignación de los encuestadores y encargados a cada zona se definió tomando en consideración la residencia de los integrantes del equipo, las experiencias de trabajo anterior, el conocimiento territorial de las comunas y la preferencia personal.

La organización del equipo de campo en zonas territoriales permitió la especialización de encuestadores y encargados de grupo, propiciando la adaptación de los estilos de persuasión y realización de entrevistas de los encuestadores a las características de los informantes. Adicionalmente, la decisión de asignar al personal considerando su lugar de residencia permitió optimizar los tiempos y recursos asociados a los desplazamientos.

118 La no respuesta provoca una disminución de la carga inicial de trabajo, debido a problemas de elegibilidad, al no contacto y al rechazo inicial. Junto con lo anterior, a medida que avanza la quincena la no respuesta puede aumentar por la interrupción de procesos de entrevista (debido a pérdidas de contacto o al rechazo por parte de los informantes de continuar su colaboración con la encuesta). Por esta razón, la estimación de cargas de trabajo se realiza en relación a la muestra objetivo y no respecto de la muestra total, la que incluyó un sobremuestreo.

Tabla 39: Distribución del personal de trabajo de campo VIII EPF

Región	Digitadores	Encuesta- dores	Encarga- dos de grupo	Encarga- dos regionales	Encarga- dos de área	Encargado nacional	Total personal
Arica y Parinacota		4	1				5
Tarapacá		4	1				5
Antofagasta		6	2				8
Atacama		3	1				4
Coquimbo		6	1				7
Valparaíso		15	4	1			20
Metropolitana	2	72	15	5 ⁽²⁾	2	1	97
O'Higgins		4	1				5
Maule		4	1				5
Ñuble ⁽¹⁾		5	1				6
Biobío		12	3	1			16
La Araucanía		5	1				6
Los Ríos		3	1				4
Los Lagos		3	1				4
Aysén		3	1				4
Magallanes		3	1				4
Nacional	2	152	36	7	2	1	200

(1) Corresponde a equipo responsable de Chillán y Chillán Viejo, comunas que compondrán la futura Región de Ñuble.

(2) Incluye encargados regionales de las zonas norte y sur del área Resto de capitales regionales.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

8.1.1.3 Perfiles de cargo, selección y condiciones contractuales del personal de trabajo de campo

Una vez definida la estructura de roles necesarios para el proceso de trabajo de campo, se elaboraron perfiles de cargo, estableciendo características deseables a nivel de formación educacional, experiencia laboral, de conocimiento informático y de atributos psicolaborales para cada perfil, según las labores a desarrollar.

El proceso de selección del personal de campo contempló las etapas de evaluación curricular¹¹⁹, evaluación psicolaboral y entrevista personal con coordinadores.

La evaluación curricular y psicolaboral de los postulantes estuvo a cargo de psicólogos del Subdepartamento de Selección y Desarrollo Organizacional del INE. La evaluación psicolaboral contempló la medición de variables de inteligencia general y rasgos de personalidad, mediante pruebas individuales y entrevistas grupales. Entre los atributos transversales considerados en esta etapa se encuentran la tolerancia a la frustración, flexibilidad ante escenarios cambiantes y capacidad de trabajo en equipo. Finalmente, en la entrevista personal con coordinadores, se evaluó la habilidad comunicativa de los postulantes y se indagó detalles respecto de su experiencia laboral.

119 En el anexo 1 se presenta una síntesis de los perfiles de cargo de trabajo de campo.

Dada la periodicidad quinquenal del estudio, el personal es contratado exclusivamente por el período que duran las labores de campo de la VIII EPF, desde la capacitación hasta el cierre de la última submuestra, sin que necesariamente cuenten con continuidad laboral en la institución¹²⁰.

El personal de trabajo de campo fue contratado por 44 horas de trabajo semanal, de lunes a viernes. Los encargados de grupo, al igual que los encargados regionales, de área y nacional, debían cumplir su tiempo contratado en jornadas diarias predefinidas. Ello se debe a la necesidad de contar con horarios fijos para facilitar las comunicaciones entre la coordinación nacional de trabajo de campo, las direcciones regionales y los encargados de grupo.

Por su parte, los encuestadores tenían flexibilidad para distribuir su jornada semanal, sin que existiera un mínimo de horas diarias de trabajo, sino solo la condición de asistir todos los días a la oficina para participar de las labores de coordinación y entrega de material para revisión del encargado de grupo¹²¹. Los trabajos en días sábado, domingos y festivos estaban permitidos y eran compensados con tiempo de descanso complementario.

La decisión de que los encuestadores contaran con flexibilidad horaria se debió a la necesidad de mayor adaptación a los horarios de disponibilidad de los informantes del estudio. Este aspecto resulta especialmente importante en capitales regionales de mayor extensión territorial, donde la población tarda más en retornar a sus viviendas debido a largos tiempos de desplazamiento, aspecto que afecta las posibilidades de contacto con informantes idóneos en horarios de oficina.

En función de la necesidad de contar con un equipo estable y altamente capacitado, todo el personal de trabajo de campo recibía una remuneración mensual fija. Esta decisión fue tomada con la finalidad de evitar la rotación de personal, especialmente de encuestadores, y desincentivar el falseamiento de información, disminuyendo la presión que genera una política de pagos por encuesta lograda¹²². A la vez, para garantizar el adecuado cumplimiento de sus funciones, los integrantes de los grupos de campo eran evaluados a través del monitoreo de indicadores de desempeño operativo y de la supervisión de la calidad de los datos producidos¹²³.

120 El período de contratación de los encargados regionales fue de un año once meses, mientras que el de los encargados de grupo fue de un año siete meses. Los encuestadores permanecieron en el proyecto durante un año cuatro meses. La única excepción respecto a la continuidad en la institución correspondió a los encargados de área y nacional de trabajo de campo, quienes son parte de la dotación permanente del Departamento de Presupuestos Familiares.

121 Si bien los encuestadores no tenían una meta mínima de hogares a encuestar, su trabajo era supervisado a través del control de cumplimiento de las horas de trabajo pactadas (mediante registro diario de asistencia), la evaluación de la calidad de los datos producidos y el monitoreo de indicadores de desempeño operativo individuales (más detalles respecto de estos indicadores se exponen en el apartado 8.1.5.2). Adicionalmente, tal como será explicado en el apartado 8.1.3.2, la metodología de trabajo contemplaba mecanismos de reasignación de cargas de trabajo, los que permitían mantener cargas equilibradas entre encuestadores de un mismo grupo de campo.

122 Uno de los elementos evaluados como críticos asociados al pago por encuesta en este tipo de encuestas donde existen altas tasas de rechazo en estratos socioeconómicos altos es que, justamente, a través del pago por encuesta, se perjudica a aquellos encuestadores que desempeñan labores de trabajo de campo en estratos socioeconómicos altos.

123 La calidad y veracidad de los datos producidos por los encuestadores se evaluó, en primera instancia, a través de la revisión realizada por los encargados de grupo (detallada en el apartado 8.1.3) y mediante los procesos de supervisión, descritos en el apartado 8.2.3. Una vez avanzado el proceso, la calidad de los datos se evaluó también a través de los indicadores reportados desde los subprocesos de digitación, codificación y validación de datos. Para mayor información de estos indicadores, ver Informe de Calidad VIII EPF (INE Chile, 2018f).

8.1.1.4 Capacitación del personal de trabajo de campo

El proceso de capacitación del personal de trabajo de campo fue diferenciado según el perfil de cargo. Durante aquel proceso, se buscó dotar a las personas de los conocimientos y herramientas necesarias para garantizar la calidad de los datos del estudio desde el ejercicio de sus funciones específicas.

El modelo de capacitación implementado durante el trabajo de campo comprendía dos momentos o etapas de formación:

- **Capacitación inicial:** consistía en la inducción respecto del proceso productivo de la encuesta, conceptos muestrales básicos, metodología de trabajo de campo y cuestionarios de recolección de datos (conceptos centrales, estructura y preguntas). Fue desarrollada a través de clases expositivas y talleres prácticos, a través de los cuales se explicaba detalladamente las labores de terreno y gabinete (en oficina) que debía ejecutar cada perfil.

Las actividades desarrolladas para la capacitación inicial variaron según el momento de ingreso del personal y el perfil de cargo. Al inicio del proyecto, se desarrolló un curso presencial para encargados de grupo y supervisores, desarrollado durante dos semanas (70 horas cronológicas) en Santiago. Este curso contemplaba, en primer lugar, una semana de instrucción conjunta sobre aspectos básicos de la encuesta, la que era impartida por integrantes de los equipos de trabajo de campo y técnico. Durante la segunda semana, cada equipo impartió sesiones de formación diferenciada para encargados de grupo¹²⁴ y supervisores, explicando las funciones específicas de cada cargo¹²⁵ y el uso de herramientas de trabajo disponibles.

Posteriormente, al ingresar el primer grupo de encuestadores del estudio, se realizó en cada región un curso presencial de dos semanas de duración (70 horas cronológicas), el cual fue guiado por una dupla docente, compuesta por un integrante del equipo técnico y un miembro del equipo trabajo de campo o del equipo de supervisión y capacitación (es decir, una dupla técnico-operativa). Además de los encuestadores, en estos cursos participaron también los encargados de grupo, con el propósito de que contaran con un reforzamiento de los conceptos y metodología de la encuesta¹²⁶.

A diferencia de lo anterior, para los encuestadores ingresados durante el período de trabajo de campo oficial (rotación de personal), la capacitación inicial contempló una semana de inducción en gabinete (44 horas cronológicas), en base a lectura de manuales y ejercicios prácticos, y otra semana consistente en un curso presencial intensivo (35 horas cronológicas), realizado en la capital nacional (Santiago), en el cual se revisaban principales contenidos de la encuesta y se aclaraban dudas de

124 La capacitación de encargados de grupo contó con contenidos levemente diferenciados, considerando que los grupos del área Resto de capitales regionales dependían administrativamente de las direcciones regionales del INE y que su localización geográfica incorporaba diferencias en labores logísticas de envío y recepción de materiales.

125 Para conocer las principales funciones de cada cargo ver anexo 1. Para mayor detalle, revisar Manual de Trabajo de Campo para Encuestadores, Tomo I (INE Chile, 2018b).

126 En el primer curso de inducción de encuestadores se capacitó un número adicional de encuestadores en cada región, dando la oportunidad de seleccionar a quienes presentaban los mejores desempeños relativos en las pruebas de evaluación de aprendizajes del curso.

los participantes. Estos cursos fueron impartidos por integrantes de la coordinación nacional de trabajo de campo y el equipo de supervisión y capacitación.

- **Capacitación continua:** comprendía el reforzamiento de contenidos de mayor complejidad de la encuesta, a través de sesiones breves de formación respecto a módulos específicos de los cuestionarios. Fueron impartidos presencialmente (en el caso de la Región Metropolitana) y a través de videoconferencias (para el Resto de las capitales regionales) por el equipo de supervisión y capacitación, empleando para ello la combinación de clases expositivas y talleres prácticos.¹²⁷

Como recursos de capacitación permanentes, los equipos contaban con manuales específicos para cada cargo y la Zona de consultas¹²⁸, que correspondía a un foro virtual disponible en el sistema informático de la encuesta, que funcionaba como medio nacional para resolver dudas técnicas y operativas del estudio.

8.1.2 Características generales del proceso de trabajo de campo

8.1.2.1 Período de recolección y distribución operativa de la muestra

El proceso de trabajo de campo oficial de la VIII EPF tuvo lugar entre el 1 de julio de 2016 y el 30 de junio de 2017. Durante este período se buscó entrevistar a los hogares residentes en una muestra total de 25.361 viviendas, distribuidas en doce submuestras mensuales de 2.113 viviendas en promedio. El proceso buscó entrevistar a todos los hogares residentes en las viviendas seleccionadas.

La muestra estuvo distribuida en un total de 2.124 manzanas, cada una de las cuales fue asignada a solo una submuestra mensual. Aquella decisión permitió controlar la cantidad de desplazamientos necesarios por parte de encuestadores y encargados de grupo. Adicionalmente, la permanencia en un territorio durante un mes permitía afianzar la relación de confianza con los informantes y sus comunidades, además de focalizar los esfuerzos de difusión del estudio¹²⁹.

Cada submuestra mensual fue subdividida en dos grupos de tamaño similar de viviendas, llamados submuestras operativas, las cuales estaban asociadas a un período específico de colaboración, correspondiente a una de las quincenas¹³⁰ del mes. De este modo, durante el trabajo de campo oficial del estudio se entrevistó un total de 24 submuestras operativas.

La asignación de las viviendas a una submuestra operativa se realizó de forma sistemática, de modo que los distintos sectores de una manzana tuviesen presencia en ambas

127 Más detalles respecto de la planificación, temas abordados y forma de implementación de la capacitación continua se exponen en el apartado 8.2.7.

128 Más información respecto a la Zona de consultas disponible en el apartado 9.1.1.

129 Para una explicación respecto a la desagregación temporal de la muestra de la encuesta en submuestras mensuales, consultar el apartado 6.1.3.3.

130 El concepto de quincena es referencial, en tanto su duración dependerá de los días contenidos en el mes de trabajo de campo. Así, en la VIII EPF existieron “quincenas” de 14 (febrero), 15 y 16 días de duración. De este modo, la quincena es equivalente al período de colaboración de los hogares en la LGI. La quincena de colaboración asignada podía ser modificada, en casos excepcionales y por razones de fuerza mayor de los informantes. No obstante, las viviendas solo podían ser entrevistadas dentro del mismo mes al que estaban asignadas originalmente. Por ejemplo, si una vivienda fue asignada a la segunda quincena de noviembre de 2016, solo podía cambiar su período de colaboración a la primera quincena del mismo mes.

quincenas¹³¹. La única excepción a este modo de distribución se aplicó a las manzanas con cinco o menos viviendas a encuestar. En estos casos, todas las viviendas de la manzana fueron asignadas a una misma submuestra operativa, con el fin optimizar los recursos monetarios y humanos asociados a traslados y tiempos de trabajo que los encuestadores destinaban a desplazamientos.

8.1.2.2 Técnica de producción de información

En cada submuestra operativa se produjo información mediante la aplicación de seis cuestionarios estructurados a cada hogar residente en las viviendas seleccionadas en la muestra. Adicionalmente, se trabajó con un formulario de control del trabajo de campo, llamado Hoja de Ruta, destinado a registrar todas las actividades realizadas diariamente en terreno en el proceso de producción de datos. El formato de aplicación de estos instrumentos fue en papel.

De los seis cuestionarios que recopilaban información de los hogares, la Libreta de Gastos Individuales consistía en un diario de autoreporte donde cada miembro del hogar de 15 años o más debía reportar sus gastos diarios durante la quincena. El resto de los cuestionarios era aplicado por el encuestador mediante entrevista directa a un informante idóneo (para mayor detalle respecto a las libretas o cuestionarios, revisar el apartado 5.1.4).

Los informantes idóneos podían variar según cada cuestionario. De este modo, para el RPH (instrumento entrevistado en la primera visita) el informante idóneo debía ser un miembro del hogar de 18 años o más¹³². A diferencia de lo anterior, para el resto de los cuestionarios el informante idóneo podía ser cualquier miembro del hogar de 15 años o más que conociera la información consultada, preferentemente el titular de los gastos o ingresos. En todos los casos, el informante idóneo debía ser una persona que contara con las capacidades físicas y/o mentales necesarias para responder el estudio.¹³³

8.1.3 Organización del proceso de trabajo de campo de una submuestra

Considerando la complejidad de las distintas labores involucradas en el trabajo de campo, al igual que en la VII EPF, las actividades a desarrollar en cada submuestra operativa fueron organizadas en tres fases¹³⁴ consecutivas:

- i. **Preparación**, previa al inicio de la quincena y destinada a generar las condiciones para el contacto de los informantes.
- ii. **Aplicación**, correspondiente al período de colaboración de los informantes y destinada a aplicar los distintos cuestionarios.

131 Las manzanas corresponden a un área geográfica delimitada por fronteras culturales (calles de uso público, por caminos, líneas férreas) o naturales (cerros, esteros, playas, ríos u otros).

132 La razón por la cual la edad requerida para el informante idóneo del RPH es mayor al resto de los instrumentos se debe a que se estableció que la autorización a la aplicación de la encuesta en el hogar debía darse por un integrante que hubiese cumplido la mayoría de edad legal (18 años).

133 Si bien la evaluación de las capacidades mentales de una persona excede las posibilidades de conocimientos de un encuestador, en los manuales y capacitaciones se sugirió como indicador la posibilidad de sostener una conversación coherente durante unos minutos.

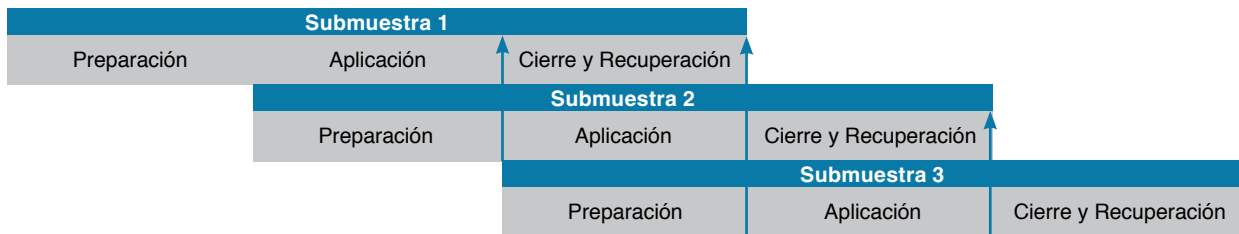
134 Las fases que aquí se identifican asociadas al trabajo de campo no deben confundirse con la mención a las fases que componen el modelo de producción estadística GSBPM.

- iii. **Cierre y recuperación**, fase posterior a la quincena donde se realiza la revisión y recuperación de información faltante.

Cada fase del trabajo de campo contempló una serie de actividades y tareas, orientadas a maximizar los niveles de contacto y cooperación de los hogares, al mismo tiempo que buscaban resguardar la calidad de los datos producidos. De este modo, si bien el período de colaboración de los hogares se extendió durante 15 días aproximadamente, el trabajo completo de la submuestra operativa se extendió por, al menos, 37 días continuados¹³⁵.

Puesto que las quincenas eran consecutivas, sin brechas ni traslapes durante el año calendario, los equipos de trabajo de campo realizaban simultáneamente actividades de las tres fases del trabajo de campo en distintas submuestras. Una representación de esta superposición de fases puede observarse en el diagrama 7.

Diagrama 7: Fases de trabajo de campo en cada submuestra VIII EPF



Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

A continuación, se presenta una síntesis de las actividades contempladas por cada fase de trabajo, así como sus características más relevantes.

8.1.3.1 Fase de preparación

Su objetivo fue generar las condiciones para que el o los hogares que habitaban las viviendas seleccionadas, aceptasen participar de la VIII EPF. Podía iniciarse desde 35 días antes del período de colaboración de los hogares (quincena). Esta fase contemplaba las siguientes actividades:

- **Precarga de la submuestra al sistema informático:** involucraba la revisión de consistencia del directorio de viviendas seleccionadas, asignación de submuestra operativa, ingreso de datos del directorio al sistema informático del estudio y asignación informática de manzanas a cada grupo de campo. Era responsabilidad de la coordinación nacional del trabajo de campo y fue realizada mensualmente.
- **Organización previa del trabajo en terreno:** contemplaba la distribución, revisión y organización de los materiales necesarios para realizar el trabajo de una submuestra operativa. Su principal responsable fue el encargado de grupo.

¹³⁵ Tal como se explicará en las siguientes páginas, la duración del trabajo de una submuestra podía variar según la quincena del mes, los plazos de cierre con los que contaban las regiones y la existencia de planes especiales que modificaban los plazos de las actividades. De este modo, una submuestra operativa podía involucrar hasta 60 días.

- **Reconocimiento de manzanas en terreno:** consistía en la verificación en terreno de las unidades de primera selección, permitiendo conocer con anticipación las características del territorio. Esta actividad incluía la difusión del estudio en el sector a encuestar, mediante el pegado de afiches¹³⁶ en espacios de alta concurrencia. Su responsable era el encargado de grupo y fue realizado mensualmente.
- **Identificación de unidades incluidas en el directorio:** se realizaba simultáneamente al reconocimiento de manzanas y consistía en la verificación de las unidades muestrales seleccionadas para una submuestra mensual. Esta actividad era realizada inicialmente por el encargado de grupo, para luego ser reiterada por el encuestador. Incluía la entrega de una carta institucional dirigida al jefe de hogar (carta de apertura¹³⁷), en la que se informaba de la selección de la vivienda para participar del estudio. La carta llegaba al menos dos semanas antes de que el encuestador acudiera a la vivienda, con el fin de propiciar la colaboración de los informantes.

En caso de detectar viviendas seleccionadas al interior de comunidades de acceso controlado (edificios, condominios, etc.), esta actividad incorporaba el contacto y entrega de cartas institucionales a administradores y/o representantes de comunidades, informando el total de viviendas en que debía aplicarse la encuesta y el período en que se realizaría el proceso.

- **Distribución y asignación de cargas de trabajo:** implicaba definir las viviendas de las que se haría responsable cada encuestador, considerando las características del sector a encuestar y localización territorial. Este proceso debía realizarse, a más tardar, ocho días antes del inicio de la quincena e involucraba asignación informática de cargas de trabajo y entrega de directorios, cartografía, material de difusión y cuestionarios a los encuestadores. El responsable de esta actividad era el encargado de grupo.
- **Contacto inicial con la unidad muestral:** consistía en la visita de los encuestadores a las viviendas asignadas, para persuadir a los hogares residentes a que participaran del estudio. Esta actividad se iniciaba con al menos siete días de anticipación al primer día de la quincena y su principal responsable era el encuestador de hogares. Durante esta actividad también estaba permitido concertar citas de apertura con los informantes para otro día cercano al inicio de la quincena de colaboración¹³⁸.

Si durante el contacto inicial se lograba persuadir a los informantes, estaba permitido realizar la entrevista del RPH, también conocida como apertura de vivienda. Adicionalmente, los dos días previos al inicio de la quincena, estaba permitida la entrega de LGI a los integrantes de 15 años o más.

136 Los afiches tenían por objetivo informar la presencia de los encuestadores en el sector, indicando el mes en que estarían aplicando encuestas en la manzana. Un ejemplo de los afiches distribuidos en el territorio se puede observar en el anexo 2.

137 Se puede observar un ejemplo de esta carta en el anexo 3.

138 Tal como se explicará en el apartado 8.1.3.2, la apertura solo podía realizarse hasta el quinto día de la quincena de colaboración.

Ante viviendas con moradores ausentes, los encuestadores debían realizar al menos cinco intentos de contacto, en distintos días y horarios, antes de cesar las visitas de entrevista a un hogar. A su vez, cuando no era posible conseguir la participación de los hogares en la encuesta debido a rechazos, la vivienda era reasignada a otro encuestador para reiniciar los intentos de entrevista. Esta acción será explicada con mayor detalle en el siguiente apartado.

Dada la normativa vigente, el INE no está facultado para entregar incentivos económicos u obsequios a los informantes¹³⁹, por lo cual las labores de persuasión eran realizadas utilizando como medios una segunda carta informativa (idéntica a la entregada previamente), un díptico informativo¹⁴⁰, la identificación de los encuestadores como funcionarios del Instituto (mediante credencial y vestuario institucional) y los argumentos sugeridos en capacitaciones y en el Manual de Trabajo de Campo para Encuestadores (INE Chile, 2018b).

8.1.3.2 Fase de aplicación

Esta fase buscó producir, junto a los hogares residentes en las viviendas seleccionadas, la información necesaria para el cumplimiento de los objetivos de la encuesta. Duraba 15, 16 o 17 días, es decir, un día más que el período de colaboración de la LGI (quincena¹⁴¹), a fin de aumentar la probabilidad de registro de los gastos individuales del último día de referencia.

Para aquellos casos en que se lograba la colaboración de los hogares, las entrevistas fueron organizadas en cuatro visitas durante la quincena de colaboración del hogar, con el fin de evitar el agotamiento de los informantes, dar apoyo a los integrantes del hogar en el registro de gastos diarios y asegurar la oportunidad de recuperar información faltante o imprecisa en los primeros instrumentos aplicados.

Cada una de esas visitas tenía períodos de realización predefinidos por un calendario de referencia¹⁴². En general, ese calendario se estructuró en base a los plazos fijados en el esquema de visitas, que se expone en la tabla 40. La secuencia de las entrevistas fue organizada en un orden que permitiese consultar desde la información más general (información sociodemográfica), pasando por la diversidad de gastos realizados por un hogar, hasta llegar a la información más sensible (ingresos y gastos financieros). De este modo, se buscó disminuir el impacto negativo de consultar información interpretada como privada por los hogares, como es el caso de los ingresos o deudas.

La primera entrevista de cada cuestionario debía realizarse en una visita específica, pudiendo recuperarse información en las visitas siguientes. Por su parte, las LGI eran entregadas y

139 Esta situación difiere respecto a lo que realizan otros Institutos de Estadística donde la entrega de incentivos monetarios ha logrado mejorar los indicadores de desempeño operativo. Al respecto, es ilustrativo revisar la experiencia del INE de España donde a partir del año 2007 introdujeron incentivos a los hogares participantes a través de la entrega de tarjetas de regalo (INE España, 2016).

140 Presentado en anexo 4.

141 Dado que las viviendas se asignaban a una quincena de recolección de información, estas podían contener 14, 15 o 16 días dependiendo si el mes contaba con 28, 30 o 31 días.

142 Un calendario de referencia consistía en la publicación de fechas de inicio y término del plazo para realizar cada actividad del trabajo de campo en una submuestra.

explicadas durante la primera entrevista, para posteriormente hacer seguimiento de su llenado durante cada visita siguiente.

Las visitas contaban con un plazo de aproximadamente cuatro días para ser realizadas. Ello buscaba generar un equilibrio entre la necesidad de llevar un avance ordenado en el trabajo de campo y la flexibilidad requerida para adaptarse a la disponibilidad de tiempo de los informantes. A medida que se realizaban las visitas, ello era registrado en la Hoja de Ruta, la cual era digitada durante el proceso, permitiendo a la coordinación nacional monitorear el avance del trabajo en cada grupo de campo.

Todas las visitas involucraban trabajo de gabinete previo a las visitas. El trabajo de gabinete consistía en la preparación del material de entrevista, registrando la información muestral de los cuestionarios a aplicar y chequeando, a partir del RPH, algunos filtros de cada libreta, los que permitían identificar qué preguntas correspondía aplicar a cada miembro del hogar. Posterior a las visitas, el trabajo de gabinete consistía en la revisión de la información producida y registro de observaciones explicativas.

Tabla 40: Esquema de visitas VIII EPF

Preparación	Aplicación				Cierre
	Visita 1	Visita 2	Visita 3	Visita 4	
Mes previo al inicio de quincena	Desde 2 días antes hasta día 5 de la quincena	Desde día 6 hasta día 9 de la quincena	Desde día 7 hasta penúltimo día de la quincena	Día siguiente al fin de la quincena	6 días siguientes al fin de la quincena
<ul style="list-style-type: none"> · Precarga de muestra en sistema informático. · Asignación de cargas de trabajo. · Reconocimientos unidades muestrales. · Entrega de cartas de apertura. · Contacto inicial. 	<ul style="list-style-type: none"> · Identificación de hogares en la vivienda. · Entrevista RPH. · Entrega y explicación de LGI. 	<ul style="list-style-type: none"> · Recuperación RPH. · Entrevista LGH. · Entrevista LGR. · Entrega Hoja Auxiliar LGR⁽¹⁾. · Seguimiento de LGI. 	<ul style="list-style-type: none"> · Recuperación LGH. · Recuperación LGR. · Entrevista LI. · Entrevista LSF. · Retiro de Hoja Auxiliar LGR. · Seguimiento de LGI. 	<ul style="list-style-type: none"> · Recuperación LI. · Recuperación de LSF. · Seguimiento y retiro de LGI. 	<ul style="list-style-type: none"> · Última revisión. · Recuperación de datos. · Orden de material. · Despacho a Digitación.

(1) La Hoja Auxiliar (entregada en la visita 2 y retirada en la visita 3) corresponde a un recurso de apoyo al registro de los gastos consultados en la Libreta de Gastos del Recuerdo (LGR). Su contenido consiste en una tabla tipo donde era posible declarar productos o servicios adquiridos por el hogar, registrando cantidad, descripción, valor total, identificación de compra por internet, tipo de establecimiento y fecha de compra (mes/año). Este instrumento se dejaba en el hogar para facilitar la declaración de gastos consultados en la LGR por parte de integrantes que no participaron de aquella entrevista.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

La información producida durante las entrevistas era revisada por los encargados de grupo tras cada visita, a fin de supervisar el avance de la fase de aplicación y detectar posibles imprecisiones u omisiones en los cuestionarios. Estas revisiones eran registradas en el Formulario de Correcciones (instrumento digital alojado en el sistema informático del estudio) y, a partir de estos datos, los encargados de grupo informaban por escrito a los encuestadores las instrucciones de corrección y solicitudes de recuperación de datos, las que debían ser ejecutadas en la siguiente visita al hogar.

En síntesis, las actividades desarrolladas durante esta fase fueron las siguientes.

- i. **Identificación de hogares en la vivienda:** consistía en la reiteración de la actividad de identificación de unidades incluidas en el directorio y en la identificación de los hogares residentes en la vivienda seleccionada. Esta actividad fue responsabilidad del encuestador de hogares, quien se apoyó en preguntas presentes en la portada del RPH para desarrollarla. Cabe aclarar que la identificación de hogares es parte de un protocolo institucional que busca resguardar la calidad de los procesos de recolección de datos en muestras de viviendas¹⁴³.
- ii. **Primera visita de entrevista al hogar:** implicó la aplicación del RPH y la entrega de las LGI a cada integrante de 15 años o más que estuviera en condiciones físicas y/o mentales aptas para responder aquel instrumento. Esta visita marca el inicio de la colaboración de los informantes y podría realizarse, a más tardar, el día 5 de la quincena.

Al finalizar esta visita, se acordaba la fecha y hora del siguiente encuentro y el encuestador informaba al hogar de la necesidad de disponer de comprobantes de gastos periódicos del hogar (cuentas de agua, luz, gas; mensualidades de establecimientos educativos, etc.) para la siguiente visita, con el fin de facilitar la entrevista de los cuestionarios que se aplicarían en esa oportunidad.

- iii. **Revisión y reasignación de cargas de trabajo:** involucró la evaluación, por parte del encargado de grupo, de la cantidad de hogares efectivamente aperturados por cada encuestador y la reasignación de viviendas¹⁴⁴ entre encuestadores de un mismo grupo de campo, en caso de producirse desequilibrios importantes en las cargas de trabajo o rechazos a participar del estudio.

La reasignación de hogares con rechazo a otro encuestador responde a la necesidad de reiniciar la interacción con los hogares, evitando posibles efectos negativos derivados de la predisposición de informantes y/o encuestadores tras un primer encuentro con resultados negativos. De este modo, la reasignación permitía que un nuevo encuestador, con características y estilos de persuasión distintos al encuestador inicial, realizara nuevos intentos de contacto, buscando adaptarse de mejor modo a las características del informante.

Para revertir situaciones de rechazo, también se emplearon formatos de cartas institucionales personalizadas, en las cuales se solicitaba reconsiderar la decisión de no colaborar con la encuesta, aludiendo a la relevancia de los resultados de la EPF para el IPC y su impacto en las decisiones de política pública. Solo cuando se producía un segundo rechazo se permitía abandonar los esfuerzos de entrevista en un hogar.

143 Al respecto revisar apartado 5.1.1 donde se revisa el concepto de hogar utilizado por la encuesta.

144 La reasignación se realizaba a nivel de viviendas y no de hogares para facilitar el vínculo con los informantes y optimizar los desplazamientos de los encuestadores. Esto quiere decir que, si en una vivienda existían dos hogares, al realizar la reasignación de las cargas, ambos hogares eran entrevistados por el mismo encuestador.

- iv. **Segunda visita de entrevista al hogar:** involucraba la recuperación de información pendiente del RPH, la aplicación de la LGH, de la LGR y el seguimiento a la autorreporte de gastos realizada en las LGI.

El seguimiento a las LGI consideraba la revisión, día por día, de los gastos registrados, la verificación de la existencia de gastos en los días en que no se produjeron registros y el apoyo en el registro de gastos no consignados previamente.

Adicionalmente, al final de esta visita se entregaba la Hoja Auxiliar LGR, formulario destinado a complementar la información declarada en LGR, a partir del autorreporte de gastos comprendidos por ese cuestionario que no hubiesen sido declarados durante la entrevista. De este modo, la Hoja Auxiliar LGR buscaba dar oportunidad a los integrantes del hogar ausentes en la entrevista para que registraran los gastos de los últimos tres, seis y doce meses consultados por ese cuestionario.

Al final de la segunda visita de entrevista se fijaba la fecha y horario de la siguiente visita, para la cual se solicitaba la disposición de comprobantes de ingreso (liquidaciones de sueldo, comprobantes de pago de pensiones, entre otros) y de gastos en servicios financieros (recibos de créditos, cuentas bancarias, etc.), para hacer más expedita la siguiente entrevista¹⁴⁵.

- v. **Tercera visita de entrevista al hogar:** comprendía la recuperación de información faltante de la LGH, además del retiro de la Hoja Auxiliar LGR y recuperación de datos faltantes de la LGR. Luego de ello, involucraba la aplicación de la LI y LSF, para, finalmente, realizar también el seguimiento de las LGI.
- vi. **Cuarta visita de entrevista y retiro de LGI en el hogar:** incluía la recuperación de datos faltantes en cualquiera de los cuestionarios antes entrevistados. Su principal actividad consistía en la revisión de los datos registrados en la LGI, complementando descripciones de gasto u otros datos que lo requirieran. Solo se permitía realizar esta visita con posterioridad al último día de la quincena.

Al retirar las LGI el encuestador informaba al hogar la posibilidad de que volviese a contactarlo, telefónica o presencialmente¹⁴⁶, para precisar alguno de los datos aportados al estudio.

8.1.3.3 Fase de cierre y recuperación

Tuvo como fin evaluar y mejorar la calidad del dato producido, para el posterior envío de los materiales a digitación. Tuvo una duración de entre seis y diez días¹⁴⁷, a partir del fin de la fase de aplicación. Las actividades desarrolladas en esta fase eran las siguientes:

145 Si bien se solicitaba y era deseable que los informantes pudiesen contar con esta información como apoyo para las entrevistas, no era condición necesaria para realizar las siguientes entrevistas, ya que se recurría a la declaración de los informantes (recordación) si no contaban con comprobantes.

146 Para realizar las recuperaciones, también fue permitido el uso de correo electrónico institucional como medio de comunicación con los informantes. No obstante, para permitir su uso se exigía la comunicación telefónica previa con el informante. Más detalles respecto de este procedimiento pueden conocerse en el Manual de Trabajo de Campo para Encuestadores (INE Chile, 2018b).

147 Se otorgaba un plazo mayor de cierre a grupos de campo con mayor concentración muestral, con el propósito de facilitar el orden en el proceso y otorgar a los encargados de grupo un tiempo suficiente para la revisión final del material.

- i. **Revisión previa al cierre de libretas:** implicó la revisión, por parte de encuestadores y encargados de grupo, de todas las libretas entrevistadas en los hogares, verificando la completitud y consistencia de la información solicitada y el cumplimiento de las instrucciones de corrección. Junto a lo anterior, esta actividad incluía el registro en las LGI de los gastos declarados mediante el pegado de boletas en estos instrumentos.

En caso de existir información faltante o inconsistente, los encargados de grupo podían solicitar la recuperación de los datos, mediante llamado telefónico o visita al hogar. Los encuestadores podían intentar la recuperación telefónica en cinco oportunidades. Si no tenían éxito tras esos intentos necesariamente debían visitar al hogar para recuperar la información.

- ii. **Visita de recuperación:** consistía en una visita adicional al hogar, con el fin de recuperar información faltante. Su realización era obligatoria en aquellos casos en que se omitiera la respuesta en más de tres preguntas o en alguna pregunta filtro.

Este proceso exigía realizar al menos tres intentos de contacto en la vivienda. Si, pese a esos esfuerzos, no era posible recuperar la información, debían agregar una observación al cuestionario con información faltante, explicando la situación.

- iii. **Envío de encuestas a digitación:** constituía la última actividad realizada en el trabajo de una submuestra e incluía la organización de los cuestionarios aplicados y la elaboración de guías de despacho por parte de los grupos de campo para la entrega del material a la coordinación nacional. Esta última recibía el material, verificaba la completitud del despacho¹⁴⁸, confirmaba su recepción en el sistema informático y, finalmente, entregaba el material al equipo de digitación.

Respecto a esta actividad, cabe aclarar que el sistema informático impedía la emisión de guías de despacho en aquellos casos que, de acuerdo a los registros del Formulario de Correcciones, mantuvieran errores sin corregir. Este control informático tenía por objetivo procurar que los encargados de grupo ingresaran una observación que explicara el motivo por el cual no fue posible recuperar la información faltante.

Todas las libretas aplicadas en una submuestra operativa eran incluidas en un solo despacho, con el fin de evitar posibles pérdidas de material. Asimismo, la entrega del material al equipo de digitación procuraba la entrega de submuestras operativas completas a nivel nacional, resguardando que todo el material estuviese físicamente en un mismo lugar, para mayor orden en la administración de las encuestas aplicadas.

8.1.4 Planes especiales implementados durante el trabajo de campo

El modo de organización del trabajo de campo, presentado en apartado anterior, buscó generar un flujo de proceso detallado y estándar, en el cual se previó una serie de

¹⁴⁸ La verificación de la completitud del despacho se realizaba a través de la lectura de los códigos de barra presentes en cada libreta, con lo cual se comprobaba que todos los cuestionarios aplicados en un hogar se hubiesen incluido en el despacho.

contingencias, propias de un proceso de producción de datos en terreno. No obstante, la existencia de hitos nacionales, tales como fiestas patrias, feriados extensos o la realización del Censo de Población y Vivienda 2017, hicieron necesario realizar ajustes al trabajo de campo de algunas submuestras, estableciendo planes especiales.

Los planes especiales fueron diseñados por la coordinación nacional del trabajo de campo, en acuerdo con el equipo de gestión central de la encuesta. Tenían como objetivo resguardar la calidad de los datos y los niveles de cooperación de los informantes, dando una respuesta estándar ante situaciones con efecto nacional que dificultaban el normal desarrollo del trabajo de campo de una o varias submuestras. Los planes especiales adoptados y sus motivos se presentan en la tabla 41.

Tabla 41: Periodos y motivos de planes especiales

Plan especial	Submuestras operativas involucradas	Mes calendario	Motivo
Fiestas patrias 2016	05 - 06	Septiembre 2016	Potencial ausencia prolongada de informantes, especialmente en estrato alto, durante periodo de apertura de submuestra operativa 06, debido a feriados irrenunciables de fiestas patrias y semana de vacaciones en algunos establecimientos educacionales.
Feridos consecutivos 2016	09 - 10	Octubre - Noviembre 2016	Potencial ausencia de informantes en periodo de cierre de submuestra 09 y apertura de submuestra 10 por dos feriados consecutivos, a continuación de último fin de semana de octubre.
Verano 2016-2017	12 - 13 - 14 - 15	Diciembre 2016 - Febrero 2017	Potencial disminución de colaboración de informantes por fiestas de fin de año (navidad y año nuevo) e inicio de periodo estival, comúnmente utilizado para vacaciones.
Censo 2017	19 - 20	Abril 2017	Realización de Censo 2017, que requirió la colaboración de todo el personal INE. Adicionalmente, feriado irrenunciable asociado al día censal implicaba la ausencia potencial de informantes durante la semana del censo, coincidente con periodo de apertura de submuestra 20.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

En cada uno de estos hitos se realizó una comunicación nacional, informando los ajustes propuestos para afrontar el período de dificultad. La implementación de los planes era apoyada por los encargados regionales, quienes intensificaban sus acciones de monitoreo en estos periodos y apoyaban la resolución de situaciones contingentes que afectaran el desarrollo del trabajo de campo. El tipo de medidas de ajuste adoptadas en los planes especiales se presentan en la tabla 42.

Tabla 42: Tipos de medidas adoptadas en planes especiales

Ajuste	Descripción	Submuestras Operativa	Plan especial
Realización simultánea de contactos iniciales de un mismo mes.	<p>Realización de los contactos iniciales de ambas submuestras del mes durante la fase de preparación de la primera quincena del mes.</p> <p>Consulta a los hogares contactados sobre viajes y otras actividades planificadas que obstaculizaran la aplicación de la encuesta en la quincena asignada⁽¹⁾.</p>	05 - 06 - 11 - 12 - 13 - 14 - 15 - 16 - 19 - 20	Fiestas Patrias 2016 Verano 2016-2017 Censo 2017
Evaluación de cambio de submuestra operativa	<p>Flexibilización de criterios de autorización de cambio de submuestra operativa de vivienda, debido a ausencias prolongadas de informantes.</p> <p>Autorización en función de evaluación de disposición a colaborar por parte del informante y disponibilidad de días suficientes de presencia en la vivienda para realización de visitas.</p> <p>Autorización acompañada de monitoreo de equilibrio en tamaño submuestras operativas.</p>	11 - 12 - 13 - 14 - 15 - 16	Verano 2016-2017
Anticipación y extensión de plazo de entrega de LGI.	<p>Autorización de la entrega de LGI en viviendas de la submuestra 20 desde 4 días antes del inicio de la quincena, es decir, desde el miércoles 12 de abril.</p> <p>A quienes entregaron LGI con anticipación, se solicitó llamar a los informantes para recordar el inicio del registro de sus gastos diarios el día 16 de abril (día 1).</p>	09 - 13 - 20	Feriados consecutivos 2016 Verano 2016-2017 Censo 2017
Extensión de plazo de apertura.	<p>Realización de aperturas de viviendas hasta el séptimo día de la quincena de colaboración (como máximo).</p> <p>Apoyo a informantes en esfuerzos de recordación para registro de gastos efectuados en días anteriores de la quincena en que no dispusieran de las LGI.</p>	06 - 9 - 13 - 20	Fiestas Patrias 2016 Feriados consecutivos 2016 Verano 2016-2017 Censo 2017

Tabla 42: Tipos de medidas adoptadas en planes especiales

Ajuste	Descripción	Submuestras Operativa	Plan especial
Extensión de plazo para cuarta visita.	Ampliación de 3 a 6 días de plazo para realización de cuarta visita, a fin de cerrar proceso de colaboración con hogares ausentes en periodo regular.	05 - 12 - 19	Fiestas Patrias 2016 Verano 2016-2017 Censo 2017
Concentración de cuarta visita en fechas específicas.	Coordinación, a nivel de grupo de campo, de jornadas intensivas dedicadas exclusivamente a cuartas visitas, con el fin de asegurar que hogares completaran en un corto plazo su colaboración con el estudio. Definición de días específicos de dedicación exclusiva del grupo de campo a labores conjuntas de contacto inicial y apertura de viviendas pendientes de la submuestra siguiente.	19	Censo 2017
Extensión de plazo para despacho de submuestra.	Asignación de hasta 4 días adicionales de plazo para el despacho de encuestas aplicadas, a fin de disponer de mayor tiempo revisión final de los cuestionarios y recuperaciones de datos.	05 - 12 - 19	Fiestas Patrias 2016 Verano 2016-2017 Censo 2017

(1) Ante esta instrucción, se solicitaba a los encuestadores ser prudentes en el modo de preguntar, evitando generar desconfianza en los informantes.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Si bien los planes especiales involucraban la posibilidad de aplicación de estas medidas, cabe aclarar que ellas no fueron aplicadas a la totalidad de las viviendas pertenecientes a las submuestras citadas en la tabla 42, sino solo a aquellas que presentaron mayores dificultades de contacto y cooperación.

8.1.5 Herramientas de gestión y monitoreo del trabajo de campo

La gestión del trabajo de campo y el monitoreo de sus resultados se basaron en la disposición de información, en tiempo real, del avance de las distintas actividades comprendidas por las fases del trabajo de campo, característica fundamental para la acción oportuna frente a contingencias o resultados insatisfactorios. Para poder efectuar esta labor, la coordinación nacional se mantuvo en constante comunicación con los encargados de grupo, apoyándolos en el trabajo cotidiano de los grupos de campo. No obstante, existían tres herramientas principales para gestionar y diagnosticar el avance del proceso: la Hoja de Ruta, los indicadores de desempeño operativo y el sistema informático, los que serán descritos en los siguientes apartados.

8.1.5.1 Hoja de Ruta

La Hoja de Ruta¹⁴⁹ es un formulario que busca caracterizar el trabajo de campo de la encuesta, registrando las tareas más importantes realizadas por los encuestadores y encargados de grupo para producir información en terreno.

El objetivo general de la Hoja de Ruta era caracterizar el trabajo realizado en cada unidad muestral y hogar a encuestar¹⁵⁰, registrando información detallada sobre:

- a. Responsables de la vivienda u hogar (módulo Identificación del encuestador y encargado de grupo)
- b. Existencia de supervisión en la vivienda u hogar (módulo Visita de supervisión del hogar)
- c. Trabajo de identificación de unidades muestrales y contacto inicial (módulo Registro de contacto con el hogar)
- d. Entrevistas aplicadas y realización de seguimiento de las LGI (módulo Registro de aplicación de libretas)
- e. Asignación y retiro de LGI a integrantes del hogar (módulo Control de Libretas de Gastos Individuales)
- f. Organización de las distintas visitas a los hogares, a través del registro de citas concertadas (módulo Agenda de citas con el informante)
- g. Otras observaciones relevantes respecto al trabajo en la vivienda u hogar (módulo Observaciones).

Entre los usos de la Hoja de Ruta, el más relevante fue la caracterización del resultado final del trabajo de campo en cada unidad muestral y hogar incluido en el estudio. Ello constituye un insumo relevante, tanto para medir el error de no respuesta total como para realizar los ajustes a los factores de expansión de la encuesta descritos en el apartado 6.2.

La Hoja de Ruta era el único instrumento digitado diariamente en las distintas regiones del país, tarea que era responsabilidad de los encargados de grupo¹⁵¹. Aquello permitió contar con información simultánea del avance del trabajo de campo en los distintos territorios, facilitando las labores de monitoreo. Específicamente, permitió la construcción de indicadores de desempeño operativo de la encuesta, mediante la clasificación de los resultados de visita.

Otro de los usos relevantes asociados a la Hoja de Ruta es que la información de los indicadores de desempeño operativo construidos a partir de la información capturada en este instrumento permitirán realizar el cálculo del tamaño muestral con sobremuestreo para

149 Formulario es presentado en el anexo 5.

150 Si bien, en principio, se registra una Hoja de Ruta para cada unidad muestral de la encuesta, en aquellos casos en que se detectaban viviendas con más de un hogar residente, se solicitaba registrar la información de los hogares en Hojas de Ruta separadas.

151 Dado que la Región Metropolitana presentaba mayores cargas de trabajo, existían en ella dos digitadores que, entre otras funciones, apoyan las labores de ingreso de datos de la Hoja de Ruta.

la próxima realización de la encuesta, así como los indicadores obtenidos de la Hoja de Ruta de la VII EPF permitieron generar aquel cálculo para la presente versión del estudio.

8.1.5.2 Indicadores de desempeño operativo

Los indicadores de desempeño operativo se analizaron durante todo el proceso de trabajo de campo, permitiendo la evaluación y oportuno seguimiento del avance de las distintas etapas del proceso. Eran contruidos a partir de la información disponible en los módulos Registro de contacto con el hogar y Control de libretas de gastos individuales de la Hoja de Ruta.

El módulo Registro de contacto con el hogar contenía información¹⁵² de cada intento de contacto realizado en las unidades muestrales. Esos intentos eran caracterizados a través de 26 códigos, que representaban distintos resultados posibles en la unidad muestral. Estos códigos eran clasificados en siete grupos, según el tipo de resultado que representaban, desde la no elegibilidad de la unidad hasta la respuesta.

La clasificación de los códigos de visita según tipo de resultado se realizaba de acuerdo al estándar propuesto por la Asociación Americana de Investigación en Opinión Pública, en su versión disponible al momento de diseñar el formulario (Apor, 2015).

Dado que en la Hoja de Ruta se registraban los resultados de visita de todos los intentos de contacto, para la clasificación se seleccionaba el código de visita que reflejara el mayor nivel de interacción alcanzado en la unidad muestral al momento de la clasificación. Al finalizar las actividades de contacto inicial y el plazo para realizar la primera visita, este código era consignado en un campo denominado Código de Disposición Final, el cual representaba el estado final del proceso de contacto.

A partir de la clasificación de unidades se calculaban indicadores de desempeño operativo, estadísticos que permitían la evaluación del proceso de trabajo de campo en un momento determinado. Es por esta razón que se distinguió entre una clasificación inicial, realizada durante la quincena de trabajo de una submuestra operativa, y una clasificación preliminar, realizada al finalizar la fase de cierre de una submuestra¹⁵³.

La distinción entre la clasificación inicial y preliminar radicaba en la definición de unidades entrevistadas y el tipo de código considerado (códigos de resultado visita o código de disposición final). La clasificación inicial se realizó considerando el mejor resultado de visita y establecía como entrevistadas aquellas unidades en que se hubiese aplicado el RPH (código de visita 81).

A diferencia de lo anterior, en la clasificación preliminar se consideró el código de disposición final y se definió como entrevistadas las unidades que tuviese aplicación del RPH y que, además, contaran con al menos una LGI con registros de gastos diarios. Esta última información se obtenía del módulo Control de libretas de gastos individuales. Los grupos de clasificación de códigos de visita, según tipo de resultado de visita y momento de evaluación se presentan en la tabla 43.

152 Fecha, hora, resultado de visita, identificación del responsable de la visita y cargo del responsable de la visita.

153 Se establece que la clasificación final del resultado de las viviendas solo es posible tras finalizar el trabajo de campo, procesamiento y análisis de datos, una vez que se realiza la evaluación de la calidad de los datos mediante la grilla técnica de mínimos de calidad (descrita en el apartado 9.4.1.1).

Este modo de clasificación de códigos, según tipo de resultado y momento en que se realizaba la evaluación, resultó de utilidad para disponer de indicadores de desempeño operativo que permitieran diagnosticar detalladamente el avance de las labores de una submuestra en cada territorio y tomar decisiones oportunas para mejorar los resultados de la encuesta. Los indicadores de desempeño operativo empleados¹⁵⁴ se presentan en la tabla 44.

Tabla 43: Clasificación de códigos de Hoja de Ruta según tipo de resultado de visita y momento de clasificación

Categoría según pertenencia a población objetivo	Categoría según nivel de contacto	Clasificación	Sigla	Código de Hoja de Ruta	
				Clasificación Inicial	Clasificación Preliminar
				(durante trabajo de submuestra operativa)	(tras cierre de submuestra operativa)
Elegible	Contactada	Entrevistada	E	81	81 con al menos una LGI
		Entrevista interrumpida (<i>break-off</i>)	BO	-	81 sin LGI
		Rechazo	R	71-72-73-74-75-76-77	71-72-73-74-75-76-77
		Otros contactados sin respuesta	O	51 ⁽¹⁾ -61-62-63-64	61-62-63-64
	No contactada	No contacto	NC	41-42-43	41-42-43
Elegibilidad desconocida			ED	31-32-33	31-32-33
No elegible			NE	21-22-23-24-25	21-22-23-24-25

(1) Si bien el código 51 "Concertó cita con el informante" no es parte de los Códigos de Disposición Final, fue considerado como no contacto en la clasificación inicial para permitir el cálculo de los indicadores de desempeño operativo sobre el total de las unidades muestrales.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

La mayor parte de estos indicadores son utilizados por organismos estadísticos a nivel internacional, tales como US. Census Bureau (2013) y Eurostat (2014), entre otros, para evaluar la calidad de los procesos de recolección de datos. La excepción la constituyen los indicadores de Desactualización del Marco Muestral y de Pérdida de Apertura, que han sido incorporados por el estudio como indicadores complementarios.

8.1.5.3 Sistema informático

La gestión del proceso de trabajo de campo se realizaba a través de un sistema informático desarrollado para la encuesta, alojado en la red interna del Instituto. Los módulos del

154 Entre los códigos de visita presentes en la Hoja de Ruta, los Códigos de reconocimiento de manzana (11 y 12) eran excluidos de la construcción de indicadores, pues solo indicaban el cumplimiento de un procedimiento (reconocimiento de manzana). Adicionalmente, respecto al código 51 (concerta cita), si bien este indicaba contacto con la unidad, representaba una situación intermedia en el proceso, que necesariamente deben complementarse con una siguiente visita, por lo que no era aceptado como un código de disposición final.

Tabla 44: Indicadores de desempeño operativo empleados en la VIII EPF

Indicador	Aspecto evaluado	Fórmula de cálculo
Desactualización del marco muestra	Calidad de información del directorio de viviendas.	$DMM = \left(\frac{(NE+ED)}{(E+BO+R+O+NC+ED+NE)} \right) \times 100$
Tasa de contacto	Capacidad de adaptación a horarios de disponibilidad de informantes.	$TC = \left(\frac{(E+BO+R+O)}{(E+BO+R+O+NC+ED)} \right) \times 100$
Tasa de cooperación	Capacidad de persuasión.	$TCoop = \left(\frac{E}{(E+BO+R+O)} \right) \times 100$
Tasa de rechazo	Impacto de negativas e insuficiencia de información en la respuesta total.	$TRec = \left(\frac{(R+BO)}{(E+BO+R+O+NC+ED)} \right) \times 100$
Pérdida de aperturas ⁽¹⁾	Dificultades de mantención de la cooperación de los hogares.	$PA = \left(\frac{BO}{(E+BO)} \right) \times 100$
Tasa de respuesta	Eficacia global del trabajo de campo.	$TR = \left(\frac{E}{(E+BO+R+O+NC+ED)} \right) \times 100$

(1) A diferencia del resto de los indicadores, solo era posible medir la Pérdida de Aperturas una vez finalizada la submuestra operativa, debido a que consideraba como entrevistadas toda unidad con apertura (con RPH completo o parcial).

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF) en base a Aapor (2015).

sistema fueron diseñados en conjunto entre la coordinación nacional de trabajo de campo y el equipo de informática del estudio, con el apoyo y coordinación técnica del Departamento de Tecnologías de la Información y Comunicaciones del INE.

El sistema informático cumplió un rol fundamental en las labores de monitoreo realizadas por la coordinación nacional de trabajo de campo, facilitando el acceso a información actualizada sobre distintos ámbitos del mismo. Las funciones más relevantes asociadas al sistema informático fueron las siguientes:

- i. **Gestión muestral:** a partir del ingreso de las bases de datos de las viviendas seleccionadas, el sistema permitía la asignación de cargas de trabajo, habilitaba la creación de Hojas de Ruta para su digitación y posibilitaba impresión de los directorios de viviendas asignadas a los encuestadores.

Junto con lo anterior, el sistema contaba con un módulo que reportaba fecha, hora y responsables de cada vivienda de la muestra, a partir del cual se supervisaba el cumplimiento de los plazos establecidos para la asignación de cargas de trabajo a encuestadores.

Adicionalmente, el sistema permitía la corrección de información de las viviendas (direcciones, observaciones de localización, etc.) y la descarga de versiones digitales de la cartografía asociada a cada manzana seleccionada.

- ii. **Digitación de Hojas de Ruta:** el sistema permitía la digitación y visualización de las Hojas de Ruta por parte de los encargados de grupo. Dicha digitación incluía la asociación informática de los cuestionarios en papel a un hogar específico, mediante la digitación de los códigos de barra para cada libreta entrevistada en un hogar. Este aspecto era clave para el despacho y posterior recepción del material aplicado, pues permitía hacer un seguimiento detallado a todos los cuestionarios en papel. La digitación de las Hojas de Ruta permitía, además, habilitar la creación de Formularios de Correcciones para cada hogar entrevistado y la generación de reportes de indicadores de desempeño operativo.

El sistema informático posibilitó la descarga de las bases de datos de cada módulo de la Hoja de Ruta. Con aquella información la coordinación nacional retroalimentaba las labores de contacto y entrevista, a partir del análisis de días y horas de visitas a terreno.

- iii. **Registro de revisiones de cuestionarios:** el sistema alojaba el Formulario de Correcciones, instrumento empleado por encargados de grupo y supervisores para registrar el resultado de las revisiones del material aplicado. Específicamente, permitía indicar la presencia de errores, omisiones o potenciales inconsistencias de información en cada libreta, identificando el módulo, pregunta y número de línea en que se presentaba el problema. Junto con ello, permitía clasificar el tipo de error, así como también ingresar instrucciones específicas de corrección. A partir de estos datos, el sistema permitía generar reportes por cada hogar encuestado e imprimirlos en papel, posibilitando la entrega de instrucciones escritas de corrección y recuperación de datos.

El sistema generaba bases de datos descargables del Formulario de Correcciones, a partir de las cuales el equipo de coordinación nacional supervisaba el cumplimiento de las labores de revisión por parte de los encargados de grupo y analizaba la calidad de los datos producidos por los encuestadores, a partir del cálculo de promedios de errores por hogar y por libreta.

- iv. **Consultas y solicitudes de autorización de medidas especiales:** El sistema permitía orientar a los grupos de campo frente a dificultades detectadas durante el proceso. Las solicitudes podrían estar referidas a problemas de identificación de unidades muestrales de primera y segunda selección, requerimientos de actualización de información domiciliaria, modificación de plazos de visitas, modificaciones de número de visitas (solicitadas por el hogar) y modificación de submuestra operativa asignada.

Estas solicitudes eran analizadas caso a caso por el equipo de coordinación nacional de trabajo de campo, el cual sugería posibles estrategias de solución ante cada situación problemática, autorizando o denegando la solicitud.

La importancia de este módulo radica en que generaba un registro de los ajustes realizados y los motivos asociados a ellos, a fin de contar con mayor información respecto del contexto en que se ejecuta el trabajo de campo en cada territorio.

- v. **Reporte de indicadores de desempeño operativo:** el cálculo de indicadores de desempeño operativo fue incorporado en el sistema informático, elaborando un módulo que reportaba resultados iniciales y preliminares a nivel nacional, y por desagregaciones a nivel de área, región, encargado de grupo y encuestador. El módulo presentaba estos datos a nivel de submuestra operativa y acumulado, es decir, con el conjunto de submuestras operativas cerradas hasta la fecha de evaluación. Finalmente, el módulo posibilitaba la descarga de los reportes y la elaboración de gráficas de los indicadores.

8.2 Metodología de supervisión

El equipo de supervisión y capacitación de la VIII EPF tuvo como función ser el principal responsable de ejecutar los mecanismos de control de calidad y retroalimentación del proceso de aplicación de la encuesta, los cuales estuvieron orientados en la mejora continua de la fase de trabajo de campo del proceso de producción estadística. Fue uno de los 5

equipos que trabajaron en la EPF y su operación se coordinaba de manera centralizada desde las oficinas centrales del Departamento de Presupuestos Familiares¹⁵⁵.

Una característica importante del equipo de supervisión y capacitación, en su ámbito de control de calidad de la información recolectada, es que se constituyó como un equipo externo y con organización separada al equipo de trabajo de campo. Esta característica buscó producir una mayor independencia del equipo, la que se traducía en mayores incentivos para la identificación de errores o aspectos a mejorar en el trabajo de los encuestadores y los grupos de campo en general. Los grupos de campo, a su vez, realizaban simultáneamente controles de calidad internos, a cargo principalmente de los encargados de grupo, los cuales se centraban en la revisión permanente del material aplicado por los encuestadores (y descrito en los apartados anteriores). El trabajo del equipo de supervisión resulta, por tanto, complementario a dichos procesos.

Entre las tareas definidas para el equipo de supervisión y capacitación, se encontraban:

- i. Asegurar que la captura de información se llevase a cabo según la metodología, en forma oportuna por las personas responsables.
- ii. Identificar problemáticas periódicas y de excepción en la aplicación de la encuesta, buscando soluciones y homogeneizando criterios para su mejora continua.
- iii. Evaluar las estrategias de comunicación de los encuestadores con el hogar informante, detectando y corrigiendo las desviaciones de los objetivos de la encuesta.
- iv. Generar Indicadores relativos a la aplicación de la encuesta, que identificaran dificultades en módulos y preguntas.
- v. Identificar y proponer los temas a reforzar en la capacitación de los encuestadores.
- vi. Retroalimentar a los grupos de campo, especialmente a encuestadores, en los temas que se identificaran como necesarios durante el proceso de supervisión.
- vii. Coordinar y apoyar los procesos de capacitación necesarios para la inducción de nuevos integrantes del equipo de trabajo de campo, durante el proceso de aplicación de la encuesta.

8.2.1 Estructura organizativa del equipo de supervisión

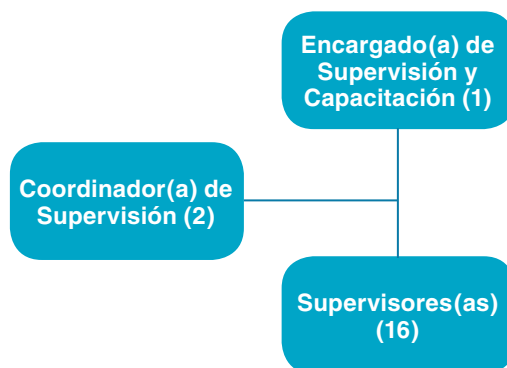
El equipo de supervisión y capacitación de la VIII EPF estuvo compuesto por 19 personas: 16 supervisores, encargados directamente del trabajo de control de calidad de la encuesta; dos coordinadores de supervisión, uno encargado de coordinar el trabajo en Región Metropolitana y otro en el área Resto de capitales regionales, y un encargado de supervisión, quien gestionaba los recursos, planificación y comunicación con el resto de los equipos de la encuesta. No obstante, además de su relación con la estructura interna del equipo de supervisión y capacitación, el supervisor se relacionaba directamente con

¹⁵⁵ El equipo de trabajo de campo, a diferencia del equipo de supervisión, contó con grupos de campo instalados en las oficinas regionales del INE, de todas las capitales regionales.

parte del equipo de trabajo de campo y mantenía comunicación con el resto de integrantes del mismo equipo.

Tal como se mencionó, el equipo de supervisión y capacitación tuvo como centro de operaciones las oficinas Departamento de Presupuestos Familiares, ubicadas en la Región Metropolitana, desde donde se desplazaba a las distintas capitales regionales del país, para realizar sus labores de supervisión. Su organización se observa en el siguiente diagrama.

Diagrama 8: Organigrama del equipo de supervisión y capacitación



Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

8.2.2 Capacitación

El proceso de capacitación de los supervisores contó principalmente con dos etapas: la inducción en temas generales de la encuesta y la inducción en aspectos específicos del trabajo de supervisión y capacitación.

- i. **Inducción en temas generales de la encuesta:** se basó esencialmente en la estructura, contenidos y actividades definidas para la capacitación de encuestadores de hogares, aunque con una metodología levemente distinta a la utilizada para capacitar a los encuestadores, en la cual eran los propios supervisores quienes estaban encargados de preparar los módulos de la capacitación, para ser expuestos a sus compañeros¹⁵⁶.

Las temáticas estaban organizadas inicialmente en torno a elementos de inducción institucional, seguidas de una descripción general de la encuesta, conceptos básicos de la EPF, descripción de las fases de preparación, aplicación y cierre de trabajo de campo, para luego abordar los módulos asociados a la correcta aplicación de las libretas. Junto a esto, y en base a la información expuesta, se realizó una serie de talleres prácticos, replicando además las evaluaciones realizadas en el proceso de capacitación de encuestadores, con fines de retroalimentación.

¹⁵⁶ La capacitación se realizó de esta forma por razones pedagógicas, ya que se buscaba que los supervisores se interiorizaran de los contenidos para lograr realizar la presentación a sus compañeros. De esta forma, se buscó controlar que los supervisores efectivamente estudiaran los contenidos de la encuesta.

El material base para esta inducción fue el Manual de Trabajo de Campo para Encuestadores, además de distintas presentaciones y ejercicios desarrollados por el departamento para la capacitación inicial de los encuestadores.

- ii. **Inducción en aspectos específicos del trabajo de supervisión y capacitación:** se centró en tres módulos generales: el diseño de supervisión, la capacitación continua y los procedimientos.

En el módulo de diseño de supervisión se trataron los fundamentos del trabajo de supervisión y capacitación, recomendaciones y estándares de trabajo con los que operaría el equipo y el modelo de supervisión planificado para la VIII EPF.

En el módulo de capacitación continua, se reforzó principalmente el rol de capacitación del equipo, a partir de los procesos de capacitación permanente que debía encabezar este durante el transcurso de la encuesta.

Finalmente, en el módulo de procedimientos, se presentó la estructura del equipo, los flujos de trabajo, los procedimientos y protocolos involucrados en el trabajo de supervisión, además de los principales temas administrativos.

Al igual que en el caso de la inducción en temas generales de la encuesta, en cada módulo se realizó una serie de talleres prácticos, complementados con instancias de ejercicio en terreno y de juego de roles, para reforzar mejor los elementos procedimentales del trabajo.

El material base utilizado para esta inducción fue el Manual de Trabajo de Campo para Supervisores, además de presentaciones y ejercicios desarrollados por la coordinación del equipo.

8.2.3 Proceso de Supervisión VIII EPF

Tomando como base la definición de Díaz de Rada (2005), se definió la supervisión como el proceso orientado a garantizar la corrección del trabajo de campo, a través de la identificación de errores que lo puedan invalidar total o parcialmente. Su objetivo fue el de asegurar que el encuestador realizara correctamente la aplicación de los instrumentos de la encuesta y así confirmar que la información recolectada fue fidedigna. Este proceso, además, recolectó datos estadísticos que dieran cuenta de la calidad del trabajo realizado por los encuestadores en terreno, para luego servir a la retroalimentación de la actividad de los encuestadores y grupos de campo.

El proceso de supervisión tuvo tres fases¹⁵⁷ para su desarrollo: planificación, supervisión y retroalimentación, tal como se detalla a continuación.

8.2.4 Fase de planificación

Se planteó un modelo de supervisión en que se asegurara, como mínimo, una visita de supervisión mensual a las regiones donde se concentraba una mayor proporción de la muestra, en específico, las regiones Metropolitana, de Valparaíso y Biobío, a las cuales se agregó luego la Región de Antofagasta, debido a complejidades propias identificadas tanto en versiones anteriores de la encuesta como en otros estudios.

¹⁵⁷ Al igual que lo puntualizado respecto a las fases de trabajo de campo, las fases que aquí se identifican asociadas al proceso de supervisión no deben confundirse con la mención a las fases que componen el modelo de producción estadística GSBPM.

Adicionalmente, se planteó, como mínimo, realizar una visita de supervisión cada dos meses al resto de las regiones del país, definiéndose estos parámetros a partir de las posibilidades de distribución del equipo y la extensión de las visitas de supervisión que, por razones técnico-operativas, se fijaron en una semana, acotándose a un grupo de campo.

Para dar cumplimiento a lo anterior, en el caso de la Región Metropolitana, se requirió una dupla de supervisores para cada visita de supervisión asignada¹⁵⁸, mientras que en el caso del resto de capitales regionales, cada visita de supervisión necesitó a un solo supervisor.

Respecto a la cobertura del trabajo de supervisión, tomando en cuenta las recomendaciones, estándares y experiencias de otros equipos del INE, se trabajó como meta de cobertura la supervisión de al menos un 20% de los hogares de la muestra, controlando además la cobertura por tipo de contacto, definiéndose la cobertura mínima de un 20% de los hogares entrevistados, además de un 20% sobre el universo de no respuesta.

Para la ejecución de estas visitas de supervisión, se definieron una serie de tareas y mecanismos de control de calidad. Estos se plasmaron en dos ámbitos de acción en los que operó la supervisión, con objetivos complementarios, a saber, supervisión de gabinete y supervisión en terreno, que mantienen los tipos de supervisión empleados en la VII EPF, y que se explicitan en los apartados 8.2.5.1 y 8.2.5.2, respectivamente.

A partir de los criterios revisados, la cantidad de visitas de supervisión quedó definida tal como se muestra en la tabla 45.

Tabla 45: Visitas de supervisión por región

Región	Total grupos de campo	Cantidad total de visitas
Arica y Parinacota	1	12
Tarapacá	1	14
Antofagasta	2	18
Atacama	1	13
Coquimbo	1	15
Valparaíso	4	52
Metropolitana	15	280
O'Higgins	1	11
Maule	1	9
Ñuble ⁽¹⁾	1	13
Biobío	3	40
La Araucanía	1	13
Los Ríos	1	10
Los Lagos	1	13
Aysén	1	10
Magallanes	1	9
Total	36	532

(1) Corresponde a equipo responsable de Chillán y Chillán Viejo, comunas que compondrán la futura Región de Ñuble.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

158 El trabajo en duplas, para los procesos de supervisión en la Región Metropolitana, comenzó a implementarse desde la submuestra operativa número 9 (noviembre de 2016), a partir de las diferencias de cargas de trabajo respecto al resto del país. En el caso de las duplas, uno de los supervisores se encargaba del trabajo de gabinete y otro del trabajo en terreno, mientras que, en el resto de capitales regionales, un mismo supervisor se encargaba de ambas funciones.

8.2.4.1 Notificación del proceso de supervisión

Los coordinadores de supervisión (o en su defecto, el encargado de supervisión) debían informar a los equipos de campo, vía correo electrónico, los detalles del trabajo de supervisión que se realizarían en cada grupo de campo. Dicha notificación se enviaba a más tardar un día hábil antes del comienzo de la supervisión en el grupo que correspondía. La información que debía contener dicha notificación era la siguiente:

- i. Submuestra a supervisar.
- ii. Fase de trabajo de campo respecto de la cual se realizaría la supervisión.
- iii. Fechas de inicio y término de la supervisión.
- iv. Solicitud de espacio físico para la realización de trabajo de gabinete.
- v. Nombre del supervisor o supervisores responsables del proceso.
- vi. Día y horario de presentación del o los supervisores.

Si bien se asumía, a modo de referencia, como submuestra a supervisar aquella que estaba en fase de aplicación, durante el trabajo de los supervisores con el grupo de campo, estos podían realizar, además, trabajos de supervisión sobre la submuestra anterior o sobre submuestras que se encontraran en ese momento en fase de preparación. Esto se definía a partir de lo que los supervisores juzgaran pertinente, según lo informado por el encargado de grupo o a partir de los hallazgos realizados por el mismo supervisor.

8.2.4.2 Presentación

Antes de iniciar su trabajo con el grupo de campo asignado, los supervisores debían presentarse ante las siguientes personas:

- i. En el caso del resto de capitales regionales, ante el director regional, o en su defecto al jefe de operaciones correspondiente.
- ii. Encargado regional correspondiente.
- iii. Encargado de grupo correspondiente.
- iv. Grupo de campo correspondiente.

La información que los supervisores debían entregar estaba relacionada con los períodos de supervisión, tipo de supervisión, metodología de la reunión de retroalimentación y la fecha estimada de entrega del informe de resultados de la supervisión.

8.2.4.3 Reunión de coordinación

Esta reunión se realizaba entre el encargado de grupo y el supervisor. En ella se daban a conocer los lineamientos de trabajo y el detalle del desarrollo del proceso, con sus aspectos operativos y técnicos a considerar, tanto en el trabajo de terreno como en gabinete.

Opcionalmente, se recomendaba que el supervisor solicitara los números telefónicos de contacto a los encuestadores y al encargado de grupo, en la medida que estos consideraran pertinente entregarlo, compartiendo a su vez el supervisor su número de contacto. Cabe aclarar que el uso de teléfonos personales en el proceso era absolutamente voluntario y, por tanto, no exigible en ninguna circunstancia por ninguna de las partes.

8.2.5 Fase de supervisión

A continuación se explican los dos tipos de supervisión aplicadas por la encuesta, supervisión de gabinete y en terreno. Respecto a la supervisión en terreno, se revisan las dos modalidades de supervisión realizada: directa e indirecta.

8.2.5.1 Supervisión de gabinete

Consistió en la revisión de los instrumentos luego de su aplicación. El principal objetivo era identificar errores tales como inconsistencias, aclaraciones u omisiones de información, entre otros, propiciando su resolución durante la aplicación y cierre de cada submuestra operativa. Se aplicó al 100% de los cuestionarios de la submuestra supervisada y su instrumento de registro fue el Formulario de Correcciones del Sistema Informático.

Esta labor de control de calidad era, en general, responsabilidad de los encargados de grupo. Sin embargo, durante los períodos en que se realizaba la visita de supervisión, el supervisor apoyaba esta revisión en aquellas libretas disponibles durante el proceso, con las mismas exigencias y atribuciones que el correspondiente encargado de grupo sobre el instrumento y las libretas, a fin de volver complementario el trabajo de ambos sobre esta modalidad.

Los instrumentos de apoyo para esta tarea fueron el Manual de Trabajo de Campo para Encuestadores y la Guía de Revisión para Encargados y Supervisores VIII EPF, documento que especificaba la categorización que debía aplicarse a los distintos errores identificados, así como el procedimiento de registro en el sistema informático.

8.2.5.1.1 Planificación de la supervisión de gabinete

El supervisor debía solicitar al encargado de grupo las libretas que se encontraran aplicadas. La entrega de dicho material debía efectuarse dentro del mismo día en que se iniciaba la visita de supervisión o, a más tardar, durante la mañana siguiente¹⁵⁹.

En caso de haber existido algún inconveniente para cumplir con la entrega de material por parte del encargado de grupo, este debía justificar por escrito la razón al encargado regional y al encargado de área correspondiente, estableciendo un nuevo horario. El supervisor, por su parte, debía dar aviso de la situación a su coordinador, mediante un correo electrónico.

El proceso de supervisión debía revisar la fecha de aplicación de las libretas, de acuerdo al calendario de referencia, respetando el contexto particular del trabajo de campo de cada encuesta y todas las excepciones indicadas por el encargado de grupo en la reunión de coordinación.

8.2.5.1.2 Recepción de material

Los supervisores, al momento de recibir los instrumentos por parte del encargado de grupo, debían registrar su recepción en la Planilla de Recepción de Material (disponible en el anexo 6) provista por la coordinación del equipo de supervisión y capacitación, firmando

¹⁵⁹ El plazo máximo de entrega de materiales para la supervisión de gabinete dependía de la hora de inicio de la visita de supervisión. Se fijaba este plazo para dar un tiempo adecuado al supervisor de revisar el material aplicado por los encuestadores.

conforme la recepción de los materiales. Al finalizar la revisión de gabinete, el supervisor debía verificar en conjunto con el encargado de grupo la entrega del material, firmando este último la devolución de los documentos. Este procedimiento permitía tener un control del flujo de materiales, evitando pérdidas de cuestionarios aplicados.

Los supervisores tenían como plazo un día para la devolución de las Hojas de Ruta y dos días para el resto de las libretas. No obstante, cuando se presentaba una visita del encuestador durante el proceso de supervisión de gabinete, el encuestador debía acudir al encargado de grupo para que este solicitara la entrega anticipada de dicho material, firmando al reverso de la Planilla de Recepción de Material, especificando los instrumentos retirados y las razones de dicho retiro.

8.2.5.1.3 Proceso de supervisión de gabinete

Una vez que se disponía del material correspondiente, los supervisores debían comenzar la supervisión de gabinete. Esta consistía en la revisión del total de libretas y Hojas de Ruta disponibles, aplicadas por los encuestadores, entre el día hábil anterior y el segundo día de la visita de supervisión (los dos primeros días del proceso de supervisión eran destinados, generalmente, a la revisión de gabinete), con el objetivo de identificar errores como inconsistencias, aclaraciones u omisiones de información, entre otros.

Dado que trabajo de supervisión de gabinete realizado por el supervisor era complementario al realizado por el encargado de grupo, durante la supervisión de gabinete debían priorizarse las libretas que no habían sido revisadas por el encargado de grupo con posterioridad a la visita de un encuestador.

El registro e identificación de los errores se realizaba en dos soportes: sobre las libretas aplicadas y en el Formulario de Correcciones del sistema informático. En las libretas, el supervisor debía marcar en los instrumentos la detección de algún error. En el Formulario de Correcciones del sistema informático, el supervisor debía identificar el hogar en revisión mediante los filtros del sistema. Una vez seleccionada esta información, debía registrar para cada error identificado la libreta, módulo y pregunta donde se identificó el error, el número de línea de RPH al que correspondía, el número de veces en que el mismo error había sido revisado, la descripción del error según el listado disponible en el formulario, observaciones que describieran de mejor manera el error, la fecha de la revisión y el resultado, en caso de tratarse de un error identificado en una revisión anterior.

Al igual que los encargados de grupo, los supervisores debían indicar a los encuestadores los aspectos que debían corregir en las libretas revisadas. Solo era posible realizar correcciones directamente en los cuestionarios cuando estos se encontraban en fase de cierre y en la medida que los manuales, el flujo de aplicación, la información de la Zona de consultas o la coherencia interna de los instrumentos entregaran evidencia suficiente para hacerlo. No obstante, aunque fuera corregido, todo error debía ser identificado igualmente en el Formulario de Correcciones, y la corrección debía realizarse con los mismos criterios descritos en el Manual de Trabajo de Campo para Encuestadores.

En el caso de encontrar errores en las unidades supervisadas, se debían ejecutar las acciones necesarias de acuerdo al grado o tipo de incidencia encontrada, entre las cuales se mencionan:

- i. Verificar el trabajo en terreno, mediante supervisión directa o indirecta.
- ii. Informar al encargado de grupo de errores encontrados.
- iii. Retroalimentar verbalmente al encuestador.

En caso de existir sospecha de falseamiento de datos o evidencia de alguna mala práctica por parte de algún encuestador, el supervisor tenía la obligación de intentar corroborar los datos con el mismo informante de la libreta correspondiente. Adicionalmente, debía supervisar de manera exhaustiva el material producido por el encuestador, incluyendo submuestras anteriores. Si la sospecha era confirmada, el supervisor debía informar inmediatamente la situación al encargado de supervisión y a los coordinadores de supervisión, para posteriormente comunicarlo al encargado de grupo y/o encargado regional. A su vez, el encargado de supervisión o, en su defecto, el coordinador correspondiente debía notificar de la situación al jefe de operaciones, encargado de área y encargada nacional de trabajo de campo¹⁶⁰.

8.2.5.2 Supervisión en terreno

A partir de lo revisado en la supervisión de gabinete, la supervisión en terreno permitió evaluar el cumplimiento de los procedimientos, la calidad y la veracidad de los datos a partir del trabajo realizado por el encuestador. Esta modalidad por su parte se realizó de dos formas: directa e indirectamente.

8.2.5.2.1 Supervisión directa

Consistía en el acompañamiento de un supervisor a un encuestador en la aplicación de los cuestionarios, con la intención de observar en detalle la interacción de este con el informante. El instrumento utilizado para estos efectos dependía de la fase en la que se realizaba la supervisión directa; durante la apertura se utilizaba el Formulario de Supervisión Directa en Fase de Apertura (disponible en el anexo 7) y, con posterioridad a la primera visita, era necesario utilizar el Formulario de Supervisión Directa en Fase de Aplicación (disponible en el anexo 8).

El Formulario de Supervisión Directa en fase de Apertura se componía de cuatro módulos: 1) Apertura, en que se observaba principalmente el protocolo de identificación de la vivienda, la presentación y la identificación del hogar; 2) Llenado de RPH, en que se observaba la aplicación de las preguntas del Registro de Personas en el Hogar y la explicación de ellas si era necesario; 3) Antecedentes generales de la entrevista, donde se observaba el cumplimiento de una serie de aspectos de protocolo durante la entrevista, y 4) Observaciones, donde debía registrarse cualquier elemento relevante detectado durante la supervisión, que no estuviera registrado en el instrumento, o bien que complementara alguna información especificada en él.

El Formulario de Supervisión Directa en Fase de Aplicación se componía de seis módulos: 1) Llenado de LGH, en que se observaba la aplicación de la Libreta de Gastos del Hogar y la explicación de preguntas si era necesario; 2) Llenado de LGR, en que se observaba la aplicación de la Libreta de Gastos del Recuerdo y la explicación de preguntas si era necesario; 3) Llenado de LI, donde se observaba la aplicación la Libreta de Ingresos y la explicación de preguntas si era necesario; 4) Llenado de LSF, en que se observaba la aplicación la Libreta de Gastos en Servicios Financieros y Seguros, la explicación de preguntas, si era necesario, y la solicitud de comprobantes por parte del encuestador; 5) Antecedentes generales de la entrevista, donde se observaba el cumplimiento de una

¹⁶⁰ Los casos en que se verificaba la existencia de falseamiento de datos en los instrumentos de recolección de información, se desvinculaba al encuestador, ya que se trata de una falta grave a la labor institucional de producción de datos.

serie de protocolos durante la entrevista, y 6) Observaciones, donde debía registrarse cualquier elemento relevante detectado durante la supervisión que no estuviera registrado en el instrumento o que complementara alguna información específica.

La supervisión directa debía realizarse como mínimo una vez a cada encuestador durante el proceso productivo, o ante la necesidad de analizar la aplicación de libretas en específico desde la información sobre errores en las mismas, por lo que se concentró principalmente durante las primeras submuestras y ante la incorporación de nuevos encuestadores. El supervisor, por tanto, disponía de la prerrogativa de incorporar supervisiones directas al proceso de supervisión, de acuerdo a las siguientes condiciones:

- i. Cumplir con la exigencia de realizar al menos una supervisión directa para cada encuestador, priorizando para ello las primeras visitas a terreno de las que se hubiese podido disponer.
- ii. Apoyar aquellos casos en que se identificaron encuestadores con altas tasas de rechazo, o con otros indicadores de desempeño operativo deficientes, identificados por el equipo de trabajo de campo, que hicieran recomendable indagar razones en terreno, priorizándolos para la realización de este tipo de supervisión, para posteriormente retroalimentarlos y contribuir a mejorar sus indicadores de apertura o mantención. Estos casos debían ser informados a los supervisores por la coordinación del equipo.
- iii. La solicitud expresa por parte del encargado de grupo, a partir de situaciones particulares que lo hicieran pertinente, las cuales debían argumentarse debidamente, mediante un correo electrónico dirigido al supervisor correspondiente, con copia a los coordinadores de supervisión, al encargado regional y encargado de área respectiva.

Para la aplicación de la supervisión directa, en cualquiera de sus dos versiones, la actitud del supervisor debía ser estrictamente de observación, evitando interferir en la dinámica de aplicación, aunque presenciara errores u omisiones por parte del encuestador. Cualquier error u omisión identificados debían ser mencionados por el supervisor una vez terminada la entrevista, en ausencia del informante, puesto que el contexto de realizar supervisión durante una aplicación, podía producir una lógica innecesariamente invasiva y, además, cualquier error u omisión podía ser corregida por el encuestador en una visita posterior, gracias a la recomendación del supervisor. Las únicas excepciones en que el supervisor podía suspender su rol de observador eran las siguientes:

- i. Si se observaba que se estaba produciendo un error en la identificación de la unidad muestral a encuestar.
- ii. Si el encuestador planteaba directamente alguna pregunta al supervisor respecto al protocolo o aplicación de instrumentos. En esta situación, el supervisor debía responder de manera precisa, a partir de la información disponible, o, en caso de desconocer la respuesta, plantear la posibilidad de ser respondida en la siguiente visita.
- iii. Si se observaba que, de no realizar alguna intervención entre encuestador e informante, era evidente que el hogar rechazaría la aplicación. Estas intervenciones debían siempre tener una lógica de mediación, para facilitar la apertura o continuidad del hogar en el estudio.

- iv. Si el encuestador alteraba de forma evidente las tareas programadas para la visita, adelantando u omitiendo labores establecidas por el esquema de visitas, sin que ello haya sido autorizado previamente.

El rol de observador culminaba una vez terminada la visita al hogar, es decir, finalizada la interacción con el informante. Era obligatorio que, una vez finalizada la visita de supervisión directa, el supervisor retroalimentara verbalmente al encuestador acerca de sus fortalezas y aspectos de mejora en el proceso de aplicación. Adicionalmente, el supervisor debía digitar en el sistema informático la totalidad de la información disponible en los instrumentos de supervisión aplicados.

En los casos de error en la identificación de unidades muestrales, o de alteraciones a las tareas definidas por el esquema de visitas, se debía reportar la situación por escrito, mediante correo electrónico al encargado de grupo, con copia al coordinador de supervisión, encargado regional y encargado de área.

8.2.5.2.2 Supervisión indirecta

Consistía en el contacto con los informantes sin la presencia del encuestador y sin aviso previo a él, con el objetivo de volver a entrevistar los aspectos principales de la encuesta aplicados hasta ese momento, además de consultar por algunos elementos del protocolo de aplicación. Esta supervisión incluía un procedimiento de conciliación de la información, en caso de encontrarse diferencias entre lo reportado por el informante y lo registrado previamente por el encuestador. El instrumento utilizado para esta modalidad es el Formulario de Supervisión Indirecta y, con las excepciones antes mencionadas (acompañamientos en primeras submuestras, a nuevos encuestadores y ante la necesidad de analizar aplicaciones a libretas específicas), debía constituir el grueso de las supervisiones de terreno planificadas.

Por su parte, la conciliación de información era resuelta mediante la aplicación de una Guía de Conciliación, desarrollada en conjunto con el equipo técnico y de trabajo de campo del departamento.

La supervisión indirecta se realizó mediante dos procedimientos: supervisión indirecta en fase de aplicación y supervisión indirecta en fase de cierre, utilizando ambas el Formulario de Supervisión Indirecta, disponible en el anexo 9. Adicionalmente, se definió una tercera modalidad de supervisión indirecta, sin procedimiento de conciliación, cuyo objetivo era corroborar información de la Hoja de Ruta en los casos en que fuese necesario chequear el correcto registro de las visitas y sus respectivos códigos, denominada Supervisión de Hoja de Ruta. Esta modalidad utilizaba el Formulario de Supervisión de Hoja de Ruta, disponible en el anexo 10.

Tanto en el caso de la supervisión indirecta en fase de aplicación como en fase de cierre, el supervisor debía visitar al hogar en un horario en que se maximizara la probabilidad de encontrar a los informantes, de acuerdo a la información provista por el encargado de grupo y a los horarios de visita registrados en la Hoja de Ruta. Además, este tipo de supervisión debía realizarse estrictamente en ausencia del encuestador y sin aviso previo a este.

Los criterios para la elección de un informante idóneo para la aplicación de la supervisión eran idénticos a los de la aplicación de las libretas, definidos en el Manual de Trabajo de Campo para Encuestadores, no siendo necesario que el contacto se realizara exactamente con la misma persona que respondió cada instrumento, pero sí con alguien que se encontrara en condiciones de chequear las respuestas. El proceso de corroborar las

respuestas se realizaba en gabinete, a través del vaciado de la información desde las libretas a los formularios de supervisión¹⁶¹.

En caso de existir sospecha de falseamiento de datos, o evidencia de alguna mala práctica por parte de algún encuestador, el supervisor tenía la obligación de intentar corroborar la información con el mismo informante de la libreta correspondiente. Adicionalmente, debía supervisar de manera exhaustiva el material producido por dicho encuestador, incluyendo, de ser posible, submuestras anteriores. Si la sospecha era confirmada, el supervisor debía informar inmediatamente la situación al encargado de supervisión y a los coordinadores de supervisión, para posteriormente comunicarlo al encargado de grupo y/o encargado regional. El encargado de supervisión o, en su defecto, el coordinador correspondiente, debía notificar la situación al jefe de operaciones, encargado de área y encargada nacional de trabajo de campo.

Una vez realizada la supervisión, el supervisor debía digitar en el sistema informático la totalidad de la información disponible en los instrumentos aplicados.

A continuación, se detallan las características generales de las modalidades de supervisión indirecta antes señaladas.

Supervisión indirecta en fase de aplicación

Esta modalidad debía aplicarse para la supervisión de submuestras que se encontraran aún en fase de aplicación. En esta se aplicaba la base del Formulario de Supervisión Indirecta (a saber, los módulos de Apertura y RPH) más un máximo de dos módulos adicionales (entre LGI, RPH, LGR, LI y LSF), los cuales debían definirse a partir de aspectos identificados en la supervisión de gabinete.

Supervisión indirecta en fase de cierre

Esta modalidad debía aplicarse para la supervisión de submuestras que se encontraran en fase de cierre. En esta se aplicaba íntegramente el Formulario de Supervisión Indirecta. Podía realizarse sobre unidades que se incorporaban a la fase de cierre durante la visita de supervisión, o bien por unidades que en instancias posteriores de validación de los datos (revisión, digitación o procesamiento) se identificaban como pertinentes de supervisar por vía indirecta.

Supervisión de Hoja de Ruta

Esta modalidad de supervisión consistía en verificar en terreno el correcto registro de los módulos de Registro de contacto con el hogar (RC) y Agenda de citas con el informante (AC). Se aplicaba cuando existían solicitudes de supervisión focalizada por parte del encargado de grupo y, principalmente, en situaciones en que el supervisor, a través del trabajo de supervisión previo, consideraba necesario verificar los datos de Hoja de Ruta, debido a identificación de registros erróneos, sospecha de malas prácticas, características de elegibilidad dudosa, no elegibilidad, o una tasa de rechazo más altas de lo normal en algún encuestador. Esta supervisión consistía en el llenado en gabinete de los módulos RC y AC, de acuerdo a lo registrado por el encargado de grupo y supervisor, para posteriormente verificar con los informantes los datos registrados en ambos módulos.

¹⁶¹ En ningún caso el supervisor debía sacar las libretas de la oficina para realizar trabajo de supervisión indirecta ni para ninguna otra labor.

8.2.5.3 Complemento de supervisión de gabinete

En aquellos hogares en que se realizaba supervisión de terreno era necesario que, una vez concluida esta, el supervisor registrara en el Formulario de Correcciones cualquier nuevo requerimiento de corrección identificado durante la visita, utilizando los mismos criterios de la supervisión de gabinete, considerando además aquellos casos en que el supervisor, por encontrarse con una mejora de información, sugiriera el reemplazo de alguna respuesta en alguna de las libretas.

Solo una vez indicados los errores detectados en terreno al encargado de grupo, el supervisor estaba en condiciones de entregar tanto el material, como el correspondiente Reporte de Correcciones al encargado de grupo, quien debía solicitar al encuestador corregirlos durante la siguiente visita.

8.2.6 Fase de Retroalimentación

Al finalizar la visita de supervisión, el supervisor debía reunirse con el encargado de grupo y hacerle entrega de una Minuta de Reunión de Retroalimentación, con la descripción y análisis de las características generales del trabajo de supervisión, fortalezas detectadas, aspectos de mejora y recomendaciones al grupo de campo. Con estos antecedentes, el encargado de grupo debía retroalimentar posteriormente a sus encuestadores. Esta minuta debía enviarse de forma posterior al encargado de grupo, con copia al encargado de área, encargado regional, encargada nacional de trabajo de campo, coordinador de supervisión correspondiente y al encargado de supervisión. Adicionalmente, el supervisor debía entregar en esta reunión el Formulario de Registro de Folios¹⁶² Supervisados (disponible en el anexo 11), que contenía un resumen de todos los hogares supervisados, con la modalidad empleada en cada libreta, para que esta información fuera registrada a su vez por el encargado de grupo en las Hojas de Ruta correspondientes.

Posterior a esto, debía realizarse una reunión entre el supervisor, el encargado de grupo y los encuestadores, donde se exponían las observaciones generales del proceso de supervisión, rescatando lo positivo del equipo y los aspectos que se debían mejorar. Una vez finalizada la reunión de retroalimentación, se realizaba la despedida protocolar con equipo de trabajo de campo, y el personal directivo y/u operativo de la dirección regional en el caso de encontrarse fuera de la Región Metropolitana.

8.2.7 Capacitación continua

Como parte de las funciones del equipo de supervisión y capacitación, y en el contexto de incorporar en los procesos nociones como la mejora continua y prevenir complicaciones relacionadas con la estacionalidad, se trabajó con una modalidad adicional de capacitación, denominada capacitación continua, la cual complementaba la ronda de capacitación inicial y las retroalimentaciones *in situ*. Este tipo de capacitación consistió básicamente en identificar, ya sea de manera preventiva o *a posteriori*, aspectos de la aplicación que pudiesen haber afectado a la calidad de los datos, para revertirlos mediante procesos de capacitación.

¹⁶² Folio refiere a una denominación operativa para los instrumentos de un hogar en específico, caracterizado por poseer un indicador único. Si bien es un término operativo, se menciona en este caso por ser parte del nombre del formulario, pero su uso práctico es equivalente al de hogar.

La capacitación continua durante la VIII EPF se organizó a partir de dos rondas y, si bien su aplicación fue responsabilidad del equipo de supervisión y capacitación del departamento, sus contenidos fueron definidos de común acuerdo entre el equipo de trabajo de campo, el equipo técnico y el equipo de supervisión y capacitación.

8.2.7.1 Definición de temas para las rondas de capacitación continua

La identificación de temas para las rondas de capacitación continua se basó principalmente en la información recolectada por los distintos equipos, mediante un trabajo de contrastes cruzados encabezado por el equipo técnico. Este trabajo consistió en la conciliación de tres fuentes de información, en específico, el Formulario de Correcciones, la Zona de consultas y el procesamiento de las submuestras digitadas.

Estas fuentes permitieron definir temas que se identificaran como convenientes de reforzar, a lo que se sumó un criterio estacional, referido a que, ante la inminencia de un momento particular del año que introdujera alguna dificultad particular a la aplicación, se definió una instancia de capacitación para resolver estas situaciones *a priori*.

A continuación, se revisan las fuentes de análisis *a posteriori* para la definición de temas.

- i. **Análisis del Formulario de Correcciones:** el Formulario de Correcciones permitía hacer seguimiento de los problemas detectados por encargados de grupo y supervisores a los distintos instrumentos de la encuesta. En este, se reportaban aquellas preguntas en las que era necesario mejorar o corregir el llenado de información, utilizando códigos de error predefinidos. De esta manera, la base de datos de este formulario entregó información respecto a las temáticas, módulos y preguntas que concentraban la mayor cantidad de errores, así como respecto al tipo de errores más frecuente (por ejemplo, respuestas en blanco o letra ilegible, entre otros).
- ii. **Identificación de temáticas:** con el propósito de identificar temáticas que necesitaran ser reforzadas en los equipos de trabajo de campo, para cada libreta se individualizaron aquellos módulos que concentraban el mayor porcentaje de errores respecto del total de errores en la libreta.
- iii. **Análisis de los tipos de error:** tras la identificación de temáticas, se analizaron cuáles fueron los tipos de errores más frecuentes por cada módulo. Esta identificación permitía preparar de mejor manera los materiales de capacitación, dando mayor énfasis a aquellos aspectos que necesitaban ser reforzados por los equipos.
- iv. **Análisis de la Zona de consultas:** se realizó un análisis de las preguntas ingresadas en la plataforma Zona de consultas, codificando las preguntas según módulos y según subtemas, priorizando las temáticas con mayor relevancia técnica a capacitar, identificadas en el análisis del Formulario de Correcciones.
- v. **Análisis de las submuestras digitadas (validación y depuración):** al igual que para el Formulario de Correcciones, tras la validación y depuración de una submuestra, se identificaron para cada libreta los tres módulos que concentraban mayor cantidad de casos inválidos. Para esto se utilizaron los archivos de revisión generados por el equipo técnico para la recuperación de datos. Dado que en estos archivos también se encontraban codificados los errores, además de permitir identificar temáticas, su análisis permitió identificar los errores más comunes.

8.2.7.2 Primera ronda de capacitación continua

A partir de la información recolectada de estas distintas fuentes, se concluyó que los principales temas que debían tratarse en una primera ronda de capacitación eran el módulo Ingresos de la propiedad y ganancia por tenencia de instrumentos financieros (IP) de la Libreta de Ingresos, y los módulos Cuenta vista y/o cuenta RUT (CV), Tarjetas de crédito bancarias (TCB) y Seguros (SE), de la Libreta de Servicios Financieros y Seguros.

La modalidad de capacitación definida fue a través de presentaciones expositivas, planeadas para sesiones de duración de máximo dos horas, que combinaban definiciones, ejemplos y ejercicios prácticos de llenado a partir de casos predefinidos.

Dado que esta capacitación debía ser abordada a nivel nacional, esta ronda se inició el martes 6 de septiembre y se extendió hasta el martes 26 de septiembre de 2016. Se emplearon distintas estrategias de capacitación entre la Región Metropolitana y el resto de las capitales regionales. En el caso de los grupos de campo de la Región Metropolitana, se realizó una capacitación expositiva de manera presencial, a razón de dos grupos por sesión, mientras que, en el caso del resto de las capitales regionales, se llevó a cabo una capacitación expositiva mediante videoconferencias.

8.2.7.3 Segunda ronda de capacitación continua

Un aspecto relevante en una nueva ronda de capacitaciones, además de los temas emergentes desde la información recolectada, era prevenir los posibles efectos del fin del año académico, por lo que este se definió como tema *a priori* para la nueva ronda. Sumado a la información recolectada de las distintas fuentes en los términos planteados anteriormente, se concluyó que los principales temas que debían tratarse en la segunda ronda de capacitación eran los módulos Educación (ED) y Condición de actividad económica (AE), del Registro de Personas en el Hogar; el módulo Gastos relativos a educación formal (EF), de la Libreta de Gastos del Hogar; Ingresos por otras pensiones y transferencias recibidas (TR), de la Libreta de Ingresos, y la Tabla de viajes realizados, de la Libreta de Gastos del Recuerdo.

Al igual que en la primera ronda, la modalidad de capacitación definida fue de presentaciones expositivas, planeada para una sesión de máximo dos horas, que combinaban definiciones, ejemplos y ejercicios prácticos de llenado a partir de casos predefinidos.

Esta ronda de capacitación se inició el martes 15 de noviembre y se extendió hasta el viernes 2 de diciembre de 2016. Para su realización, se emplearon nuevamente distintas estrategias de capacitación entre la Región Metropolitana y el resto de las capitales regionales (capacitación expositiva de manera presencial a razón de dos grupos por sesión en Región Metropolitana y capacitación expositiva mediante videoconferencias en el resto de capitales regionales).

9. METODOLOGÍA DE PROCESAMIENTO DE DATOS

A continuación, se desarrollan los subprocesos incorporados en la fase de procesamiento de datos (brevemente descrita en el apartado 7.2.5). Se describen los subprocesos de digitación y codificación, para luego describir la metodología de preparación de las bases de datos, la que contempla los subprocesos de validación; recuperación; depuración; imputación, ajuste temporal y creación de nuevas variables; cálculo de ponderaciones; consolidación de las bases de datos y cálculo de agregaciones.

Si bien el procesamiento informático no se encuentra descrito en la estructura del estándar GSBPM, se incluye como parte del procesamiento de los datos, en su rol de soporte transversal al proceso de producción de la encuesta.

9.1 Procesamiento informático

Dentro del proceso de diseño de la encuesta, se buscó que el sistema informático de la VIII EPF, además de ser una herramienta para el ingreso de información recopilada por el equipo de trabajo de campo y de participar en el proceso de producción de la base de datos, tuviera como objetivo dar apoyo a los distintos equipos de la encuesta en diversas tareas informativas y de gestión. Para ello fue necesario reunir los requerimientos de los diferentes usuarios, con la finalidad de realizar el diseño y posterior construcción de una plataforma informática en ambiente web, que diera respuesta a las necesidades específicas de los equipos involucrados en el estudio.

En esta sección se describen los módulos pertenecientes al sistema informático de gestión de la VIII EPF y se presenta brevemente el proceso de entrega de bases de datos al equipo de análisis de la encuesta.

9.1.1 Descripción general del sistema informático de la VIII EPF

El sistema desarrollado buscó ser una herramienta de apoyo a los distintos equipos de la encuesta, para lo cual se construyeron diferentes módulos. Estos módulos permitieron administrar y controlar los procesos de la encuesta, diferenciados por el perfil de cargo involucrado en la encuesta.

El anexo 12 presenta los módulos del sistema con los respectivos perfiles de acceso.

A continuación, se describen brevemente los módulos que conforman el sistema informático de la VIII EPF:

Módulo Autenticación y menú: el sistema contó con un control de acceso, basado en el uso de nombre de usuario y contraseña. Cada usuario accedía a los distintos módulos según su perfil de cargo.

Módulo Directorio: permitía ingresar un directorio individual o de forma masiva, además de realizar búsqueda de códigos de hogares según los filtros seleccionados, información que era posible exportar (archivo formato .xls). Junto a lo anterior, el módulo otorgaba la opción de revisar, habilitar y deshabilitar la información de los directorios, seleccionando previamente el mes, región, provincia y comuna.

Módulo Monitoreo del sistema: permitía el control de los usuarios del sistema. Mostraba información respecto a los accesos de los usuarios, errores del sistema ocurridos durante su uso, además de permitir asignar y definir perfiles de sistema y consultar información respecto a estas.

Módulo Administración: permitía administrar usuarios y perfiles, además de controlar los niveles en los que se desplegaba la información (administración de menú). Permitía administrar las temáticas del módulo de zona de consulta, además de entregar y controlar información relacionada con el manejo de los códigos de barra y guías de despacho utilizadas en el control del envío y recepción de material.

Módulo Informaciones: permitía ingresar y mostrar noticias asociadas a la encuesta. Además, el usuario podía realizar descarga de bases internas (formulario de correcciones, consultas de usuario, solicitudes y guías de despacho) en formato Excel.

Módulo Zona de consultas: correspondía a un foro nacional donde se ingresaban consultas para tener respuestas estandarizadas y entregadas en un corto período de tiempo. El módulo permitía ingresar y visualizar consultas realizadas por los usuarios y su respuesta. Por otro lado, los responsables de ciertas temáticas podían dar respuesta a las preguntas ingresadas de forma pública o privada. Además, el sistema guardaba un registro de las preguntas realizadas, según temática.

Módulo Trabajo de campo: este módulo permitía administrar las tareas asociadas al trabajo de campo de la encuesta y se componía de las siguientes secciones:

Sección Operación: permitía realizar solicitudes relacionadas con la información de viviendas a encuestar, siendo posible ingresar y responder solicitudes, ver solicitudes realizadas, entre otras acciones. Además, en esta sección se digitaba la Hoja de Ruta y se ingresaba las revisiones de cuestionarios aplicados en el formulario de correcciones. Por último, en esta sección se realizaban las operaciones asociadas a las guías de despacho (emisión y recepción de estas).

La sección operación permitía controlar la asignación de las cargas de trabajo de los encuestadores, permitiendo realizar reasignaciones entre los perfiles de encargados de grupo y encuestadores, así como cambios en las asignaciones de cargas de trabajo.

Finalmente, la sección contenía una biblioteca de trabajo de campo, desde donde podían descargarse los documentos asociados a una submuestra (planos de ubicación y croquis).

Sección Asignación de folios: permitía asignar las viviendas seleccionadas a los distintos perfiles de cargo del equipo de trabajo de campo, además de visualizar reportes de asignación de trabajo de campo y de asignaciones de vivienda realizadas a un encuestador.

Sección Desempeño operativo: permitía visualizar una serie de reportes de trabajo de campo segmentados por región, zona, grupo y encuestador, siendo posible seleccionar distintos tipos de reporte y submuestra, además de visualizar resultados acumulados. Permitía, además, generar gráficos, de acuerdo a indicadores y niveles de desagregación seleccionados.

Sección Supervisión: esta sección fue diseñada para el uso del equipo de supervisión y permitía la digitación en el sistema informático de los formularios asociados a los distintos tipos de supervisión. La sección daba la posibilidad de guardar, eliminar datos e imprimir los formularios.

Además de poder ingresar los formularios mencionados, también era posible cargar y descargar reportes de supervisión, ver directorios, descargar bases de datos asociadas a los cuatro tipos de supervisión y visualizar los reportes de indicadores de supervisión.

Módulo Digitación: este módulo permitió gestionar la digitación de las encuestas de la VIII EPF, posibilitando la asignación de cargas de digitación y revisión de digitación, además del ingreso de datos de las libretas utilizadas para recopilar información en la VIII EPF.

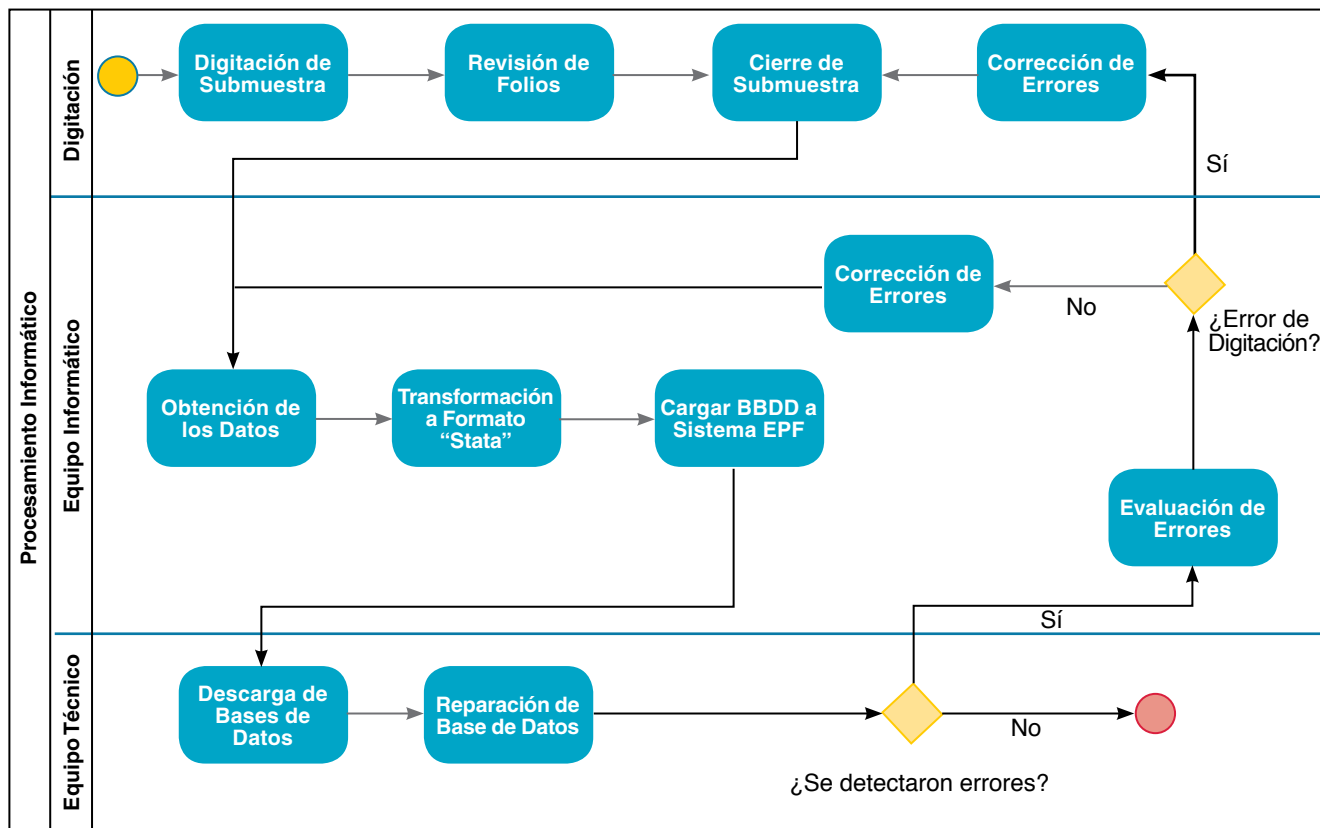
Módulo Capacitación: contenía una biblioteca virtual, que mostraba y permitía la descarga y subida de diversos archivos de capacitación tales como videos informativos, capacitaciones, libretas, manuales e instructivos. También era posible modificar nombres y eliminar archivos y carpetas de capacitación.

9.1.2 Generación de bases de datos para procesamiento

El sistema informático fue utilizado para dar soporte al proceso de producción estadística de la encuesta. A continuación se describe el proceso informático que acompañó al proceso de digitación de datos y generación de las bases de datos que se entregaban al equipo de análisis, para continuar con el procesamiento de la información (etapas contenidas en la codificación y metodología de preparación de bases de datos explicadas en los apartados 9.3 y 9.4, respectivamente).

El proceso informático de las bases se representa en el siguiente diagrama:

Diagrama 9: Proceso informático para la generación de bases de datos para procesamiento



Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

A continuación se describen brevemente los roles de los participantes del proceso:

Equipo de digitación: el equipo de digitación introducía al sistema informático los datos recolectados por el equipo de trabajo de campo, según lo descrito en el apartado 9.1. Una vez digitado el material asociado a una submuestra operativa, se procedía a realizar el cierre de esta, para la obtención de los datos en los formatos acordados entre el equipo de informática y el equipo técnico de la encuesta.

Equipo informático: luego del cierre de la digitación de cada submuestra operativa, el equipo informático procedía a procesar la información, almacenándola en el formato solicitado por el equipo técnico y en archivos compatibles con el *software* Stata. Posteriormente, los archivos se ponían a disposición del equipo técnico a través de los módulos correspondientes.

Equipo técnico: este equipo descargaba los archivos y realizaba distintos procesos para preparar la información para su posterior procesamiento. Si se encontraba algún error en la preparación de la información, se reportaba al equipo informático, el cual analizaba la raíz del error y determinaba si debía solucionarlo directamente, solicitar la corrección a digitación o trabajar en conjunto con este equipo para resolverlo. Una vez solucionados los errores, se volvía a procesar la información y se subía nuevamente al sistema *web*, para la descarga por parte del equipo técnico. Cuando la preparación de la información no presentaba errores, se daba por terminado el proceso.

9.2 Metodología de digitación

La digitación fue un proceso incorporado al procesamiento de los datos, que consistió en ingresar al sistema informático de la VIII EPF toda la información recopilada por trabajo de campo¹⁶³ de forma centralizada y de la manera más fidedigna posible.

Para llevar a cabo este proceso fue necesario constituir el equipo de digitación, cuya labor consistió en realizar el ingreso de los datos recopilados por el estudio, procurando cumplir los estándares de calidad y resguardo de la información que se establecieron como departamento.

Al definir la estructura organizacional del equipo de digitación, a diferencia de la VII EPF, se incorporó el perfil de cargo de revisor de digitación. De esta manera, el equipo estuvo compuesto por 20 digitadores, 2 revisores de digitación y una coordinadora, que funcionaron de forma centralizada, junto al resto de los equipos, en las oficinas del departamento, ubicadas en la Región Metropolitana. Esto permitió, por una parte, aprovechar economías de escala respecto a la distribución del material a digitar y, por otra, contar con una comunicación directa y de continua retroalimentación con los equipos de trabajo de campo y equipo técnico.

¹⁶³ Se digitó la información capturada en todos los formularios salvo en la Hoja de Ruta, que fue digitada por el equipo de trabajo de campo.

Diagrama 10: Organigrama del equipo de digitación VIII EPF

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

9.2.1 Proceso de capacitación

La capacitación del equipo de digitación consistió en una inducción inicial de dos semanas de duración, en la que se abarcaron los siguientes ítems: 1) características generales de la encuesta con sus objetivos y usos, 2) proceso de producción de la encuesta e 3) inducción y funcionamiento de los módulos de digitación del sistema informático de la VIII EPF. Para realizar esta capacitación se utilizaron presentaciones, material para digitar empleado para ejercicios prácticos y un manual del equipo de digitación como apoyo a la inducción.

Respecto a la capacitación asociada al ingreso de nuevo personal durante el proceso de ejecución de la encuesta (es decir, en casos en que se producía rotación), la capacitación inicial tuvo un carácter más reducido y combinó la entrega de material de apoyo y material a digitar. Hubo una fase de adaptación al equipo, que implicó una menor carga de trabajo en las primeras semanas de trabajo de las personas que se integraban al equipo, además de un proceso de acompañamiento e integración por parte de un digitador experimentado, quien guiaba el trabajo de estas y resolvía posibles dudas.

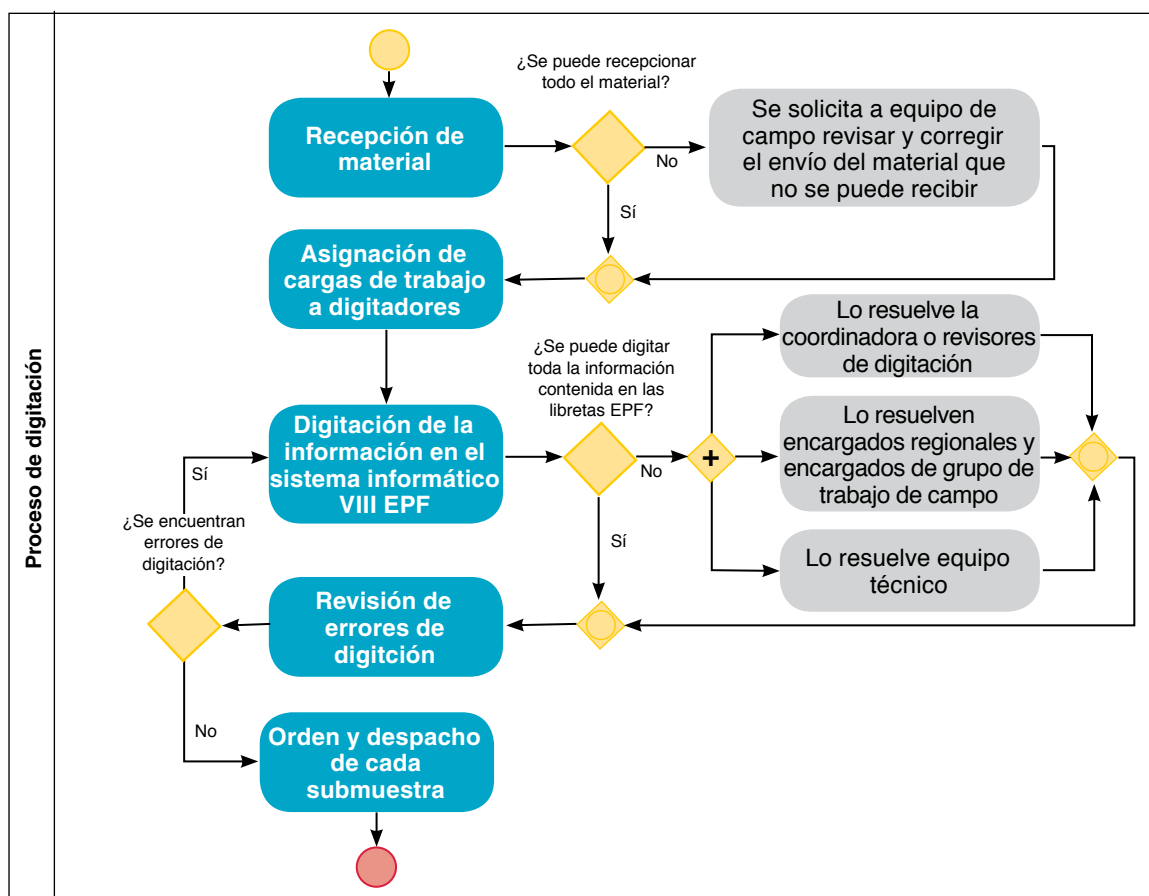
Respecto al proceso de capacitación continua, cuando se detectaban errores en la aplicación de instrucciones o criterios de digitación, se realizaba una retroalimentación (individual o grupal dependiendo de la temática) en donde se entregaban o reforzaban lineamientos para mejorar el proceso de digitación de encuestas. Además, cada vez que se terminaba de digitar una submuestra operativa, se realizaba una reunión de equipo para reforzar los criterios que se debían utilizar, aclarar dudas y tratar temas administrativos.

Descripción general del proceso de digitación

El proceso de digitación se desarrolló de forma continua, utilizando como insumo el material generado por el equipo de trabajo de campo, el cual era despachado al equipo de digitación tras el cierre de cada submuestra operativa. Este proceso se dividió en 5 etapas: 1) recepción de material, 2) asignación de cargas de trabajo, 3) digitación de la información en el sistema informático, 4) revisión de errores de digitación y 5) orden y despacho del material.

El diagrama que se presenta a continuación grafica el flujo asociado al proceso de digitación.

Diagrama 11: Flujograma del proceso de digitación VIII EPF



Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

A continuación, se realiza una revisión de las distintas etapas relacionadas con la digitación de encuestas.

9.2.2.1 Recepción de material

Esta etapa consistió en revisar el material recibido desde el equipo de trabajo de campo. La recepción se realizaba utilizando lectores digitales de códigos de barra incorporados en cada libreta, con el fin de asegurar que el contenido de cada guía de despacho¹⁶⁴ fuera correcto y coincidiera con lo indicado en el sistema de gestión.

Si se detectaba errores en la recepción de información, se retroalimentaba al equipo de trabajo de campo, a través de los encargados regionales o encargados de grupo, con el fin de corregir los errores y proceder a rectificar la recepción de material.

La etapa de recepción de material estuvo a cargo de la coordinadora del equipo, junto a los revisores de digitación.

¹⁶⁴ La guía de despacho contenía información de todas las libretas asociadas a los hogares. Dicha información se ordenaba a nivel de manzana (unidades primarias de muestreo).

9.2.2.2 Asignación de cargas de trabajo

A medida que se realizaba la recepción del material, este se asignaba a los digitadores, considerando las siguientes variables: cantidad de libretas totales a digitar, cantidad de Libretas de Gastos Individuales (LGI) a digitar y comuna (esta se rotaba entre submuestras).

En promedio, la carga de trabajo de cada digitador fue de 224 libretas totales y 51.154 campos digitados aproximados por submuestra operativa (35 hogares y 62 LGI). Este era el parámetro aproximado para asignar las cargas de trabajo por digitador.

Es importante mencionar que cada guía de despacho se asignaba a un solo digitador, salvo en aquellos casos en que la guía contenía muchos hogares y debía ser asignada a dos digitadores. Todas las libretas asociadas a un hogar estaban contenidas en una sola guía de despacho. Una vez que los digitadores avanzaban en su labor, se revisaban las cargas de trabajo del equipo, con el fin de realizar eventuales ajustes y redistribuciones de cargas.

Las cargas de trabajo eran asignadas de forma paulatina, buscando mantener un equilibrio entre los digitadores, a la vez que procurando que ningún digitador se quedara sin carga de trabajo. Los ajustes en las cargas de trabajo buscaban optimizar los tiempos de digitación, resguardando que los cuestionarios terminaran de digitarse antes de finalizar el plazo contemplado para el ingreso de datos de cada submuestra operativa.

La etapa de asignación de cargas de trabajo estuvo a cargo de la coordinadora del equipo, junto a los revisores de digitación.

9.2.2.3 Digitación de información en el sistema informático

Cada digitador tuvo acceso al sistema informático de la VIII EPF, donde debía digitar su carga de trabajo. Cada perfil del equipo podía acceder a distintos módulos del sistema, según sus tareas a desarrollar. De este modo, los digitadores tuvieron acceso a los módulos de digitación de libretas RPH, LGI, LGR, LI y LSF para aquellos hogares que les eran asignados. Adicionalmente, tenían acceso al reporte “campos digitados” y al módulo de consulta de usuarios.

Cada digitador debía ingresar en el sistema informático de la VIII EPF toda la información contenida de las libretas RPH, LGI, LGR, LGH, LI y LSF de la forma más fidedigna posible. Para ello se les instruyó, considerando las indicaciones específicas respecto de cómo ingresar la información en el sistema.

Durante el proceso de digitación fueron detectados algunos problemas asociados al registro de información. Dichos problemas fueron canalizados a través de la jefatura del equipo, con el fin de solucionarlos con el equipo correspondiente.

Los errores detectados se asociaron a las siguientes fuentes:

- i. **Errores de tipo informático:** estos errores estaban asociados a caídas de los servidores, eventuales interrupciones del servicio de internet o al funcionamiento de ciertas validaciones del sistema de digitación.
- ii. **Errores de trabajo de campo:** este error se asociaba a deficiencias o errores generados en la fase de recolección de información. Este tipo de errores se asociaba, por ejemplo, a la falta de traspaso de boletas a las tablas de gasto de la LGI, errores de recolección o bien a errores en el llenado de información de la HR, que el sistema vinculaba automáticamente al sistema de digitación de la encuesta.

Respecto al orden de digitación, siempre se comenzaba por el ingreso de datos del RPH, ya que dicha libreta se vincula con el resto de los instrumentos de recolección, lo que permitía precargar saltos y ciertas respuestas del resto de las libretas, facilitando así la digitación.

9.2.2.4 Revisión de errores de digitación

En cada submuestra operativa se realizó un proceso de revisión de la información digitada, considerando una revisión superior al 10% de las libretas digitadas. Para realizar esta revisión, se seleccionaba cada hogar con sus respectivas libretas, comparando los registros ingresados en el sistema con aquellos contenidos en los formularios físicos.

Los errores se tipificaron en cuatro tipos, descritos en la tabla 46. Aquellos digitadores que presentaban un alto nivel de errores en la digitación, se les devolvía el material para que fuese corregido. Asimismo, se focalizaba la revisión de cada submuestra en el material de aquellos digitadores cuyo trabajo había presentado la mayor cantidad de errores en las revisiones anteriores.

Tabla 46: Tipos de errores de digitación.

Tipificación de errores	Descripción del error
Información faltante (preguntas abiertas)	Refiere a la falta de digitación parcial o total de información (campos de texto).
Inconsistencia de registro	Refiere a inconsistencias entre la información digitada y aquella contenida en los formularios (campos numéricos o de texto).
Campo vacío	Refiere al uso incorrecto de códigos asociados a campos obligatorios sin información.
Información faltante en LGI	Refiere a la falta de digitación de las líneas de gasto de la LGI

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Si bien la digitación de la Hoja de Ruta estuvo a cargo del equipo de trabajo de campo, el equipo de digitación incluyó su revisión, retroalimentando a los primeros respecto a los errores identificados.

La revisión de errores fue realizada por los revisores de digitación con apoyo de la coordinadora, sumando, en algunos casos, a aquellos digitadores más capacitados o con mejor desempeño (quienes tenían menos errores y digitaban más rápido). La información generada en esta etapa se utilizaba como fuente de información para retroalimentar al equipo de digitación.

9.2.2.5 Orden y despacho del material

Después de digitar y revisar cada submuestra, se procedía a realizar la etapa de orden y despacho del material. En esta etapa se verificaba que todas las libretas estuviesen efectivamente digitadas (para ello se utilizó un reporte de estado entregado por el sistema informático). Posteriormente, se efectuaba la extracción de las bases de datos del sistema, la cual era coordinada por los equipos informático y técnico del departamento, para que se prosiguiera con el resto de los subprocesos asociados a la fase de análisis.

En esta etapa se enviaba el material a almacenaje, a disposición del equipo técnico, ya que eventualmente se podía volver a este en el subproceso de depuración. Esta tarea fue realizada por los revisores de digitación.

9.2.3 Herramientas de gestión de digitación

A continuación, se describen los módulos más utilizados del sistema informático de la VIII EPF en el proceso de digitación.

- i. **Folios asignados para digitación:** listado de todos los hogares de la submuestra seleccionada con sus respectivas libretas. Este módulo permitía consultar respecto a los hogares asignados para digitar, identificando el digitador responsable y aquellos hogares que debían ser asignados.
- ii. **Estado de libretas:** este reporte entregaba un listado con todas las libretas de cada hogar y el estado de digitación¹⁶⁵ en que se encontraban. A partir de la revisión de este módulo se controlaba que los digitadores fuesen digitando su material según los plazos estipulados. Al finalizar cada submuestra, se verificaba que todas las libretas estuviesen digitadas.
- iii. **Campos digitados:** mostraba las libretas asignadas y la cantidad de campos digitados¹⁶⁶ por digitador, en cada submuestra. Este reporte sirvió para revisar el avance de cada digitador, contrastando la cantidad de hogares y libretas asignadas con la cantidad de registros digitados. A partir de esta información se realizaba la revisión de las cargas de trabajo y eventuales reasignaciones de cargas entre digitadores.
- iv. **Folios asignados para revisión:** reporte que presentaba información respecto a los hogares asignados para revisión, identificando responsable y el tiempo que demoró su revisión (indicaba el intervalo de tiempo en que se asigna y finaliza la revisión de las libretas).
- v. **Revisiones:** módulo que entregaba información respecto a los errores identificados por los revisores.

9.3 Metodología de codificación

En el marco de la VIII EPF, la codificación fue un subproceso que se realizó de forma posterior al trabajo de campo y digitación de los datos. Este consistió en asignar códigos, a partir de clasificadores establecidos, a todos los bienes, servicios, establecimientos¹⁶⁷, establecimientos educacionales y ocupaciones registrados en la encuesta.

El propósito de este apartado es describir el procedimiento realizado por el equipo de codificación, los mecanismos implementados para resguardar la calidad de este subproceso y los clasificadores utilizados por la encuesta (CIUO88/CIUO08.CL, CCIF, Clasificador de establecimientos, Rol Base de Datos (RBD), Clasificador de establecimientos de educación superior).

¹⁶⁵ Se trabajó con los siguientes 4 estados: sin digitar, en digitación, terminada y en revisión.

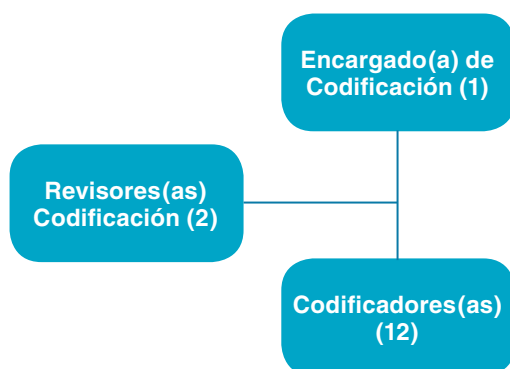
¹⁶⁶ Campos digitados: se cuantificaba la cantidad de celdas de respuestas digitadas y no la cantidad de caracteres digitados.

¹⁶⁷ La clasificación de establecimientos es un insumo para el proceso de levantamiento de precios de la nueva canasta del IPC.

9.3.1 Estructura organizacional del equipo de codificación

El equipo de codificación estuvo compuesto por doce codificadores, dos revisores de codificación y un encargado de codificación. Este equipo estuvo ubicado en su totalidad en la Región Metropolitana.

Diagrama 12: Organigrama del equipo de codificación VIII EPF



Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

9.3.2 Proceso de capacitación

La capacitación del equipo de codificación se dividió en las siguientes etapas:

Proceso de inducción: proceso de capacitación realizado a la llegada de nuevos codificadores, en el que se les explicaba las funciones, tareas y responsabilidades del codificador, así como también los conceptos y terminología de los clasificadores.

Proceso de reforzamiento: proceso de capacitación continua realizado durante el transcurso de la codificación, en el cual las temáticas sustanciales y complejas sobre los manuales fueron reforzadas para su mejor comprensión, poniendo énfasis en los criterios internos de codificación.

La capacitación y el reforzamiento fueron desarrollados a nivel interno en el equipo de codificación. El reforzamiento se implementó asociado a la necesidad de robustecer el conocimiento de temáticas complejas a codificar, a partir de las interrogantes que emanan durante la codificación de las libretas, manejo de criterios y uso de los clasificadores.

9.3.3 Descripción general del proceso de codificación

9.3.3.1 Planificación del proceso de codificación

El subproceso de codificación se organizó en tres etapas fundamentales: 1) codificación, 2) revisión cruzada entre pares y 3) revisión y consolidación de bases codificadas. A continuación, se describen brevemente las etapas:

Codificación: esta etapa consistió en asignar códigos estandarizados a la información generada en la etapa de recolección, utilizando los clasificadores definidos por la encuesta. Estuvo constituida, además, por dos tareas: ejecutar un protocolo de codificación

automatizada¹⁶⁸ y realizar una codificación manual. Esta etapa duraba aproximadamente cinco días por submuestra operativa y variaba según la calidad y cantidad de información contenida en la base de datos asignada.

Revisión cruzada entre pares: consistió en una revisión, entre codificadores, de los códigos asignados a sus pares. Duraba aproximadamente tres días.

Revisión y consolidación de las bases de datos codificadas: esta etapa se subdividía en dos tareas: revisión realizada por revisores, en la que se comparaban los códigos asignados en la fase de codificación versus el código registrado en la revisión cruzada, además de algunas revisiones generales, que permitían mejorar la precisión de la codificación. Esta tarea duraba aproximadamente cuatro días.

Posteriormente, la encargada de codificación efectuaba una revisión sobre aquellas discrepancias de códigos que no pudieron ser resueltas por los revisores, así como de algunos casos especiales que podían variar según la contingencia o eventuales ajustes a los criterios de clasificación utilizados.

Las cargas de trabajo eran asignadas a cada codificador de forma equitativa, según número de registros a codificar, y su distribución se realizaba con posterioridad a la codificación automatizada. En promedio, a cada codificador se le asignaron aproximadamente 4.700 registros¹⁶⁹ por submuestra operativa.

9.3.3.2 Calidad y desempeño de codificación

La calidad en la codificación fue uno de los ejes fundamentales del equipo y por ello se utilizaron herramientas que buscaban recopilar información para poder identificar errores, controlar y mejorar la calidad de los resultados. Por una parte, se identificaron errores externos, provenientes del equipo de trabajo de campo; por ejemplo, información deficiente para realizar una buena codificación. En estos casos dichas situaciones se sistematizaban y retroalimentaban a los equipos de campo, con el fin de mejorar la calidad de los registros entregados por dicho equipo.

Por otra parte, se realizó una identificación de errores y posibilidades de mejora internos al equipo de codificación, para lo cual se recurrió a constantes retroalimentaciones, que pretendían mejorar la calidad de la codificación aplicada.

Con el fin de mejorar la gestión interna del equipo, se introdujeron indicadores de desempeño, se sistematizaron criterios de codificación y, por último, se realizaron recodificaciones en bases de datos de los bienes, servicios, establecimientos, establecimientos educacionales y ocupaciones.

Finalmente, se sistematizaron y tipificaron las glosas problemáticas que dificultaban la codificación de los datos. Este registro buscó generar información que permitiera describir

168 La codificación automatizada fue un proceso de asignación de códigos a través de procesamientos informáticos estandarizados en base a información presente en los cuestionarios. Dicho proceso se aplicó a la base donde se codificaban bienes, servicios y establecimientos (base CCIF). Para los casos de las bases de establecimientos educacionales (base RBD) y de ocupaciones (base CIUO) los registros se codificaron de forma manual.

169 En la base CCIF un registro podía contener información de bienes o servicios a codificar, además del establecimiento en el cual fue adquirido. En muchas ocasiones, el establecimiento era codificado de forma automatizada, por lo que se debía codificar el registro del bien o servicio, o bien se daba la combinación contraria.

las formas en que los hogares reportan las glosas y así poder implementar mejoras para la próxima EPF.

9.3.3.3 Etapas del subproceso de codificación

A continuación se describen las distintas etapas que conformaban el subproceso de codificación.

9.3.3.3.1 Recepción y revisión de bases de datos provenientes del equipo de análisis

El objetivo de esta etapa fue proveer a los codificadores el material necesario para realizar la codificación de información. En ella la encargada de codificación recibía y revisaba la base de datos proveniente del equipo de análisis, verificando la cantidad de registros y que las variables entregadas fueran las necesarias para realizar el proceso de codificación. Cuando una base de datos no cumplía con los estándares requeridos, era devuelta al equipo de análisis para su modificación.

Una vez obtenidas y revisadas las bases de datos, se aplicaba la codificación automatizada, para luego extraer aquellos registros que no habían sido codificados. Estos últimos eran ordenados, distribuidos y asignados a las cargas de trabajo de cada codificador.

9.3.3.3.2 Codificación

El objetivo de esta etapa fue realizar la codificación de los registros asignados en cada base de datos. En primer lugar, se codificaban todos los registros de ocupaciones, establecimientos educacionales, establecimientos, bienes y servicios consignados en la encuesta, utilizando para ello los clasificadores CIUO, CCIF, Clasificador de Establecimientos INE y Clasificador de Establecimientos Educacionales.

En caso de encontrar registros complejos que dificultaran la codificación, el codificador debía recurrir a la encargada de codificación y/o revisores de codificación, para realizar las consultas pertinentes. Así también, para los bienes y servicios registrados de manera deficiente o incompleta, se buscaba información adicional en internet ¹⁷⁰.

Una vez codificados la totalidad de los registros asignados y resueltas las dudas sobre los registros complejos que eventualmente hubieran surgido, el codificador entregaba las bases de datos a los revisores de codificación, quienes eran los encargados de verificar la coincidencia entre la carga entregada (al inicio del proceso) y la recibida en esta instancia.

9.3.3.3.3 Resolución de dudas de codificación

El objetivo de esta etapa, que se realizaba en paralelo al proceso de codificación, era resolver dudas de los codificadores en cuanto al uso e implementación de los clasificadores. Cuando se identificaban errores reiterativos en la descripción de los registros, que imposibilitaban una codificación adecuada, los revisores o la encargada de codificación tenían la tarea de elaborar informes de errores que se enviaban al equipo de trabajo de campo, como una forma de retroalimentación.

¹⁷⁰ El recurso de la búsqueda en internet se utilizó para resguardar la correcta identificación del producto, servicio o establecimiento, y así reducir los casos en que se utilizaran códigos residuales debido a marcas o de establecimientos poco conocidos. Los hallazgos recurrentes se introducían en los instrumentos de apoyo a la codificación.

9.3.3.3.4 Proceso de revisión cruzada entre pares (poscodificación)

El objetivo de esta etapa era disminuir la cantidad de errores en el proceso de codificación. Para ello se intercambiaban las bases de datos codificadas entre pares y se procedía a realizar una revisión del trabajo realizado por otro compañero. Cuando existían dudas, los codificadores debían consultar a la encargada de codificación o a los revisores. La revisión cruzada se efectuaba sobre la totalidad de registros codificados.

9.3.3.3.5 Consolidación de bases y envío a equipo de análisis

El objetivo de esta etapa era enviar el material codificado al equipo de análisis. Una vez finalizados los procesos de codificación y revisión cruzada, la encargada de codificación, junto a los revisores debía consolidar la base de datos que era entregada al equipo de análisis.

Cuando existían dudas sobre algún registro y no se podía resolver con la información disponible, se enviaba a revisar la libreta física, para verificar si en ella existía alguna información adicional que permitiera codificar. Solo en casos excepcionales¹⁷¹, según importancia del registro, se enviaba a recuperar información con el equipo de trabajo de campo¹⁷².

Finalmente, se verificaba que todos los registros se encontraran codificados, que viniera la misma cantidad de registros entregada inicialmente por parte del equipo de análisis, que los códigos cumplieran con el largo y la forma correspondiente, etc. Una vez finalizada la etapa de revisión, se enviaban las bases consolidadas al equipo de análisis.

9.3.3.3.6 Recodificación por base ex post

El objetivo de esta etapa fue depurar las bases de datos, para disminuir la cantidad de errores no muestrales producidos en la fase de codificación.

Luego de entregadas las bases al equipo de análisis y una vez que este había avanzado en el proceso de análisis de datos, se producía una instancia en la que se solicitaba la revisión de un subconjunto acotado de registros, debido a la detección de errores, inconsistencias u omisiones. También se realizaron recodificaciones de algunos productos de forma semestral o bien, por necesidades de los equipos de precios de la institución, a cargo del IPC (Departamentos de Estadísticas y Estudios de Precios).

La encargada de codificación tenía la tarea de recibir estas bases de datos y distribuir la carga de recodificación (tanto propia como para los revisores de codificación). Posteriormente, se procedía a recodificar las bases, corrigiendo los errores detectados. La revisión final y la consolidación de las bases estaba a cargo de la encargada de codificación, quien consolidaba las bases recodificadas y las enviaba al equipo de análisis.

171 Los casos excepcionales eran evaluados por la encargada de codificación o bien por la encargada del equipo técnico, siendo ellas quienes definían si se debía solicitar la recuperación al equipo de trabajo de campo (TDC). El contacto con TDC se realizaba de forma indirecta, ya que era el equipo de análisis el encargado de enviar los registros a consulta.

172 Las necesidades tanto de revisiones de libreta como de recuperaciones con trabajo de campo se podían detectar en los distintos procesos del equipo de codificación (codificación, revisión cruzada y consolidación de base de datos). No obstante, para el envío de esta información, se consolidaban los datos y era enviados por el encargado de codificación.

9.3.4 Clasificadores utilizados

9.3.4.1 Clasificación Internacional Uniforme de Ocupaciones (CIUO-08)

El clasificador CIUO es un sistema de clasificación de ocupaciones concebido por la Organización Internacional del Trabajo (OIT, 1991). Específicamente, es un instrumento para establecer una clasificación nacional y facilitar comparaciones internacionales de las ocupaciones habituales de la población económicamente activa de un país, permitiendo también presentar los datos de una manera clara y sencilla.

El clasificador CIUO para ambas versiones tiene una estructura jerárquica de cuatro niveles (tabla 47), los cuales se establecen según similitud de tareas y especialización de las competencias requeridas.

Tabla 47: Niveles CIUO88 y CIUO-08.CL

Categorías	Desagregación	Elementos por categoría	Elementos por categoría CIUO-08.CL
		CIUO-88	
Gran grupo	1 Dígito	10	10
Subgrupo principal	2 Dígitos	28	42
Subgrupo	3 Dígitos	116	126
Grupo primario	4 Dígitos	390	138

Fuente: Instituto Nacional de Estadísticas (INE), CIUO-08.CL

Debido a las características de la VIII EPF, la variable ocupación se clasificó hasta nivel de subgrupo principal (dos dígitos). Para efectos de publicación, los datos son entregados tanto a nivel de gran grupo como de subgrupo principal.

9.3.4.1.1 Adaptación del clasificador CIUO

La codificación de ocupaciones durante la VIII EPF se realizó en base al clasificador CIUO-88 (OIT, 1991). Sin embargo, durante ese período se trabajó como institución en la actualización del mismo, tomando como insumo la versión internacional CIUO-08. Para desarrollar este trabajo, se conformó una mesa técnica redactora, compuesta por los departamentos de Infraestructura Económica (Sección Nomenclatura), Estudios Laborales, Estadísticas de Hogares y Presupuestos Familiares de la institución. Además, contó con las revisiones y aportes de una mesa institucional. Esta mesa tomó como decisión que la actualización de la CIUO-08 sería una adaptación¹⁷³ a la realidad nacional, que en términos generales mantiene la comparabilidad internacional a dos dígitos¹⁷⁴, siendo los principales cambios incorporados a nivel de tres y cuatro dígitos, los que no son utilizados por la encuesta.

173 Las opciones eran adoptar, adaptar o utilizar una clasificación nacional que no se basara en CIUO.

174 Se plantea que en términos generales mantiene comparabilidad a dos dígitos, pero existen algunas excepciones a saber:

Los carabineros fueron trasladados desde los grandes grupos 1 y 5 al gran grupo 0 junto con las Fuerzas Armadas por una necesidad de anonimización de los datos.

Los paramédicos y los técnicos en educación de párvulos fueron trasladados desde el gran grupo 2 al gran grupo 3, considerando que en la realidad nacional las tareas desarrolladas por estas ocupaciones se encuentran vinculadas al nivel de competencias requerido por este último gran grupo.

Las ocupaciones de los profesionales de la medicina tradicional y alternativa fueron trasladadas desde el gran grupo 2 al gran grupo 3, debido a la complejidad de realizar la distinción de estas ocupaciones en los niveles profesionales y técnicos, tal como lo proponía CIUO-08.

Considerando lo anterior, como departamento, y en conjunto con la mesa técnica redactora, se decidió aplicar la versión adaptada CIUO-08.CL a las ocupaciones obtenidas por la VIII EPF. El traspaso de lo codificado con CIUO-88 a CIUO-08.CL se realizó con posterioridad al cierre del trabajo del equipo de codificación. En la tabla 48 se agrega la estructura de la clasificación CIUO-08.CL (mientras que en el anexo 13 se presenta la versión de la CIUO-88 utilizada por la encuesta).

Tabla 48: Clasificación CIUO-08.CL según Gran grupo y Subgrupo principal

Gran grupo	Subgrupo principal	Glosa gran grupo y subgrupos principales
1		Directores, gerentes y administradores
	11	Directores ejecutivos, personal directivo de la administración pública y miembros del poder ejecutivo y de los cuerpos legislativos.
	12	Directores y gerentes de servicios comerciales y administrativos.
	13	Directores, gerentes y administradores de producción y operaciones.
	14	Directores, gerentes y administradores de hoteles, restaurantes, comercios y otros servicios.
2		Profesionales, científicos e intelectuales
	21	Profesionales de las ciencias y de la ingeniería.
	22	Profesionales de la salud.
	23	Profesionales de la enseñanza.
	24	Profesionales de negocios y administración.
	25	Profesionales de tecnología de la información y las comunicaciones.
	26	Profesionales en derecho, ciencias sociales y culturales.
3		Técnicos y profesionales de nivel medio
	31	Técnicos de las ciencias y la ingeniería.
	32	Técnicos de la salud.
	33	Técnicos en operaciones financieras y administrativas.
	34	Técnicos de servicios jurídicos, sociales y culturales.
	35	Técnicos de la tecnología de la información y las comunicaciones.
	36	Técnicos en educación de párvulos.
4		Personal de apoyo administrativo
	41	Oficinistas.
	42	Empleados en trato directo con el público.
	43	Empleados contables y encargados del registro de materiales.
	44	Otro personal de apoyo administrativo.
5		Trabajadores de los servicios y vendedores de comercios y mercados
	51	Trabajadores de los servicios a las personas.
	52	Vendedores.
	53	Trabajadores de los cuidados personales.
	54	Personal de los servicios de protección y seguridad (excluye policías).
6		Agricultores y trabajadores calificados agropecuarios, forestales y pesqueros
	61	Agricultores y trabajadores calificados de explotaciones agropecuarias con destino al mercado.
	62	Trabajadores forestales calificados, pescadores y cazadores.
	63	Trabajadores agropecuarios, pescadores, cazadores y recolectores de subsistencia.

Tabla 48: Clasificación CIUO-08.CL según Gran grupo y Subgrupo principal

Gran grupo	Subgrupo principal	Glosa gran grupo y subgrupos principales
7		Artesanos y operarios de oficios
	71	Operarios de la construcción (excluyendo electricistas).
	72	Operarios de la metalurgia y operarios de máquinas herramientas; mecánicos de vehículos, maquinarias, aviones y bicicletas.
	73	Artesanos y operarios de las artes gráficas.
	74	Trabajadores especializados en electricidad y electrónica.
	75	Operarios de procesamiento de alimentos, de la confección, ebanistas y otros oficios.
8		Operadores de instalaciones y máquinas; ensambladores
	81	Operadores de instalaciones fijas y máquinas.
	82	Ensambladores.
	83	Conductores de vehículos y operadores de equipos pesados y móviles.
9		Ocupaciones elementales
	91	Limpiadores y auxiliares de aseo.
	92	Obreros agropecuarios, pesqueros y forestales.
	93	Obreros de la minería, la construcción, la industria manufacturera y el transporte.
	94	Cocineros de comidas rápidas y ayudantes de cocina.
	95	Trabajadores ambulantes de servicios y vendedores ambulantes (excluyendo comida de consumo inmediato).
	96	Recolectores de desechos y otras ocupaciones elementales.
0		Otros grupos no identificados

Fuente: Instituto Nacional de Estadísticas (INE), CIUO-08.CL.

9.3.4.2 Clasificación del Consumo Individual por Finalidades (CCIF)

La CCIF es un sistema de clasificación funcional del Sistema de Cuentas Nacionales (SCN), que es utilizado a nivel internacional y que se encuentra a cargo de Naciones Unidas (ONU, 2001). La CCIF busca asignar códigos estandarizados a los bienes y servicios de consumo final de los hogares, instituciones sin fines de lucro que sirven a los hogares (ISFLSH) y el Gobierno. Esta clasificación es utilizada principalmente por las EPF, el Sistema de Cuentas Nacionales e IPC.

Esta clasificación consta de catorce divisiones; las primeras doce corresponden a la clasificación de gasto de consumo final de los hogares, mientras que las restantes, división 13 y división 14, refieren al gasto de consumo final de las instituciones sin fines de lucro que sirven a los hogares y del Gobierno general, respectivamente.

Para efectos de la VIII EPF se utilizaron solo las divisiones relativas al consumo final de los hogares, es decir, de la división 1 a la división 12 (ver tabla 49). Durante el proceso de codificación, se trabajó de manera temporal con una división residual, denominada división 15¹⁷⁵: gastos no desglosados de diferentes divisiones y otros bienes y servicios N.C.P.

175 Esta división temporal no es publicada ni puesta a disposición de la ciudadanía, ya que los gastos registrados temporalmente en dicha división son redistribuidos en el resto de los registros según lo especificado en el apartado 9.4.6.1.

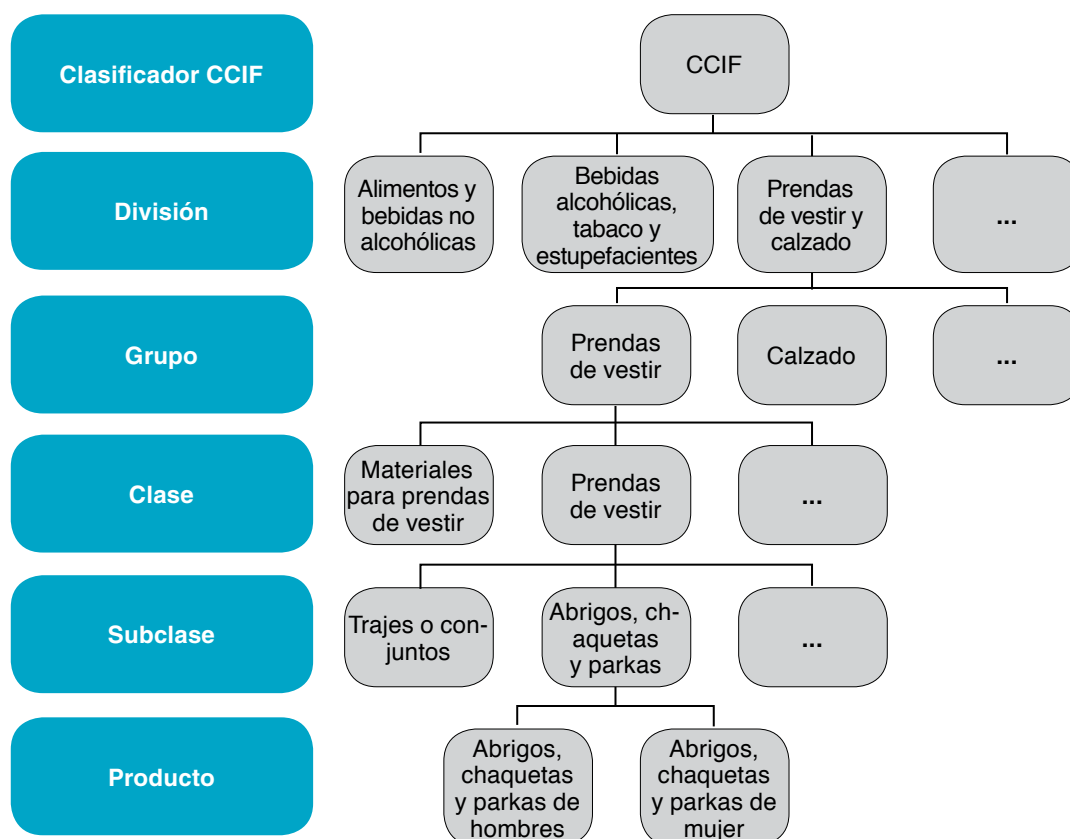
Tabla 49: Divisiones CCIF-VIII EPF

Divisiones	Glosa de división
División 1	Alimentos y bebidas no alcohólicas.
División 2	Bebidas alcohólicas, tabaco y estupefacientes.
División 3	Prendas de vestir y calzado.
División 4	Alojamiento, agua, electricidad, gas y otros combustibles.
División 5	Muebles, artículos para el hogar y para la conservación del hogar.
División 6	Salud.
División 7	Transporte.
División 8	Comunicaciones.
División 9	Recreación y cultura.
División 10	Educación.
División 11	Restaurantes y hoteles.
División 12	Bienes y servicios diversos.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

El mayor nivel de agregación de la CCIF son las divisiones (12), las que se forman de grupos. Los grupos están formados, a su vez, por clases y estas se componen por subclases. Las subclases son el resultado de la agregación de los productos. Los niveles de mayor agregación (división, grupo y clase) permiten realizar comparaciones internacionales, en cambio, los niveles de subclase y producto responden a la adaptación realizada por cada país. A continuación, se presenta un ejemplo del funcionamiento esquemático del clasificador:

Diagrama 13: Niveles del clasificador CCIF



Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

El clasificador distingue, además, a nivel de clase, entre bienes y servicios. A su vez, clasifica los bienes según durabilidad en durables, semidurables y no durables.

9.3.4.2.1 Adaptación del clasificador

El clasificador CCIF fue adoptado hasta nivel de clase según la propuesta de Naciones Unidas. Para los niveles de subclase y producto se generó una adaptación nacional.

Para el desarrollo de los niveles subclase y producto se utilizaron como insumos la Clasificación de Gastos por Finalidades (ONU, 2001), la adaptación utilizada en la VII EPF y la canasta de bienes y servicios del Subdepartamento de Estadísticas de Precios al Consumidor, estableciéndose un clasificador preliminar, que fue revisado tanto por entidades internas como externas del INE. En términos generales, el clasificador CCIF utilizado durante la VIII EPF es similar al utilizado en la versión anterior de la encuesta, sin embargo, se realizaron algunas modificaciones a nivel de subclase y producto, de acuerdo a los usos que se le dará a estos datos.

La codificación de bienes y servicios se realizó sobre registros provenientes de las Libretas de Gastos Individuales (LGI), Libreta de Gastos del Recuerdo (LGR), Libreta de Gastos del Hogar (LGH), Libreta de Ingresos (LI) y Libreta de Gastos en Servicios Financieros y Seguros (LSF).

Códigos residuales: la división 15 “Gastos no desglosados de diferentes divisiones y otros bienes y servicios N.C.P.”, y los códigos residuales incorporados en las distintas subclases, fueron el resultado de una decisión operativa frente a casos de registros deficientes, principalmente para productos no desglosados en las libretas, pero también para productos que no poseían una clasificación específica. Según el nivel de información entregada en la encuesta, se procedía a clasificar en alguno de estos códigos.

Los productos sin desglosar pertenecientes a una misma subclase se clasificaban en el residual correspondiente a dicha subclase. Si los productos no desglosados son de distintas subclases, pero de una misma división, los productos se consignaban en la división 15, en el código asociado a la división (tabla 50). También se contemplaron códigos para gastos agregados correspondientes a divisiones distintas, pero que son afines o complementarias. Por último, si correspondían a divisiones distintas y no podían ser clasificados en los residuales de divisiones afines o complementarias, el gasto se asignaba en el residual de todas las clases denominado Gastos en bienes y servicios no desglosados N.C.P. (código 15.1.1.01.26).

Finalmente, los registros de los productos aquí codificados temporalmente se distribuyen según el peso de la división en el gasto total, salvo aquellos registros que, si bien fueron registrados en la encuesta, no corresponden a gasto de consumo final de los hogares y, por lo tanto, quedan fuera del marco teórico conceptual del estudio.

9.3.4.3 Clasificador de establecimientos

El clasificador de establecimientos sistematiza diversos puntos de venta de bienes o servicios, describiendo 15 categorías de establecimientos y dos residuales. El primer residual corresponde a “otros establecimientos N.C.P.”, y agrupa establecimientos que no se pudieron clasificar en las categorías anteriores, pero que son identificables en su tipo, mientras que el segundo denota los establecimientos que no cuentan con suficiente información para ser clasificados.

Tabla 50: Códigos residuales división 15

Divisiones	Código	Glosa división residual 15
División 1	15. 1.1.01.01	Gastos no desglosados en alimentos y bebidas no alcohólicas.
División 2	15. 1.1.01.02	Gastos no desglosados en bebidas alcohólicas, tabaco y estupefacientes.
División 3	15. 1.1.01.03	Gastos no desglosados en prendas de vestir y calzado.
División 4	15. 1.1.01.04	Gastos no desglosados en alojamiento, agua, electricidad, gas y otros combustibles.
División 5	15. 1.1.01.05	Gastos no desglosados en muebles, artículos para el hogar y para la conservación del hogar.
División 6	15. 1.1.01.06	Gastos no desglosados en salud.
División 7	15. 1.1.01.07	Gastos no desglosados en transporte.
División 8	15. 1.1.01.08	Gastos no desglosados en comunicaciones.
División 9	15. 1.1.01.09	Gastos no desglosados en recreación y cultura.
División 10	15. 1.1.01.10	Gastos no desglosados en educación.
División 11	15. 1.1.01.11	Gastos no desglosados en restaurantes y hoteles.
División 12	15. 1.1.01.12	Gastos no desglosados en bienes y servicios diversos.
Divisiones 1 y 2	15. 1.1.01.13	Gastos no desglosados correspondientes a las divisiones 1 y 2.
Divisiones 1 y 11	15. 1.1.01.14	Gastos no desglosados correspondientes a las divisiones 1 y 11.
Divisiones 1 y 5	15. 1.1.01.15	Gastos no desglosados correspondientes a las divisiones 1 y 5.
Divisiones 1 y 12	15. 1.1.01.16	Gastos no desglosados correspondientes a las divisiones 1 y 12.
Divisiones 2 y 11	15. 1.1.01.17	Gastos no desglosados correspondientes a las divisiones 2 y 11.
Divisiones 5 y 12	15. 1.1.01.18	Gastos no desglosados correspondientes a las divisiones 5 y 12.
Divisiones 8 y 9	15. 1.1.01.19	Gastos no desglosados correspondientes a las divisiones 8 y 9.
Cuotas por participación en actividades recreativas sin desglosar	15. 1.1.01.20	Gastos no desglosados correspondientes a pagos de cuotas por participación en actividades recreativas (cuotas para: asados, fiestas, celebraciones, paseos, etc.).
Donaciones⁽¹⁾	15. 1.1.01.21	Gastos en donaciones a instituciones sin fines de lucro (bomberos, iglesias, fundaciones, etc.).
Limosnas⁽¹⁾	15. 1.1.01.22	Gastos en limosnas.
Compra de divisas⁽¹⁾	15. 1.1.01.23	Gastos en compras de divisas.
Descuentos⁽¹⁾	15. 1.1.01.24	Descuentos en bienes y servicios N.C.P.
Automóvil dado en parte de pago⁽¹⁾	15. 1.1.01.25	Automóviles dados en parte de pago.
Productos no desglosados de distintas divisiones N.C.P.	15. 1.1.01.26	Gastos en bienes y servicios no desglosados N.C.P.

(1) Estos productos utilizados en la división residual no corresponden a gastos de consumo final de los hogares, por lo tanto, no son incorporados en las bases de datos publicadas por la encuesta.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

La codificación de establecimientos responde a un uso interno de la encuesta, relacionado con obtener información respecto al tipo de establecimiento donde los hogares realizan sus gastos en bienes y servicios.

Tabla 51: Clasificador de establecimientos

Código	Glosa teórica
1	Supermercados.
2	Almacenes tradicionales.
3	Ferias libres.
4	Restaurantes y bares.
5	Distribuidoras – mayoristas.
6	Vegas – mercados.
7	Comercio ambulante.
8	Multitiendas.
9	Ferreterías y multiferreterías.
10	Farmacias.
11	Clínicas.
12	Hospitales públicos y consultorios.
13	Internet.
14	Extranjero.
15	Tiendas especializadas.
16	Otros establecimientos N.C.P.
17	Establecimientos imposibles de identificar en su tipo.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

9.3.4.4 Clasificador de establecimientos educacionales

El clasificador de establecimientos educacionales utilizado por la encuesta corresponde al listado de establecimientos de educación básica, media y superior reconocidos por el Ministerio de Educación (Mineduc) como tales.

En específico, para el caso de los establecimientos que imparten educación básica y media, todos ellos poseen un rol único de identificación denominado Rol Base de Datos (RBD), el cual se compone de un código de hasta 5 dígitos, junto a un dígito verificador del mismo. Este código es el símil del RUT en los establecimientos educacionales y se encuentra estandarizado a nivel país.

9.4 Metodología de preparación de bases de datos

A continuación, se describe el proceso de preparación de base de datos, que fue desarrollado por el equipo técnico de la encuesta. Este trabajo utilizaba como insumo los datos digitados y codificados, etapas que fueron explicadas en los apartados anteriores.

El propósito de la preparación de bases de datos fue verificar la validez de los datos digitados, editándolos cuando era necesario y preparando así la información para sus análisis finales.

La preparación de la base de datos se enmarcó dentro de la fase de procesamiento y abarcó los siguientes seis subprocesos: 1) validación, 2) recuperación, 3) depuración, 4) imputación y creación de nuevas variables, 5) ajuste temporal y cálculo de ponderadores y 6) consolidación de las bases de datos y cálculo de agregaciones.

Es relevante mencionar, tal como se mencionó en el apartado 7.2, que si bien los subprocesos se describen de manera secuencial, en la práctica el flujo de trabajo no lo es. De hecho, existe una serie de instancias en las que se combinan tanto temporalmente como en términos de flujo de procesos, tareas que pertenecen a distintas etapas e incluso distintas fases del proceso de producción. Los flujos pueden retroceder a etapas que, en una lectura lineal, parecerían previas, para luego avanzar a etapas que se presentan como más avanzadas. En ese sentido, los subprocesos que se describen a continuación, dan cuenta de qué es lo que se hizo en cada momento, pero no necesariamente del orden en el que las tareas se llevaron a cabo.

9.4.1 Validación

El subproceso de validación buscaba comprobar que la información contenida en las bases provenientes de digitación cumpliera con ciertos criterios que aseguraran la calidad de los datos. En caso de que alguno de dichos criterios no se cumpliera, se identificaba el error, para su posterior corrección. Cabe señalar que durante la validación se excluyeron de la base de datos ciertos hogares que no cumplían con estándares mínimos de calidad para formar parte de la muestra, establecidos mediante la implementación de un procedimiento de validación denominado *grilla técnica de mínimos de calidad*.

Este subproceso consta de cinco etapas:

- i. **Revisión de validadores informáticos:** esta etapa confirmaba que las bases de datos descargadas del sistema informático cumplieran con los criterios de calidad mínimos acordados con el equipo de informática. Consistió en una validación básica en la cual se evaluaron cuatro aspectos: 1) existencia de la variable, 2) tipo de variable, 3) existencia de datos en variables obligatorias y 4) rango de la variable. En caso de que no se cumpliera alguno de los criterios establecidos, se generaba un requerimiento de corrección al equipo informático.
- ii. **Aplicación de grilla técnica 1:** se evaluaba la calidad de la información asociada a cada hogar, a través de la asignación de un puntaje. Si el hogar no cumplía con un puntaje mínimo establecido, era excluido de la base de datos. Cabe señalar que ciertos casos ubicados cerca del puntaje de corte se mantenían en el proceso de validación, ya que eventualmente podía recuperarse información valiosa que permitiera mejorar el puntaje del hogar, haciendo que este superase el puntaje mínimo establecido, una vez finalizada la validación y así pasar a formar parte de la base de datos final¹⁷⁶.
- iii. **Conversión de unidad monetaria:** todos los valores de ingreso y gasto registrados en una unidad monetaria distinta del peso eran convertidos a pesos chilenos. Si bien este proceso era parte del subproceso de depuración, se llevó a cabo en este momento por criterios prácticos.

¹⁷⁶ En el próximo apartado se ofrece una descripción acabada de la grilla técnica utilizada en la VIII EPF.

iv. **Validación interna:** utilizando determinadas reglas de validación, cada base de datos fue validada para identificar inconsistencias. Las reglas de validación correspondían a:

a. **Criterios de flujos:** se verificaba que el flujo de la encuesta fuese coherente con el diseño de las libretas. Dado que el flujo de la encuesta variaba de acuerdo a la clase de informante y al tipo de respuestas que este entregaba, era posible que se produjeran errores durante la aplicación del instrumento. Su detección es relevante, pues en muchos casos de ello depende la decisión de incluir o no cierta información en la base de datos final. Asimismo, la identificación de este tipo de errores daba cuenta de información faltante que era necesario recuperar.

b. **Criterios de rango:** se verificaba que las respuestas se encontrasen dentro de los rangos predeterminados de respuesta, de acuerdo a la variable y al instrumento. Para variables discretas, estas debían tomar alguno de los códigos establecidos para la pregunta o los valores de no respuesta válidos¹⁷⁷. Para variables continuas, como el gasto o el ingreso, no se generaron validaciones de rango, sino que se depuraron estadísticamente, a través de un proceso posterior (revisión de valores atípicos u *outliers*). Los rangos de validación no siempre dependen solamente de una variable, sino de una combinación de las mismas, por lo tanto, estas validaciones de rango se realizaron tomando en cuenta la interacción de las distintas variables, tanto al interior de cada instrumento como entre instrumentos.

iv. **Aplicación de grilla técnica 2:** se evaluó nuevamente la calidad de la información asociada a cada hogar. El objetivo de volver a aplicar la grilla técnica fue identificar si aquellos hogares que se encontraban cercanos al punto de corte lograban superar dicho puntaje mínimo de calidad después de los procesos de recuperación y depuración de datos.

El subproceso de validación tuvo como resultado dos productos. En primer lugar, bases de datos en las cuales era posible identificar todos los casos que presentaban algún tipo de error de los evaluados en este subproceso. En segundo lugar, un archivo de revisión que reunía los errores detectados en cada una de las libretas. Dicho archivo constituía el principal insumo para el subproceso de recuperación que se describe en el apartado 9.4.2.

9.4.1.1 Grilla técnica de mínimos de calidad

La grilla técnica de calidad de la VIII EPF correspondió a un método de evaluación de mínimos de calidad y cantidad de los datos entregados por cada hogar que participó en la encuesta, evaluando, de esta forma, la información entregada en cada instrumento. Estos requerimientos mínimos de calidad fueron definidos en concordancia con los objetivos de la encuesta, estableciendo distintas ponderaciones para cada libreta y sus preguntas prioritarias¹⁷⁸.

El objetivo de la grilla técnica fue establecer una distinción entre hogares entrevistados con información completa o parcial, de hogares que fueron clasificados como rechazo, debido

177 Los valores de no respuesta válidos establecidos en la encuesta fueron -88 (no sabe) y -99 (no responde).

178 El anexo 14 presenta la grilla técnica de mínimos de calidad con sus respectivos puntajes a evaluar por libreta.

a que entregaron información insuficiente respecto a los objetivos del estudio (*break-off*¹⁷⁹). En otras palabras, la grilla técnica permitía identificar qué nivel de no respuesta al ítem en las preguntas relevantes era aceptable para clasificar a un hogar como entrevistado y qué nivel de no respuesta parcial en estas preguntas hacía necesario clasificar un hogar como rechazo.

La importancia de esta definición radicaba en que los principales estimadores de la encuesta deben ser calculados con los hogares clasificados como entrevistados. Por lo tanto, la grilla permitía excluir del análisis a hogares que podrían presentar un importante nivel de subdeclaración de información¹⁸⁰, con la finalidad de no sesgar el cálculo de dichos estimadores, es decir, la grilla jugó un relevante rol en la reducción del sesgo de subdeclaración de información (gasto e ingreso).

La grilla de mínimos de calidad asignaba a cada hogar (considerando todos los cuestionarios que debían contestarse) un puntaje entre 1 y 100 puntos, en los que se evaluaban ciertos módulos y preguntas de los seis instrumentos. A cada criterio evaluado se le asignó distintas ponderaciones. En cada libreta se identificaron preguntas prioritarias, que fueron evaluadas de acuerdo a la completitud y consistencia de la información proporcionada por los hogares y no debido a los montos de gasto e ingreso entregados en los cuestionarios. Bajo este criterio de asignación de puntajes, la grilla de la VIII EPF estableció un mínimo de 76 puntos de exigencia¹⁸¹ para ser considerado como hogar entrevistado. Los hogares que no alcanzaron los 76 puntos, pero que entregaron información para el estudio fueron clasificados como *break-off* y fueron excluidos del cálculo de los estimadores finales de la encuesta.

Uno de los criterios más importantes para evaluar la calidad de la información de un hogar correspondió al registro de los gastos diarios que capturaba la LGI, ya que dicha libreta captura un alto porcentaje del gasto de los hogares¹⁸². Debido a lo anterior, se le asignó el puntaje más alto en relación a los criterios evaluados por la grilla, valorando que todos los integrantes elegibles¹⁸³ del hogar hubiesen respondido un mínimo de cuatro días de registro, con al menos un día de fin de semana. Adicionalmente, del total de miembros del hogar que debían contestar la LGI, se distinguió a aquel clasificado como administrador de gastos del hogar¹⁸⁴ respecto de aquellos integrantes elegibles, ajustando la evaluación para que el puntaje de la LGI de un administrador de gastos ponderase el doble que el puntaje de un integrante que no lo era al interior de cada hogar. Esta distinción se realizó así considerando que no todos los miembros del hogar aportan de igual manera al gasto total de este¹⁸⁵.

179 Para una explicación al respecto revisar el estándar propuesto por Aapor (2015, pág. 23).

180 Al no entregar respuestas a ciertas preguntas/cuestionarios importantes para la encuesta.

181 El puntaje de corte corresponde a un criterio normativo de exigencia.

182 Según datos de la VII versión de la encuesta, aproximadamente un 42% del total del gasto capturado corresponde a información proveniente de este cuestionario.

183 Corresponde a los miembros de 15 años o más que están capacitados para responder la Libreta de Gastos Individuales.

184 El administrador de gastos del hogar corresponde a la persona que habitualmente ha realizado las compras del hogar en los últimos seis meses. Puede existir más de un administrador del gasto en el hogar.

185 Según la VIII EPF, la LGI de un administrador de gastos del hogar posee un gasto en promedio un poco más elevado que el doble del gasto reflejado en las LGI del resto de los integrantes del hogar.

El método de la grilla técnica fue aplicado dos veces durante el proceso de validación de las bases de datos, ya que se buscaba conseguir dos objetivos. Por un lado, focalizar el trabajo de recuperación en aquellos hogares en donde se estimase que era posible alcanzar un nivel mínimo de información de acuerdo al estándar establecido por esta grilla. En segundo lugar, este método de evaluación buscaba dirimir entre unidades entrevistadas parcialmente y unidades que, por su nivel de no respuesta al ítem, debiesen ser clasificadas como entrevistas interrumpidas o *break-off*.

La primera aplicación de la grilla generó los siguientes niveles de clasificación de los hogares:

- i. **Supera los criterios de cantidad y calidad:** el hogar cumplía satisfactoriamente con el mínimo de los criterios establecidos para ser incluido en el análisis.
- ii. **No supera los criterios de cantidad y calidad, pero permanece en el proceso:** el hogar no cumplía de manera satisfactoria con el mínimo de criterios establecidos para ser incluido en el análisis. No obstante, el puntaje de la grilla era cercano al puntaje de corte, por lo cual era factible recuperar datos faltantes en el proceso de revisión y recuperación de la encuesta.
- iii. **No supera los criterios de cantidad y calidad:** el hogar no cumplía de manera satisfactoria con el mínimo de criterios establecidos para ser incluido en el análisis. La cantidad de datos faltantes impedía incluirlos *a priori* en las siguientes etapas de validación y depuración.

La segunda aplicación de la grilla se realizaba una vez finalizados los subprocesos de recuperación y depuración de datos. De esta forma, si existían errores de digitación o se recuperaba información al volver a contactar al hogar, su información era evaluada sobre los datos corregidos. El objetivo de esta aplicación era determinar el código de disposición final de los hogares contactados.

El resto de los procesos descritos a continuación fueron realizados solo con aquellos hogares que cumplieron los estándares establecidos por la grilla técnica de mínimos de calidad y que, por tanto, forman parte de la base de datos final de la encuesta.

9.4.2 Recuperación

Este subproceso tuvo por objetivo mejorar la calidad de los datos provenientes de los subprocesos de digitación y de trabajo de campo, mediante revisión física de libretas, lectura de observaciones y recuperación en terreno. El insumo más importante de este subproceso correspondía al archivo de revisión (elaborado durante la etapa de validación). El producto principal correspondía a un archivo de revisión actualizado, que contenía los resultados de los esfuerzos realizados por mejorar la calidad de los datos en las distintas instancias de recuperación.

Este subproceso se compuso de cuatro etapas:

- i. **Consolidación de archivo de revisión:** dado que las solicitudes de recuperación podían provenir de los subprocesos de codificación, validación y depuración, se hizo necesario, en primer lugar, reunir toda esta información en un archivo estandarizado.
- ii. **Revisión de libretas:** esta etapa constituía el primer esfuerzo por mejorar la calidad de los datos. Consistía en la revisión física de las libretas, con el fin de encontrar información relevante respecto a los datos inválidos. Dentro de

los casos que no lograban corregirse, se identificaron aquellos que debían ser recuperados por el equipo de trabajo de campo.

- iii. **Recuperación en trabajo de campo:** correspondía a la última instancia para mejorar la calidad de la información recolectada. Esta etapa se realizaba bajo dos modalidades, dependiendo de la libreta y/o de la naturaleza del error. La primera modalidad consistía en recuperación de gabinete realizada por el encuestador (por ejemplo, en los casos de letra ilegible), mientras que la segunda modalidad consistía en volver a contactar al hogar, para solicitar información faltante o consultar respecto a inconsistencias.
- iv. **Consolidación de bases de datos:** una vez que se llevaban a cabo los distintos esfuerzos por mejorar la información, los datos recuperados eran nuevamente validados y se procedía a realizar la consolidación de las bases de datos del subproceso de recuperación.

9.4.3 Depuración

Este subproceso tuvo por objetivo mejorar la consistencia y completitud de los datos de las unidades muestrales, mediante la corrección de la información de las distintas bases de datos, de acuerdo con los protocolos predefinidos. Los principales insumos del subproceso fueron las bases de datos validadas y el resultado principal fueron las bases de datos depuradas para imputar datos.

El subproceso estuvo compuesto de cinco etapas:

- i. **Depuración prerrecuperación:** se inició después de la etapa de validación interna y consistió en la depuración de datos que resultaron inválidos y que no se enviaron al subproceso de recuperación. Esta edición de datos se realizó mediante la revisión de observaciones y a través de ciertas depuraciones automáticas, en los casos en los que era evidente la solución del error. La decisión de no enviar todos los datos inválidos al subproceso de recuperación obedeció a la necesidad de disminuir la carga del equipo encargado de dicha tarea.
- ii. **Depuración de bases de datos:** una vez terminada la recuperación en libretas, no todos los datos con error lograban ser depurados, lo que hizo necesario llevar a cabo algún procedimiento para que dichos datos se validaran. En ese sentido, el objetivo de esta etapa era que todos los datos tuvieran un valor válido, a excepción de aquellos casos que debían recuperarse en trabajo de campo. Lo anterior implicaba que todos los datos inválidos que no lograban modificarse en el subproceso de recuperación recibían algún tratamiento automático definido previamente. Algunos ejemplos de ello son: el llenado de campos vacíos con el valor -88 (no responde), corrección de fechas en la LGI y LGR para generar consistencia con el período de referencia, corrección de gastos de la vivienda secundaria en función de lo declarado en RPH y modificación en datos de LGH para no superar la periodicidad anual, entre otros.
- iii. **Análisis de gastos duplicados:** debido a las características que tiene la EPF, existía la posibilidad de que un informante declarase más de una vez un mismo gasto (por ejemplo, el pago de servicios básicos registrados en la LGH y LGI). Es por ello que esta etapa tuvo como objetivo identificar, mediante el uso de ciertos criterios, aquellos casos en que se duplicaba el registro del gasto para su posterior tratamiento. Al finalizar esta etapa, se obtenían bases de datos en las cuales cada registro tenía un identificador de gastos duplicados.

- iv. **Traspaso de gastos:** en esta etapa, en primer lugar, se aplicó un tratamiento a los casos que fueron identificados como duplicados en la etapa anterior, lo cual implicaba eliminar algunos registros. En segundo lugar, se trasladaron aquellos casos que, dada su naturaleza, no correspondían a la libreta en la que fueron registrados originalmente. La necesidad de trasladar dichos gastos se explicaba por la periodicidad que teóricamente debiese tener cada gasto. El caso de la LGI es ilustrativo: si bien al informante se le solicitaba registrar todos los gastos realizados durante la submuestra operativa de trabajo de campo, los gastos en bienes durables, como electrodomésticos o muebles, no fueron ajustados temporalmente con el resto de registros de la LGI (periodicidad intramensual), sino que fueron trasladados a la LGR, para aplicarles el ajuste temporal correspondiente a los 3, 6 o 12 meses, según correspondiera¹⁸⁶.
- v. **Análisis de valores atípicos (outliers):** esta etapa tuvo el objetivo de detectar la existencia de datos atípicos en las distribuciones de gastos e ingresos. En el caso de ingresos, se generaron grupos utilizando las variables CSE y sexo, para luego, dentro de cada grupo, analizar la distribución. En el caso de gastos, el análisis se llevó a cabo a nivel de productos.

Cabe señalar que, considerando la cantidad y dispersión en el registro de los gastos, los métodos utilizados variaron según la libreta. A continuación, se describe el procedimiento realizado en cada uno de los instrumentos.

- a. **LGI:** se analizó el valor unitario de cada registro utilizando la desigualdad de Chebyshev. Adicionalmente, se revisaron aquellos casos que se ubicaban en el percentil 1 o en el percentil 99 de la distribución de gastos.
- b. **LI:** se utilizó el método Hadi¹⁸⁷. Adicionalmente, se etiquetaron como *outliers* los valores menores a 1.000.
- c. **LGR:** se utilizó el método Hadi. Adicionalmente, se etiquetaron como *outliers* los casos que están bajo cierto percentil de la base de gastos.
- d. **LSF:** se utilizó el método Hadi.
- e. **LGH:** se utilizó el método Hadi. Adicionalmente, se etiquetaron como *outliers* los valores menores a 500.

Una vez detectados los valores atípicos (*outliers*), estos se reunían en un único archivo y para ser sometidos al subproceso de recuperación descrito en el apartado 9.4.2¹⁸⁸. Ello quiere decir que los datos recuperados eran validados y luego incorporados en las distintas bases de gastos.

9.4.4 Imputación, ajuste temporal y creación de nuevas variables

El presente subproceso se compone de las etapas de Imputación de datos, ajuste temporal al mes de referencia y creación de nuevas variables. En los siguientes apartados se da cuenta de estas etapas.

186 Los algoritmos de ajuste temporal utilizados por la encuesta son explicados en el apartado 9.4.4.2

187 El método de Hadi permite incorporar distintas variables de contexto y evaluar así el comportamiento de los registros (ingresos o gastos) en función de características socioeconómicas de los informantes.

188 Es importante recordar que el flujo de trabajo no es lineal.

9.4.4.1 Imputación de datos

La ausencia de datos es un problema inevitable con el cual debe lidiar cualquier encuesta. La falta de información suele clasificarse en dos tipos: 1) **no respuesta total**, que ocurre cuando en el trabajo de campo no se logró recabar información sobre las unidades seleccionadas¹⁸⁹ y 2) **no respuesta parcial**, que ocurre cuando se cuenta con información del hogar seleccionado, pero por algún motivo no fue posible recolectar información de algunos ítems (existen datos *missing*). La no respuesta total en la VIII EPF se intentó atenuar mediante la incorporación de sobremuestreo en el diseño muestral (ver apartado 6.1.3.3), estrategias de trabajo de campo (ver apartado 8.1.3) y correcciones en los factores de expansión (ver apartado 6.2.3), mientras que la no respuesta parcial fue atendida mediante métodos de imputación.

A pesar de los distintos resguardos y estrategias aplicadas por los equipos de campo para buscar la participación de los hogares, y entendiendo que la respuesta de los hogares era preferible a cualquier método de imputación de datos, la experiencia internacional, nacional e institucional muestra que la existencia de datos faltantes es una situación habitual en las encuestas de hogares.

La presencia de información faltante tiene serias implicancias para la interpretación de los resultados de una encuesta, tales como la disminución del tamaño muestral, la potencial presencia de sesgo en las estimaciones y la falta de significancia estadística de los mismos, ante lo que se recomienda imputar los datos faltantes, para no inducir a errores de interpretación y disminuir el peligro de sesgos en las encuestas¹⁹⁰ (Eurostat, 2013).

Antes de aplicar métodos de imputación de datos, se debe estudiar si la no respuesta sigue algún tipo de patrón respecto a variables en particular que pudiesen sesgar los resultados de la encuesta. Si la no respuesta no es sistemática, la aleatoriedad asegurará que los datos imputados no estén sesgados. Al contrario, si la no respuesta es sistemática, se debe tener presente esta información, con el fin de implementar procesos de imputación que tengan en cuenta esta característica.

Para la VIII EPF se realizaron imputaciones para las variables de ingreso y de gasto. Para estos dos procesos, se probaron distintas metodologías antes de determinar, a través de análisis distribucionales, el método a utilizar. Es fundamental tener en consideración que el comportamiento de la no respuesta puede variar entre las variables de una misma encuesta, de modo que no es conveniente adoptar *a priori* el mismo procedimiento de imputación para todas las variables (Medina & Galván, 2007). Cada situación es diferente, por lo que se hace necesario llevar a cabo análisis diferenciados según las variables que se pretenda imputar.

El proceso de imputación de datos se organizó en tres etapas: 1) proceso de preparación, en el que se caracterizó la no respuesta de la variable a imputar, se testeó su aleatoriedad y se modeló su comportamiento; 2) proceso de aplicación, en el cual se implementó el o los métodos definidos para cada variable a imputar, y 3) proceso de análisis, en el que se

189 O bien la información recolectada no cumple los mínimos técnicos de calidad impuestos por la encuesta (*break-off*).

190 Dado que el objetivo tras la realización de una encuesta es estimar parámetros poblacionales específicos, a partir de un estimador cuyo valor se obtiene de una muestra aleatoria de la población, la imputación de datos puede ser vista como un método utilizado para la corrección de la no respuesta parcial.

estudió el desempeño del o los métodos, mediante ejercicios de simulaciones y a través del análisis de los datos obtenidos después de la imputación. En el caso de las pruebas de los distintos métodos de imputación, en la tercera etapa se calcularon distintos estadísticos para evaluar los métodos.

Estimación del arriendo imputado para la vivienda principal y secundaria

El arriendo imputado corresponde a la estimación del gasto en arriendo que debería realizar un hogar si no fuera propietario de su vivienda o si esta no le fuera cedida de forma gratuita¹⁹¹. En la VIII EPF, el cálculo de arriendo imputado se realiza basado en la respuesta de dos preguntas realizadas directamente al hogar:

- i. Si Ud. tuviera que arrendar esta vivienda a un tercero, ¿cuánto cobraría por arrendársela mensualmente?
- ii. ¿Cuánto estaría dispuesto/a a pagar por el arriendo de una vivienda de características similares a la que habita?

Estas preguntas se realizaron en cuestionarios diferentes (la primera en RPH y la segunda en LGH) con el fin de realizarlas en días diferentes. En ambas preguntas se especificó que el valor debía estimarse sin considerar mobiliario. Se esperaba que ambas preguntas presentaran sesgos contrapuestos. La primera de ellas se esperaba que tuviese sesgo a la sobrestimación, mientras que la segunda a la subdeclaración. Al preguntar por las dos se buscó que, en promedio, estos sesgos se compensaran.

En el caso de tener ambas respuestas, el arriendo imputado se estimó como el promedio simple de ambas preguntas. En el caso de tener solo una de ellas, el arriendo imputado tomó el valor declarado para la pregunta que se encontraba contestada. Cuando no se tenía respuesta para ninguna de las dos preguntas, el arriendo imputado se estimó a partir del promedio de arriendo imputado declarado por el resto de los hogares de la manzana, en caso de no existir donantes a ese nivel, se amplió el nivel de búsqueda de donantes (distrito censal y comuna), manteniendo la clasificación socioeconómica de la manzana (CSE) a la que pertenecía el hogar con información faltante.

9.4.4.1.2 Imputación de Libreta de Gastos Individuales

La metodología utilizada en la VII EPF para imputar la no respuesta parcial de las Libreta de Gastos Individuales (LGI) fue a través de un Factor de No Respuesta (INE Chile, 2014).

Este ajuste también es utilizado en otros países y resuelve de manera satisfactoria el problema de la no respuesta parcial en las LGI que contengan información. Por ejemplo, Australia utiliza un método combinado para la corrección de la no respuesta; para las libretas incompletas utiliza un método similar al Factor de No Respuesta, mientras que para las libretas no contestadas utiliza *Hot-Deck* (ABS, 2011). Por su parte, Canadá sólo utiliza el factor de no respuesta para las libretas parcialmente contestadas (Lapierre, 2015).

¹⁹¹ Las categorías que se consideran como parte del usufructo son: propia (totalmente pagada y propia, pagándose), cedida por un pariente o amigo, ocupada de hecho, propiedad en litigio y herencia o sucesión compartida. Por tanto, no se calcula arriendo imputado para aquellos hogares que pagan arriendo (con o sin contrato) o si la vivienda fue cedida por el trabajo.

Una de las mayores desventajas del factor de ajuste por no respuesta es que no está diseñado para tratar la no respuesta de libretas completas de LGI, es decir, libretas completamente vacías sin información.

Considerando lo anterior, en la VIII EPF se tomó la decisión de imputar tanto aquellas libretas que tenían información, pero no completas (libretas parciales), como aquellas libretas no contestadas (libretas sin información)¹⁹². Ello supone una diferencia importante respecto a la versión anterior de la encuesta, en la que solo se imputaron libretas parcialmente completadas, es decir, todas aquellas que tenían al menos un día de registro, completando así los días que se debían registrar. Esta modificación metodológica se hace cargo del aumento de la no respuesta de la LGI respecto a la versión previa¹⁹³ y, al mismo tiempo, recoge las buenas prácticas de países como Australia, Canadá, España, Italia, Reino Unido, Sudáfrica y otros, que imputan libretas completas (ONS, 2017; Statistics South Africa, 2012; INDEC, 2006; ABS, 2011; Lapierre, 2015; INE España, 2017; ISTAT, Italia, 2011).

Tabla 52: Estadística de la no respuesta parcial en gastos individuales a nivel de personas que debían contestar la libreta

Libretas LGI	Cantidad	Porcentaje
Libretas LGI esperadas⁽¹⁾	37.718	100
Libretas LGI contestadas (completas y parciales)	30.930	82,0
Diferencia (No contestadas⁽²⁾)	6.788	18,0

(1) Para el cálculo de las LGI esperadas, se restan aquellos códigos asociados a informantes que no debían contestar este instrumento (informantes no elegibles).

(2) Si todos los días se encontraban sin registro, la LGI se asumía como no contestada

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

En la tabla 52, se puede observar que, en términos de libretas, la no respuesta de libretas en LGI es de un 18%, mientras que a nivel de días (tabla 53) la no respuesta parcial aumenta a un 25,8%, considerando los días sin registro tanto de las libretas con información parcial como de aquellas sin información. Considerando que las LGI capturan la mayor parte del gasto en los hogares, es de suma importancia abordar la no respuesta por medio de procesos de imputación.

Tabla 53: Estadística de la no respuesta parcial en gastos individuales a nivel de días, considerando aquellas personas que debían contestar la libreta

Categorías	N° de días	Porcentaje
Días con registro⁽¹⁾	425.619	74,2
Días con gastos	284.319	49,6
Días sin gastos	141.300	24,6
Días sin registro	148.018	25,8
Total de días	573.637	100

(1) Los "Días con Registro" corresponden a la suma de los "Días con gasto" y "Días sin gasto".

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

192 Dada la particularidad de seis libretas para el estudio de la EPF, la existencia de libretas incompletas en su totalidad sigue siendo no respuesta parcial en aquellos hogares que aprueban la grilla técnica de mínimos de calidad. Los hogares que contienen información, pero son rechazados por la grilla son considerados como no respuesta total. Al respecto, ver el apartado 9.4.1.1.

193 La VII EPF presentó un nivel de no respuesta de libretas sin nada de información de 16,2% (INE Chile, 2014b, pág. 118)

El método de imputación seleccionado para la VIII EPF fue *Hot-Deck*. En términos generales, el método de *Hot-Deck* reemplaza cada valor perdido por la respuesta observada de una unidad *similar*, que en la literatura suele denominarse *donante* o *vecino cercano*. Cabe señalar que tanto la definición de semejanza como la forma de seleccionar a los donantes no son cuestiones exentas de discusión. De hecho, existe una gran variedad de mecanismos para seleccionar a los donantes (Andridge & Little, 2010). En ese sentido, más que un método en particular, el *Hot-Deck* es una familia de técnicas basadas en la idea común de que es posible reemplazar los datos faltantes de una unidad por los de una unidad de características similares (individuo, hogar, registro, dependiendo de la unidad de análisis).

Este tipo de métodos es ampliamente utilizado en encuestas de hogares a nivel nacional e internacional. Respecto a la experiencia internacional en encuestas de presupuestos familiares, este tipo de imputación es utilizado por países como Australia, Reino Unido, Argentina y Sudáfrica, entre otros (ONS, 2017; Statistics South Africa, 2012; INDEC, 2006; ABS, 2011). A nivel nacional, en tanto, la imputación por *Hot-Deck* es utilizada por la Encuesta Suplementaria de Ingresos (ESI) y por la encuesta Casen. En ambos, este método es usado para la corrección de la no respuesta parcial de ingresos de la ocupación y de jubilaciones¹⁹⁴ (INE España, 2017; CEPAL, 2016).

En el caso de la VIII EPF, la forma de seleccionar a los donantes fue a través de la creación de grupos de individuos similares o clusters, mediante variables que cumplieran con una o más de las siguientes características:

- i. **Presentar una alta correlación con el gasto:** dado que lo que se imputó es el gasto (monto y estructura), fue necesario que las variables utilizadas para construir los grupos o clusters tuviesen una correlación alta con la variable a imputar. En este sentido, se priorizó la correlación con el monto del gasto. Dicho de otro modo, para la metodología empleada, no era necesario que los donantes compartieran todas las características del receptor, sino aquellas que se relacionaran más con el gasto.
- ii. **Explicar la propensión a responder la LGI:** la no respuesta constituye un problema, desde el punto de vista del sesgo, cuando las propensiones a responder de las personas son muy disímiles entre sí. Cuando todos los individuos tienen la misma probabilidad de responder, la no respuesta no genera sesgo¹⁹⁵, lo cual es una situación extremadamente poco común en el contexto de encuestas de hogares. En el caso concreto de la VIII EPF, se tiene certeza de que ciertos perfiles de informantes tienen una probabilidad más alta de responder que otros, lo cual implicaría que algunos segmentos queden subrepresentados. Para que la metodología de imputación fuera exitosa, se buscó que las personas dentro de un mismo cluster tuvieran una probabilidad similar de responder. Es por ello que se requirió utilizar variables que buscasen explicar la no respuesta al momento de construir los clusters. Además, la existencia de sobremuestreo según niveles socioeconómicos ayudó a obtener respuestas de distintos perfiles de unidades considerando distintas probabilidades de respuesta, lo que ayudó a mejorar el proceso de imputación de LGI.

194 En el caso de la Casen, el método *Hot-Deck* fue utilizado para imputar jubilaciones y rentas vitalicias.

195 En la literatura esto se conoce como una no respuesta tipo MCAR (por sus siglas en inglés missing completely at random). Los autores Little y Rubin (2002) hacen una detallada explicación del efecto de la no respuesta.

- iii. ***Estar dentro de las variables utilizadas en el diseño muestral:*** el diseño muestral estableció distintas probabilidades de selección para las unidades primarias y secundarias de muestreo. Imputar por el método de *Hot-Deck* sin considerar que las unidades tienen distintas probabilidades de selección podría generar distorsiones en el diseño muestral. En términos prácticos, el método de imputación no debería elegir con la misma probabilidad a dos donantes que en el diseño tienen probabilidades diferentes de selección. La manera de lidiar con este problema fue incluir dentro de las variables del *Hot-Deck*, aquellas que se utilizan en el diseño de la muestra, de acuerdo a las recomendaciones de la literatura (Andridge & Little, 2009).

El mejor donante es aquel que comparte todas las características del receptor, sin embargo, no siempre existe en la muestra alguien que cumpla ese criterio. Por lo tanto, el método buscaba donantes para los casos perdidos, partiendo desde el máximo nivel de exigencia, para luego flexibilizar el criterio de similitud, hasta que se lograba encontrar a la persona más parecida. Para realizar este ejercicio de búsqueda de donantes, se generó una matriz de exigencia para la elección del vecino cercano (en el anexo 15 se puede consultar la matriz de exigencia definida). El funcionamiento de la matriz consistió en establecer un primer nivel de imputación con las variables definidas, para luego ir flexibilizándolas, comenzando por aquellas que presentaban una menor correlación con el gasto. Cabe señalar que en aquellos casos en que existía más de un candidato para llevar a cabo la donación (clúster formado por tres o más individuos), el método realizaba una selección aleatoria del donante¹⁹⁶.

En la tabla 54 se presentan algunos estadísticos en relación al proceso de imputación de LGI. Se puede observar que el gasto promedio por LGI sin considerar el ajuste por imputación es inferior en \$5.491 respecto del gasto ajustado por no respuesta (imputación *Hot-Deck*), donde el promedio de gasto por libreta aumenta de \$174.968 a \$180.459. Esto se debe a que la imputación se hizo cargo de la falta de respuesta de todos los días sin registro, incluidas las libretas parcialmente llenadas.

Considerando los resultados de los procesos de imputación, se observa que existe una brecha al revisar el gasto promedio de las libretas completas en comparación al total de libretas imputadas. Esta comparación sitúa al gasto promedio por libreta post imputación por debajo del de las libretas completas, situación que es esperable, ya que los administradores de gastos tienen una mejor respuesta en comparación al resto de los integrantes del hogar¹⁹⁷.

196 Este proceso de selección aleatoria se controló a través de la implantación de "semillas" en la programación computacional con el fin de hacer replicable el proceso.

197 La evidencia de este resultado también es parte de la experiencia de la VII EPF (INE Chile, 2014b). Los administradores realizan un gasto que en promedio alcanza a aproximadamente el doble del gasto del resto de integrantes del hogar.

Tabla 54: Comparación de estadística descriptiva entre datos imputados y sin imputar, a nivel muestral

Comparación Libretas de Gastos Individuales	N° Obs.	Promedio	Mínimo ⁽¹⁾	Máximo	SD	CV	p25	p50	p75
LGI completas y parciales sin ajuste	30.930	174.968	0	3.190.127	195.042	1,11	46.919	116.357	236.660
Sólo LGI completas	22.587	187.483	0	3.190.127	202.643	1,08	53.146	128.840	253.826
LGI post imputación	37.718	180.459	0	3.190.127	199.671	1,11	50.084	120.814	243.312

(1) El gasto mínimo en de un infórmate capturado en la LGI es de 0 pesos, dado que los informantes pudieron declarar durante todos los días de la quincena un gasto de 0 pesos (días sin gasto). Dicho registro es válido y difiere de un día sin registro, donde no se tiene información si tuvo gasto o no (no respuesta).
Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

El efecto de la imputación en el gasto promedio por hogar se observa en la tabla 55, con un aumento en \$96.212 pesos por hogar. Esto tiene gran relevancia, ya que, considerando la magnitud de la no respuesta en LGI, no realizar un proceso de imputación implicaría sesgar el gasto promedio de forma considerable (subestimación). Por otro lado, al observar el efecto de la imputación de LGI en la estructura de gastos de los hogares, se constata que los mayores aumentos relativos se encuentran en la división 1 y la división 11. Esto ocurre debido a que la LGI captura la mayor parte de los gastos en estas dos divisiones, por lo que la imputación de esta libreta implicó un aumento el gasto promedio por hogar en estos apartados, al igual que su participación en el presupuesto del hogar.

Tabla 55: Estructura por división del clasificador de CCIF del gasto de consumo final de los hogares pre y post imputación de LGI, a nivel expandido (excluye arriendo imputado)

División CCIF	Glosa	Total de hogares antes del ajuste		Total de hogares después del ajuste	
		Gasto promedio por hogar	% del gasto respecto al total	Gasto promedio por hogar	% del gasto respecto al total
1	Alimentos y bebidas no alcohólicas.	171.024	16,7	209.983	18,7
2	Bebidas alcohólicas, tabaco y estupefacientes.	14.562	1,4	19.636	1,8
3	Prendas de vestir y calzado.	38.220	3,7	38.690	3,5
4	Alojamiento, agua, electricidad, gas y otros combustibles.	158.210	15,4	160.692	14,3
5	Muebles, artículos para el hogar y para la conservación ordinaria del hogar.	66.721	6,5	70.778	6,3
6	Salud.	79.475	7,7	84.791	7,6
7	Transporte.	158.879	15,5	170.238	15,2
8	Comunicaciones.	58.351	5,7	59.209	5,3
9	Recreación y cultura.	69.941	6,8	76.167	6,8
10	Educación.	72.604	7,1	72.596	6,5
11	Restaurantes y hoteles.	58.315	5,7	73.096	6,5
12	Bienes y servicios diversos.	79.411	7,7	86.050	7,7
Total		1.025.713	100	1.121.925	100

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

9.4.4.1.3 Imputación de ingresos

Al igual que en la versión anterior de la encuesta, para la imputación de ingresos en la VIII EPF se probaron métodos alternativos¹⁹⁸. Finalmente, del mismo modo que en la imputación de gasto, se seleccionó el método *Hot-Deck*, debido a los resultados obtenidos.

En el caso de los ingresos, la corrección de la falta de respuesta se concentra en las principales fuentes de ingreso¹⁹⁹, a saber, los ingresos laborales de la actividad principal y aquellos provenientes de jubilaciones o pensiones de vejez. La proporción de no respuesta en este caso fue similar a la versión previa de la encuesta²⁰⁰. Las tasas de no respuesta parcial según las diferentes fuentes se pueden observar en la tabla 56.

Tabla 56: Estadística de la no respuesta parcial de ingresos, a nivel personas

Categoría en la ocupación principal	N° Obs.	Datos faltantes		N total
		n	%	
Sueldos y salarios (dependientes)	14.794	1.498	9,2	16.292
Asalariados (excluye honorarios)	13.861	1.395	9,1	15.256
Asalariados (honorarios)	933	103	9,9	1.036
Independientes	5.536	422	7,1	5.958
Empleadores	465	40	7,9	505
Cuenta propia	5.071	382	7	5.453
Jubilados	5.230	185	3,4	5.415

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Los trabajadores asalariados son definidos como aquellos que indican una relación de subordinación y dependencia en su relación laboral, lo que no requiere necesariamente un contrato de por medio. Lo anterior se materializa en una regularidad en el pago de su salario, tenencia de jefatura y cumplimiento de horario, entre otros.

En el caso de la imputación de ingresos, los trabajadores que entregan boletas de honorarios se trabajaron de forma independiente al resto de los asalariados, ya que tienen características diferentes que fueron consideradas al momento de imputar datos faltantes. Esta categoría de trabajadores realiza cotizaciones de salud y previsión de forma voluntaria y por montos diferentes a los obligatorios para la categoría de asalariados con contrato.

Considerando los trabajadores que entregan boletas de honorarios, la ley N° 20.255²⁰¹ de la reforma previsional chilena indica que aquellos trabajadores que entreguen boletas

198 El detalle de los resultados de los distintos modelos y de la toma de decisión se incluirán en documento de trabajo INE a publicarse en fecha posterior a la publicación de la VIII EPF, el cual también incluirá los ejercicios realizados para la imputación de Libretas de Gastos Individuales.

199 Las mismas categorías fueron corregidas por no respuesta en la VII EPF. Respecto a esta información consultar INE Chile (2014b)

200 Mientras que la no respuesta de los dependientes fue 8,93 en la VII EPF, para la VIII EPF dicha cifra se elevó a 9,13%.

201 La ley 20.255 está disponible en: <https://www.leychile.cl/Navegar?idNorma=1086935#20160>

deberán comenzar a cotizar de manera obligatoria desde el año 2018²⁰². Considerando lo anterior, para la IX EPF, probablemente el proceso de imputación considerará el grupo de asalariados en su conjunto sin la presente distinción.

Para seleccionar las variables correlacionadas con los ingresos que fueron utilizadas en los distintos métodos de imputación, se utilizaron los siguientes dos criterios:

- i. **Variables teóricas:** variables determinantes del ingreso, de acuerdo a diversos estudios realizados desde la academia. La selección de las variables escolaridad, edad, edad al cuadrado y sexo se fundan en el modelo de capital humano, teoría del ciclo de vida y, más actualmente, el modelo de señalización (Weiss, 1995; Mincer, 1974). Se ha demostrado la correlación entre el ingreso y la escolaridad mediante el modelo de capital humano y señalización. A su vez, se ha constatado la pertinencia de emplear la variable edad como *proxy* de experiencia y la edad al cuadrado, la cual suaviza los retornos de la actividad que van disminuyendo con el tiempo. Adicionalmente, la brecha salarial entre hombre y mujeres ha sido estudiada y confirmada, existiendo estudios que la demuestran tanto a nivel nacional como internacional.

Por otro lado, la inclusión de variables como clasificación socioeconómica de la manzana (CSE) y variables geográficas se debe al diseño muestral de la encuesta, en el que las probabilidades de selección de una manzana están condicionadas por estas variables, por lo que es relevante su inclusión como forma de control (Andridge & Little, 2009).

En el caso de las jubilaciones, la evidencia empírica muestra una estrecha relación entre la edad y el monto ahorrado (no medido en la encuesta), aspecto que motiva la inclusión de esta variable. Asimismo, se ha constatado que las jubilaciones se relacionan con el desarrollo de la vida laboral, de modo que las mujeres cuentan con menos cotizaciones y jubilan a menor edad, por lo que el valor de su pensión se correlaciona con su sexo (CEPAL, 2016).

- ii. **Variables empíricas:** variables presentes en las VIII EPF que muestran alta correlación con el ingreso y significancia como predictor de modelos de ingresos. Se evaluaron distintas variables que pudiesen tener correlación con el ingreso, entre las cuales fueron seleccionadas las siguientes: CIUO²⁰³, Clasificación Internacional de la Situación en el Empleo (CISE), Sustentador principal y Presencia de menores de 6 o 15 años en el hogar. Para el caso de las personas que reciben ingresos por jubilación y/o pensiones de vejez se incorporó como variable el sistema previsional al cual estaba afiliada la persona.

A partir de las variables seleccionadas, al igual que en la imputación de gasto, se definieron distintos conjuntos de individuos, agrupados según sus valores en las variables definidas (clúster de imputación). Dado que la búsqueda de individuos similares en torno al criterio

202 Se puede encontrar una breve explicación en la forma de cotización SII – SPENSIONES:

http://www.sii.cl/portales/renta/2016/cotizacion_previsional.htm

<https://www.spensiones.cl/portal/orientacion/580/w3-propertyvalue-9358.html>

203 Se considera que la Clasificación Internacional Uniforme de Ocupaciones (CIUO), estandariza ocupaciones como variable categórica; además, cuenta con un componente teórico, ya que existe evidencia de la relación del puesto de trabajo y el ingreso.

de agrupación de todas las variables definidas genera un proceso exigente de similitud, a la vez, se implementaron reglas para establecer un orden de flexibilización de las variables utilizadas para la generación de los grupos. Para realizar este proceso, se consideraron tres elementos importantes, que además ayudaron a priorizar la construcción de los niveles de exigencia de los grupos (respecto a lo cual se crearon matrices de exigencia): 1) variables correlacionadas con el ingreso, 2) variables que explicaban la estructura del ingreso y 3) variables relacionadas con la propensión a responder.

La definición del orden en que las variables serían relajadas y agrupadas en las matrices de exigencia fue realizada considerando su nivel de correlación con el ingreso, su criterio de inclusión (variables teóricas y empíricas) y el criterio de eliminación de variables. Para generar la matriz de exigencia, las variables se agrupaban y se relajaban hasta que se eliminaba aquella variable menos correlacionada. Una vez que esto ocurría, se volvían a desagrupar las variables para comenzar a aplicar el proceso nuevamente (sin la variable eliminada). Entre el anexo 16 y el anexo 20 se puede consultar por las matrices de exigencia definidas para la imputación de ingresos.

Respecto a la selección del informante donante de datos, esta fue aleatoria dentro del nivel en el que se categorizaba²⁰⁴ a los distintos individuos con información faltante.

El método utilizado es similar al implementado en la versión anterior de la encuesta, sin embargo, el planteamiento de selección aleatoria y el criterio de flexibilización de variables fueron actualizados según las pruebas realizadas para la VIII EPF. En la siguiente tabla se pueden observar los estadísticos de distribución para las fuentes de ingresos desagregadas que se imputaron.

Tabla 57: Resultados muestrales del ingreso bruto promedio por persona según fuentes de ingreso imputadas por método *Hot Deck* aleatorio

Por fuente de ingreso	N° Obs.	Promedio	Mínimo	Máximo	SD	CV	p25	p50	p75
Sueldos y salarios	14.794	803.657	0	29.250.000	992.733	1,24	315.000	500.000	885.396
Imputado HD	16.292	798.758	0	29.250.000	980.752	1,23	318.000	500.000	879.000
Empleadores	465	1.685.414	0	15.000.000	1.867.476	1,11	500.000	1.200.000	2.000.000
Imputado HD	505	1.657.239	0	15.000.000	1.835.391	1,11	490.000	1.180.000	2.000.000
Cuenta Propia	5.071	416.196	0	16.308.604	703.685	1,69	100.000	250.000	455.387
Imputado HD	5.453	420.482	0	16.308.604	709.320	1,69	100.000	250.000	469.950
Jubilados	5.230	278.328	789	5.672.000	334.783	1,20	109.490	167.809	280.000
Imputado HD	5.415	278.361	789	5.672.000	333.859	1,20	109.220	167.867	280.000

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

204 Al igual que en la metodología de gasto, se estableció una semilla de imputación para poder hacer el proceso replicable.

En la tabla 57, se evidencian los principales puntos de la distribución, antes y después de la imputación *Hot-Deck*, en cada fuente de ingresos que fue imputada. En general, el rango intercuartil y la mediana presentan leves cambios con la imputación. De mismo modo, los coeficientes de variación se mantienen estables después del proceso de imputación realizado.

Se observa que los comportamientos de los estadísticos descriptivos estudiados no presentan grandes variaciones entre los datos observados e imputados. Esto se debe, en parte, a que se está imputando dentro de rangos aceptables de no respuesta, tal como indica la tabla 56 (Medina & Galván, 2007).

Finalmente, en la tabla 58 se presentan los resultados de la imputación de ingresos, en los que se compara la estructura del ingreso disponible promedio del hogar con y sin imputación según fuentes de ingreso.

Al comparar la estructura del ingreso disponible promedio, se observa que la imputación realizada lo incrementa en un 8,85%. Específicamente, la imputación produce un aumento de un 10,23% en los ingresos de la ocupación y de un 3,84% en los ingresos por jubilaciones o pensiones de vejez. No obstante, se estima que la imputación en estas partidas no altera en demasía la estructura, considerando que el peso relativo de cada partida es similar al presente en la estructura del ingreso sin imputar.

Tabla 58: Resultados de la imputación de ingresos en la estructura del ingreso promedio del hogar por tipo de fuentes

Fuentes de ingreso	Estructura del ingreso disponible			
	Observada	%	Imputada por <i>Hot-Deck</i> aleatoria	%
Total	1.096.407	100	1.193.456	100
De la ocupación	918.847	83,81	1.012.888	84,87
Sueldos y salarios	688.108	62,76	766.187	64,2
Empleadores	41.935	3,83	46.065	3,86
Cuenta propia	143.259	13,07	155.089	12,99
Ingresos de otros trabajos	45.546	4,15	45.546	3,82
De otras fuentes	177.559	16,19	180.568	15,13
Jubilaciones o pensiones de vejez	78.380	7,15	81.389	6,82
Rentas de la propiedad ⁽¹⁾	39.203	3,58	39.203	3,29
Transferencias netas ⁽²⁾	36.691	3,35	36.691	3,07

(1) Incluye arriendos e intereses

(2) Se consideran transferencias netas, al descontar las emitidas de aquellas recibidas. Las transferencias monetarias provienen y se emiten de diferentes fuentes como el Estado, Instituciones Sin Fines de Lucro (ISFL) y otros hogares.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

9.4.4.1.4 Imputación de la estructura del gasto en servicios financieros

Tal como se explicó en el apartado dedicado a la descripción de los instrumentos, la LSF se componía de ocho módulos. La mayoría de ellos, a su vez, se encontraban subdivididos en dos componentes: 1) caracterización del crédito, en términos de monto solicitado, plazo, cuota e institución que otorga el crédito, y 2) captura de los gastos en servicios financieros, donde se consulta sobre intereses, comisiones, cobros de administración y seguros asociados. Para responder el primer set de consultas (zona blanca de los módulos de la LSF), los informantes no requerían tener comprobantes, mientras que para responder el segundo set de preguntas (zona ennegrecida de los módulos de la LSF), se requería que los informantes tuviesen acceso a comprobantes²⁰⁵.

Cuando no se tenía información directa sobre los gastos en servicios financieros para créditos (zona ennegrecida), se contaba con la caracterización del crédito (zona blanca), lo que permitió la estimación del gasto financiero total de este. Con esa estimación era posible hacer una desagregación entre los componentes de los gastos en servicios financieros, a saber, 1) gasto en interés, 2) comisiones, 3) cobros de administración y 4) seguros asociados (en otras palabras, imputar la estructura del gasto). Para ello se utilizó la estructura de los componentes, presente en los registros con información completa, según institución financiera (bancaria, no bancaria, comercial, automotora, etc.). La estimación de los gastos financieros de acuerdo a esta metodología se realizó solamente para los créditos hipotecarios, créditos de consumo y automotrices, en los cuales se contaba con información suficiente para generar esta desagregación.

Es importante destacar que los créditos hipotecarios, de consumo y automotrices “donantes” para la imputación (que podían donar datos a los registros sin información desagregada) debían estar completamente caracterizados, es decir, debían contar con la información respecto a la institución financiera, monto, plazo y cuota.

Los casos capturados para los que se contó con información de la caracterización del crédito, pero no se capturó información desagregada de los gastos en servicios financieros, correspondieron al 51,8% en créditos hipotecarios y 81,8% en créditos de consumo²⁰⁶. Para estos casos, se llevó a cabo la estimación del gasto financiero total y la desagregación de este. Su distribución se puede observar en la tabla 59. Si bien el porcentaje de créditos con información faltante respecto a su estructura es elevado, se cuenta con información para estimar el gasto y solo se está imputando la desagregación.

Tabla 59: Estadística de la no respuesta parcial del gasto financiero desagregado, a nivel créditos

Tipo de crédito	N° obs.	Datos faltantes		N Teórico
		n	%	
Crédito hipotecario	447	480	51,8	927
Créditos de consumo	395	1.776	81,8	2.171

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

205 Los dos módulos que capturaron el gasto, considerando sólo el primer set de preguntas, fueron los módulos de 1) Cuenta Vista y/o Cuenta RUT y 2) Seguros.

206 La cantidad de créditos a los que equivalen los anteriores porcentajes son 479 y 1.776, respectivamente.

Uno de los principales elementos a destacar es el cambio de la estructura posterior a la imputación (tabla 60). Principalmente, cambia el peso relativo de los gastos en comisiones y gastos de administración. Esto ocurre debido a la antigüedad de los créditos imputados, ya que, si el crédito fue solicitado previamente a 2015, no corresponde imputar gastos en administración a la desagregación del gasto financiero total del crédito, ya que aquellos se cobran en el primer año. De este modo, en aquellos casos para la imputación de la estructura de gasto financiero total del crédito, este se distribuye entre interés, comisión y seguros asociados. A continuación, se presenta la estructura del gasto financiero total, con y sin imputación, de acuerdo a la desagregación del mismo.

Tabla 60: Resultados de la imputación de la desagregación del gasto financiero

Desglose del gasto	Gasto financiero total del crédito	
	Observada	Posterior a imputación
Crédito hipotecario	100	100
Interés	83,5	82,8
Comisión	4,6	1,7
Administración	2	4,7
Seguros asociados	9,9	10,7
Créditos de consumo	100	100
Interés	82,2	78,7
Comisión	6,3	7,9
Seguros asociados	11,4	13,4

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

9.4.4.2 Ajuste temporal (ajuste al mes de referencia)

Dado que los instrumentos de recolección de información tienen distintos períodos de referencia²⁰⁷, los gastos se deben ajustar temporalmente para llevar todos los gastos a la misma unidad temporal. En el caso de la EPF, los gastos se mensualizan, es decir, se calculan a la escala de un mes tipo.

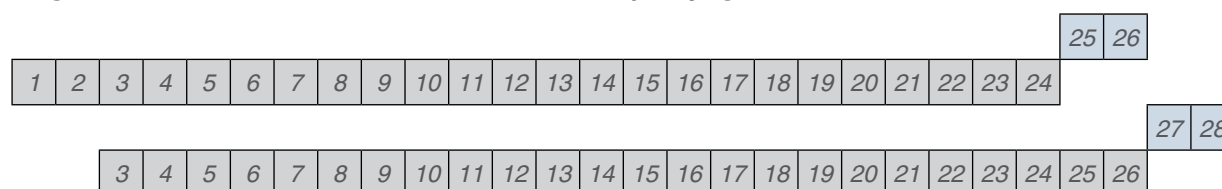
Para comprender la manera en la que se lleva a cabo el ajuste temporal, además de tener en consideración el período de colaboración de los hogares, es útil pensar en el período de recuerdo de la LGR. Ello da como resultado un total de 48 quincenas: 24 de trabajo de campo y 24 de recuerdo.

A modo de ejemplo, en el diagrama 14 las quincenas 25 y 26 corresponden al primer mes de trabajo de campo oficial (julio de 2016), mientras que las submuestras 1 a la 24 dan cuenta del período de recuerdo de la LGR. Análogamente, las quincenas 27 y 28 representan al

²⁰⁷ Gastos individuales, 15 días; gastos del recuerdo, 3, 6 o 12 meses hacia atrás; gastos en servicios financieros y seguros, el mes anterior; gastos del hogar, mensuales o con otras frecuencias como trimestrales, semestrales, etc.

segundo mes de recolección (agosto de 2016), mientras que las submuestras 3 a la 26 corresponden al período de recuerdo de un año.

Diagrama 14: Periodos de referencia de los meses de julio y agosto



Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Para la **primera submuestra operativa** del trabajo de campo (julio de 2016), se tienen los siguientes períodos de recuerdo:

- i. 3 meses: quincena 19 – 24
- ii. 6 meses: quincena 13 – 24
- iii. 12 meses: quincena 1 – 24

Para la **segunda submuestra** del trabajo de campo (julio de 2016) se tienen los siguientes períodos de recuerdo:

- i. 3 meses: quincena 19 – 25
- ii. 6 meses: quincena 13 – 25
- iii. 12 meses: quincena 1 – 25

9.4.4.2.1 Ajuste temporal de la LGI

En el caso de la LGI, el factor de ajuste temporal es igual al total de días del mes de trabajo de campo, dividido por el número de días de la quincena, lo cual da como resultado un número cercano a 2.

$$Ajuste = \frac{n^{\circ} \text{ días del mes}}{n^{\circ} \text{ días de la quincena}}$$

La idea detrás de esto es que la quincena de trabajo de campo captura la mitad del gasto que un hogar realiza en un mes. Por ende, es necesario multiplicar dicho gasto por un factor cercano a 2.

9.4.4.2.2 Ajuste temporal en LGH

En la libreta de gastos del hogar existen preguntas que implícitamente contemplan una periodicidad mensual y otras en las que es posible que los informantes registren una periodicidad específica. Ello genera tres ajustes temporales diferentes:

- i. Gastos consultados de manera mensual: no se requiere ajuste
- ii. Gastos en educación: suma del gasto realizado en el año, dividido por 12.

$$Ajuste = \frac{\text{valor del gasto } n^{\circ} \text{ veces que realiza el gasto}}{12}$$

- iii. Gastos con otras periodicidades: valor del gasto, dividido por el número de meses que abarca la boleta.

$$Ajuste = \frac{\text{valor del gasto}}{n^{\circ} \text{ de meses que abarca la boleta}}$$

9.4.4.2.3 Ajuste temporal de la LGR

El ajuste temporal de la LGR es igual al número de días del mes de trabajo de campo, dividido por el número de días del período de recuerdo. El resultado se utiliza como un multiplicador de los gastos.

$$Ajuste = \frac{n^{\circ} \text{ días del mes}}{n^{\circ} \text{ días de recuerdo}}$$

Teniendo en consideración los períodos de recuerdo mostrados anteriormente, a modo de ejemplo, se presentan los ajustes temporales para la primera submuestra operativa:

- i. **3 meses:** el factor es igual a 0,34 ya que el numerador sería 31 que corresponde al número de días que tiene julio, dividido por la suma de los días comprendidos entre las quincenas 19 a la 24, que da como resultado 91.
- ii. **6 meses:** el factor es igual a 0,17, ya que el numerador se mantiene en 31, pero ahora es dividido por la suma de los días comprendidos entre las quincenas 13 a la 24, que lo que da como resultado 181.
- iii. **12 meses:** el factor es igual a 0,085, siendo el numerador 31, y dividido por la suma de los días comprendidos entre las quincenas 1 a la 24, que da como resultado 365.

De acuerdo con esta metodología de cálculo, se obtienen ajustes cercanos a 3, 6 y 12. Cabe señalar que para la segunda quincena de cada mes, la suma de días es un poco mayor que para la primera. El efecto de ello es que el gasto se divide por un número mayor, lo cual compensa el hecho de que en la segunda quincena las personas registren más gastos, como resultado de tener un período de recuerdo más largo (las primeras dos semanas del mes).

En el anexo 21 se presentan los algoritmos de ajuste al mes de referencia.

9.4.4.3 Creación de nuevas variables

Esta etapa del proceso tiene como objetivo la generación de variables que consoliden la información capturada respecto al gasto e ingreso de los hogares. Estas variables, que no fueron capturadas de forma explícita en los cuestionarios que componen el estudio, son presentadas en las publicaciones de la encuesta. En este subproceso se generan variables a nivel hogar, derivadas de unidades más desagregadas de recolección (persona, producto).

Para la variable de gasto, se debe considerar que la información fue capturada en cinco de las seis libretas EPF, según la temática del gasto consultado (de acuerdo a lo especificado en el capítulo 5). En esta etapa, la información contenida en diferentes niveles de información (registros de compra, hogares, personas e instrumentos financieros) fue transformada a una unidad de información común, dejando una observación por hogar para cada producto

o servicio en que se realizó gasto. En este subproceso se generan las variables asociadas al gasto por hogar (que es la unidad de análisis de la encuesta) ²⁰⁸.

En el caso de los ingresos, fue necesario también generar variables nuevas, derivadas de la aplicación de cálculos a partir de las variables de captura de información. Así, por ejemplo, en esta etapa se calcularon las ganancias promedio mensuales asociadas a la tenencia de cuentas de ahorro, depósitos a plazo, fondos mutuos, Ahorros Previsionales Voluntarios, entre otros instrumentos financieros. Para la creación de estas variables se siguieron los siguientes pasos: 1) el monto registrado por el informante se dividió entre doce; 2) el monto anterior se multiplicó por la tasa de interés de captación del Banco Central, asumiendo que el interés se devenga una vez al año y manteniendo la unidad monetaria, y 3) el monto anterior se divide entre doce, para obtener la ganancia mensual estimada.

Una vez que se generaron las mensualizaciones de los distintos componentes del ingreso en distintas variables derivadas, se generaron nuevas variables a nivel persona y a nivel hogar, asociadas al ingreso de estas unidades, las cuales fueron incluidas en las bases publicadas.

Es importante mencionar que, dado que el GSBPM no tiene una estructura lineal estricta, las nuevas variables de ingreso aquí detalladas son utilizadas en los procesos de imputación de gastos descritos en el apartado 9.4.4.1.

En esta etapa, también se crean las variables necesarias para la trazabilidad de la imputación de ingresos, haciendo la distinción entre los casos observados y aquellos que fueron imputados. Estas variables se crearon en función de la información observada y, en el caso de los imputados, de aquellos montos que corrigen la no respuesta parcial, de acuerdo al método descrito en el apartado 9.4.4.1. En la Base de Personas publicada, el usuario podrá diferenciar los datos observados, de aquellos que fueron imputados.

9.4.5 Cálculo de ponderadores (factor de expansión)

En este subproceso se calculan los ponderadores de la encuesta, que hacen posible la expansión de los datos muestrales a la población objetivo. El subproceso de cálculo de ponderadores se expone en el apartado 6.2, sobre desarrollo de los factores de expansión, mientras que las medidas de precisión son descritas en el apartado 6.3.6 del presente documento.

9.4.6 Consolidación de las bases de datos y cálculo de agregaciones

Este subproceso consta de dos etapas. En primer lugar, se consolidan las bases de gastos y personas y, en segundo lugar, se lleva a cabo el cálculo de agregaciones (gasto promedio por hogar y gasto promedio por división, entre otros).

Respecto a la primera etapa, esta se compone de dos tareas, a saber, distribuir los gastos codificados en la división 15 (división residual) y realizar el proceso de quintilización de los hogares según su nivel de ingreso.

²⁰⁸ En el subproceso anterior se realizó el ajuste temporal de los gastos, con el fin de llevarlos a un período común de referencia mensual.

Respecto a la segunda etapa, en ella se lleva a cabo el cálculo de agregaciones (gasto promedio por hogar y gasto promedio por división, entre otros) y se aplican las medidas de precisión (coeficientes de variación, intervalos de confianza, etc.), trabajo necesario para el posterior análisis de los datos. Esta segunda etapa es desarrollada en el apartado 6.3, sobre estimadores y medidas de precisión de la encuesta y en el Informe de principales resultados VIII EPF (INE Chile, 2018e).

A continuación, se describen las tareas asociadas a la etapa de consolidación de bases de datos.

9.4.6.1 Distribución de la división 15

Como se mencionó en el apartado 9.3.4.2, para la codificación de gastos de la VIII EPF, se trabaja con una división residual en el clasificador CCIF, la división 15 (ver tabla 50).

Los registros clasificados dentro de los primeros 12 códigos de esta división corresponden a bienes y servicios no desglosados de cada una de las 12 divisiones del clasificador. Por su parte, existen códigos que agrupan gastos sin desglosar de 2 divisiones. Así, por ejemplo, un código reúne gastos agregados de las divisiones 1 y 2, mientras que otro hace lo propio con las divisiones 1 y 11.

La metodología para distribuir estos gastos contempla procedimientos diferenciados según el código y la libreta. Para los códigos agregados de una y dos divisiones se utilizan las siguientes pautas como metodología de desagregación:

- i. LGI: se construyen ponderadores de acuerdo a la participación que cada división tiene en el gasto promedio de los hogares, considerando únicamente los registros de las libretas de gastos individuales, es decir, excluyendo del cálculo al resto de las libretas. A su vez, para los códigos que reúnen 2 divisiones se calcula un ponderador que da cuenta del peso relativo de las 2 divisiones que componen el agregado. Posteriormente, se distribuyen estos gastos agregados de acuerdo a los ponderadores definidos previamente. Por su parte, los agregados de más de 2 divisiones se reparten entre los códigos no desglosados, según su peso relativo en la libreta. En la base final, cada uno de los agregados pasa a formar un grupo en la división correspondiente, con el nombre de “Gastos no desglosados en...”.
- ii. Resto de libretas: se construyen ponderadores de acuerdo a la participación que cada división tiene en el gasto promedio de los hogares, pero a diferencia del caso anterior, se utiliza el total de libretas para realizar el cálculo. Una vez hecho esto, se procede del mismo modo que en el caso de la LGI para distribuir el gasto.

Para el código 15.1.1.01.26 (gastos en bienes y servicios no desglosados) se lleva a cabo un procedimiento diferenciado según libreta:

- i. LGI y LI: se distribuye en las 12 divisiones
- ii. LGH: se distribuye en las divisiones 3, 6 y 12.
- iii. LGR: Se distribuye en todas las divisiones, a excepción de la 6 y la 7.

9.4.6.2 Quintilización

Un modo de lectura y análisis frecuente de la información de las encuestas de hogares que capturan ingreso es a través de la construcción de quintiles.

Los quintiles corresponden a una variable de orden que indica en qué 20% de la distribución de los ingresos (del hogar o *per cápita*) se encuentra el hogar encuestado. El orden se genera de forma ascendente en relación con el ingreso respecto al que se construyen. Cabe aclarar que no corresponde a una variable de estratificación socioeconómica, ya que para lograr una estratificación se deben incorporar al análisis otras dimensiones. El primer quintil de ingresos representa al 20% de hogares de menores ingresos, mientras que el quinto quintil representa al 20% de los hogares de mayores ingresos.

En el caso de la VIII EPF, en el conjunto de tabulados oficiales se incorpora la quintilización de los hogares, tanto por ingreso disponible del hogar como por ingreso disponible *per cápita* promedio mensual del hogar para los diferentes niveles geográficos de estimación.

Para la construcción de quintiles que denoten un ordenamiento de acuerdo a la realidad nacional, se genera un nuevo vector de ingresos que, en la mayoría de los casos, corresponde a la variable de ingreso elegida (disponible del hogar o *per cápita*). No obstante, considerando las particularidades de esta encuesta respecto a desajustes entre ingresos y gastos²⁰⁹ (OCDE, 2013; OIT, 2003c; Brewer & O'Dea, 2012), a los hogares ubicados en el primer percentil de ingresos (sin considerar a aquellos hogares que declararon no haber recibido ingresos)²¹⁰, para efectos de la quintilización, se les imputa al nuevo vector de ingresos el valor del gasto total del hogar o *per cápita*, según corresponda.

Los desajustes de gasto e ingreso para el grupo descrito podrían deberse a una eventual subdeclaración de ingresos. Si bien esta situación se genera a lo largo de toda la distribución de ingresos, para la estimación de los indicadores principales se genera un mayor efecto en los quintiles más bajos, lo que podría alterar la estructura de la canasta de bienes y servicios del IPC²¹¹. Por esta razón, se considera necesario construir la quintilización a partir de un nuevo vector de ingresos con las características antes descritas.

209 Dado que la EPF considera distintos períodos de referencia para la captura de información de gastos e ingresos, es habitual encontrar desbalances entre estas variables.

210 En el caso de la VIII EPF, este monto se fijó en \$80.000.

211 Ese efecto podría ser mayor en la estimación de la canasta que conforma IPC, ya que se asignaría un consumo mayor a en los quintiles más bajos.

10. METODOLOGÍA DE ANÁLISIS

Esta fase busca interpretar, explicar y validar con fuentes externas los resultados oficiales obtenidos en la encuesta. Su principal producto son las bases de datos totalmente validadas y listas para ser publicadas. Adicionalmente, en esta fase se elaboran una serie de documentos metodológicos que acompañan la publicación oficial del material de la EPF.

Cabe aclarar que las fases del procesamiento y análisis pueden ser iterativas y paralelas (UNECE, 2016). El análisis puede revelar mayor entendimiento de los datos, lo que puede llevar a notar que se requiere procesamiento adicional de estos. En el caso de la VIII EPF, las actividades que conforman las fases de procesamiento y análisis comienzan antes de que la fase de recolección termine. Esto permite la compilación de resultados provisionales para los equipos relacionados con la actualización de la canasta de bienes y servicios del IPC.

La fase de análisis de la VIII EPF estuvo compuesta por los siguientes seis subprocesos:

- i. **Análisis descriptivo:** consistió en explorar los datos a nivel agregado, con el objeto de detectar comportamientos no esperados en las distribuciones de gasto e ingreso. Tal como se ha señalado en apartados anteriores, el flujo de trabajo consideró la posibilidad de volver a etapas anteriores. Esto fue de gran relevancia en este subproceso, ya que era la última instancia de mejorar la calidad de los datos antes de la publicación. En caso de detección de problemas, las bases podían volver al subproceso de depuración, para ser revisadas. A partir de ello podían ocurrir dos situaciones: se constataba la existencia de un error y se modifican los datos o se determinaba que no existían errores, de modo que los datos se validaban. En ambas situaciones, la base volvía a la fase de análisis para continuar con el flujo de trabajo.

En el caso de los ingresos, se realizó un análisis detallado de los hogares que registraron ingresos bajos (posible subdeclaración), con el objeto de encontrar algún error o explicación para los mismos. Una vez realizada esta tarea, se analizó el ingreso promedio por hogar y la distribución del ingreso según fuentes (trabajo asalariado, independiente, ingresos de la propiedad), con el fin de identificar si existían comportamientos inconsistentes en los datos.

Respecto al gasto, se analizó el gasto promedio por hogar y por libreta, observando su distribución según división, según libreta y división, según división y quintil de ingreso (*per cápita* y del hogar) y según grupos de la CCIF. De este modo, se buscaba detectar problemas con los resultados de las divisiones de gasto.

Adicionalmente, se observó cómo se distribuían el ingreso y el gasto según ciertas variables consideradas relevantes, tales como educación, sexo, ocupación, jefatura de hogar, administrador de gastos y sustentador principal.

- ii. **Comparación con fuentes externas:** con el objeto de validar los resultados, se contrastaron los resultados de la encuesta con información proveniente de fuentes externas. Que los datos de gasto e ingreso entregaran información consistente con los resultados de dichas fuentes de comparación fue interpretado como una señal de coherencia de la encuesta con otras estadísticas oficiales del país (INE

Chile, 2015b). Para la comparación se utilizaron fundamentalmente las siguientes fuentes:

- a. **VII EPF:** la VII EPF es la comparación más cercana, en el sentido de que la metodología empleada en ambas versiones es muy similar. Un aspecto fundamental a considerar fue la identificación de los cambios ocurridos en la distribución de gasto por división. Se debía determinar si estos eran producto de los ajustes metodológicos o de modificaciones reales en el patrón de consumo de los hogares.
- b. **Cuentas Nacionales:** dado que el marco conceptual de la EPF es el Sistema de Cuentas Nacionales, esta es una fuente de comparación importante para la encuesta. En esta comparación era importante evaluar las diferencias de los resultados de la VIII EPF respecto a la distribución de gasto por división entregada por Cuentas Nacionales y explorar sus causas.
- c. **CASEN/ESI:** en lo relativo a ingresos, las fuentes de contraste utilizadas fueron Casen y ESI. Se puso especial atención a posibles sesgos que pudiesen haber en el ingreso promedio por hogar. Además, se compararon los ingresos según su fuente: trabajo asalariado, independiente, ingresos de la propiedad, etc. Dado que las metodologías de recolección y procesamiento no son idénticas en las tres encuestas, los resultados podían diferir, sin embargo, lo esperable era que no existiesen grandes diferencias a nivel de estimaciones agregadas.
- d. **Otras fuentes de datos nacionales:** se utilizaron fuentes de otras encuestas y registros administrativos nacionales para analizar resultados de la encuesta. Ejemplo de ello son las bases de establecimientos de Mineduc, para comparar la estructura de estudiantes de la población encuestada por dependencias administrativas de los establecimientos educacionales, ventas de automóviles registrados por la Asociación Nacional Automotriz de Chile (ANAC), entre otras.
- e. **Encuestas de ingresos y gastos de otros países:** en cuanto a gastos, se comparó la estructura de gastos con encuestas de ingreso y gastos de otros países en base a dos grupos: la VIII EPF comparada con otros países de la región (Uruguay, Colombia y México) y la VIII EPF comparada con países de otras regiones al momento de tener un PIB similar al Chile actual y que presentaran cierta cercanía cultural con el país. (España y Portugal).

Cabe señalar que, al igual que en el análisis descriptivo, la base podía volver a la fase anterior si se observa un comportamiento inesperado en los datos.

- iii. **Interpretación de datos:** su propósito fue interpretar la información y reunir insumos para la publicación oficial. Teniendo en consideración los análisis previos, el subproceso de interpretación debía dar cuenta de los cambios experimentados por el patrón de consumo de los hogares en Chile. En ese sentido, resultaba importante discriminar entre los cambios que se explicaban por modificaciones metodológicas y aquellos que realmente se producían por cambios en los patrones de consumo de los hogares. Adicionalmente, se identificaron las variables que estaban asociadas a diferencias en los patrones de consumo y se caracterizó a los hogares.

- iv. **Preparación de tabulados:** una vez que las bases han sido completamente validadas, se elaboraron los tabulados seleccionados en el plan de tabulación para la publicación oficial de la encuesta. En este subproceso se produjo una parte importante de los resultados, alojados en el micrositio de la encuesta. Algunos de los tabulados más relevantes son: gasto promedio por hogar, ingreso promedio por hogar, participación del gasto e ingreso según quintiles, gasto desagregado según división, gasto desagregado según los niveles funcionales de la CCIF, entre otros. Cabe señalar que en el Manual de Usuarios (INE Chile, 2018d) se proporcionan indicaciones respecto al método de cálculo de los principales tabulados de la encuesta.
- v. **Preparación de la publicación y controles de difusión:** a partir de las conclusiones extraídas en los subprocesos anteriores, se elaboraron los documentos oficiales que acompañaron la publicación de las bases de datos. En este subproceso se llevó a cabo la anonimización de los datos. Al respecto, cabe señalar que la VIII EPF, al igual que todos los productos estadísticos producidos por el INE, basa sus mecanismos de confidencialidad en la Política de Protección de Datos Personales institucional. Esta política establece que los datos producidos por el Instituto no pueden hacer referencia expresa a personas que participen en sus estudios. En consecuencia, se excluyeron de las bases finales aquellas variables que permitían identificar unidades de análisis (personas, viviendas) que participaron en el estudio o que implicaran un riesgo elevado de identificación de las mismas (variables de ubicación de manzana por ejemplo). Adicionalmente, se modificó las bases de datos de acuerdo con las áreas de estimación determinadas en el diseño muestral.
- vi. **Cierre de resultados:** tuvo como propósito verificar la consistencia de los cinco subprocesos anteriores y corroborar que la información estuviera en condiciones de ser difundida. En este subproceso se realizaron tareas de revisión de consistencia de los datos, introducción de advertencias de uso y revisión de resultados en instancias técnicas internas a la institución, entre otras actividades.

11. GLOSARIO

Adaptación de un clasificador: se habla de adaptación cuando un clasificador internacional (o de referencia) se modifica a nivel de estructura, para que sus categorías queden acorde a la realidad nacional (o a alguna realidad específica).

Adopción de un clasificador: se habla de adopción cuando se utiliza directamente el clasificador internacional (o de referencia), sin realizar modificaciones a nivel de estructura.

Ajuste temporal: algoritmo mediante el cual los gastos de las distintas libretas son transformados a la unidad temporal mes.

Arriendo imputado: corresponde a la estimación del gasto en arriendo que debería realizar un hogar si no fuera propietario de su vivienda o si esta no le fuera cedida de forma gratuita.

Bien o servicio de consumo: corresponde a un bien o servicio utilizado por los hogares para la satisfacción directa de sus necesidades.

Clasificación de Consumo Individual por Finalidades (CCIF): corresponde a una clasificación funcional del SCN que estandariza los bienes y servicios capturados por la encuesta. Se utiliza este clasificador desagregado hasta nivel de producto (quinto nivel de desagregación). Los primeros tres niveles del clasificador se adoptaron según el clasificador internacional (hasta nivel de clase), mientras que los siguientes dos niveles fueron adaptados a la realidad nacional (subclase y producto). La adaptación fue realizada por el INE con el apoyo de instituciones externas.

Clasificador: es una estructura jerárquica que agrupa los sucesos de un fenómeno según determinados criterios, estableciendo categorías de clasificación únicas e irrepetibles. Además, poseen atributos definitorios que son similares entre ellos (internamente homogéneos), pero que, al mismo tiempo, permiten diferenciarlos de las demás categorías (externamente heterogéneos).

Clasificación Internacional Uniforme de Ocupaciones (CIUO): estandariza ocupaciones, empleándose a un nivel de desagregación de subgrupos principales (dos dígitos). Se utiliza el clasificador de 1988 y una adaptación nacional de su versión 2008.

Codificación Automatizada: corresponde a la asignación de códigos a través de procesos informáticos estandarizados, en base a información presente en los cuestionarios.

Código de Disposición Final: código que representa el estado final del proceso de contacto con las unidades muestrales.

Confiabilidad: la fiabilidad se refiere a la proximidad del valor verdadero del parámetro que se desea estimar con los valores estimados posteriormente (OCDE, 2007). Se relaciona con la probabilidad de que el verdadero valor del parámetro estimado en la población se sitúe próximo a la estimación de un estudio (Eurostat, 2014).

Conciliación de resultados de reentrevista: procedimiento mediante el cual se resuelven las discrepancias entre la información registrada por el encuestador en las libretas y la obtenida por un supervisor mediante supervisión indirecta.

Diseño muestral: corresponde a la elección del tipo de muestreo y tamaño muestral a utilizar. Ambos son elegidos considerando el cumplimiento de los objetivos de la encuesta.

En el caso de la VIII EPF se ha utilizado un diseño muestral probabilístico, estratificado y bietápico.

Encuesta de Presupuestos Familiares (EPF): encuesta socioeconómica que busca identificar y caracterizar la estructura del gasto de consumo final de los hogares urbanos, en las capitales regionales de Chile y algunas de sus zonas conurbadas, con un período de referencia de un año.

Entrevistas interrumpidas (*break-off*): corresponde a una subcategoría del rechazo en la que se establece comunicación con los hogares y estos entregan parte de la información requerida, no obstante, los datos reportados no cumplen con los criterios de suficiencia para considerar la unidad como entrevistada.

Error muestral: proviene del hecho de entrevistar solo a una parte de la población y no a la población completa. Por tanto, todas las encuestas tienen un error muestral.

Error de estimación: la diferencia entre un valor estimado y el verdadero valor de un parámetro o, a veces, de un valor a predecir (OCDE, 2007), en relación con la confiabilidad de la estimación.

Error no muestral: este tipo de error agrupa a todas las fuentes que se originan por los problemas prácticos de realizar una encuesta. En esta categoría se incluyen errores de cobertura (ej.: información incompleta del proceso de enumeración de las viviendas), errores de medición (ej.: preguntas formuladas de manera ambigua) y errores de procesamiento (ej.: error al digitar o codificar datos), entre otros. Todos estos errores afectan la precisión y confiabilidad de las estimaciones que se realizan a partir de las encuestas.

Estimación: se refiere a la inferencia sobre el valor numérico de valores de población desconocidos a partir de datos incompletos, como una muestra (OCDE, 2007).

Estimación puntual: es una de las dos bases principales de estimación en el análisis estadístico. La estimación de puntos intenta dar el mejor valor estimado individual de un parámetro, en comparación con la estimación de intervalo, que procede especificando un rango de valores. Dado que la estimación puntual está rodeada por una banda de error, la distinción entre estos dos métodos a veces es difusa y en la interpretación a menudo equivale a la misma cosa (OCDE, 2007).

Estimador: corresponde a una regla, función o método para estimar un parámetro de una población (OCDE, 2007).

Factor de expansión: se interpreta como la cantidad de unidades en la población que representa una unidad en la muestra. De acuerdo con el diseño muestral, el factor de expansión se calcula en función del total de manzanas y viviendas seleccionadas en cada estrato, así como también del número de manzanas totales en el estrato (según información reportada por el Marco Muestral de Viviendas) y del número de viviendas particulares enumeradas.

Gasto en consumo final: gasto en bienes y servicios. Se excluyen los gastos en activos fijos en forma de viviendas o de objetos valiosos, así como los gastos en consumo intermedio. Estos últimos corresponden a bienes y servicios que los hogares adquieren para el desarrollo de sus actividades productivas. Solo aquellos bienes y servicios adquiridos para actividades no productivas forman parte de su gasto de consumo final.

Gasto de consumo adquirido: acto mediante el cual se adquiere el derecho de propiedad sobre un bien o el derecho a la prestación de un servicio, independiente de si el consumo de ese bien se efectúa durante el período de referencia de la encuesta.

Grilla técnica de calidad: corresponde a un método de evaluación de mínimos de calidad y cantidad de los datos entregados por cada hogar que participó en la VIII EPF, evaluando de esta forma la información entregada en cada instrumento. Estos requerimientos mínimos de calidad fueron definidos en concordancia con los objetivos de la encuesta, estableciendo distintas ponderaciones para cada libreta y sus preguntas prioritarias.

Hogar: una o más personas, con o sin lazos de parentesco, que comparten domicilio y satisfacen sus necesidades de alimentación y otros bienes esenciales a través de un presupuesto común.

Imputación de datos: método que completa los datos perdidos de una encuesta mediante algún procedimiento estadístico.

Indicadores de desempeño operativo: estadísticos para el monitoreo del trabajo de campo. Permite la evaluación del avance y son construidos a partir de la información disponible en los módulos Registro de contacto con el hogar y Control de Libretas de Gastos Individuales de la Hoja de Ruta.

Índice de Precios al Consumidor (IPC): es un indicador económico que mide la variación de los precios de una canasta de bienes y servicios representativa de los gastos de los hogares urbanos, cuya cobertura geográfica corresponde a todas las capitales regionales y sus zonas conurbadas dentro de las fronteras del país.

Informante idóneo: corresponde a un miembro del hogar, de 15 años o más de edad, que se encuentra en condiciones físicas o mentales para responder la encuesta. Si bien un informante de 15 años o más puede suministrar información para la encuesta, para autorizar la participación del hogar en el estudio, se requería siempre la autorización de un mayor de edad (de 18 o más años).

Ingresos del trabajo: ingresos generados por una actividad económica laboral. Estos incluyen ingresos del trabajo asalariado, ingresos del trabajo independiente y otros ingresos del trabajo.

Ingresos de la propiedad: corresponde al flujo de ingresos netos que se obtienen de tenencia de activos financieros. Incluye el retorno de dichos activos (interés ganado de los depósitos), recibos de ingresos por inversiones, anualidades y otros pagos regulares de seguros de vida y fondos privados de pensiones (excluidos de la seguridad social), además de las rentas netas recibidas por el arriendo de activos no financieros, tales como recursos naturales, tierras, edificios, equipos, patentes o derechos de autor.

Ingresos de autoprovisión de servicios producidos por el hogar: son aquellos ingresos asociados a la producción doméstica de servicios para el autoconsumo del hogar. Incluye el valor de la prestación del servicio de alojamiento en las viviendas ocupadas por sus propietarios (arriendo imputado), el valor de los servicios domésticos no remunerados y el valor de los servicios aportados por el consumo de bienes durables. Los ingresos provenientes de los servicios domésticos no remunerados y del consumo de bienes durables fueron excluidos de esta encuesta y han sido tradicionalmente excluidos de la medición Cuentas Nacionales (ONU, 2001).

Ingreso disponible del hogar (incluye arriendo imputado): corresponde a la sumatoria de todas las fuentes de ingreso del hogar que fueron percibidos en el mes anterior a la aplicación de la encuesta. Se excluyen las transferencias emitidas (se restan a los ingresos totales del hogar) e incluye el arriendo imputado para aquellos hogares propietarios de su vivienda o que usufructúan de ella y el arriendo imputado estimado por la vivienda secundaria (para los hogares con vivienda secundaria). Esta variable se encuentra expresada en pesos chilenos.

Ingreso primario: corresponde a la suma de los ingresos del trabajo, de la propiedad y de auto-provisión de servicios producidos por el hogar.

Ingreso total del hogar (incluye arriendo imputado): corresponde a la sumatoria de todas las fuentes de ingreso de todos los integrantes del hogar, percibidas en el mes anterior a la aplicación de la encuesta. Se calcula como la suma entre el ingreso primario y las transferencias corrientes recibidas (excluyendo las transferencias recibidas en especie).

Marco de muestreo: información cartográfica y del número de viviendas del cual es posible seleccionar una muestra. En el caso de la VIII EPF se utiliza el Marco Muestral Maestro de manzanas del INE.

Modelo Genérico del Proceso Estadístico (Generic Statistical Business Process Model - GSBPM): es el modelo genérico internacional asociado a procesos de producción estadística. El GSBPM describe y define el conjunto de procesos necesarios para producir estadísticas oficiales. Según la Comisión Económica de las Naciones Unidas para Europa el GSBPM

Provee un marco estándar y una terminología armonizada para ayudar a las organizaciones estadísticas a modernizar sus procesos de producción estadística, así como a compartir métodos y componentes. Además, el GSBPM puede ser usado para integrar estándares en datos y metadatos, como modelo para la documentación de procesos, para la armonización de infraestructuras de cómputo estadístico y para proveer un marco de referencia para mejoras y evaluaciones del proceso de calidad. (UNECE, 2016, pág. 4)

Muestra con reemplazo: tipo de muestreo que se utiliza para solucionar el problema de rechazo de la encuesta y que considera la selección de viviendas de reemplazo para ser utilizadas solo cuando alguna unidad rechaza la encuesta.

No respuesta: diferencias entre el valor estimado y el valor verdadero, debido a que no se produce información para todos los tópicos consultados en todas las unidades seleccionadas para la recolección de datos (Eurostat, 2014). De este modo, es posible distinguir entre la no respuesta total, es decir, la ausencia de información para una unidad muestral completa, y la no respuesta al ítem, producida cuando existe ausencia de datos para una o más de las variables consultadas en una unidad muestral.

No respuesta parcial o al ítem: respuesta que tiene lugar cuando las personas participan de las entrevistas, no obstante existen ciertas preguntas o temas que no están dispuestos a contestar. En la literatura, este tipo de no respuesta también se denomina no respuesta al ítem.

No respuesta total: unidades que rechazan participar en el estudio (ver definición de rechazo).

Población objetivo de la EPF: hogares (una o más personas) que residen habitualmente en las viviendas particulares ocupadas de las capitales regionales del país (incluyendo sus zonas conurbadas más importantes).

Porcentaje: un porcentaje es un tipo especial de proporción donde la razón se multiplica por una constante, 100, de modo que la relación se expresa por 100 (OCDE, 2007).

Precisión: aproximación de cálculos o estimaciones a los valores exactos o verdaderos que las estadísticas tenían la intención de medir (OCDE, 2007). Es por ello que cuanto más precisión se exige a la estimación de un parámetro, menor deberá ser la dispersión de las estimaciones de los indicadores principales de la encuesta (Vivanco M., 2005).

Proporción: una proporción es un tipo especial de razón en la que el denominador incluye el numerador. Un ejemplo es la proporción de muertes de hombres que se calcula considerando el número de muertes de hombres dividido por el total de muertes de hombres y mujeres (es decir, la población total) (OCDE, 2007).

Quintil de Ingresos: es un valor de posición de la variable de ingresos (del hogar o *per cápita*). Los cinco valores o quintiles de ingreso se determinan al ordenar todos los datos²¹² de acuerdo con sus valores observados en la variable de ingreso, desde aquel con menores ingresos al con mayores ingresos, para luego separar las unidades estadísticas en cinco grupos de igual tamaño. Cada grupo representa así el 20% de los valores observados de la variable ingreso en la muestra. Así los quintiles son los valores umbrales que separan estos grupos, y por extensión pasaron a designar los grupos que separan. El primer quintil se refiere entonces al 20% de las unidades con menores ingresos observados, el segundo quintil se refiere al 20% de las unidades siguientes en la escala de ingresos y el quinto quintil, al 20% de las unidades con mayores ingresos. De esa manera, los quintiles conforman una transformación de la variable ingresos de una variable de razón a una ordinal en cinco niveles. Es importante señalar que la quintilización de los ingresos no constituye en sí misma una variable de estratificación socioeconómica, ya que para lograr una estratificación socioeconómica se deberían incorporar al análisis otras dimensiones.

Rechazo: corresponde a aquellos hogares seleccionados para participar en el estudio, pero que declinan entregar información o bien entregan datos que no cumplen con los criterios necesarios para considerar aquellas unidades como entrevistadas.

Recuperación de datos: etapa de la fase de procesamiento que tiene por objetivo mejorar la calidad de los datos provenientes de las fases de digitación y de trabajo de campo, mediante revisión física de libretas, lectura de observaciones y recuperación en terreno.

Sistema de Cuentas Nacionales (SCN): “el Sistema de Cuentas Nacionales (SCN) es el conjunto normalizado y aceptado internacionalmente de recomendaciones relativas a la elaboración de mediciones de la actividad económica de acuerdo con convenciones contables estrictas, basadas en principios económicos” (ONU *et al.*, 2016, pág. 1)

Sobremuestreo: inclusión en la muestra total de una mayor cantidad de unidades estadísticas que la cantidad definida como muestra objetivo o teórica. Se utiliza para prevenir el impacto de la no respuesta en un estudio probabilístico y así lograr mayor consistencia entre muestra objetivo y muestra efectiva.

Submuestra operativa: conjunto de viviendas a encuestar en un período de aproximadamente 15 días, durante el cual los hogares participan en la encuesta.

²¹² Un dato estadístico es la dupla constituida de la unidad estadística y del valor asociado por la medición a esta unidad.

Supervisión directa: corresponde al acompañamiento de un supervisor a un encuestador en la aplicación de los instrumentos, con el objeto de observar en detalle la interacción de este con el entrevistado.

Supervisión indirecta: se refiere al contacto con los informantes sin la presencia del encuestador, y sin aviso previo a él, en el que se realiza una reentrevista de los aspectos principales de la encuesta que hayan sido aplicados hasta el momento.

Tasa: una tasa se refiere a la ocurrencia de eventos en un intervalo de tiempo específico (OCDE, 2007).

Tasa de respuesta: indica la proporción del total de unidades elegibles o de elegibilidad desconocida (incluido el sobremuestreo) donde se obtuvo datos suficientes para realizar cálculos oficiales. Esta tasa se obtiene dividiendo el número de hogares con información suficiente por la suma de las unidades elegibles y de elegibilidad desconocida.

Transferencias corrientes emitidas: son flujos monetarios emitidos por los hogares que incluyen impuesto a la renta, contribuciones a la seguridad social, pago por pensiones alimenticias, transferencias a hogares (dentro y fuera del país) y transferencias a instituciones sin fines de lucro que sirven a los hogares.

Transferencias corrientes recibidas: son flujos monetarios recibidos por los hogares que no generan relaciones de reciprocidad respecto a quienes las emiten. Estas incluyen los ingresos asociados a las jubilaciones, pensiones, otros beneficios en efectivo asociados a la seguridad social o al empleo, beneficios de asistencia social en efectivo del Gobierno, transferencias desde otros hogares (dentro y fuera del país), transferencias en efectivos recibidas desde instituciones sin fines de lucro que sirven a los hogares. Este tipo de transferencias excluye las transferencias sociales en especies que los Gobiernos y las instituciones sin fines de lucro que sirven a los hogares emiten a los hogares.

Validación de datos: subproceso de la fase del procesamiento de datos que comprueba que la información contenida en las bases provenientes de digitación cumpla con ciertos criterios que aseguren la calidad de los datos.

Valor outlier: en español, valor atípico. Corresponde a una observación que es numéricamente distante del resto de los datos. Las estadísticas derivadas de los conjuntos de datos que incluyen valores atípicos son frecuentemente engañosas.

12. BIBLIOGRAFÍA

- Aapor. (2015). Final dispositions of case codes and outcome rates for surveys. 8th Edition. Recuperado el 16 de Abril de 2015, de https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions2015_8theditionwithchanges_April2015_logo.pdf.
- ABS. (2011). Household Expenditure Survey. Summary of Results. Canberra.
- Andridge, R. R., & Little, R. (2010). A Review of Hot Deck Imputation for Survey Non-response. *International Statistical Review*, 78(1), 40-64.
- Andridge, R., & Little, R. (2009). The Use of Sample Weights in Hot Deck Imputation. *J Off Stat*, 25(1), 21-36.
- Araneda, J., Pinheiro, A., Rodríguez, L., & Rodríguez, A. (2016). Consumo aparente de frutas, hortalizas y alimentos ultraprocesados en la población chilena. *Revista Chile Nutrición*, 43(3). Obtenido de <https://scielo.conicyt.cl/pdf/rchnut/v43n3/art06.pdf>
- Bradburn, N. (2010). Recall Period in Consumer Expenditure Surveys Program. NORC/ University of Chicago. Obtenido de https://www.bls.gov/cex/methwrkshp_pap_bradburn.pdf
- Brewer, M., & O'Dea, C. (2012). Measuring living standards with income and consumption: evidence from the UK. Institute for Social and Economic Research. Obtenido de <https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2012-05.pdf>
- Brick, J. M. (2013). Unit nonresponse and weighting adjustments: A critical review. *Journal of Official Statistics*, 329-353.
- Budlender, D. (2003). The Debate about Household Headship. Obtenido de *Social Dynamics: A journal of African studies*: <http://www.statssa.gov.za/publications/DiscussHouseholdHead/DiscussHouseholdHead.pdf>
- CEPAL. (2011). Recomendaciones para la ronda de los censos de la década de 2010 en América Latina. Santiago: CEPAL. CELADE. Obtenido de http://repositorio.cepal.org/bitstream/handle/11362/55111/1/S1100525_es.pdf
- CEPAL. (2015). Informe sobre la Aplicación de la Nueva Metodología de Medición de la Pobreza en Chile 2006-2013. Comisión Económica para América Latina y el Caribe, División de Estadísticas. Obtenido de http://observatorio.ministeriodesarrollosocial.gob.cl/documentos/Cepal_Informe_Aplicacion_Nueva_Metodologia_Medicion_Pobreza_Chile_2006-2013.pdf
- CEPAL. (2016). La brecha de género en jubilaciones y pensiones. Los casos de Argentina, Brasil, Chile y Uruguay. Obtenido de Amarante, Verónica; Colacce, Maira; Manzi, Pilar: https://repositorio.cepal.org/bitstream/handle/11362/40650/2/S1600929_es.pdf
- CEPAL. (2016). Medición de la pobreza en Chile, Casen 2015. Obtenido de http://observatorio.ministeriodesarrollosocial.gob.cl/casen-multidimensional/casen/docs/Medicion_de_la_pobreza_en_Chile_2015.pdf
- CMP. (2014). Informe Final. Comisión para la Medición de la Pobreza. Santiago de Chile. Obtenido de http://www.cl.undp.org/content/dam/chile/docs/pobreza/undp_cl_pobreza_comision_pobreza_2014.pdf
- Crovetto, M., & Uauy, R. (2012). Evolución del gasto en alimentos procesados en la población del Gran Santiago en los últimos 20 años. *Revista Médica de Chile*, 140, 305-312. Obtenido de <https://scielo.conicyt.cl/pdf/rmc/v140n3/art04.pdf>

- Crovetto, M., & Uauy, R. (2014). Cambios en el consumo aparente de lácteos, bebidas azucaradas y jugos procesados en el Gran Santiago. Chile. 1987-2007. *Revista Médica de Chile*, 142, 1530-1539. Obtenido de <https://scielo.conicyt.cl/pdf/rmc/v142n12/art06.pdf>
- Crovetto, M., Uauy, R., Martins, A., Moubarac, J., & Montero, C. (2014). Disponibilidad de productos alimentarios listos para el consumo en los hogares de Chile y su impacto sobre la calidad de la dieta (2006-2007). *Revista Médica de Chile*, 142, 850-858. Obtenido de <https://scielo.conicyt.cl/pdf/rmc/v142n7/art05.pdf>
- Deville, J.-C., & Särndal, C.-E. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 87:418, 376-382.
- Díaz de Rada, V. (2005). *Manual de Trabajo de Campo en la Encuesta*. Madrid: Centro de Investigaciones Sociológicas.
- DIESA. (1984). Improving concepts and methods for statistics and indicators on the situation of women . Obtenido de Department of International Economic and Social Affairs: https://unstats.un.org/unsd/publication/SeriesF/SeriesF_33E.pdf
- Eurostat. (2003). Household Budget Surveys in the EU, Methodology and recommendations for harmonisation. Obtenido de http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-BF-03-003-__-N-EN.pdf
- Eurostat. (2013). Handbook on precision requirements and variance estimation for ESS household surveys. Obtenido de <http://ec.europa.eu/eurostat/documents/3859598/5927001/KS-RA-13-029-EN.PDF>
- Eurostat. (2014). European Statistical System Handbook for Quality Reports. Obtenido de <http://ec.europa.eu/eurostat/documents/3859598/6651706/KS-GQ-15-003-EN-N.pdf>
- Eurostat. (2015). Household Budget Survey 2010 Wave EU Quality report - VERSION 2. Obtenido de http://ec.europa.eu/eurostat/documents/54431/1966394/LC142-15EN_HBS_2010_Quality_Report_ver2+July+2015.pdf/fc3c8aca-c456-49ed-85e4-757d4342015f
- Guerrero Herrera, M. (2003a). Metodología de Clasificación Socioeconómica de los Hogares Chilenos. Instituto Nacional de Estadísticas (INE), Departamento de Metodología Estadística, Santiago de Chile.
- Guerrero Herrera, M. (2003b). Método PRINCALS para la Clasificación Socioeconómica del Censo 2002. Instituto Nacional de Estadísticas Chile, Departamento de Metodología Estadística. Obtenido de http://historico.ine.cl/canales/menu/publicaciones/revistas_economicas/23/miguelguerrerorev%2023.pdf
- Hoffmeyer-Zlotnik, J., & Warner, U. (2009). Private Household Concepts and their Operationalisation in Cross-National Social Surveys. Mannheim: GESIS-Leibniz-Institute for the Social Sciences. Obtenido de <https://www.stat-d.si/mz/mz6.1/hoff.pdf>
- INDEC . (2006). Encuesta Nacional de Gastos de los hogares 2004/2005 .
- INE Chile. (2013). Metodología VII EPF. Instituto Nacional de Estadísticas de Chile. Obtenido de <http://historico.ine.cl/epf/files/documentacion/METODOLOGIA.pdf>
- INE Chile. (2013b). Metodología Índice de Precios al Consumidor (IPC). Base Anual 2013. Santiago, Chile: Instituto Nacional de Estadísticas. Obtenido de <http://www.ine.cl/estad%C3%ADsticas/precios/ipc>
- INE Chile. (2014a). Medición y análisis de la falta de respuesta total en la VII Encuesta de Presupuestos Familiares.

INE Chile. (2014b). Métodos de imputación VII EPF: Gastos diarios e ingresos de la actividad laboral principal y jubilaciones. Instituto Nacional de Estadísticas de Chile. Obtenido de VII Encuesta de Presupuestos Familiares: http://historico.ine.cl/epf/files/documentacion/metodos_de_imputacion_VII_EPF.pdf

INE Chile. (2015a). Grupo de Trabajo: Estandarización de la Identificación de Hogar (Documento interno). Instituto Nacional de Estadísticas de Chile. Obtenido de Departamento de Calidad Estadística.

INE Chile. (2015b). Código de buenas prácticas para las estadísticas chilenas. Instituto Nacional de Estadísticas de Chile, Santiago.

INE Chile. (2016a). Jefatura de hogar: usos del concepto, historia, críticas y expresión en los indicadores. Instituto Nacional de Estadísticas de Chile. Santiago: INE. Obtenido de http://historico.ine.cl/genero/files/estadisticas/pdf/documentos/jefatura_hogar.pdf

INE Chile. (2017). Encuesta Suplementaria de Ingresos (ESI) Manual Conceptual y Metodológico. Instituto Nacional de Estadísticas de Chile. Obtenido de http://www.ine.cl/docs/default-source/ingresos-y-gastos/esi/ingreso-de-hogares-y-personas/antecedentes-metodol%C3%B3gicos/manual_conceptual_y_metodologico_esi.pdf?sfvrsn=8

INE Chile. (2018a). Lineamiento para la documentación de metodologías. Instituto Nacional de Estadísticas de Chile, Documento interno.

INE Chile. (2018b). Manual de Trabajo de Campo para Encuestadores, Tomo I. Santiago: Instituto Nacional de Estadísticas de Chile.

INE Chile. (2018c). Manual de Codificación VIII EPF. Instituto Nacional de Estadísticas de Chile, Santiago de Chile.

INE Chile. (2018d). Manual de Usuarios de las Bases de Datos VIII EPF. Santiago: Instituto Nacional de Estadísticas.

INE Chile. (2018e). Principales Resultados VIII Encuesta de Presupuestos Familiares. Instituto Nacional de Estadísticas de Chile.

INE Chile. (2018f). Informe de Calidad VIII Encuesta de Presupuestos Familiares. Santiago: Instituto Nacional de Estadísticas.

INE España. (2016). Metodología 2016 Encuesta de Presupuestos Familiares. Instituto Nacional de Estadísticas de España. Obtenido de <http://www.ine.es/metodologia/t25/t2530p45816.pdf>

INE España. (2017). Encuesta de Presupuestos Familiares.

ISTAT, Italia. (2011). I consumi delle famiglie. Anno 2009. Istituto Nazionale Di Statistica.

Lapierre, C. (2015). Data Collection Using a Diary: The Experience of the Survey of Household Spending. SSC Annual Meeting, June 2015.

Little, R. J. (1986). Survey Nonresponse Adjustments for Estimates of Means. *International Statistical Review*, 54(2), 139-157. Obtenido de <http://www.jstor.org/stable/1403140>

Little, R. J., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. John Wiley & Sons.

MECOVI. (2004). Plataforma del Proyecto. Obtenido de Banco Interamericano de Desarrollo (BID), el Banco Mundial y la Comisión Económica para América Latina y el Caribe (CEPAL): <https://www.cepal.org/deype/mecovi/>

Medina, F., & Galván, M. (2007). *Imputación de datos: teoría y práctica*. Santiago: CEPAL. División de estadística y proyecciones económicas.

Mincer, J. A. (1974). Schooling, Experience, and Earnings. Obtenido de National Bureau of Economic Research (NBER): <http://papers.nber.org/books/minc74-1>

Observatorio Social. (2015). Nueva Metodología de Medición de la Pobreza por Ingresos y Multidimensional. Santiago: Ministerio de Desarrollo Social. Obtenido de http://observatorio.ministeriodesarrollosocial.gob.cl/documentos/Nueva_Metodologia_de_Medicion_de_Pobreza.pdf

OCDE. (2007). Glossary of Statistical Terms. Organisation for Economic Cooperation and Development, Nueva York.

OCDE. (2013). OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth. París: Organización para la Cooperación y el Desarrollo Económico. Obtenido de http://www.oecd-ilibrary.org/economics/framework-for-statistics-on-the-distribution-of-household-income-consumption-and-wealth_9789264194830-en

OIT. (1991). Clasificación Internacional Uniforme de Ocupaciones: CIUO-88. Organización Internacional del Trabajo. Naciones Unidas.

OIT. (2003a). Informe de la Decimoséptima Conferencia Internacional de Estadísticos del Trabajo (CIET). Obtenido de Organización Internacional del Trabajo: http://www.ilo.org/wcmsp5/groups/public/-dgreports/-stat/documents/meetingdocument/wcms_087570.pdf

OIT. (2003b). Informe I: Estadísticas de ingreso y gasto de los hogares. Decimoséptima Conferencia Internacional de Estadísticas del Trabajo. Obtenido de <http://www.ilo.org/public/spanish/bureau/stat/download/mels-r-1.pdf>

OIT. (2003c). Informe II: Estadísticas de ingreso y gasto de los hogares. Decimoséptima Conferencia Internacional de Estadísticas del Trabajo. Ginebra: OIT. Obtenido de <http://www.ilo.org/public/spanish/bureau/stat/download/17thicls/r2hies.pdf>

ONS. (2017). Living Costs and Food Survey . Office for National Statistics.

ONU. (2001). Clasificación de Gastos por Finalidades. Informes Estadísticos, Naciones Unidas, Departamento de Asuntos Económicos y Sociales, División de Estadística, Nueva York. Obtenido de <https://unstats.un.org/unsd/cr/registry/regdnld.asp?Lg=3>

ONU. (2001). Ingresos y gastos de consumo de los hogares en el marco del SCN y en encuestas de hogares. (H. Camelo, Ed.) CEPAL. Obtenido de https://repositorio.cepal.org/bitstream/handle/11362/4718/1/S01010054_es.pdf

ONU. (2008). Principles and recommendations for population and housing censuses - Revision 2. New York: ONU. Department of Economic and Social Affairs. Obtenido de https://unstats.un.org/unsd/publication/seriesm/seriesm_67rev2e.pdf

ONU. (2009). Diseño de muestras para encuestas de hogares. Directrices Prácticas (Vol. Serie FN°98.). Obtenido de https://unstats.un.org/unsd/publication/seriesf/Seriesf_98s.pdf

ONU. (2011). Canberra Group Handbook on Household Income Statistics. Geneva: United Nations Economic Commission for Europe. Obtenido de https://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canberra_Handbook_2011_WEB.pdf

ONU et al. (2016). Sistema de Cuentas Nacionales 2008. New York: Naciones Unidas, Comisión Europea, OCDE, FMI, Banco Mundial. Obtenido de <https://unstats.un.org/unsd/nationalaccount/docs/SNA2008Spanish.pdf>

Statistics South Africa. (2012). Income and Expenditure of Households 2010/2011. Pretoria.

- U.S. BLS. (2004). Temporal Patterns of Survey Response Rates and Reporting Rates in the U.S. Consumer Expenditure Interview and Other. U.S. Bureau of Labor Statistics; Cho, Moon; Eltinge, John; Steinberg, Barry. Obtenido de <https://www.bls.gov/cex/cesrvymethscho1.pdf>
- U.S. BLS. (2006). Analysis of Recall Effect on the Reporting of Expenditures For the Consumer Expenditure Interview Survey. U.S. Bureau of Labor Statistics; Steinberg, Barry; Chopova, Borianna; Tan, Lucilla; Ogden, Jared; Bahizi, Pierre, Washington, DC, United States. Obtenido de <https://www.bls.gov/osmr/pdf/st060150.pdf>
- U.S. Census Bureau. (2003). Census Bureau Standard: Pretesting Questionnaires and Related Materials for Surveys and Censuses. U.S. Census Bureau. Obtenido de <https://www.census.gov/srd/pretest-standards.pdf>
- U.S. Census Bureau. (Julio de 2013). U.S. Census Bureau Statistical Quality Standards. Obtenido de https://www.census.gov/content/dam/Census/about/about-the-bureau/policies_and_notices/quality/statistical-quality-standards/Quality_Standards.pdf
- UNECE. (2016). Modelo Genérico del Proceso Estadístico (GSBPM). Comisión Económica de las Naciones Unidas para Europa.
- Valliant, R. D. (2012). Practical Tools for Designing and Weighting Survey Samples. . New York, USA: Springer.
- Vivanco, M. (2005). Muestreo Estadístico. Diseño y Aplicaciones. Santiago, Chile: Editorial Universitaria.
- Weiss, A. (1995). Human Capital vs. Signalling Explanations of Wages. Obtenido de <https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.9.4.133>
- Wolter, K. (2007). Statistics for Social and Behavioral Sciences, Introduction to Variance Estimation. New York, USA: Springer.

13. ANEXOS

Anexo 1: Perfiles de cargo equipo de trabajo de campo

Perfil de cargo	Dimensión evaluada			Principales funciones
	Formación educacional	Experiencia laboral	Otros conocimientos	
Encargado/a de área	Licenciatura del área de las ciencias sociales, preferentemente Sociología o Antropología, otorgada por una institución de educación del Estado o reconocida por este.	Deseable experiencia en producción y análisis de datos para investigación social, preferentemente encuestas socioeconómicas, durante 2 años continuos o 6 estudios distintos. Deseable experiencia desempeñando funciones de coordinación y supervisión de equipos de trabajo.	Office nivel medio/avanzado. MS Visio nivel básico. Deseable conocimientos en software estadísticos (SPSS, Stata) nivel básico.	<ul style="list-style-type: none"> • Coordinar, supervisar y retroalimentar el trabajo de los encargados regionales del área. • Dar seguimiento a los indicadores de desempeño operativo y calidad de los datos de la encuesta, promoviendo acciones de mejoramiento de los mismos. • Gestionar los recursos materiales y humanos requeridos para el normal funcionamiento del trabajo de campo en su área de responsabilidad. • Resolver situaciones excepcionales del trabajo de campo, de acuerdo a los protocolos establecidos por la encuesta. • Proveer información al equipo de supervisión y facilitar la coordinación de éstos con los grupos de campo de su área. • Capacitar, en conjunto con encargada nacional, al personal de su área antes y durante el período de trabajo de campo.
Encargado/a regional	Licenciatura del área de las ciencias sociales, administración, o educación, preferentemente Sociología, Antropología, Trabajo Social, Pedagogía E. Media, Geografía, Administración de empresas, Administración Pública o Contador Auditor, otorgada por una institución de educación del Estado o reconocida por este.	Deseable experiencia en producción y análisis de datos para investigación social, preferentemente encuestas socioeconómicas, durante 1 año o en 4 estudios distintos. Deseable 2 años de experiencia desempeñando funciones de coordinación y supervisión de equipos de trabajo.	Office nivel medio/avanzado. Deseable conocimientos en software estadísticos (SPSS, Stata) nivel básico.	<ul style="list-style-type: none"> • Coordinar y supervisar el trabajo de los encargados/as de grupo de su zona regional. • Gestionar los recursos materiales y humanos requeridos para el normal funcionamiento del trabajo de campo en su zona regional. • Resolver situaciones excepcionales del trabajo de campo, de acuerdo a los protocolos establecidos por la encuesta. • Coordinar y apoyar a los equipos de trabajo de campo en las labores de difusión local de la encuesta.

Perfil de cargo	Dimensión evaluada			Principales funciones
	Formación educacional	Experiencia laboral	Otros conocimientos	
Encargado/a de grupo	Licenciatura del área de las ciencias sociales, administración, comunicación o educación, preferentemente Sociología, Antropología, Trabajo Social, Asistencia Social, Periodismo, Pedagogía E. Media, Geografía, Administración de empresas, Administración Pública o Contador Auditor, otorgada por una institución de educación del Estado o reconocida por este.	Experiencia en producción de datos para investigación social, preferentemente encuestas de hogares, durante un año o en 4 estudios distintos. Experiencia en coordinación de equipos de trabajo, preferentemente de actividades en terreno, durante al menos un año.	Office nivel medio.	<ul style="list-style-type: none"> • Gestionar los recursos necesarios para el buen desempeño del equipo (materiales de oficina, cuestionarios, vehículos, etc.). • Velar por el cumplimiento administrativo del grupo de campo (control de horarios de trabajo y de cumplimiento de cargas de trabajo asignadas). • Identificar las unidades de primera selección (manzanas) en terreno, difundiendo en ellas la realización de la encuesta. • Coordinar las labores de los encuestadores, tanto en gabinete como en terreno, de acuerdo a los protocolos de la encuesta. • Revisar el material aplicado, evaluando la completitud y consistencia de la información producida.
Encuestador	Licenciatura del área de las ciencias sociales, administración, o educación, preferentemente Sociología, Antropología, Trabajo Social, Pedagogía E. Media, Geografía, Administración de empresas, Administración Pública o Contador Auditor, otorgada por una institución de educación del Estado o reconocida por este.	Deseable experiencia en producción de datos para investigaciones sociales, preferentemente encuestas, durante al menos 1 año o 4 estudios distintos.	Office nivel básico.	<ul style="list-style-type: none"> • Contactar a los hogares residentes en las viviendas seleccionadas y persuadirlos de que participaran del estudio. • Aplicar los cuestionarios de la encuesta mediante entrevistas presenciales en las viviendas. • Realizar seguimiento al llenado de la Libreta de Gastos Individuales por parte de los integrantes del hogar de 15 años o más. • Revisar el material aplicado, complementando la información producida y, cuando era necesario, recuperando datos faltantes en posteriores visitas a los hogares.

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF)

Anexo 2: Afiche VIII EPF



**¡CONTESTE
CON
CONFIANZA!**

Todos los datos que entregue serán
tratados de forma **confidencial** y
estarán protegidos por el
Secreto Estadístico
(artículo 29, Ley 17.374)

VIII EPF ENCUESTA DE
PRESUPUESTOS
FAMILIARES

Durante..... / encuestaremos este sector

La Encuesta de Presupuestos Familiares (EPF) es la encuesta que
recopila información base para elaborar la canasta de productos y
servicios que componen el Índice de Precios al Consumidor (IPC).

Para mayor información ingrese a www.ine.cl/epf

INE
Instituto Nacional de Estadísticas - Chile
www.ine.cl

Anexo 3: Ejemplo de carta de apertura de viviendas VIII EPF



Santiago, junio de 2016

SUBDIRECCIÓN TÉCNICA
Departamento de Presupuestos Familiares

Señor(a)
Jefe(a) de Hogar
Presente

Estimado(a) señor(a):

Tengo el agrado de comunicar a usted que el Instituto Nacional de Estadísticas (INE) se encuentra desarrollando la **VIII Encuesta de Presupuestos Familiares (EPF)**, en el marco de la generación de estadísticas oficiales del país.

Este estudio tiene como finalidad identificar la estructura y características del gasto de consumo final de los hogares urbanos de Chile, información fundamental para la actualización de la canasta de bienes y servicios del Índice de Precios al Consumidor (IPC). Además, los resultados del estudio constituyen información relevante para la actualización y cálculo de la línea de pobreza del país.

Ante esto, su vivienda ha sido seleccionada en forma aleatoria para participar en esta importante encuesta. La información proporcionada será fundamental para disponer de estadísticas más actualizadas y que reflejen mejor la realidad del país.

Dicha información será utilizada en forma estrictamente confidencial, protegida por el **Secreto Estadístico**, según lo dispuesto en la **Ley N° 17.374**. Aquella ley faculta al INE para solicitar información a cualquier persona natural o jurídica que habita en el territorio nacional.

Agradeciendo su valiosa cooperación para el cumplimiento de este proceso de gran relevancia para el país, le saluda atentamente,



XIMENA CLARK NÚÑEZ
Directora Nacional
Instituto Nacional de Estadísticas



INFORMACIÓN DE IMPORTANCIA

Quien le ha entregado esta carta es un encuestador profesional del Instituto Nacional de Estadísticas, quien deberá identificarse exhibiendo la credencial respectiva de la Institución. Esta persona será la responsable de responder a todas las preguntas que usted o algún miembro de su hogar requiera. Además, coordinará las fechas y horarios en que nuestro equipo visitará su hogar para el desarrollo de este proceso.

La encuesta tiene una duración de 15 días, período en que se pedirá a los miembros del hogar de 15 años o más que registren todos sus gastos diarios en una libreta. Junto con ello, se solicitará al hogar que responda preguntas, contenidas en otros 5 cuestionarios, sobre gastos de carácter mensual, como agua, luz, arriendo -entre otros-, así como también los gastos realizados en los últimos meses, de acuerdo a las especificaciones del proyecto. Por último, se consultará sobre los ingresos de los integrantes de su hogar, información que complementará los datos sobre los gastos realizados.

Para efectuar este trabajo, un encuestador INE visitará cuatro veces su hogar durante estos 15 días, con la finalidad de apoyar en el llenado de las libretas de gastos diarios, resolver dudas y aplicar los otros cuestionarios. Cada una de estas visitas demorará, en promedio, 30 minutos.

Si desea tener mayor información sobre este proyecto, puede comunicarse a los siguientes teléfonos: 232201500- 232201521 o escribir un correo electrónico a epf@ine.cl. Además, puede obtener más información de esta encuesta y verificar la identidad del encuestador en www.ine.cl/epf

Muchas gracias por su colaboración.

Anexo 4: Díptico VIII EPF

EL ENCUESTADOR DE HOGARES

Para cumplir con el objetivo de la VIII EPF, se cuenta con un equipo altamente capacitado de encuestadores de hogares. Estas personas son las encargadas de recopilar información coherente y verídica de los distintos tipos de gastos e ingresos en que incurren los hogares seleccionados.

¿CÓMO IDENTIFICO A LOS ENCUESTADORES?

Los encuestadores de hogares contarán con la indumentaria necesaria para ser reconocidos por toda la población: camiseta o blusa celeste con el logo del INE, chaqueta o parka con el logo del INE y la credencial que los identifica como funcionarios de la institución.

Para su seguridad y la de todos los encuestados, podrá consultar el RUT del encuestador en el sitio **www.ine.cl** en la sección "**Verificación Encuestador**" y comprobar si es o no un funcionario del INE. Además, se podrá acceder al verificador desde un Smartphone con el código QR presente en la credencial de cada encuestador. Para cualquier duda o consulta se puede contactar al correo **eine@ine.cl**

Su vivienda fue visitada por un encuestador del INE el				A las				hrs.
						:		
	/							/

Nombre del encuestador:	
RUT del encuestador:	
Nombre de jefe directo:	
Teléfono de jefe directo:	

**¡SU VIVIENDA
HA SIDO
ELEGIDA!**

Y representará aproximadamente a 260 familias como la suya.



VII
ENCUESTA DE
PRESUPUESTOS
FAMILIARES



Su vivienda ha sido seleccionada para responder la octava Encuesta de Presupuestos Familiares (EPF) que el INE realiza. Invitamos a usted y a toda su familia a participar de este proceso fundamental para conocer la estructura de los gastos e ingresos de los hogares en Chile.

¿QUÉ ES EL INE?



El Instituto Nacional de Estadísticas (INE) es uno de los organismos públicos con más trayectoria en Chile. Se fundó en 1843 y actualmente tiene como misión producir y difundir las estadísticas oficiales de Chile, proporcionando información confiable y accesible a los usuarios, para la toma de decisiones y un mayor conocimiento de la realidad del país.

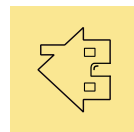
¿QUÉ ES LA ENCUESTA DE PRESUPUESTOS FAMILIARES?



La Encuesta de Presupuestos Familiares (EPF) es un estudio que permite conocer las modificaciones en la estructura del gasto de los hogares del país, como también sus nuevas pautas de consumo.

Desde 2011 se realiza cada cinco años y su información es la base para elaborar la canasta de productos y servicios que mide el Índice de Precios al Consumidor (IPC). Aquel indicador mide la variación de los precios de diversos productos. La EPF es, además, utilizada como la principal fuente de información para la actualización de las líneas de incidencia y de pobreza del país.

¿CUÁNDO Y DÓNDE SE REALIZARÁ ESTA ENCUESTA?



La VIII EPF se desarrollará entre el 1 julio de 2016 y el 30 de junio de 2017. Su cobertura contempla las áreas urbanas de cada una de las quince capitales regionales del país y algunas de sus zonas conurbadas, además de las comunas de Chillán y Chillán Viejo. En esta oportunidad se encuestarán más de 13.000 viviendas en todo Chile. Previamente, en junio de 2016, se realizará una Marcha Blanca de la encuesta.



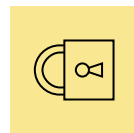
¿POR CUÁNTO TIEMPO ENCUESTARÁN MI HOGAR?

La encuesta tiene una duración de 15 días, período en que los miembros del hogar deberán entregar información sobre sus gastos e ingresos en seis libretas. Se solicitará al hogar que informe sobre aquellos gastos de carácter mensual, como agua, luz, arriendo, entre otros, como también los gastos realizados en los últimos 3, 6 y 12 meses.

Cada miembro del hogar de 15 años o más recibirá un formulario especial donde registrar sus gastos diarios. Para efectuar este trabajo, un encuestador INE visitará cuatro veces su hogar durante estos 15 días, con la finalidad de apoyar en el llenado de las libretas y resolver dudas. Cada visita demorará, en promedio, 30 minutos. Para realizar estas visitas, el encuestador se adaptará a su disponibilidad de tiempo.

La Encuesta de Presupuestos Familiares se rige por estándares internacionales de alta calidad y que utilizan los países pertenecientes a la Organización para la Cooperación y el Desarrollo Económico (OCDE). Responda con confianza, pues su información se encuentra protegida y resguardada por ley.

¿LA INFORMACIÓN QUE YO ENTREGUE SERÁ PUBLICADA?



La información entregada será totalmente confidencial y se utilizará exclusivamente para este estudio. Para su tranquilidad, estos antecedentes están protegidos por el Secreto Estadístico, según lo dispuesto en la Ley N° 17.374, misma ley que faculta al INE a solicitar información a todas las personas.

Anexo 5: Hoja de Ruta VIII EPF

Los datos que el INE solicita son
CONFIDENCIALES
El artículo 29, Ley 17.374, establece el
SECRETO ESTADÍSTICO.



HOJA DE RUTA DEL ENCUESTADOR (HR)

NÚMERO DE FOLIO
—

INFORMACIÓN DE MARCO MUESTRAL									
SUB-MUESTRA	REGIÓN	PROV.	COMUNA	ESTRATO	DISTRITO CENSAL	ZONA CENSAL	CÓDIGO DE MANZANA	VIVIENDA	HOGAR

IDENTIFICACIÓN DE LA VIVIENDA				
NOMBRE DE CALLE O CAMINO	NÚMERO	BLOCK	PISO	DPTO. O PIEZA

IDENTIFICACIÓN DEL/LA ENCUESTADOR/A Y ENCARGADO/A DE GRUPO (ID)		
Registre su nombre y código respectivo, además de la información referente al encargado de grupo. Anote con letra clara solo el primer nombre y apellido paterno.		
NOMBRE ENCUESTADOR/A INICIAL (01)	APELLIDO ENCUESTADOR/A INICIAL (02)	CÓDIGO EPF (03)
NOMBRE ENCARGADO/A (04)	APELLIDO ENCARGADO/A (05)	CÓDIGO EPF (06)

VISITA DE SUPERVISIÓN DEL HOGAR (SU)									
INDIQUE SI EL HOGAR FUE SUPERVISADO DURANTE LA SUBMUESTRA (01)				1. SI → Pase a responder la siguiente pregunta 2. No					
SEÑALE QUÉ ÁREAS FUERON SUPERVISADAS (02) (Anote 1. Si o 2. No)				T. Campo	HR	RPH	LGR	LGH	LI

VIII ENCUESTA DE PRESUPUESTOS FAMILIARES

CÓDIGOS DE RECONOCIMIENTO DE MANZANA
(uso exclusivo del encargado de grupo)

11	Entrega carta de forma directa
12	Entrega carta de forma indirecta (buzón o casilla, conserje, etc.)

CÓDIGOS DE VISITA AL HOGAR

UNIDADES NO ELEGIBLES	
21	Vivienda de verano o uso temporal
22	Vivienda desocupada (ej.: arriendo, venta, etc.)
23	Vivienda colectiva (ej.: hotel, hospital, cárcel, asilo de ancianos, dormitorio militar, de trabajo, internado, etc.)
24	Edificación para uso no habitacional (ej.: local comercial, empresa, oficina de gobierno, organización, etc.)
25	Edificación no calificable (ej.: construcción, demolición, incendiada, destruida o erradicada)
UNIDADES DE ELEGIBILIDAD DESCONOCIDA	
31	Se desconoce el uso de la propiedad
32	No encontró la dirección
33	Área peligrosa o de difícil acceso

UNIDADES ELEGIBLES	
NO CONTACTO (códigos 4)	
41	Vivienda ocupada sin moradores presentes
42	Se impide acceso a la vivienda (ej.: existe portón automático o el administrador, conserje, junta de vigilancia niega acceso)
43	Informante no ubicable, se contacta a otra persona (ej.: menor de edad, familiar, trabajador de la vivienda, amigo, etc.)
CONTACTO INICIAL (códigos 5)	
51	Concertó cita con el informante
NO ENTREVISTADO POR OTROS MOTIVOS (códigos 6)	
61	No se logró comunicación con el informante por problemas de idioma
62	Informante impedido físico y/o mental para contestar
63	Muerte del informante o algún familiar
64	Otro motivo (especifique en observaciones)
RECHAZO (códigos 7)	
71	No tiene tiempo
72	Se niega, no cree en la confidencialidad
73	Se niega, preguntas muy personales
74	Se niega, no cree en las encuestas
75	Se niega, no obtiene beneficios por contestar encuestas
76	Se niega, le aburre
77	Se niega sin dar razones
ENTREVISTA (códigos 8)	
81	Entrevista lograda

REGISTRO DE CONTACTO CON EL HOGAR (RC)

Recuerde que el módulo de registro de contactos con el hogar debe ser llenado en el reconocimiento de manzana y durante el PROCESO DE APERTURA DEL HOGAR. Pasado este periodo, el módulo NO DEBE SER MODIFICADO.
Registre información sobre la fecha, hora, código de visita y código del responsable de cada visita realizada al hogar. Además, señale si la visita la realizó 1. ENCUESTADOR/A o 2. ENCARGADO/A DE GRUPO.

N° (01)	FECHA			HORA hh/mm (05)	CÓDIGO DE VISITA (06)	CÓDIGO EPF RESPONSABLE VISITA (07)	RESPONSABLE VISITA 1. Encuestador/a 2. Encargado/a (08)	CÓDIGO DE DISPOSICIÓN FINAL (09)
1				:				<p>- Recuerde que el código de disposición final corresponde al código que refleje el mayor nivel de interacción logrado. Debe anotar el MAYOR DE LOS CÓDIGOS DE VISITA que ha registrado anteriormente en la pregunta RC06 "Código de visita".</p> <p>- Complete el recuadro sólo al finalizar EL PROCESO DE APERTURA DE LOS HOGARES.</p> <div style="border: 1px solid black; width: 40px; height: 40px; margin: 10px auto;"></div>
2				:				
3				:				
4				:				
5				:				
6				:				
7				:				
8				:				
9				:				
10				:				
11				:				
12				:				
13				:				
14				:				

SOLO PARA INFORMANTES QUE NO PARTICIPAN DE LA ENCUESTA
(registran código de visita 6 o 7)

Esta pregunta se debe aplicar sólo a aquellos informantes que expresan un RECHAZO (cód. 7) o no participan de la encuesta por OTROS MOTIVOS (cód. 6). Responda esta pregunta de acuerdo a las características observables del INFORMANTE con el que se realiza el contacto.

EDAD ESTIMADA DEL INFORMANTE (10)	SEXO DEL INFORMANTE (11)
<div style="border: 1px solid black; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center;"> <input type="checkbox"/> </div> 1. BAJO 30 AÑOS 2. 30-64 AÑOS 3. 65 AÑOS O MÁS 4. NO SE PUEDE DETERMINAR	<div style="border: 1px solid black; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center;"> <input type="checkbox"/> </div> 1. HOMBRE 2. MUJER

SOLO PARA HOGARES APERTURADOS
(registran código de visita 81)

Responda esta pregunta al FINALIZAR EL PROCESO DE APLICACIÓN de la encuesta en el hogar y sólo para aquellos HOGARES que han sido APERTURADOS (implica la aplicación de al menos RPH).

REGISTRE EL ESTADO FINAL DEL PROCESO DE APLICACIÓN (12)

<div style="border: 1px solid black; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center;"> <input type="checkbox"/> </div> 1. EL HOGAR TERMINA LA APLICACIÓN 2. EL HOGAR NO TERMINA LA APLICACIÓN, SE PIERDE EL CONTACTO 3. EL HOGAR NO TERMINA LA APLICACIÓN, SE NIEGA A CONTINUAR 4. EL HOGAR NO TERMINA LA APLICACIÓN, POR MUERTE O ENFERMEDAD (ALGÚN FAMILIAR) 5. EL HOGAR NO TERMINA LA APLICACIÓN, POR OTROS MOTIVOS
--

2 HOJA DE RUTA DEL ENCUESTADOR/A

REGISTRO DE APLICACIÓN DE LIBRETAS (RL)																
Registre la fecha, hora de inicio y hora de término para cada libreta aplicada. Para el caso de la revisión de LGI, indique el tiempo aproximado que tardó en revisar los cuestionarios en cada visita de aplicación. Para los hogares donde revisó más de un cuestionario durante la visita, sume los tiempos de revisión de LGI.																
N° (01)	LIBRETA (02)	1° APLICACIÓN					2° APLICACIÓN					3° APLICACIÓN				
		FECHA			HORARIO INICIO hh/mm (06)	HORARIO TÉRMINO (07)	FECHA			HORARIO INICIO hh/mm (11)	HORARIO TÉRMINO (12)	FECHA			HORARIO INICIO hh/mm (16)	HORARIO TÉRMINO (17)
DÍA dd (03)	MES mm (04)	AÑO aa (05)	DÍA dd (08)	MES mm (09)			AÑO aa (10)	DÍA dd (13)	MES mm (14)			AÑO aa (15)				
1	RPH				:	:				:	:				:	:
2	LGH				:	:				:	:				:	:
3	LGR				:	:				:	:				:	:
4	LI				:	:				:	:				:	:
5	LSF				:	:				:	:				:	:
6	LGI				:	:				:	:				:	:
CÓDIGOS ASIGNACIÓN DE GASTOS INDIVIDUALES (LGI)																
91	Entrega LGI de forma personal															
92	Entrega LGI de forma indirecta (a través de un miembro del hogar)															
93	No entrega LGI, rechaza contestar															
94	No entrega LGI, miembro del hogar no ubicable en el período de referencia (ej.: motivos vacacionales, trabajo, estudios, salud, etc.)															
95	No aplica contestar LGI, miembro del hogar es menor de 15 años															
96	No aplica contestar LGI, miembro del hogar se encuentra impedido físico y/o mental para contestar															
CONTROL DE LIBRETAS DE GASTOS INDIVIDUALES (CL)																
Registre TODOS LOS MIEMBROS DEL HOGAR de acuerdo al orden del RPH. Al momento de ENTREGAR la libreta, registre el código de asignación de LGI para cada persona, luego señale al administrador de gastos del hogar anotando 1. Sí o 2. No, y la fecha de entrega. En el momento del CIERRE del proceso, debe indicar si la LGI fue retirada anotando 1. Sí o 2. No. Además, debe registrar la fecha de retiro, días de registro y revisión de LGI, anotando 1. "Sí, al menos revisó una vez" o 2. "No, nunca pudo revisar la libreta".																
N° Línea (RPH)	CÓDIGO ASIGNACIÓN (01)	ADM. DE GASTO 1. Sí o 2. No (02)	FECHA DE ENTREGA			RETIRO DE LGI 1. Sí o 2. No (06)	FECHA DE RETIRO			NÚMERO DÍAS DE REGISTRO (10)	REVISIÓN DE LGI 1. Sí o 2. No (11)					
			DÍA dd (03)	MES mm (04)	AÑO aa (05)		DÍA dd (07)	MES mm (08)	AÑO aa (09)							
1																
2																
3																
4																
5																
6																
7																
8																
9																
10																
11																
VIII ENCUESTA DE PRESUPUESTOS FAMILIARES											3					

[illegible]

Anexo 6: Planilla de recepción de material



Planilla de Recepción de Material

Los datos que el INE solicita son **CONFIDENCIALES**
El artículo 29, Ley 17.374, establece el **SECRETO ESTADÍSTICO**.

Región Supervisada	
Fecha de entrega de libretas a supervisores	
Fecha de devolución de libretas a encargado de grupo	
Cantidad de Folios supervisados	

Cargo	Nombre	Código EPF
Supervisor		
Encargado de Grupo		

	Folio	HR	RPH	LGR	LGH	LI	LSF	LGI
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								



Los datos que el INE solicita son **CONFIDENCIALES**
El artículo 29, Ley 17.374, establece el **SECRETO ESTADÍSTICO**.

	Folio	HR	RPH	LGR	LGH	LI	LSF	LGI
36								
37								
38								
39								
40								
41								
42								
43								
44								
45								
46								
47								
48								
49								
50								
51								
52								
53								
54								
55								
56								
57								
58								
59								
60								

Observaciones

	Firma Encargado de Grupo	Firma Supervisor
Entrega de Libretas		
Devolución de Libretas		

Anexo 7: Formulario de supervisión directa en fase de apertura



Los datos que el INE solicita son **CONFIDENCIALES**
El artículo 29, Ley 17.374, establece el **SECRETO ESTADÍSTICO**.

Formulario de Supervisión Directa - Apertura

NÚMERO DE FOLIO

■

Información de marco muestral

Sub-muestra	Región	Prov.	Comuna	Estrato	Distrito censal	Zona censal	Código de manzana	Vivienda	Hogar

Identificación de la vivienda

Nombre de calle o camino	Número	Block	Piso	Dpto. o pieza

Datos de la Visita

Código E. de Grupo	Código Encuestador	Código Supervisor	Fecha			Estado (1. Lograda; 2. Sin contacto; 3. Concertó cita; 4. Rechazo; 5. Unidad no elegible / elegibilidad desconocida; 6. No lograda por otros motivos)	Hora Inicio (HH:MM)	Hora Término (HH:MM)	N° línea informante	Nombre de pila informante
			Día	Mes	Año					
							:	:		

Apertura

A. Identificación de la vivienda y presentación

	1. Sí; 2. No	Observaciones (Obligatorio en caso de no observarse la situación)
El encuestador saluda y se presenta ante el informante con cordialidad		
El encuestador se identifica mostrando su credencial o la porta en un lugar visible		
El encuestador utiliza vestimenta institucional (ejemplo: jockey, chaqueta/parka, camisa, etc.)		
El encuestador informa que la vivienda fue seleccionada al azar para participar en la VIII EPF		
El encuestador comunica claramente las temáticas involucradas y los usos de la encuesta		
El encuestador entrega la carta informativa acerca de la encuesta		
El encuestador informa sobre la duración de la colaboración y el número de visitas al hogar		
El encuestador informa sobre el tiempo de duración de las visitas al hogar		
El encuestador menciona el resguardo de la información vía secreto estadístico, en caso que el informante manifieste desconfianza o temor		
El encuestador informa sobre posibles visitas de supervisión		
El encuestador explica de manera clara el proceso de llenado de LGI		
El encuestador entrega las LGI que corresponden		
El encuestador entrega el material de difusión y los materiales para llenado de LGI		

B. Identificación del hogar		
	1. Sí; 2. No	Observaciones (Obligatorio en caso de no observarse la situación)
Realiza de forma literal la pregunta H2: De las personas que residen normalmente aquí, ¿Todas comparten los gastos para alimentación?		
(Solo si se identifica más de un grupo con presupuesto separado para alimentación) ¿El encuestador realizó el ejercicio de indagación sobre acceso independiente?		

Llenado de RPH			
	Realiza las preguntas de manera textual	Explica o detalla si es necesario	Observaciones (Obligatorio en caso de no observarse siempre alguna situación)
	1. Siempre; 2. La mayoría de las veces; 3. Algunas veces; 4. Nunca		
Miembros del Hogar (MH)			
Educación (ED)			
Condición de Actividad Económica (AE)			
Salud y Previsión (SP)			
Vivienda Principal (VP)			
Vivienda Secundaria (VS)			

Antecedentes generales de la entrevista		
	1. Sí; 2. No	Observaciones (Obligatorio en caso de no observarse la situación)
El encuestador utiliza un vocabulario y tono adecuado durante la entrevista		
El encuestador es capaz de revertir alguna situación de rechazo del hogar o abordar tranquilamente una situación imprevista (ej: interrupciones, visitas, llamadas telefónicas, etc.)		
El encuestador maneja conceptualmente el instrumento y conoce los saltos de éste		
El encuestador responde con claridad y argumentos válidos las preguntas del informante		
El encuestador utiliza el ritmo adecuado para la realización de la entrevista		
El encuestador deja observaciones pertinentes en los instrumentos cuando corresponde		
El encuestador se despide cordialmente del informante		
El encuestador concierta día, horario y lugar de próxima visita		
El encuestador completa las libretas durante la entrevista		
El encuestador utiliza lápiz pasta azul para llenar las libretas		
El encuestador registra correctamente la información de la visita en la Hoja de Ruta		

Observaciones

Anexo 8: Formulario de supervisión directa en fase de aplicación



Formulario de Supervisión Directa - Aplicación

Los datos que el INE solicita son **CONFIDENCIALES**
El artículo 29, Ley 17.374, establece el **SECRETO ESTADÍSTICO**.

NÚMERO DE FOLIO

■

Información de marco muestral

Sub-muestra	Región	Prov.	Comuna	Estrato	Distrito censal	Zona censal	Código de manzana	Vivienda	Hogar

Identificación de la vivienda

Nombre de calle o camino	Número	Block	Piso	Dpto. o pieza

Datos de la Visita

Código E. de Grupo	Código Encuestador	Código Supervisor	Fecha			Estado (1. Lograda; 2. Sin contacto; 3. Concertó cita; 4. Rechazo; 5. Unidad no elegible / elegibilidad desconocida; 6. No lograda por otros motivos)	Hora Inicio (HH:MM)	Hora Termino (HH:MM)	Nº línea informante	Nombre de pila informante
			Día	Mes	Año					
							:	:		

Llenado de LGH

¿Aplicó LGH en esta visita?

- ☐ 1. Si —> Continuar
☐ 2. No —> Pasar a Llenado de LGR

	Realiza las preguntas de manera textual	Explica o detalla si es necesario	Observaciones (Obligatorio en caso de no observarse siempre alguna situación)
	1. Siempre; 2. La mayoría de las veces; 3. Algunas veces; 4. Nunca		
Gastos periódicos relativos a la vivienda principal (VP)			
Gastos periódicos relativos a la vivienda secundaria (VS)			
Gastos periódicos relativos a telecomunicaciones del hogar (TC)			
Gastos periódicos relativos al derecho de aseo municipal del hogar (AM)			
Gastos periódicos relativos al uso de automóviles con fines domésticos en el hogar (VD)			
Gastos periódicos relativos a servicio doméstico contratado por el hogar (SD)			
Otros gastos periódicos (OG)			
Gastos relativos a educación formal por persona (EF)			
Gastos relativos a otros cursos de capacitación o talleres con fines			
Otros gastos relativos a educación no relacionado a cursos (OE).			

Llenado de LGR				
¿Aplicó LGR en esta visita?				<input type="checkbox"/> 1. Sí —> Continuar <input type="checkbox"/> 2. No —> Pasar a Llenado de LI
	Realiza las preguntas de manera textual	Explica o detalla si es necesario	Observaciones (Obligatorio en caso de no observarse siempre alguna situación)	
	1. Siempre; 2. La mayoría de las veces; 3. Algunas veces; 4. Nunca			
Periodo del recuerdo de 3 Meses				
Periodo del recuerdo de 6 Meses				
Periodo del recuerdo de 12 Meses				
Llenado de LI				
¿Aplicó LI en esta visita?				<input type="checkbox"/> 1. Sí —> Continuar <input type="checkbox"/> 2. No —> Pasar a Llenado de LSF
	Realiza las preguntas de manera textual	Explica o detalla si es necesario	Observaciones (Obligatorio en caso de no observarse siempre alguna situación)	
	1. Siempre; 2. La mayoría de las veces; 3. Algunas veces; 4. Nunca			
Ingresos del Trabajo Asalariado (TA)				
Ingresos de Trabajo Independiente (TI)				
Otros ingresos del trabajo (OT)				
Ingresos por Jubilaciones (JU)				
Ingresos por Otras Pensiones y Transferencias Recibidas (TR)				
Otros Ingresos de Carácter No Habitual (NH)				
Ingresos de la Propiedad y Ganancia por tenencia de instrumentos financieros (IP)				
Transferencias Emitidas (TE)				
Llenado de LSF				
¿Aplicó LSF en esta visita?				<input type="checkbox"/> 1. Sí —> Continuar <input type="checkbox"/> 2. No —> Pasar a Antecedentes Generales
	Realiza las preguntas de manera textual	Explica o detalla si es necesario	Solicita comprobantes	Observaciones (Obligatorio en caso de no observarse siempre alguna situación)
	1. Siempre; 2. Generalmente; 3. Ocasionalmente; 4. Nunca			
Créditos hipotecarios (CH)				
Tarjetas de crédito de establecimientos comerciales (TE)				
Créditos de consumo y crédito automotriz (CCA)				
Cuenta corriente y línea de crédito (PB)				
Cuenta vista y/o cuenta rut (CV)				
Tarjetas de crédito bancarias (TCB)				
Créditos universitarios (CU)				
Seguros (SE)				
2 SUPERVISIÓN DIRECTA - APLICACIÓN				

[illegible]

Anexo 9: Formulario de supervisión indirecta



Formulario de Supervisión Indirecta

Los datos que el INE solicita son **CONFIDENCIALES**
El artículo 29, Ley 17.374, establece el **SECRETO ESTADÍSTICO**.

NÚMERO DE FOLIO

Información de marco muestral

Sub-muestra	Región	Prov.	Comuna	Estrato	Distrito censal	Zona censal	Código de manzana	Vivienda	Hogar

Identificación de la vivienda

Nombre de calle o camino	Número	Block	Piso	Dpto. o pieza

Datos de la Visita

Código E. de Grupo	Código Encuestador	Código Supervisor	Fecha			Estado (1. Lograda; 2. Sin contacto; 3. Concertó cita; 4. Rechazo; 5. Unidad no elegible / elegibilidad desconocida; 6. No lograda por otros motivos)	Hora Inicio (HH:MM)	Hora Término (HH:MM)	N° línea informante	Nombre de pila informante
			Día	Mes	Año					
							:	:		

Apertura

(Sólo si el informante estuvo presente durante la apertura. En caso contrario, complete con -77)

	1. Si; 2. No	Observaciones
¿El encuestador utilizó vestimenta con el logo del INE (ejemplo: jockey, chaqueta/parka, camisa, etc.)?		
¿El encuestador informó claramente las temáticas involucradas y los usos de la encuesta?		
¿El encuestador entrega una carta informativa acerca de la encuesta?		
¿El encuestador informó sobre la duración de la colaboración y el número de visitas al hogar?		
¿El encuestador informa sobre el tiempo de duración de las visitas al hogar?		
¿El encuestador mencionó el resguardo de la información vía secreto estadístico, en caso que el informante manifestara desconfianza o temor?		
¿El encuestador informó sobre posibles visitas de supervisión?		
¿El encuestador explica de manera clara el proceso de llenado de las libretas para registrar diariamente sus gastos?		
¿El encuestador entregó el material de difusión de la encuesta y material para llenado de las libretas para registrar diariamente sus gastos?		

RPH						
		N° línea informante libreta	N° línea informante supervisión	Nombre informante supervisión	Coincide con informante de la libreta (1. Si; 2. No)	
Considerando que la entrevista con el encuestador EPF fue realizada durante la primera/segunda quincena del mes de _____						
Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación	
Identificación del Hogar	H1. ¿Cuántas personas residían normalmente aquí? (Número de personas)					
	H2. De las personas que residían normalmente aquí, ¿todas compartían los gastos para alimentación? (1. Si; 2. No)					
	H3. Entonces, contando el de usted ¿Cuántos grupos de personas tenían gasto separado para la alimentación? (Número de Grupos)					
MH	MH01. ¿Quién es el/la jefe/a de este hogar? (Nombre de pila)					
	MH03. Indique el parentesco con el/la jefe/a de hogar 1. Jefe/a de hogar 2. Esposo/a o cónyuge 3. Conviviente civil (AUC) 4. Conviviente de hecho o pareja 5. Hijo/a 6. Hijo/a del cónyuge o pareja 7. Hermano/a 8. Padre o madre 9. Cuñado/a 10. Suegro/a 11. Yerno/Nuera 12. Nieto/a 13. Abuelo/a 14. Otro pariente 15. Otro no pariente 16. Servicio doméstico y parientes de éste 17. Pensionistas	1				
		2				
		3				
		4				
		5				
		6				
		7				
		8				
		9				
		10				
		11				
	MH07. Durante los últimos 6 meses, ¿qué persona fue la que más aportó al presupuesto de este hogar? (Número de línea)					
	MH08. ¿Quién es la persona que habitualmente realiza las compras del hogar? (ej.: vestuario, calzado, electrodomésticos, alimentos, bebidas, etc.) (Número de línea. Si existe más de uno, separar distintas personas usando comas)					

2 SUPERVISIÓN INDIRECTA

Módulo	Pregunta	Respuesta registrada		Respuesta supervisión		Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
		ED01	ED03	ED01	ED03		
ED	ED01. Considerando que la educación formal tiene cuatro niveles: párvulo, básica, media y superior y que a la vez existen otros cursos complementarios fuera de los niveles formales, tales como preuniversitarios, cursos de idiomas (no conducentes a título), capacitaciones, diplomados sin licenciatura, cursos de arte o música, etc. ¿Se encuentra estudiando actualmente?						
	1. Si estudia, en establecimientos de educación formal						
	2. Si estudia, en otro tipo de cursos complementarios						
	3. Si estudia, en ambas (alternativas 1 y 2)						
	4. No estudia actualmente						
	ED03. Indique el nivel educacional más alto alcanzado (Si ignora curso o nivel, anotar -99).						
	1. Nunca ha estudiado en establecimientos de educación formal						
	2. Ninguno, está iniciando su educación formal						
	3. Educación especial (diferencial)						
	4. Sala cuna						
	5. Jardín infantil						
6. Prekinder o kinder							
7. Educación básica							
8. Educación primaria (sistema antiguo)							
9. Educación media científico-humanista							
10. Educación media técnica profesional							
11. Educación humanidades (sistema antiguo)							
12. Educación técnica comercial, industrial o normalista (sistema antiguo)							
13. Técnico nivel superior (carrera de 1 a 3 años)							
14. Profesional (carrera de 4 años o más)							
15. Postítulo (incluye diplomados que requieren licenciatura)							
16. Magister							
17. Doctorado							

SUPERVISIÓN INDIRECTA 3

Considerando que la entrevista con el encuestador EPF fue realizada durante la primera/segunda quincena del mes de _____

Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
AE	AE01. Durante la semana previa a la entrevista, es decir entre el ____ y el ____ de _____, ¿tuvo un trabajo, de al menos una hora, por el cual recibió o recibirá pago en dinero o en especies? 1. Sí → registre y continúe con Resultado 2. No → pasar a AE02	AE01-03	AE01-03		
	1				
	2				
	3				
	AE02. Independiente de lo que señaló anteriormente, durante esa misma semana: • ¿Realizó alguna venta o negocio? (no considerar venta de los bienes del hogar) • ¿Realizó alguna actividad agrícola, minera, artesanal u otra, para vender a futuro? • ¿Prestó servicio a algún familiar? 1. Sí → registre y continúe con Resultado 2. No → pasar a AE03	4			
	5				
	6				
	7				
	8				
	AE03. Aunque no trabajó durante esa semana, ¿tuvo en ese periodo algún empleo, negocio o actividad del cual estuvo ausente temporalmente por licencia, huelga, enfermedad, vacaciones u otra razón? 1. Sí 2. No	9			
	10				
11					
VP	VP03. La vivienda que ocupa este hogar es: 1. Propia (pagada totalmente) 2. Propia (pagando a plazo) 3. Arrendada (con contrato) 4. Arrendada (sin contrato) 5. Cedida por trabajo o servicio 6. Cedida por un pariente o amigo 7. Ocupada de hecho 8. Propiedad en litigio 9. Herencia o sucesión compartida				
Considerando que la entrevista con el encuestador EPF fue realizada durante la primera/segunda quincena del mes de _____					
Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
VS	VS01. Durante el mes _____ (mes anterior al de aplicación de entrevista), ¿el hogar ha tenido a su disposición otra vivienda además de la vivienda principal? 1. Sí 2. No → Pasar a LGI				

4 SUPERVISIÓN INDIRECTA

Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
VS	VS02. La vivienda secundaria que ocupa este hogar es: 1. Propia pagada totalmente 2. Propia pagándose 3. Arrendada (con contrato) → Pasar a LGI 4. Arrendada (sin contrato) → Pasar a LGI 5. Cédida por trabajo o servicio → Pasar a LGI 6. Cédida por un pariente o amigo → Pasar a LGI 7. Ocupación de hecho → Pasar a LGI 8. Propiedad en litigio → Pasar a LGI 9. Herencia o sucesión compartida				
	No lea las alternativas al informante. Si el hogar señala que el principal uso es descanso, recreación, vacacional y usos similares, continuar el módulo. VS03. ¿Cuál es el principal uso de esta vivienda? 1. Descanso, recreación o vacacional 2. Arriendo (percibe pago de arriendo por ella) → Pasar a LGI 3. Otro uso				
	VS04. Considerando los doce meses previos, es decir, entre los meses de _____ y _____ ¿por cuánto tiempo tuvo el hogar a su disposición la vivienda? (Número de meses)				
LGI					
		1. Si; 2. No	Observaciones		
Durante cada visita ¿El encuestador revisó las libretas donde los miembros del hogar registraban sus gastos, que se encontraban disponibles?					
Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
CL (HR)	CL01. ¿De qué manera recibió cada miembro del hogar la libreta de gastos individuales? (donde se registran los gastos diarios)	1			
	91. La recibió directamente del encuestador (Entrega LGI de forma personal)	2			
	92. La recibió a través de un miembro del hogar (Entrega LGI de forma indirecta)	3			
	93. No la recibió, pues rechazó contestar (No entrega LGI, rechaza contestar)	4			
	94. No la recibió, pues no estaba ubicable durante el periodo de aplicación (No entrega LGI, miembro del hogar no ubicable en el periodo de referencia)	5			
	95. No la recibió, pues es menor de 15 años (No aplica contestar LGI, miembro del hogar es menor de 15 años)	6			
	96. No la recibió, pues se encuentra impedido física o mentalmente para contestar (No aplica contestar LGI, miembro del hogar se encuentra impedido físico y/o mental para contestar)	7			
	88. No Sabe	8			
		9			
		10			
		11			
SUPERVISIÓN INDIRECTA 5					

Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación			
TC	TC01. En su vivienda principal o secundaria, ¿tuvo servicios de telecomunicaciones contratados en forma de pack? (Pack dúo, triple pack, entre otros) (1. Sí; 2. No)							
ED	(Realizar preguntas sólo a estudiantes) EF02. ¿Debió realizar algún pago por concepto de matrícula? (1. Sí; 2. No) EF10. ¿Posee alguna beca que le descuenta parte del arancel, colegiatura o mensualidad; o está adscrito a gratuidad? (1. Sí; 2. No) (Realizar pregunta sólo a estudiantes de educación superior) EF14. ¿Posee algún crédito de educación superior? (1. Sí; 2. No)	EF02	EF10	EF14	EF02	EF10	EF14	
		1						
		2						
		3						
		4						
		5						
		6						
		7						
		8						
		9						
		10						
		11						

LGR

LGR Aplicada	<input type="checkbox"/> 1. Sí → Continuar <input type="checkbox"/> 2. No → Pasar a LI	N° línea informante libreta	N° línea informante supervisión	Nombre informante supervisión	Coincide con informante de la libreta (1. Sí; 2. No)

Considerando que la entrevista con el encuestador EPF fue realizada durante la primera/segunda quincena del mes de _____

Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
Tabla 16	¿En los seis meses previos, es decir, entre los meses de _____ y _____, usted o algún miembro del hogar ha realizado algún viaje con fines vacacionales, de descanso, por motivos de salud y/o de estudios? (1. Sí; 2. No)				
Tabla 19	¿En los doce meses previos, es decir, entre los meses de _____ y _____, usted o algún miembro del hogar ha adquirido vehículos nuevos? INCLUIR: - Adquisición de vehículos para transporte personal, automóviles, furgones, camionetas, jeep, suv, bicicletas, motos de agua, motocicletas, cuádrimotos, etc. (1. Sí; 2. No)				
Tabla 19	(Sólo si han adquirido vehículos nuevos) ¿Entregó algún vehículo usado en parte de pago? (sólo en automotora o concesionarias) (1. Sí; 2. No)				

7

Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
Tabla 21	¿En los doce meses previos, es decir, entre los meses de _____ y _____ usted o algún miembro del hogar, ha debido ser hospitalizado en centros de salud, clínicas, hospitales u otro? INCLUIR: como hospitalizados, aquellos que hayan sido ingresados al menos una noche. (1. Si; 2. No)				

LI

LI Aplicada <input type="checkbox"/> 1. Sí → Continuar <input type="checkbox"/> 2. No → Pasar a LSF	N° línea informante libreta	N° línea informante supervisión	Nombre informante supervisión	Coincide con informante de la libreta (1. Si; 2. No)

Considerando que la entrevista con el encuestador EPF fue realizada durante la primera/segunda quincena del mes de _____

Módulo	Pregunta	Respuesta registrada		Respuesta supervisión		Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
TA	TA03 Durante el mes de _____ ¿Realizó pagos o tuvo descuentos por concepto de cotizaciones en salud? 1. Sí 2. No TA05. Durante el mes de _____ ¿Realizó pagos o tuvo descuentos por concepto de cotizaciones previsionales? 1. Sí 2. No	TA03	TA05	TA03	TA05		
		1					
		2					
		3					
		4					
		5					
		6					
		7					
		8					
		9					
		10					
		11					

Módulo	Pregunta	Respuesta registrada			Respuesta supervisión			Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
		TI02	TI04	TI10	TI02	TI04	TI10		
TI	TI02. Durante el mes de _____ ¿Realizó pagos por concepto de cotizaciones de salud? 1. Sí 2. No TI04. Durante el mes de _____ ¿Realizó pagos por concepto de cotizaciones previsionales? 1. Sí 2. No TI10. Su empleo principal, ¿lo realizó para su propio negocio, empresa o actividad por cuenta propia o como actividad profesional independiente? 1. Negocios por cuenta propia 2. Actividad profesional de forma independiente	1							
		2							
		3							
		4							
		5							
		6							
		7							
		8							
		9							
		10							
		11							

Módulo	Pregunta		Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
JU	JU01. El mes _____, ¿usted obtuvo ingresos por jubilación y/o pensión de vejez? 1. Sí 2. No	1				
		2				
		3				
		4				
		5				
		6				
		7				
		8				
		9				
		#				
		#				

Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación
TR	TR08. El mes _____, ¿recibió dinero desde el Estado y/o de alguna institución sin fines de lucro? 1. Sí, desde el Estado Por ejemplo: subsidio único de cesantía, subsidio único familiar, becas de estudio u otros subsidios del Estado. 2. Sí, desde alguna institución sin fines de lucro. 3. Sí, desde ambas 4. No	1			
		2			
		3			
		4			
		5			
		6			
		7			
		8			
		9			
		10			
		11			
IP	IP05. ¿Usted participa de alguna sociedad de personas? 1. Sí 2. No IP09. ¿Posee cuentas de ahorro o depósitos a plazo? 1. Sí 2. No	IP05	IP09	IP05	IP09
		1			
		2			
		3			
		4			
		5			
		6			
		7			
		8			
		9			
		10			
		11			
TE	TE09. El mes de _____, ¿transfirió especies a otros hogares? Considere solo productos nuevos 1. Sí 2. No	1			
		2			
		3			
		4			
		5			
		6			
		7			
		8			
		9			
		10			
		11			

LSF												
LSF Aplicada	<input type="checkbox"/>	1. Sí —————> Continuar 2. No —————> Finalizar	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="font-size: small;">N° línea informante libreta</th> <th style="font-size: small;">N° línea informante supervisión</th> <th style="font-size: small;">Nombre informante supervisión</th> <th style="font-size: small;">Coincide con informante de la libreta (1. Sí; 2. No)</th> </tr> <tr> <td style="height: 20px;"></td> <td></td> <td></td> <td></td> </tr> </table>	N° línea informante libreta	N° línea informante supervisión	Nombre informante supervisión	Coincide con informante de la libreta (1. Sí; 2. No)					
N° línea informante libreta	N° línea informante supervisión	Nombre informante supervisión	Coincide con informante de la libreta (1. Sí; 2. No)									
Considerando que la entre vista con el encuestador EPF fue realizada durante la primera/segunda quincena del mes de _____												
Módulo	Pregunta	Respuesta registrada	Respuesta supervisión	Resultado (1. Coincide; 2. Mejora de información; 3. Error de registro; 4. Verifique información; 5. No disponible; 6. Otro)	Observación							
CH	SF1a. El mes de _____, ¿usted o algún integrante del hogar pagó o debió haber pagado al menos un crédito hipotecario? (1. Sí; 2. No)											
TE	SF2a. El mes de _____, ¿usted o algún integrante del hogar pagó o debió haber pagado al menos una tarjeta de crédito de establecimiento (1. Sí; 2. No)											
CCA	SF3a. El mes de _____, ¿usted o algún integrante del hogar pagó o debió haber pagado al menos un crédito de consumo o un crédito automotriz? (1. Sí; 2. No)											
PB	SF4a. El mes de _____, ¿usted o algún integrante del hogar pagó o debió haber pagado al menos una cuenta corriente y/o líneas de crédito? (1. Sí; 2. No)											
CV	SF5a. El mes de _____, ¿usted o algún integrante del hogar tuvo en uso al menos una cuenta vista y/o cuenta RUT? (1. Sí; 2. No)											
TCB	SF6a. El mes de _____, ¿usted o algún integrante del hogar pagó o debió haber pagado tarjetas bancarias nacionales e internacionales? (1. Sí; 2. No)											
CU	SF7a. Durante el mes de _____, ¿Usted o algún integrante del hogar tuvo algún crédito universitario vigente? Recuerde que la frecuencia del cobro puede ser mensual, trimestral, semestral o anual. (1. Sí; 2. No)											
SE	SF8a. Durante el mes de _____, ¿Usted o algún integrante del hogar tuvo algún un seguro vigente? Recuerde que la frecuencia del cobro puede ser mensual, trimestral, semestral o anual. (1. Sí; 2. No)											
En caso de indicar tener algún servicio financiero, indagar lo siguiente:		1. Sí; 2. No	Observaciones									
Durante la entrevista ¿El encuestador le solicitó comprobantes de los servicios financieros que posee su hogar? (Créditos, cuentas corrientes, líneas de crédito, cuentas vista, cuentas RUT, tarjetas de crédito o seguros)												
SUPERVISIÓN INDIRECTA 11												

[illegible]

Anexo 10: Formulario de supervisión de Hoja de Ruta



Formulario de Supervisión Hoja de Ruta

Los datos que el INE solicita son **CONFIDENCIALES**
El artículo 29, Ley 17.374, establece el **SECRETO ESTADÍSTICO**.

NÚMERO DE FOLIO

■

Información de marco muestral

Sub-muestra	Región	Prov.	Comuna	Estrato	Distrito censal	Zona censal	Código de manzana	Vivienda	Hogar

Identificación de la vivienda

Nombre de calle o camino	Número	Block	Piso	Dpto. o pieza

Datos de la Visita

Código E. de Grupo	Código Encuestador	Código Supervisor	Fecha			Estado (1. Lograda; 2. Sin contacto; 3. Concertó cita; 4. Rechazo; 5. Unidad no elegible / elegibilidad desconocida; 6. No lograda por otros motivos)	Hora Inicio (HH:MM)	Hora Término (HH:MM)	N° línea informante	Nombre de pila informante
			Día	Mes	Año					
							:	:		

Registro de Contacto con el Hogar
 Verifique la información del módulo RC de Hoja de Ruta

N° (01)	FECHA			HORA hh/mm (05)	CÓDIGO DE VISITA HOJA DE RUTA (06)	CÓDIGO DE VISITA SUPERVISOR	CÓDIGO CORRESPONDE 1. SI 2. NO	OBSERVACIONES
	DÍA dd (02)	MES mm (03)	AÑO aa (04)					
1				:				
2				:				
3				:				
4				:				
5				:				
6				:				
7				:				
8				:				
9				:				
10				:				
11				:				
12				:				

[illegible]

Instituto Nacional de Estadísticas • Chile

Anexo 12: Perfiles y módulos asociados, sistema informático VIII EPF

Nivel 1	Nivel 2	Nivel 3	Administrador	Analista zona A	Digitador	Supervisor	Apoyo supervisión	Analista técnico	Encargado levantamiento	Encargado supervisión	Encargado codificación	Secretaría	Encargado área	Encargado regional	Encargado grupo	Encuestador	Encargado digitación	Revisor digitación	Digitador regional	Encargado equipo técnico	Revisor codificación	Dirección regional	Codificador	Dirección nacional	Analista zona B
Directorio.	Mantenedor directorio.		X						X	X			X	X											
	Log acceso de usuario.		X																						
Monitoreo del sistema.	Log errores de aplicación.		X																						
	Log de asignaciones.		X																						
	Log de asignaciones por usuario.		X																						
Administración.	Administrar usuarios.		X																						
	Administrar perfiles.		X																						
	Administrar menú.		X																						
	Mantenedor temáticas.		X																						
	Administrar contratos.		X																						
	Administrar códigos de barra.		X																						
	Mantenedor guía de despacho.		X						X				X												
	Mantenedor códigos de barra.		X						X				X												
Informaciones.	Noticias.		X	X		X	X		X	X	X		X	X	X	X	X				X	X		X	X
	Descarga bases interna.		X	X			X		X	X	X		X							X					X
Zona de consultas.	Consultas de usuario.		X	X	X	X	X		X	X	X		X	X	X	X				X	X	X		X	X
	Responder preguntas.		X	X					X	X			X	X						X					X
	Mis preguntas.		X	X	X	X	X		X	X	X		X	X	X	X				X	X	X		X	X
Trabajo de campo	Operación.	Ingresar Hoja de Ruta.	X	X		X	X		X	X			X	X	X		X	X	X	X					X
		Formulario de correcciones.	X			X	X		X	X			X	X	X										
		Emitir guía de despacho.	X						X				X	X	X					X					
		Recibir guía de despacho.	X						X				X	X											
		Solicitudes.	X						X				X	X	X										
		Biblioteca TDC.	X			X	X		X	X			X	X	X								X		X
		Control de asignaciones.	X						X				X	X									X		X
		Reasignación ER EG.	X										X												
	Asignación folios.	Encargado regional.	X						X				X												
		Encargado de grupo.	X						X					X											
		Encuestador.	X						X						X										
		Asignación de cargos.	X						X				X	X	X		X								
		Reporte asignación Tdc.	X			X	X		X				X	X											
	Desempeño operativo.	Asignación encargado de grupo.	X											X	X										
		Indicadores TdC.																							
		Indicadores enc. grupo.																							
		Indicadores regionales.																							
	Supervisión.	Indicadores.	X	X			X		X	X			X	X	X					X		X		X	X
		Directa - apertura.	X			X	X		X																
		Directa - aplicación.	X			X	X		X																
		Indirecta.	X			X	X		X																
		Ver directorio.	X			X	X		X	X			X	X											
		Descarga bases.	X			X			X											X					
		Hoja de Ruta.	X			X	X		X																
		Cargar reportes.	X			X			X																
		Descargar reportes.	X			X	X		X	X			X	X	X										
		Indicadores.	X			X			X													X		X	
Digitación	Gestión Digitación.	Asignación digitación.	X														X	X							
		Asignación revisión.	X														X	X							
		Directorio.	X			X	X										X								
		Recibir guía de despacho.	X														X	X							
		Reporte de asignación.	X														X	X							
		Reporte de digitación.	X		X												X	X							
		Descarga bases digitación.	X	X																X					X
	Registro de Personas en el Hogar.	Registro de Personas en el Hogar.	X		X				X								X	X							
		Libreta de Gastos Individuales.	X		X												X	X							
		Libreta de Gastos del Hogar.	X		X				X								X	X							
		Libreta de Gastos del Recuerdo.	X		X				X								X	X							
		Libreta de Ingresos.	X		X												X	X							
		Libreta de Servicios Financieros y Seguros.	X		X				X								X	X							
Capacitación	Biblioteca virtual		X	X		X	X		X	X	X		X	X	X	X				X	X	X	X	X	X

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Anexo 13: Clasificación CIUO-88 según Gran grupo y Subgrupo principal

Gran grupo	Subgrupo principal	Glosa Gran grupo y Subgrupos principales
1		Miembros del poder ejecutivo y de los cuerpos legislativos y personal directivo de la administración pública y de empresas.
	11	Miembros del poder ejecutivo y de los cuerpos legislativos y personal directivo de la administración pública.
	12	Directores de empresa.
	13	Gerentes de empresa.
2		Profesionales, científicos e intelectuales.
	21	Profesionales de las ciencias físicas, químicas y matemáticas y de la ingeniería.
	22	Profesionales de las ciencias biológicas, la medicina y la salud.
	23	Profesionales de la enseñanza.
	24	Otros profesionales científicos e intelectuales.
3		Técnicos y profesionales de nivel medio.
	31	Técnicos y profesionales de nivel medio de las ciencias físicas y químicas, la ingeniería y afines.
	32	Técnicos y profesionales de nivel medio de las ciencias biológicas, la medicina y la salud.
	33	Maestros e instructores de nivel medio.
	34	Otros técnicos y profesionales de nivel medio.
4		Empleados de oficina.
	41	Oficinistas.
	42	Empleados en trato directo con el público.
5		Trabajadores de los servicios y vendedores de comercios y mercados.
	51	Trabajadores de los servicios personales y de los servicios de protección y seguridad.
	52	Modelos, vendedores y demostradores.
6		Agricultores y trabajadores calificados agropecuarios y pesqueros.
	61	Agricultores y trabajadores calificados de explotaciones agropecuarias, forestales y pesqueras con destino al mercado.
	62	Trabajadores agropecuarios y pesqueros de subsistencia.
7		Oficiales, operarios y artesanos de artes mecánicas y de otros oficios.
	71	Oficiales y operarios de las industrias extractivas y de la construcción.
	72	Oficiales y operarios de la metalurgia, la construcción mecánica y afines.
	73	Mecánicos de precisión, artesanos, operarios de las artes gráficas y afines.
	74	Otros oficiales, operarios y artesanos de artes mecánicas y de otros oficios.
8		Operadores de instalaciones y máquinas y montadores.
	81	Operadores de instalaciones fijas y afines.
	82	Operadores de máquinas y montadores.
	83	Conductores de vehículos y operadores de equipos pesados móviles.
9		Trabajadores no calificados.
	91	Trabajadores no calificados de ventas y servicios.
	92	Peones agropecuarios, forestales, pesqueros y afines.
	93	Peones de la minería, la construcción, la industria manufacturera y el transporte.
10		Otros grupos no identificados.

Fuente: Instituto Nacional de Estadísticas (INE), CIUO-88.

Anexo 14: Módulos evaluados por la grilla técnica de mínimos de calidad VIII EPF

El siguiente anexo especifica la distribución de puntajes por libreta y la unidad de análisis para su evaluación. Se consideró como unidad mínima de análisis el hogar, al que se evaluó y asignó puntaje de forma individual, no obstante, se evaluaron a todos los integrantes que componían cada hogar.

Libretas	Ítems a considerar por módulo o pregunta	Unidad de análisis	Puntaje
Registro de Personas en el Hogar	Parentesco con el jefe de hogar.	Nº integrantes	15
	Edad.	Nº integrantes	
	Sexo.	Nº integrantes	
	Nivel de educación e identificación de los estudiantes.	Nº integrantes	
	Condición de actividad económica.	Nº integrantes	
Libreta de Gastos Individuales	Libretas contestadas con un mínimo de 4 días de registro y al menos 1 día de registro de fin de semana.	Nº integrantes elegibles	29
Libreta de Gastos del Hogar	Gastos básicos del Hogar (agua y electricidad).	Hogar	25
	Módulo de Telecomunicaciones.	Hogar	
	Módulo de Educación (si corresponde)	Nº estudiantes	
Libreta de Gastos del Recuerdo	Tablas 1 a la 21.	Hogar	16
Libreta de Ingresos	Asalariados (Remuneración neta/Monto bruto honorarios).	Nº ocupados dependientes	10
	Cuenta Propia (Monto líquido o promedio mensual).	Nº ocupados independientes	
Libreta de Gastos en Servicios Financieros y Seguros	Crédito hipotecario.	Nº titulares	5
	Crédito de consumo y automotriz.	Nº titulares	
	Tarjetas de crédito bancarias.	Nº titulares	
	Seguros.	Nº titulares	
	Cuentas corrientes y líneas de crédito.	Nº titulares	
	Cuentas vista y /o cuenta RUT.	Nº titulares	
Puntaje total	100		

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Anexo 15: Matriz de exigencia para la elección del vecino cercano para la imputación de días faltantes de Libretas de Gastos Individuales

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
Submuestra operativa	Submuestra operativa	Submuestra operativa	Submuestra operativa	Submuestra operativa	Submuestra operativa	Submuestra operativa	Mes
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar
Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad
Edad	Edad	Edad	Edad	Edad	Edad	Edad	Edad
CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)
CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)
Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
RPC	RPC	RPC	RPC	RPC	RPC	Región	Región
Manzana	Manzana	Manzana	Manzana	Manzana			
FOLIO	FOLIO	FOLIO	FOLIO				
Número de personas en el hogar	Número de personas en el hogar	Tramo de la cantidad de personas en el hogar					
NRO_RPH							

Nivel 9	Nivel 10	Nivel 11	Nivel 12	Nivel 13	Nivel 14	Nivel 15	Nivel 16
Mes	Mes	Mes	Agrupación de meses	Agrupación de meses (2)			
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar
Escolaridad	Escolaridad	Niveles educativos	Niveles educativos	Niveles educativos	Niveles educativos	Niveles educativos	Niveles educativos
Edad	Tramos de edad (1)	Tramos de edad (1)	Tramos de edad (1)	Tramos de edad (1)	Tramos de edad (1)	Tramos de edad (2)	Tramos de edad (2)
CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)	CIUO (1 dígito)		
CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	CISE (2 categorías)	Ocupado
Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Macrozona	Macrozona	Macrozona	Macrozona	Macrozona	Macrozona	Macrozona	Macrozona

Nivel 17	Nivel 18	Nivel 19	Nivel 20	Nivel 21	Nivel 22	Nivel 23
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	
Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar	Ingreso disponible del hogar
Niveles educativos	Niveles educativos	Niveles educativos	Niveles educativos			
Tramos de edad (2)	Tramos de edad (2)	Tramos de edad (2)		Tramos de edad (2)		
Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos	Administrador de gastos
CSE	CSE					
Macrozona						

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Anexo 16: Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de trabajadores asalariados

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)
Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad
Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)
Edad	Edad	Edad	Edad	Edad	Tramo de edad (1)	Tramo de edad (2)	Tramo de edad (3)
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Manzana	Comuna	Región	Macrozona	Zona	Zona	Zona	Zona
Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar

Nivel 9	Nivel 10	Nivel 11	Nivel 12	Nivel 13	Nivel 14	Nivel 15	Nivel 16
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)
Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)	Tramo de escolaridad (2)	Escolaridad	Escolaridad	Escolaridad	Escolaridad
Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
CISE (3 categorías)	CISE (3 categorías)	CISE (3 categorías)	CISE (3 categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)
Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Edad	Tramo de edad (1)	Tramo de edad (2)	Tramo de edad (3)
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Zona	Zona	Zona	Zona	Zona	Zona	Zona	Zona
Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar	Presencia de menores de 6 años en el hogar				

Nivel 17	Nivel 18	Nivel 19	Nivel 20	Nivel 21	Nivel 22	Nivel 23	Nivel 24
CIUO (2 Dígitos)	CIUO (1 Dígito)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)
Tramo de escolaridad (1)	Tramo de escolaridad (2)	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)
Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
CISE (3 categorías)	CISE (3 categorías)						
Tramo de edad (3)	Tramo de edad (3)	Edad	Tramo de edad (1)	Tramo de edad (2)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Zona	Zona	Zona	Zona	Zona	Zona	Zona	Zona

Nivel 25	Nivel 26	Nivel 27	Nivel 28	Nivel 29	Nivel 30	Nivel 31	Nivel 32
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)			
Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)	Escolaridad	Escolaridad	Escolaridad
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Tramo de edad (1)	Tramo de edad (2)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Edad	Tramo de edad (1)	Tramo de edad (2)
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Zona	Zona	Zona	Zona	Zona	Zona	Zona	Zona

Nivel 33	Nivel 34	Nivel 35
Escolaridad	Tramo de escolaridad (1)	Escolaridad
Sexo	Sexo	Sexo
Tramo de edad (3)	Tramo de edad (3)	
CSE	CSE	CSE
Zona	Zona	

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Anexo 17: Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de trabajadores honorarios

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)
Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal
Escolaridad	Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)	Tramo de escolaridad (2)
Edad	Tramo de edad (1)	Tramo de edad (2)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
CSE	CSE	CSE	CSE	CSE	CSE	CSE
Macrozona	Zona	Zona	Zona	Zona	Zona	Zona

Nivel 8	Nivel 9	Nivel 10	Nivel 11	Nivel 12	Nivel 13	Nivel 14
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)				
Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)	Escolaridad	Escolaridad	Escolaridad	Escolaridad
Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Edad	Tramo de edad (1)	Tramo de edad (2)	Tramo de edad (3)
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
CSE	CSE	CSE	CSE	CSE	CSE	CSE
Zona	Zona	Zona	Zona	Zona	Zona	Zona

Nivel 15	Nivel 16	Nivel 17	Nivel 18
Tramo de escolaridad (1)	Escolaridad	Escolaridad	Escolaridad
Tramo de edad (3)	Tramo de edad (1)	Tramo de edad (1)	
Sexo	Sexo		
CSE	CSE		
Zona			

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Anexo 18: Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de trabajadores de negocios por cuenta propia

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)
CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)	Tramo de escolaridad (2)
Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal
Edad	Edad	Tramo de edad (1)	Tramo de edad (2)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Región	Zona	Zona	Zona	Zona	Zona	Zona	Zona
Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar

Nivel 9	Nivel 10	Nivel 11	Nivel 12	Nivel 13	Nivel 14	Nivel 15	Nivel 16
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)	CIUO (2 Dígitos)	CIUO (2 Dígitos)
CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)	Tramo de escolaridad (2)	Escolaridad	Escolaridad
Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal		
Edad	Tramo de edad (1)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (1)	Tramo de edad (2)
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Zona	Zona	Zona	Zona	Zona	Zona	Zona	Zona

Nivel 17	Nivel 18	Nivel 19	Nivel 20	Nivel 21	Nivel 22	Nivel 23	Nivel 24
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)
CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)				
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)	Tramo de escolaridad (2)	Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (1)	Tramo de edad (2)	Tramo de edad (3)	Tramo de edad (3)
Zona	Zona	Zona	Zona	Zona	Zona	Zona	Zona

Nivel 25	Nivel 26	Nivel 27	Nivel 28	Nivel 29	Nivel 30	Nivel 31	Nivel 32
CIUO (1 Dígito)							
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Tramo de escolaridad (2)	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad
Tramo de edad (3)	Edad	Tramo de edad (1)	Tramo de edad (2)	Edad	Tramo de edad (1)		
CSE	CSE	CSE	CSE	CSE	CSE	CSE	
Zona	Zona	Zona	Zona				

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Anexo 19: Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de profesionales independientes

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)	Escolaridad
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)	CIUO (1 Dígito)	CIUO (1 Dígito)	CIUO (2 Dígitos)
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal
CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)
Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	Presencia de menores de 15 años en el hogar	
Edad	Tramo de edad (1)	Tramo de edad (2)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Edad
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Comuna	Zona	Zona	Zona	Zona	Zona	Zona	Zona

Nivel 9	Nivel 10	Nivel 11	Nivel 12	Nivel 13	Nivel 14	Nivel 15	Nivel 16
Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)	Escolaridad	Escolaridad	Tramo de escolaridad (1)	Escolaridad
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)	CIUO (1 Dígito)	CIUO (2 Dígitos)	CIUO (1 Dígito)	CIUO (1 Dígito)	CIUO (2 Dígitos)
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	
CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)	CISE (todas categorías)				
Tramo de edad (1)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Edad
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Zona	Zona	Zona	Zona	Zona	Zona	Zona	Zona

Nivel 17	Nivel 18	Nivel 19	Nivel 20	Nivel 21	Nivel 22	Nivel 23	Nivel 24
Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)	Tramo de escolaridad (2)	Escolaridad	Escolaridad	Escolaridad
CIUO (2 Dígitos)	CIUO (2 Dígitos)	CIUO (1 Dígito)	CIUO (1 Dígito)	CIUO (1 Dígito)			
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Tramo de edad (1)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (3)	Edad	Tramo de edad (1)	Tramo de edad (2)
CSE	CSE	CSE	CSE	CSE	CSE	CSE	CSE
Zona	Zona	Zona	Zona	Zona	Zona	Zona	Zona

Nivel 25	Nivel 26	Nivel 27	Nivel 28	Nivel 29	Nivel 30	Nivel 31
Escolaridad	Tramo de escolaridad (1)	Escolaridad	Tramo de escolaridad (1)	Escolaridad	Escolaridad	Escolaridad
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	
Tramo de edad (3)	Tramo de edad (3)	Tramo de edad (1)	Tramo de edad (3)			
CSE	CSE	CSE	CSE	CSE		
Zona	Zona					

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Anexo 20: Matriz de exigencia para la elección del vecino cercano para la imputación de ingresos de jubilaciones o pensiones de vejez

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Escolaridad	Tramo de escolaridad (1)
Tramo de edad para jubilados (1)	Tramo de edad para jubilados (2)	Tramo de edad para jubilados (2)	Tramo de edad para jubilados (2)	Tramo de edad para jubilados (2)	Tramo de edad para jubilados (2)	Tramo de edad para jubilados (2)
Sistema previsional	Sistema previsional	Sistema previsional	Sistema previsional	Sistema previsional	Sistema previsional	Sistema previsional
Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal	Sustentador principal
CSE	CSE	CSE	CSE	CSE	CSE	CSE
Manzana	Manzana	Comuna	Región	Macrozona	Zona	Zona

Nivel 8	Nivel 9	Nivel 10	Nivel 11	Nivel 12	Nivel 13
Sexo	Sexo	Sexo	Sexo	Sexo	Sexo
Tramo de escolaridad (2)	Escolaridad	Tramo de escolaridad (1)	Escolaridad	Escolaridad	Escolaridad
Tramo de edad para jubilados (2)	Tramo de edad para jubilados (1)	Tramo de edad para jubilados (2)	Edad	Tramo de edad para jubilados (1)	
Sistema previsional	Sistema previsional	Sistema previsional			
Sustentador principal					
CSE	CSE	CSE	CSE	CSE	CSE
Zona	Zona	Zona	Zona	Zona	Zona

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Anexo 21: Algoritmos de ajuste al mes de referencia VIII EPF

Si bien para efectos de la construcción del factor de ajuste temporal, las submuestras del trabajo de campo de la encuesta van desde la submuestra 25 a la 48, para referirse a las mismas quincenas, se trabaja a nivel de equipos considerando submuestras operativas de 1 a 24.

La tabla 61 muestra la cantidad de días que abarca cada una de las submuestras asociándolas al mes de referencia.

Tabla 61: Número de días que abarca cada una de las submuestras operativa

Período comprendido	N° de días de la submuestra (ts)	N° de submuestra (s)	N° de submuestra operativa de referencia	Mes de referencia (m)
01/07/15-15/07/15	15	1	-	-
16/07/15-31/07/15	15	2	-	
01/08/15-15/08/15	15	3	-	-
16/08/15-31/08/15	16	4	-	
01/09/15-16/09/15	16	5	-	-
17/09/15-30/09/15	15	6	-	
01/10/15-15/10/15	14	7	-	-
16/10/15-31/10/15	14	8	-	
01/11/15-15/11/15	15	9	-	-
16/11/15-30/11/15	16	10	-	
01/12/15-15/12/15	15	11	-	-
16/12/15-31/12/15	15	12	-	
01/01/16-15/01/16	15	13	-	-
16/01/16-31/01/16	16	14	-	
01/02/16-14/02/16	15	15	-	-
15/02/16-29/02/16	15	16	-	
01/03/16-15/03/16	16	17	-	-
16/03/16-31/03/16	15	18	-	
01/04/16-15/04/16	16	19	-	-
16/04/16-30/04/16	15	20	-	
01/05/16-15/05/16	15	21	-	-
16/05/16-31/05/16	15	22	-	
01/06/16-15/06/16	16	23	-	-
16/06/16-30/06/16	15	24	-	
01/07/16-15/07/16	15	25	1	jul-16
16/07/16-31/07/16	15	26	2	
01/08/16-16/08/16	15	27	3	ago-16
17/08/16-31/08/16	16	28	4	
01/09/16-15/09/16	16	29	5	sep-16
16/09/16-30/09/16	15	30	6	
01/10/16-15/10/16	15	31	7	oct-16
16/10/16-31/10/16	14	32	5	
01/11/16-15/11/16	15	33	8	nov-16

Tabla 61: Número de días que abarca cada una de las submuestras operativa

Período comprendido	N° de días de la submuestra (ts)	N° de submuestra (s)	N° de submuestra operativa de referencia	Mes de referencia (m)
16/11/16-30/11/16	16	34	10	
01/12/16-15/12/16	15	35	11	dic-16
16/12/16-31/12/16	15	36	12	
01/01/17-16/01/17	15	37	13	ene-17
17/01/17-31/01/17	16	38	14	
01/02/17-14/02/17	15	39	15	feb-17
15/02/17-28/02/17	15	40	16	
01/03/17-15/03/17	16	41	17	mar-17
16/03/17-31/03/17	15	42	18	
01/04/17-15/04/17	16	43	19	abr-17
16/04/17-30/04/17	15	44	20	
01/05/17-16/05/17	15	45	21	may-17
17/05/17-31/05/17	15	46	22	
01/06/17-15/06/17	16	47	23	jun-17
16/06/17-30/06/17	15	48	24	

Fuente: Instituto Nacional de Estadísticas (INE), VIII Encuesta de Presupuestos Familiares (EPF).

Número de días que abarca cada una de las submuestras de referencia de la muestra

Considerando que la muestra es representativa a nivel mensual, y que además, los distintos instrumentos de recolección de información tienen asociados distintos períodos de referencia, se requiere ajustar los gastos a un período de referencia común para presentar la información (en el caso de la EPF, mes de referencia).

Todos los gastos recopilados en los distintos instrumentos son mensualizados y llevados al mes de referencia de cada submuestra de trabajo de campo según los algoritmos que a continuación se describen.

Algoritmos de ajuste del gasto de un hogar al mes de referencia.

Los siguientes algoritmos explican la conformación del gasto mensual del hogar k , en el mes de referencia m a partir de cada submuestra s .

El hogar k , se presenta clasificado dentro de un estrato h , manzana i , vivienda j .

Gasto mensual en el producto c , del hogar k , en el mes de referencia m .

Sea s = submuestra de referencia para el trabajo de campo de la VIII EPF, definido por:

$s = \{1, 2, 3, \dots, 48\}$ siendo el trabajo de campo oficial de la encuesta el período comprendido entre las submuestras 25 y 48.

Sea t_s = número de días de la submuestra.

Sea m = mes de referencia del trabajo de campo de cada submuestra.

$$m = s - D \cdot (s \cdot 0,5 + 11,5) - (1 - D) \cdot (s \cdot 0,5 + 12) \text{ , donde } D = \begin{cases} 1, & \text{si } s \text{ es impar} \\ 0, & \text{EOC} \end{cases}$$

$m = \{1, 2, 3, \dots, 12\}$, donde junio de 2016 corresponde a $m=1$.

$$G_{mkc} = F_q \cdot G_{mkc}^q + \left[\sum_l F_{rl} \cdot G_{mkc}^{rl} \right] + F_o \cdot G_{mkc}^o + F_e \cdot G_{mkc}^e + G_{mkc}^m$$

Donde $\{3, 6, 12\}$ (meses)

Siendo,

$$F_q = \frac{t_s + D \cdot t_{s+1} + (1 - D) \cdot t_{s-1}}{t_s} \text{ ; donde } D = \begin{cases} 1, & \text{si } s \text{ es impar} \\ 0, & \text{EOC} \end{cases}$$

$$F_q = \frac{\text{n}^\circ \text{ de días del mes } m}{\text{n}^\circ \text{ de días de la quincena } t} \text{ ; factor de ajuste que lleva los gastos quincenales de la}$$

submuestra s al mes m .

$$F_{rl} = \frac{t_s + D \cdot t_{s+1} + (1 - D) \cdot t_{s-1}}{t_s^l} = \frac{t_s + D \cdot t_{s+1} + (1 - D) \cdot t_{s-1}}{(1 - D) \cdot t_{s-2 \text{ } l-1}}$$

$$F_{rl} = \frac{\text{n}^\circ \text{ de días del mes } m}{\text{n}^\circ \text{ de días del recuerdo } l(t_s^l)} \text{ ; factor de ajuste que lleva los gastos del período de}$$

recuerdo l al mes m .

$$F_o = \frac{1}{\text{n}^\circ \text{ de meses que abarca la boleta}} \text{ ; factor de ajuste que lleva los gastos del hogar}$$

que se pagan de forma periódica no mensual al mes m .

$$F_e = \frac{\text{n}^\circ \text{ de veces que realiza el pago durante el año académico}}{12} \text{ ; factor de ajuste que}$$

lleva los gastos de educación del hogar que se pagan de forma periódica mensual o no al mes de referencia.

G_{mkc}^q = Gasto quincenal en el producto c , del hogar k en la submuestra del mes m .

G_{mkc}^{rl} = Gasto de recuerdo del periodo " l " en el producto c , del hogar k en el mes m .

G_{mkc}^o = Gasto del hogar en el producto c , del hogar k en el mes m que el hogar realiza de forma periódica con una periodicidad distinta a la mensual.

G_{mke}^e = Gastos en educación en el producto c , del hogar k en el mes m que el hogar puede realizar de forma mensual o con una periodicidad distinta.

G_{mke}^m = Gasto mensual en el producto c , del hogar k en el mes m .

Para obtener el gasto total mensualizado del hogar k , en el mes de referencia m , debemos realizar la sumatoria de todos los productos consumidos por el hogar en el mes de referencia.

$$G_{mk} = \sum_{c=1}^n G_{mke}$$

Donde c es cada uno de los bienes y servicios adquiridos por el hogar k en el mes de referencia.

c	CCIF
1	01.1.1.01.01
2	01.1.1.01.02
⋮	⋮
n	12.7.1.08.03