

ENCUESTA FINANCIERA DE HOGARES 2024



Informe Metodológico



Septiembre 2025



Contenido

	Pág.
I. Resumen ejecutivo	3
II. Introducción	4
III. Cuestionario EFH 2024	5
III.1. Definición de hogar y persona a entrevistar	5
III.2. Tópicos del Cuestionario	6
III.3. Modificaciones al Cuestionario	10
IV. Diseño Muestral	14
IV.1. Características del diseño muestral	14
IV.2. Muestra Panel	15
IV.3. Muestra Refresco	15
IV.3.1. Marco muestral	16
IV.3.2. Estratificación	16
IV.3.3. Tamaño muestral	17
IV.3.4. Primera etapa de muestreo: selección de comunas	18
IV.3.5. Segunda etapa de muestreo: selección de manzanas	19
IV.3.6. Tercera etapa de muestreo: selección de viviendas	22
V. Trabajo de Campo	23
V.1. Prueba piloto	23
V.2. Resultados del trabajo de campo	24
V.3. Duración de la entrevista	25
VI. Factores de Expansión	26
VI.1. Factores de expansión de la muestra refresco EFH 2024	26
VI.1.1. Primera etapa de muestreo: comunas	26
VI.1.2. Segunda etapa de muestreo: manzanas	27
VI.1.3. Tercera etapa de muestreo: viviendas	27
VI.1.4. Probabilidad total y factor de expansión directo	28
VI.1.5. Ajuste por sub o sobre representación	28
VI.2. Factores de expansión muestra panel	29
VI.3. Factores de expansión de corte transversal EFH 2024	31

VI.4.	Calibración.....	32
VII.	No respuesta parcial e imputaciones.....	34
VII.1.	Estimación Imputación MICE en la EFH 2024.....	34
VII.2.	Análisis de valores faltantes.....	35
VII.3.	Transformación de variables y valores iniciales.....	36
VII.4.	Evaluación y consistencia de las imputaciones.....	36
VIII.	Estimación de la varianza.....	37
VIII.1.	Métodos de estimación de varianza en encuestas complejas.....	37
VIII.1.1.	Método de linealización de Taylor.....	37
VIII.1.2.	Métodos de réplicas.....	38
VIII.2.	Estimación de varianza en la EFH 2024.....	38
IX.	Referencias.....	40
Anexo N°1:	Error relativo de la muestra de refresco de la EFH 2024.....	43
Anexo N°2:	Lista de comunas EFH 2024.....	44
Anexo N°3:	Reglas de Rubin.....	45
	Media de una variable imputada.....	45
	Varianza de una variable imputada.....	45
	Intervalos de confianza.....	46
	Regresiones y significancia estadística.....	46
Anexo N°4:	Análisis de valores faltantes EFH 2024.....	48
Anexo N°5:	Transformación Box-Cox.....	56

I. Resumen ejecutivo

La Encuesta Financiera de Hogares (EFH), realizada por el Banco Central de Chile desde 2007, tiene como objetivo generar información detallada sobre el balance financiero de los hogares y hacer seguimiento de su evolución en el tiempo. El análisis de estos datos permite comprender en profundidad el comportamiento financiero de los hogares, lo que resulta relevante para los objetivos de estabilidad financiera y de precios del Banco Central.

Este documento presenta una descripción general de los principales aspectos metodológicos de la versión 2024 de la EFH, cuya estructura se detalla a continuación.

En primer lugar, se presenta el cuestionario de la EFH 2024, con los contenidos de cada módulo que conforma la encuesta y las definiciones clave que determinan el hogar y la persona entrevistada. También se describen las principales modificaciones respecto al cuestionario de la EFH 2021, que fue adaptado para su aplicación telefónica en el contexto de la pandemia.

En segundo lugar, se describe el diseño muestral implementado en la EFH 2024. Al tratarse de una encuesta tipo panel rotativo, cada ronda incluye una muestra panel (hogares entrevistados en la ronda anterior) y una muestra de refresco (hogares entrevistados por primera vez en la ronda actual). Se detalla el diseño muestral de la muestra de refresco y se resume el proceso de contacto con la muestra panel previo al levantamiento de la EFH 2024.

En tercer lugar, se presentan las principales características del trabajo de campo y sus resultados. El proceso culminó con 4.649 entrevistas exitosas: 1.801 de la muestra panel y 2.848 de la muestra de refresco.

A continuación, se explica el cálculo de los factores de expansión, necesarios para que las estadísticas derivadas de las 4.649 encuestas sean representativas de los hogares urbanos a nivel nacional. El cálculo se realiza por separado para cada muestra y posteriormente se calibra a la población total. Para obtener estimaciones robustas, se detalla el proceso de estimación de varianzas, junto con la metodología empleada y sus principales resultados.

Finalmente, se expone la metodología de imputación aplicada y sus resultados más relevantes.

II. Introducción

Los hogares son actores fundamentales tanto en el sistema financiero como en la economía en general, y sus decisiones financieras están estrechamente relacionadas con otras decisiones económicas, como las vinculadas al mercado laboral.

El análisis de las finanzas de los hogares permite comprender con mayor profundidad su comportamiento dentro del sistema financiero y desarrollar herramientas para monitorear su situación financiera. Ambas dimensiones son relevantes para los objetivos de estabilidad financiera y control de precios del Instituto Emisor.

En este contexto, desde 2007 el Banco Central de Chile realiza la Encuesta Financiera de Hogares (EFH), cuyo propósito es generar información detallada sobre el balance financiero de los hogares.

Entre 2007 y 2024 se han realizado nueve rondas de la EFH. Las ediciones de 2007, 2011-2012, 2014, 2017, 2021 y 2024 tienen representatividad a nivel nacional urbano, mientras que las de 2008, 2009 y 2010 se limitaron a la zona urbana de la Región Metropolitana.

La EFH, al igual que otras encuestas financieras¹, presenta características que requieren un tratamiento metodológico específico. En primer lugar, se aplica bajo un esquema de panel rotativo, con un diseño muestral complejo que incorpora el sobremuestreo de los grupos de mayor riqueza. En segundo lugar, busca recopilar información sensible para los hogares, como la información financiera. Por último, los niveles esperados de no respuesta hacen que la planificación y la estrategia de contacto con los entrevistados sean determinantes para la calidad de los resultados.

El sexto levantamiento nacional de la EFH se llevó a cabo durante el segundo semestre de 2024, logrando un total de 4.649 entrevistas a hogares urbanos distribuidos en todas las regiones del país.

Este documento expone el marco conceptual y metodológico de la EFH 2024, junto con la descripción del proceso operativo que la respalda. La estructura del documento es la siguiente: la sección III describe el cuestionario utilizado en esta versión y las modificaciones respecto a la ronda anterior; la sección IV detalla el diseño muestral implementado; y la sección V presenta los resultados del trabajo de campo. Las secciones VI, VII y VIII abordan los aspectos técnicos relacionados con el cálculo de los factores de expansión, la estimación de varianzas y el tratamiento de la información faltante, respectivamente.

¹ Por ejemplo, en España (véase Bover, 2004, 2008) y en Estados Unidos (véase Kennickell y Woodburn, 1997).

III. Cuestionario EFH 2024

Esta sección presenta el Cuestionario EFH 2024, comenzando por la identificación de la unidad de análisis —el hogar— y de la persona entrevistada dentro de cada uno. A continuación, se describen en detalle los módulos que componen el cuestionario, y se resumen las principales modificaciones introducidas en esta versión respecto a la ronda anterior.

III.1. Definición de hogar y persona a entrevistar

En la EFH, se define hogar como: i) un grupo de personas que viven en una misma vivienda y comparten un presupuesto de alimentación, o ii) una persona que vive sola y se provee de alimentación². También se considera parte del hogar a los integrantes que se encuentren transitoriamente fuera de la vivienda, siempre que la ausencia no supere los seis meses. Esto incluye, por ejemplo, personas que estén de vacaciones, realizando trabajos temporales, estudiando fuera del hogar, entre otros casos.

Tras el primer contacto con el hogar, se debe identificar a la persona más idónea para responder la encuesta en representación del hogar. En primer lugar, se intenta entrevistar al jefe de hogar. Si no está disponible, se busca a otra persona que tenga el mayor conocimiento sobre las finanzas, ingresos, gastos y deudas del hogar. Si esta persona tampoco está disponible y el hogar tiene deudas, se entrevista a quien posea la mayor deuda en monto. En caso contrario, se selecciona a la persona con más activos financieros y/o inmobiliarios. Si ninguna de las anteriores está disponible, se opta por quien aporte el mayor ingreso al hogar. En todos los casos, la persona entrevistada debe tener 18 años o más.

En el caso de la muestra panel, se verifica si la persona entrevistada en la ronda anterior (reportante) continúa residiendo en el hogar. Si está disponible, se mantiene como entrevistado. En caso contrario, se identifica a otra persona calificada para responder la encuesta, siguiendo los criterios previamente descritos.

² La definición de hogar de la EFH obedece a estándares internacionales, ver Organización para la Cooperación y el Desarrollo Económico OCDE, “Glosario de términos estadísticos” (<https://stats.oecd.org/glossary>).

III.2. Tópicos del Cuestionario

El cuestionario de la EFH 2024 está compuesto por **15 módulos temáticos**, que recogen información a distintos niveles: individual (para cada integrante del hogar), del hogar en su conjunto, o exclusivamente del entrevistado. La siguiente tabla resume el nombre de cada módulo y el nivel de información que se levanta:

Tabla 1
Descripción de módulos del cuestionario y nivel de información levantada

Módulo	Nivel de Información
A. Composición del hogar	Individual
B. Educación	Individual
C. Situación laboral	Individual
D. Medios de pago	Hogar
E1. Activos reales y deuda hipotecaria (vivienda principal)	Hogar
E2. Activos reales y deuda hipotecaria (otras propiedades)	Hogar
F. Deudas no hipotecarias	Hogar
G. Carga financiera y restricciones de crédito	Hogar
H. Vehículos y otros activos reales	Hogar
I. Activos financieros, pensiones y seguros	Hogar/Entrevistado
J. Ingresos relacionados	Individual
L. Otros ingresos/Ingresos no relacionados	Hogar
K1. Expectativas, percepciones y alfabetización financiera	Entrevistado
K2. Decisiones financieras del hogar	Hogar
K3. Gasto	Hogar

A continuación, se describe brevemente el contenido de cada módulo:

A. Composición del hogar e identificación de entrevistado

Identifica a todos los miembros del hogar, al jefe de hogar y la relación de parentesco entre ellos. Se recoge información sobre género, edad, nacionalidad y estado civil (solo para personas de 15 años o más).

B. Educación

Recoge los antecedentes educacionales de todos los miembros del hogar de 15 años o más. El entrevistado responde por todos. Se consulta por el nivel educacional alcanzado, asistencia actual a establecimientos educativos y pagos de colegiatura.

C. Situación laboral

Indaga la situación laboral de cada miembro del hogar de 15 años o más. Se consulta si trabajó la semana anterior, si está desocupado o inactivo, y si tiene actividades secundarias remuneradas. También se pregunta por jubilación, tiempo sin trabajar, características del empleo actual y afiliación previsional.

D. Medios de pago

Investiga la tenencia y uso de medios de pago a nivel de hogar: cuentas corrientes, cuentas vistas, tarjetas de crédito bancarias y no bancarias, efectivo, cheques, pagos en línea, entre otros.

E. Activos reales y deuda hipotecaria

Se divide en dos submódulos:

- **E1:** Vivienda principal. Tipo, propiedad, fecha de compra, precio, valor estimado, estructura de financiamiento, subsidios, créditos y características del crédito vigente.
- **E2:** Otras propiedades. Hasta tres propiedades de mayor valor. Año de adquisición, valor estimado, financiamiento y deuda asociada.

F. Deudas no hipotecarias

Incluye tarjetas de crédito, líneas de crédito, créditos de consumo, automotrices, educacionales, prendarios, entre otros. Se identifican las tres deudas de mayor cuantía y sus características.

G. Percepción de la carga financiera y restricciones al crédito

Evalúa la percepción del hogar sobre su nivel de endeudamiento, capacidad de enfrentar contingencias, situaciones de impago y acceso al crédito.

H. Vehículos y otros activos reales

Cuantifica activos como vehículos, maquinaria, obras de arte, ganado, entre otros.

I. Activos financieros, pensiones y seguros

Caracteriza la tenencia y montos en instrumentos financieros, previsionales y seguros, incluyendo ahorro informal, hábitos de ahorro y seguros voluntarios. Además, se incorpora la medición de los flujos de inversión realizados en el último año, con el fin de captar la actividad y movimientos recientes de los activos.

J. Ingresos relacionados

Para integrantes del hogar con trabajo remunerado o mayores de 15 años. Se declara el ingreso mensual por ocupación principal y actividades secundarias.

L. Otros ingresos/Ingresos no relacionados

Captura ingresos no asociados a ocupaciones remuneradas: pensiones, subsidios, ganancias de capital, indemnizaciones, arriendos, entre otros.

Adicionalmente, en la versión 2024 de la EFH se incorporaron tres módulos nuevos:

K1. Expectativas y percepciones

Este módulo está dirigido exclusivamente al entrevistado y tiene como objetivo recoger información sobre sus expectativas y percepciones ante distintos escenarios que puedan afectar sus finanzas personales. Además, incorpora preguntas estandarizadas internacionalmente para medir educación financiera.

- **Expectativas laborales:** se consulta a los entrevistados ocupados sobre la probabilidad de perder su empleo antes de una fecha futura específica, y a quienes buscan trabajo, sobre la probabilidad de encontrar empleo en ese mismo horizonte temporal. Para trabajadores independientes, se pregunta por la probabilidad de no poder ejercer su actividad principal por motivos ajenos a su control.
- **Expectativas de retiro:** se pregunta a qué edad el entrevistado espera dejar de trabajar definitivamente.
- **Aversión al riesgo e impaciencia:** se incluyen preguntas experimentales sobre la disposición a renunciar a parte de un premio de lotería a cambio de recibirlo antes, y sobre la aceptación de riesgos en escenarios de incertidumbre (por ejemplo, juegos de moneda con diferentes pagos posibles). Estas preguntas permiten medir la tolerancia al riesgo y las preferencias intertemporales.
- **Capacidad de respuesta ante emergencias:** se consulta si el hogar podría reunir el equivalente a seis meses de ingresos en caso de emergencia y, en caso afirmativo, a qué fuentes recurriría, tales como ahorros, bancos, familiares, venta de activos, etc.
- **Alfabetización financiera (Big3):** el módulo finaliza con tres preguntas estándar sobre interés compuesto, interés real y diversificación de riesgo, ampliamente utilizadas en la literatura internacional para evaluar conocimientos financieros básicos.

K2. Decisiones financieras en el hogar

Este módulo se aplica únicamente a hogares no unipersonales, es decir, aquellos compuestos por dos o más integrantes. Su objetivo es recolectar información sobre la distribución de roles y la dinámica de toma de decisiones financieras dentro del hogar, desde una perspectiva estructural y no condicionada a eventos específicos.

- **Preguntas K12 y K13:** indagan sobre cómo se gestionan los ingresos dentro del hogar. Se consulta si los ingresos de todos los miembros del hogar son considerados como una sola fuente de ingresos, es decir se consideran totalmente comunes o parcialmente comunes, o en caso contrario si los ingresos de cada miembro cada uno de ellos los administra de manera individual. Esto permite entender la proporción del ingreso que se considera “ingreso del hogar” y, por tanto, analizar con mayor precisión las dinámicas de endeudamiento y las posibles vulnerabilidades financieras familiares.
- **Pregunta K14:** explora cómo el hogar enfrenta problemas de pago de deudas personales de alguno de sus miembros, distinguiendo entre soluciones individuales, apoyo parcial o responsabilidad compartida.
- **Preguntas K15 y K16:** se centran en la toma de decisiones sobre compras cotidianas y compras de alto valor (como muebles, automóviles o electrodomésticos), identificando quién toma la decisión (el entrevistado, la pareja o ambos por igual).
- **Preguntas K17 a K19:** dirigidas a parejas que lideran el hogar, abordan la toma de decisiones sobre endeudamiento y el uso de ahorros comunes, permitiendo analizar la corresponsabilidad y los posibles sesgos de género en la gestión financiera familiar.

K3. Gastos

El objetivo principal de este módulo es recolectar información detallada sobre el gasto de los hogares, una dimensión que hasta ahora no estaba presente en la EFH. La incorporación de datos de consumo, junto con la información financiera, permitirá un mejor entendimiento de:

- i) la reacción de los hogares ante shocks económicos o políticas que afecten su presupuesto, y
- ii) la situación o vulnerabilidad financiera de los hogares.

Se solicita al entrevistado que declare el gasto del hogar en distintos ítems y horizontes temporales, para minimizar la omisión involuntaria de categorías relevantes. La estructura de las preguntas está inspirada en instrumentos internacionales como la EPS (España), HILDA (Australia) y HFCS (Europa), y los ítems definidos para los horizontes de 3 y 6 meses son subconjuntos de las categorías de la Encuesta de Presupuestos Familiares (EPF) de Chile.

- **Gasto promedio mensual:** se pregunta por el gasto promedio mensual en compras de supermercado, ferias libres, almacenes y negocios de barrio, así como en servicios básicos

como agua, electricidad, gas, leña, internet, televisión pagada, telefonía y gastos comunes de edificio o condominio.

- **Gasto en transporte:** se solicita el gasto promedio mensual en transporte público, combustibles y peajes/tag.
- **Gasto en salud:** se pregunta por el monto total gastado en los últimos 3 meses en consultas médicas (incluyendo dentales, psicológicas, kinesiológicas, etc.), exámenes/procedimientos médicos ambulatorios y medicamentos.
- **Gasto en bienes durables y otros rubros:** se indaga sobre el monto total gastado en los últimos 3 meses en electrodomésticos pequeños, materiales y servicios para la conservación de la vivienda, vestuario y calzado, y equipos de telefonía móvil (K25). Para los últimos 6 meses, se pregunta por el gasto total en electrodomésticos grandes y muebles.
- **Balance de ingresos y gastos:** se consulta si, en los últimos 6 meses, el gasto del hogar ha sido mayor, igual o menor que sus ingresos. Si el gasto ha sido mayor, se pregunta cómo se cubrió la diferencia (endeudamiento, venta de activos, uso de ahorros, ayuda de familiares, aumento de ingresos o reducción de gastos).

III.3. Modificaciones al Cuestionario

En el marco del proceso continuo de mejora de la EFH, cada ronda contempla una revisión exhaustiva tanto del cuestionario como de los procedimientos asociados a su aplicación. Habitualmente, este proceso parte del cuestionario utilizado en la ronda anterior, sobre el cual se implementan las modificaciones necesarias. Sin embargo, dado que en 2021 fue necesario adaptar y reducir el cuestionario para su implementación telefónica, en la elaboración del cuestionario 2024 se consideraron como referencia tanto el cuestionario de 2017 como el de 2021.

El análisis y revisión de los cuestionarios se realizó considerando dos dimensiones principales: por una parte, el examen de las tasas de no respuesta y la calidad de las respuestas obtenidas en las rondas 2017 y 2021; por otra, la identificación de preguntas susceptibles de ser eliminadas para permitir la incorporación de los nuevos módulos K1, K2 y K3, manteniendo así una duración promedio del cuestionario similar a la de rondas previas.

A continuación, se detallan los principales cambios por módulo:

A. Composición del hogar e identificación de entrevistado

En términos generales, este módulo se mantiene respecto a la versión 2017. Se actualizaron las alternativas de nacionalidad conforme a los cambios implementados en la ronda 2021, y se incorporó una nueva pregunta que indaga sobre la relación de parentesco entre cada miembro del hogar y el entrevistado. Esta modificación permite una mejor identificación de los vínculos familiares y facilita el análisis de la estructura interna del hogar.

B. Educación

Este módulo se mantiene sin modificaciones respecto a la versión aplicada en 2017. Se conservan las mismas preguntas y estructura.

C. Situación laboral

Este módulo retoma la estructura aplicada en 2017, incorporando algunos ajustes menores. Se añade una pregunta sobre si el miembro del hogar ha trabajado alguna vez, dirigida a quienes actualmente no se encuentran ocupados ni jubilados, con el fin de filtrar adecuadamente las preguntas de búsqueda de empleo para quienes nunca han trabajado. Además, al igual que en la versión 2021, se incluye la pregunta sobre horas semanales trabajadas. Finalmente, se ajustan los universos de aplicación de ciertas preguntas para asegurar su consistencia con la naturaleza de la información solicitada, extendiendo criterios ya implementados en 2021: (1) la pregunta sobre relación contractual no se aplica a empleadores ni a trabajadores por cuenta propia, y (2) la pregunta sobre boleta de honorarios solo se consulta a empleadores, trabajadores por cuenta propia y trabajadores dependientes sin contrato.

D. Medios de pago

En este módulo, respecto a 2017, se eliminan las mismas las preguntas que fueron eliminadas en 2021, relacionadas con la frecuencia y la razón de uso de cada uno de los medios de pago utilizados por el hogar.

Asimismo, se conservan los tres medios de pago incorporados en 2021: pagos en línea (por ejemplo, Paypal, Webpay, MercadoPago, etc.), tarjetas de prepago físicas o virtuales y tarjetas asociadas a criptomonedas.

Finalmente, se eliminan las preguntas referidas a servicios financieros digitales, con el objetivo de optimizar la extensión y relevancia del módulo.

E. Activos reales y deuda hipotecaria

Tanto para la vivienda principal (E1) como para otras propiedades (E2), la versión 2024 del cuestionario retoma la estructura y contenido de la versión 2017, recuperando preguntas y formatos que, por la modalidad de aplicación, habían sido eliminados o ajustados en la ronda 2021.

Respecto a la versión 2017, se introducen las siguientes modificaciones:

- Se incorpora una pregunta sobre la valoración de la vivienda principal para arrendatarios, solicitando la estimación del precio de venta, mientras que anteriormente solo se consultaba a los propietarios.

- Se eliminan las preguntas sobre el nombre del banco o institución del crédito hipotecario, tanto para la vivienda principal como para otras propiedades (cambio ya implementado en 2021).
- Se elimina la pregunta sobre si el crédito hipotecario fue pactado en pesos o en UF, tanto para la vivienda principal como para otras propiedades (cambio ya implementado en 2021).
- Se incorpora una nueva pregunta sobre el tipo de tasa del crédito hipotecario (fija, variable o mixta), tanto para la vivienda principal como para otras propiedades.

F. Deudas no hipotecarias

La versión 2024 de este módulo retoma la estructura y el contenido de la versión 2017, recuperando preguntas y formatos que, por la modalidad de aplicación, habían sido eliminados o ajustados en la ronda 2021. Se eliminan, para todas las deudas, las preguntas que consultan sobre el nombre del banco o institución asociada a deudas en tarjetas de crédito, líneas de crédito y préstamos, cambio que ya había sido implementado en 2021.

G. Percepción de la carga financiera y restricciones al crédito

Este módulo no presenta cambios significativos respecto a la versión 2017. Solo se incorporan dos nuevas preguntas: (1) los motivos por los cuales el hogar tuvo que hacer frente a gastos inesperados en los últimos 12 meses y (2) el monto destinado a cubrir dicha situación.

H. Vehículos y otros activos reales

No hay cambios en este módulo respecto a la versión 2017.

I. Activos financieros, pensiones y seguros

La versión 2024 de este módulo retoma la estructura y el contenido de la edición 2017, reincorporando preguntas y formatos que habían sido eliminados o ajustados en la ronda 2021 debido a la modalidad de aplicación.

Preguntas incorporadas:

- Se incluye una pregunta sobre el **monto invertido en cada tipo de activo financiero durante los últimos 12 meses**, con el objetivo de caracterizar de manera más precisa el comportamiento de inversión de los hogares.
- Se añade una pregunta sobre el **saldo que el hogar mantiene como ahorro en cuenta corriente o cuenta vista**, lo que permite capturar información más detallada sobre la liquidez disponible.

Preguntas eliminadas:

- Se eliminan las preguntas relativas a la **entidad financiera que gestiona el activo financiero**, así como la pregunta sobre **otros activos financieros**.

- También se suprimen las preguntas sobre la **frecuencia con que el hogar ahorra**, y la **tenencia de seguros**.
- Cabe señalar que estas eliminaciones ya habían sido incorporadas en la versión 2021, por lo que su exclusión en 2024 responde a la consolidación de los ajustes realizados en la ronda anterior.

J. Ingresos relacionados

No hay cambios en este módulo respecto a la versión 2017.

L. Otros ingresos/Ingresos no relacionados

No hay cambios en este módulo respecto a la versión 2017.

IV. Diseño Muestral

El uso del muestreo probabilístico en investigaciones estadísticas con encuestas les confiere buenas propiedades estadísticas a sus resultados, el hecho de que cada unidad de la población tenga una probabilidad conocida y distinta de cero de ser seleccionada, permite tener muestras representativas y realizar inferencias válidas sobre la población objetivo.

IV.1. Características del diseño muestral

La **población objetivo** de la EFH 2024 corresponde a los hogares urbanos que habitan en viviendas particulares en todo el territorio nacional. La cobertura es nacional urbana, con **representatividad a nivel de macrozonas geográficas**³ y **estratos de riqueza**⁴ definidos según el avalúo fiscal de la vivienda.

Al igual que en la ronda anterior, la EFH 2024 fue diseñada como un **panel rotativo de dos periodos**. La muestra se compone de dos grupos: hogares entrevistados previamente en la ronda 2021 (**muestra panel**) y hogares entrevistados por primera vez en 2024 (**muestra refresco**).

El diseño muestral tiene dos objetivos principales:

1. Garantizar una muestra aleatoria y representativa de la población urbana, incorporando un sobremuestreo de los hogares de mayor riqueza para mejorar la precisión en el análisis de variables clave.
2. Mantener el componente panel, incluyendo hogares que participaron en la EFH 2021, lo que permite mantener la representatividad de la población objetivo a lo largo del tiempo, especialmente en las variables de interés.

El diseño muestral de la EFH 2024, al igual que en versiones anteriores, es **probabilístico, estratificado y tri-etápico**. Los estratos se definen por la intersección de macrozonas y estratos de riqueza.

- **Primera etapa:** selección de comunas.
- **Segunda etapa:** selección de manzanas dentro de las comunas seleccionadas.
- **Tercera etapa:** selección de viviendas dentro de cada manzana seleccionada.

³ Las macrozonas corresponden a agrupaciones regionales en la EFH para fines de diseño y análisis muestral. La macrozona **Norte** incorpora las regiones XV, I, II, III y IV; la macrozona **Centro** incorpora las regiones V, VI, VII, XVI y VIII; la macrozona **Sur** incorpora las regiones IX, XIV, X, XI y XII; la macrozona **RM** comprende la Región Metropolitana de Santiago.

⁴ Esta variable corresponde a una agrupación realizada según la distribución del avalúo fiscal de la propiedad, obtenida del marco muestral (Catastro de Bienes Raíces). El marco muestral se dividió en tres estratos de riqueza según el avalúo fiscal: el **estrato 1** comprende los percentiles 1 al 50, el **estrato 2** los percentiles 51 al 80 y el **estrato 3** los percentiles 81 al 100.

Este enfoque garantiza representatividad geográfica y socioeconómica, y permite aplicar el sobremuestreo en hogares de mayor riqueza. Además, el diseño es comparable con el utilizado en la *Survey of Consumer Finances* (SCF) de Estados Unidos y con la *Household Finance and Consumption Survey* (HFCS) coordinada por el Banco Central Europeo.

El **tamaño de muestra** objetivo de la EFH 2024 se mantiene igual al de las rondas 2021 y 2017, con aproximadamente 4.500 hogares, de los cuales 2.800 corresponden a la muestra de refresco y 1.700 a la muestra panel. A continuación, se describen en detalle las características de cada tipo de muestra.

IV.2. Muestra Panel

La muestra panel de la EFH 2024 se compone de los 2.855 hogares que participaron como muestra de refresco en la EFH 2021. Durante el último trimestre de 2023, se realizó un proceso de contacto con estos hogares para agradecer su participación en la EFH 2021 y, al mismo tiempo, verificar su disposición para participar en la ronda 2024. Como resultado de este proceso, 432 hogares no pudieron ser contactados, 68 hogares rechazaron la entrevista y 195, aunque actualizaron sus datos, finalmente rechazaron participar en la EFH 2024. Así, la muestra panel disponible para la EFH 2024 quedó conformada por 2.155 hogares, con el objetivo de lograr alrededor de 1.700 entrevistas efectivas (ver Tabla 2).

Tabla 2
Muestra panel objetivo por estratos de riqueza y macrozonas

Macrozonas	Estratos de riqueza			Total
	1	2	3	
Norte	75	53	86	214
Centro	185	132	210	527
Sur	208	149	238	595
RM	127	91	146	364
Total	595	425	680	1.700

Dado que la muestra panel de la EFH 2024 se deriva de la muestra de refresco de la EFH 2021, hereda de ésta su diseño muestral. Para conocer detalles adicionales sobre el diseño original, se recomienda consultar el documento metodológico de la EFH 2021.

IV.3. Muestra Refresco

Como ya se mencionó, siguiendo con la metodología utilizada en rondas anteriores de la EFH, el diseño de la muestra de refresco de la EFH 2024 es probabilístico, estratificado, y tri-etápico, donde los estratos vienen dados por las intersecciones de las macrozonas y los estratos de riqueza. En la

primera etapa se seleccionan comunas, en la segunda etapa se sortean manzanas dentro de las comunas ya seleccionadas, y finalmente en la tercera y última etapa, se sortean viviendas dentro de cada manzana.

IV.3.1. Marco muestral

El marco muestral de la EFH se construyó a partir del Catastro de Bienes Raíces (CBR) del Servicio de Impuestos Internos (SII), actualizado a diciembre de 2023, el cual tiene cobertura nacional y contiene información de avalúo fiscal, comuna, manzana y dirección para aproximadamente 8,1 millones de bienes raíces. De estos, 5,7 millones tienen uso habitacional y, en adelante, se denominan viviendas.

Adicionalmente, se empleó información proveniente de las proyecciones de población del Instituto Nacional de Estadísticas (INE) base 2017 a diciembre de 2023⁵. Esto permitió identificar las comunas predominantemente urbanas⁶, que corresponden a 151 de las 346 comunas del país. Una vez identificadas las comunas urbanas en el CBR 2023, el marco muestral para la EFH 2024 quedó conformado por un total de 4.993.318 viviendas que cumplen con los requisitos establecidos en la definición de la población objetivo.

Finalmente, para evitar entrevistar en la muestra de refresco de la ronda 2024 a hogares que hayan sido visitados o hayan participado en la ronda inmediatamente anterior, se identificaron y excluyeron del proceso de selección a aquellas manzanas que ya habían sido seleccionadas para las dos rondas previas de la EFH (2017 y 2021).

IV.3.2. Estratificación

Uno de los atributos del CBR es que incluye el avalúo fiscal de las propiedades, utilizado como *proxy* de la riqueza de los hogares, facilitando así la definición de estratos y en consecuencia el sobremuestreo de los hogares de mayor riqueza.

Con el propósito de realizar el sobremuestreo de los hogares más ricos, se estratifican las viviendas del CBR según su avalúo fiscal. La clasificación de las viviendas se genera con base en la segmentación de la distribución de los avalúos fiscales por macrozona, así, el estrato 1 comprende los percentiles 1 a 50, el estrato 2 los percentiles 51 a 80 y el estrato 3 los percentiles 81 a 100 de los avalúos fiscales.

⁵ <https://www.ine.gob.cl/estadisticas/sociales/demografia-y-vitales/proyecciones-de-poblacion>.

⁶ Se refiere a aquellas comunas que según el Censo de Población y Vivienda poseen menos de 25% de población rural.

IV.3.3. Tamaño muestral

El tamaño muestral objetivo de la muestra de refresco de la EFH 2024 se definió en aproximadamente 2.800 hogares, distribuidos por macrozona y estrato de riqueza. La evaluación de precisión de las estimaciones se realizó utilizando la tenencia de deuda bancaria como variable de interés y el criterio de error relativo⁷, tanto a nivel total como en cada una de las celdas formadas por el cruce de los estratos de riqueza y las macrozonas.

El análisis consideró un sobremuestreo del 100% en el estrato de mayor riqueza, así como un incremento en la participación de la macrozona Sur, con el objetivo de mejorar la precisión de las estimaciones para este dominio de estudio. La Tabla 3 muestra la distribución de la muestra de refresco según macrozona y estrato.

Tabla 3
Distribución muestra de refresco de la EFH 2024 según macrozona y estrato
(hogares y porcentaje del total de la muestra)

Macrozona	Estrato			Total
	1	2	3	
Norte	132	89	150	371
	4,7%	3,2%	5,3%	13,2%
Centro	337	251	362	950
	12,0%	8,9%	12,9%	33,7%
Sur	184	130	207	521
	6,5%	4,6%	7,3%	18,5%
RM	356	238	381	975
	12,6%	8,4%	13,5%	34,6%
Total	1.009	708	1.100	2.817
	35,8%	25,1%	39,0%	100,0%

*Cifras secundarias en porcentaje aproximadas.

La Tabla 4 muestra el error relativo esperado para la muestra de refresco, desagregado por macrozona y estrato de avalúo fiscal⁸. Los resultados indican que el error relativo para la muestra total es de 8,0%, lo que, según la clasificación del INE (2025), se considera “Muy bueno”. Por

⁷ El error relativo es la máxima desviación porcentual que se puede esperar entre el parámetro estimado y el verdadero parámetro poblacional con probabilidad de (1 – α). Este estadístico permite homogenizar las unidades de medida de varios indicadores, al convertirse en un porcentaje relativo al estimador. La fórmula utilizada del error relativo en el caso de una proporción es:

$$e = 1,96 \sqrt{\frac{def f * p(1 - p)}{Np^2}}$$

donde *def f* corresponde al efecto diseño, *p* al valor puntual de la variable de referencia, *N* es el tamaño de muestra a evaluar y 1,96 corresponde al percentil de la distribución normal asociado a un nivel de confianza (1 – α) de 95%. Para más detalles sobre el cálculo y uso del error relativo ver INE (2025)

⁸ El Anexo N°1 muestra el error relativo para estratos generados a partir del cruce de estrato de riqueza y macrozona.

macrozona, el indicador se clasifica como “Bueno” (por debajo del 20%), siendo la macrozona Norte la que presenta el valor más elevado (18,5%). Por estrato, se observa que los estratos 1 y 2 presentan un error relativo considerado como “Bueno”, mientras que el estrato 3 alcanza un nivel de error considerado como “Muy bueno”. Este último resultado está asociado al sobremuestreo aplicado a dicho estrato, lo que permite una mejor representación de los fenómenos en este grupo de la población.

Tabla 4
Error relativo de la muestra de refresco de la EFH 2024 por macrozona y estrato

Macrozona	Estrato	Proporción TDB en 2021	Deff en 2021	Muestra EFH 2024	Error relativo en 2024
1. Norte		37,7%	1,99	371	18,5%
2. Centro		31,1%	2,18	950	14,0%
3. Sur		34,3%	1,16	521	12,8%
4. RM		36,9%	3,52	975	15,4%
	Estrato 1	25,3%	2,72	1.009	17,5%
	Estrato 2	40,1%	2,21	708	13,4%
	Estrato 3	51,6%	1,31	1.100	6,5%
TOTAL		34,9%	2,53	2817	8,0%

IV.3.4. Primera etapa de muestreo: selección de comunas

Entre 2007 y 2014, la muestra de comunas se mantuvo sin modificaciones. En 2017 se actualizó atendiendo a los criterios demográficos y operativos para mejorar la eficiencia muestral. En 2021 se introdujo una modificación adicional para garantizar dos aspectos técnicos clave:

- 1. La aleatoriedad en la selección de las comunas (UPM).
- 2. La obtención de factores de expansión directos para la primera etapa del muestreo.

Para la ronda 2024, se mantiene el diseño implementado en 2021, con un tamaño de muestra objetivo de 77 comunas, seleccionadas del universo de 151 comunas urbanas.

En cada macrozona del dominio de estudio, las 151 comunas urbanas se dividen en dos tipos:

- **Comunas auto-representadas:** aquellas que corresponden a capitales regionales y zonas conurbadas (62 comunas cumplen esta condición).
- **Comunas co-representadas:** el resto de las comunas urbanas (89 comunas en total).

Las 62 comunas auto-representadas se incluyen forzosamente en la muestra (probabilidad de selección igual a 1). Para alcanzar el objetivo de una muestra de 77 UPM, las 15 comunas co-representadas restantes se distribuyen proporcionalmente al número de comunas que posee cada

macrozona. La Tabla 5 muestra la distribución final de la muestra de comunas entre auto-representadas y co-representadas por macrozona.

Tabla 5
Distribución de comunas en la muestra de refresco de la EFH 2024 según macrozona (número de comunas)

Macrozona	Comunas Urbanas					
	Auto-representadas		Co-representadas		Total	
	Población	Muestra	Población	Muestra	Población	Muestra
Norte	7	7	15	3	21	10
Centro	15	15	53	9	68	24
Sur	5	5	14	2	19	7
RM	35	35	7	1	42	36
Total	62	62	89	15	151	77

Una vez distribuido el total de 15 comunas co-representadas por macrozona, la selección de comunas en cada macrozona se realizó de manera independiente, aplicando la técnica de muestreo con probabilidades desiguales sin reemplazo, conocida como muestreo PPS (*Probabilities Proportional to Size*). La probabilidad de selección de cada comuna en cada macrozona está dada por el número de viviendas que contiene, utilizando el algoritmo de Hanurav-Vijayan.

La muestra final de comunas abarca un total de 4,1 millones de unidades habitacionales, lo que representa aproximadamente el 83% de las viviendas urbanas del territorio nacional.

En el Anexo N°2 se presenta el detalle de la muestra de comunas seleccionada para la muestra de refresco de la EFH 2024 y la correspondiente a la muestra de panel resultante del proceso de fidelización de los hogares que pertenecían a la muestra de refresco de la ronda 2021.

IV.3.5. Segunda etapa de muestreo: selección de manzanas

La unidad secundaria de muestreo (USM) se define como una sub-manzana, es decir, un conjunto de viviendas dentro de una misma manzana geográfica que pertenecen al mismo estrato de avalúo fiscal. Por ejemplo, si una manzana geográfica contiene viviendas de dos estratos diferentes, se descompone en dos sub-manzanas distintas. Para simplificar, en adelante las sub-manzanas se denominarán simplemente como manzanas o USM.

Partiendo del tamaño de muestra objetivo inicial de 2.800 viviendas, distribuidas por macrozona y estrato de riqueza según lo descrito en la Tabla 4, para la selección de las USM se establecieron las siguientes condiciones:

- Para garantizar la factibilidad de estimar varianzas a nivel de manzana, el tamaño de muestra de viviendas por manzana es de 4.
- En cada intersección de macrozona, comuna y estrato de riqueza, se debe seleccionar al menos una (1) manzana.

Considerando estas condiciones, se realizó la afijación inicial del tamaño de muestra de manzanas ($m_{0,z,e}$) por cada intersección macrozona-estrato, utilizando los pesos iniciales ajustados por el sobremuestreo del estrato 3 y de la macrozona Sur, de acuerdo con la siguiente fórmula:

$$m_{0,z,e} = w_{0,z,e} * \frac{n_0}{4}$$

Donde:

n_0 : es el tamaño de muestra inicial objetivo de 2.800 viviendas.

$w_{0,z,e}$: es el peso inicial de la intersección de la macrozona z y el estrato e.

Las siguientes tablas se describen los pesos y tamaños iniciales de la muestra de manzanas.

Tabla 6

Afijación de los pesos de la muestra de refresco de la EFH 2024 según macrozona y estrato (porcentaje del total)

Zona	Estrato 1	Estrato 2	Estrato 3	Total
Norte	4,6%	3,3%	5,3%	13,2%
Centro	11,7%	8,3%	13,3%	33,3%
Sur	6,5%	4,6%	7,4%	18,4%
RM	12,3%	8,7%	14,0%	35,0%
Total	35,0%	25,0%	40,0%	100,0%

Tabla 7

Afijación de muestra inicial de manzanas según macrozona y estrato (porcentaje del total)

Zona	Estrato 1	Estrato 2	Estrato 3	Total
Norte	32	24	37	93
Centro	82	58	93	233
Sur	45	32	52	129
RM	86	61	98	245
Total	245	175	280	700

Una vez obtenida la afijación inicial de la muestra de manzanas por macrozona y estrato, se procedió a la asignación de la muestra de manzanas en las comunas, de forma proporcional al número de manzanas contenidas en cada comuna de la muestra, con la restricción de que cada intersección comuna-estrato debía tener al menos una manzana seleccionada:

$$m_{0,z,c,e} = \text{round} \left(m_{0,z,e} * \frac{M_{z,c,e}}{\sum_{\forall c} M_{z,c,e}} \right)$$

Donde:

$m_{0,z,c,e}$: es el tamaño de muestra inicial de manzanas en la macrozona z, para la comuna c y el estrato e.

$M_{z,c,e}$: es el número de total de manzanas de la macrozona z, comuna c y estrato e.

El tamaño de muestra final de manzanas se determinó mediante la siguiente regla:

$$m_{z,c,e} = \begin{cases} 1 & ; \text{ si } m_{0,z,c,e} = 0 \\ m_{0,z,c,e} & ; \text{ si } m_{0,z,c,e} > 0 \end{cases}$$

De las condiciones anteriores se desprende que la muestra total de USM es de 725 manzanas. En esta segunda etapa se aplicó el mecanismo de muestreo aleatorio sistemático con selección proporcional al tamaño, utilizando como variable de ordenamiento en cada comuna el avalúo fiscal promedio de la manzana. Esto asigna una probabilidad de selección proporcional al avalúo promedio, permitiendo recorrer todas las manzanas de cada intersección estrato-comuna desde las de menor a mayor valor de avalúo promedio y evitando la concentración de la muestra.

Antes de la selección, se censuraron alrededor de 2.700 manzanas que pertenecían a las muestras de la EFH 2021 y 2017, para evitar seleccionar en la última etapa viviendas que hayan participado en las dos rondas anteriores. Se consideraron los cambios que esto ocasiona en los totales de manzanas por comuna para ajustar los factores de expansión directos, garantizando así la reproducción íntegra de la población objetivo. La distribución de las USM se muestra en la Tabla 8.

Tabla 8
Distribución de manzanas (USM) de la muestra de refresco de la EFH 2024
según macrozona y estrato
(unidades y porcentaje del total)

Macrozona	Estrato de riqueza						Total	
	1		2		3			
	Manzanas	%	Manzanas	%	Manzanas	%	Manzanas	%
Norte	33	4,6%	23	3,2%	38	5,2%	94	13,0%
Centro	85	11,7%	63	8,7%	94	13,0%	242	33,4%
Sur	47	6,5%	33	4,6%	53	7,3%	133	18,3%
RM	89	12,3%	64	8,8%	103	14,2%	256	35,3%
Total	254	35,0%	183	25,2%	288	39,7%	725	100,0%

IV.3.6. Tercera etapa de muestreo: selección de viviendas

La tercera etapa del muestreo corresponde a la selección de las unidades últimas de muestreo (UUM), es decir, las viviendas. Estas fueron seleccionadas mediante un muestreo aleatorio sistemático en cada una de las 725 manzanas, previamente ordenadas según el valor del avalúo fiscal. Se seleccionaron cuatro viviendas por manzana; en caso de que una manzana tuviera cuatro viviendas o menos, se seleccionaron todas. Así, la muestra total de refresco para la EFH 2024 quedó conformada por 2.817 viviendas.

V. Trabajo de Campo

El levantamiento de la EFH 2024 fue adjudicado a la empresa IPSOS Chile mediante un proceso de licitación pública.

V.1. Prueba piloto

Con el objetivo de preparar el trabajo de campo de la EFH 2024, se llevó a cabo una prueba piloto destinada a replicar el proceso completo de levantamiento de la encuesta. Esta instancia permitió evaluar el flujo, la consistencia y la extensión del cuestionario; revisar el material de capacitación y los recursos de apoyo para los encuestadores; y probar tanto el protocolo de contacto como el procedimiento de reemplazo de la muestra refresco. El propósito final fue anticipar y corregir posibles dificultades que pudieran surgir durante el levantamiento definitivo.

La prueba piloto se realizó en el mes de mayo, con la participación de 24 encuestadores. Se entrevistaron 109 hogares, de los cuales 97 correspondían a muestra refresco y 12 a muestra panel.

De la aplicación de esta prueba piloto se concluyó lo siguiente:

1. El número de encuestadores con la capacidad y motivación necesarias para llevar a cabo un estudio de la complejidad muestral y temática de la EFH es limitado. La mayoría de las encuestas fueron realizadas por solo cuatro encuestadores.
2. La disposición de los hogares a participar en encuestas ha disminuido respecto a rondas anteriores de la EFH, ya que además de las dificultades tradicionalmente atribuibles de este tipo de estudios, tales como, (a) temática sensible de la encuesta y (b) extensión del cuestionario, se observaron otros factores como (c) inseguridad y desconfianza para abrir las puertas de la vivienda y (d) un menor sentido de colaboración con las instituciones.

Como resultado del piloto, se implementaron las siguientes acciones para mejorar el trabajo de campo:

1. Fortalecimiento del material de apoyo a los encuestadores, con el fin de brindar mayor seguridad y fomentar la participación de los hogares, los elementos adicionales incorporados para la ronda 2024 fueron:
 - a. Se enviaron cartas a las municipalidades solicitando su colaboración en la difusión del despliegue de encuestadores en sus comunas, a través de sus sitios web institucionales y redes sociales.
 - b. Se publicó un aviso en la página principal del Banco Central, con un enlace a la sección de la EFH para obtener más información y validar la identidad de los encuestadores.

- c. Se mejoró la calidad de la carta de presentación, incorporando un sello de agua para aumentar su veracidad.
 - d. Se capacitó a los operadores de la central telefónica del Banco para resolver dudas de las personas que se contactaran por esta vía. Para los días y horarios no hábiles, se implementó una grabación que confirmaba la existencia de la encuesta y entregaba información sobre cómo obtener más detalles en la página web del Banco.
 - e. Se difundió información relevante sobre el levantamiento en las redes sociales del Banco.
2. Se hizo una revisión del cuestionario, eliminando preguntas y optimizando el flujo, para que la duración mediana fuese de 35 minutos.

V.2. Resultados del trabajo de campo

El trabajo de campo de la EFH 2024 se desarrolló entre el 22 de julio de 2024 y el 13 de enero de 2025, abarcando un total de 26 semanas.

El objetivo para la muestra panel era alcanzar al menos 1.700 entrevistas. Para ello, se visitaron 2.080 de los 2.155 hogares que, en 2023, habían actualizado su información de contacto y manifestado interés en participar en la siguiente ronda. Como resultado, se lograron 1.801 entrevistas.

Respecto a la muestra refresco, el objetivo era realizar 2.817 entrevistas. La muestra aleatoria seleccionada además incorporó hasta 4 reemplazos por cada unidad original, que fueron administradas por el Banco conforme al protocolo establecido. En total, se visitaron 4.724 hogares, obteniéndose 2.848 entrevistas.

Durante el trabajo de campo participaron 141 encuestadores, con un promedio de 33 encuestas realizadas por encuestador. No obstante, se observó una marcada asimetría en el rendimiento: 10 encuestadores realizaron el 52% del total de encuestas, con un promedio de 240 encuestas cada uno, mientras que 65 encuestadores realizaron 10 encuestas o menos a lo largo de las 26 semanas de levantamiento.

La tasa de respuesta fue de un 71,5%, presentando diferencias entre la muestra panel (87,2%) y la muestra de refresco (64,1%). La tasa de rechazo fue de un 11,3%, también mostrando diferencias entre muestra refresco, donde la tasa de rechazo fue de un 14,7%, y la muestra panel, donde la tasa de rechazo fue 3,7%.

V.3. Duración de la entrevista

El cuestionario de la EFH 2024 contiene un total de 855 preguntas. No obstante, debido a la estructura de filtros del instrumento, ningún hogar responde la totalidad de ellas. El hogar que respondió la mayor cantidad de preguntas alcanzó aproximadamente el 40% del cuestionario, es decir, cerca de 340 preguntas. En promedio, los hogares respondieron 240 preguntas, lo que equivale al 28% del recorrido completo del cuestionario.

La duración promedio de la entrevista fue de 33 minutos, mientras que la duración mediana fue de 29,6 minutos. Solo en el 10% de los hogares la entrevista superó los 55 minutos, registrándose un tiempo máximo de 2 horas y media.

Tabla 9
Duración promedio y mediana EFH 2024 por módulos (minutos)

	Módulo	N° Preguntas	Promedio	Mediana
A	Composición del hogar	7	2,0	1,3
B	Educación	5	1,2	0,7
C	Situación laboral	15	4,0	2,9
D	Medios de pagos	23	2,1	1,5
E1	Activos reales y deuda hipotecaria (vivienda principal)	50	2,8	1,8
E2	Activos reales y deuda hipotecaria (otras propiedades)	137	1,3	0,3
F	Deudas no hipotecarias	342	4,9	1,7
G	Carga financiera y restricciones de crédito	33	2,0	1,0
H	Vehículos y otros activos reales	19	1,2	0,5
I	Activos financieros, pensiones y seguros	64	3,3	2,2
J	Ingresos relacionados	10	1,3	0,8
L	Otros ingresos/Ingresos no relacionados	81	2,0	1,4
K1	Expectativas, percepciones y alfabetización financiera	26	4,3	3,2
K2	Decisiones financieras del hogar	8	1,6	0,9
K3	Gasto	27	4,6	3,6
M	Datos de contacto	8	3,0	1,6
	Total	855	33,0	29,6

VI. Factores de Expansión

Los diseños de muestras probabilísticas, como el utilizado en la EFH 2024, permiten calcular errores muestrales, intervalos de confianza y otros indicadores estadísticos esenciales para evaluar la precisión de los resultados. Además, constituyen la base metodológica para la aplicación de factores de expansión, los cuales ajustan los resultados muestrales para reflejar adecuadamente la estructura de la población total, asegurando así la validez y comparabilidad de los indicadores producidos.

En su forma más simple, los factores de expansión corresponden al inverso de la probabilidad de inclusión de cada elemento en la muestra. Su objetivo es permitir la obtención de estimaciones insesgadas de los parámetros poblacionales de interés, por lo que representan un componente fundamental en el análisis de los datos, especialmente cuando la probabilidad de selección de las distintas unidades varía según lo establecido por la complejidad del diseño muestral.

La EFH 2024 emplea un diseño complejo, en el que coexisten dos muestras independientes (panel y refresco), y se combinan distintos tipos de muestreo (estratificado y multietápico). Considerando lo anterior, además del cálculo de los factores de expansión directos o de diseño, es necesario:

- Realizar ajustes a los factores de expansión directos (por no repuesta y por cobertura).
- Ensamblar los factores correspondientes a las muestras de ambos diseños.
- Calibrar los factores a la población objetivo, conforme al diseño muestral.

A continuación, se detallan las distintas etapas de construcción para los factores de expansión, tanto para la muestra de refresco como para la muestra de panel.

VI.1. Factores de expansión de la muestra refresco EFH 2024

El diseño muestral de la muestra refresco en la EFH 2024 es probabilístico, estratificado y tri-etápico. Los estratos se definen por la intersección entre las macrozonas geográficas y los estratos de riqueza, mientras que las etapas de selección corresponden a la elección de comunas, manzanas y viviendas.

VI.1.1. Primera etapa de muestreo: comunas

Para la selección aleatoria de las comunas se empleó la técnica del muestro con probabilidades desiguales sin reemplazamiento, conocida como PPS (*Selection with Probabilities Proportional to Size Measures*), donde la probabilidad de selección está determinada por el número de viviendas en cada comuna.

La probabilidad de selección de la comuna c_h en la macrozona geográfica m del estrato de riqueza e se define como⁹:

$$P(c_h) = \begin{cases} 1 & ; \text{ sí } c_h \text{ es auto-representada (A)} \\ n_{m,e}^B \frac{N_{h,m,e}}{\sum_{\forall h \in B} N_{h,m,e}} & ; \text{ sí } c_h \text{ es co-representada (B)} \end{cases}$$

Donde:

$N_{h,m,e}$: número de viviendas en la comuna h en la macrozona m del estrato e .

$n_{m,e}^B$: número de comunas a seleccionar del grupo co-representado.

VI.1.2. Segunda etapa de muestreo: manzanas

Las manzanas fueron seleccionadas mediante muestreo aleatorio sistemático con selección proporcional al tamaño (PPS_SYS), utilizando como variable de ordenamiento el avalúo fiscal promedio de cada manzana dentro de cada celda macrozona-comuna-estrato.

Así, la probabilidad de selección de la manzana m_i en la comuna h para el estrato e se define como¹⁰:

$$P(m_i) = n_{h,e} \cdot \frac{\bar{x}_{i,h,e}}{\sum_{\forall i}^{N_{h,e}} \bar{x}_{i,h,e}}$$

Donde:

- $n_{h,e}$: número de manzanas a seleccionar en la comuna h y estrato e .
- $\bar{x}_{i,h,e}$: avalúo promedio de las viviendas de la manzana i .
- $N_{h,e}$: total de manzanas de la comuna h y estrato e .

El procedimiento se ejecutó utilizando el algoritmo de Hanurav -Vijayan¹¹.

VI.1.3. Tercera etapa de muestreo: viviendas

Para la selección de viviendas se aplicó muestreo aleatorio simple sin reemplazo, seleccionando cuatro viviendas por manzana. En caso de que la manzana tuviera cuatro o menos viviendas, todas fueron seleccionadas.

⁹ Si bien el sorteo de las comunas es sin reemplazos, a efectos descriptivos se muestran las fórmulas de sorteo con reemplazo dada la complejidad de representar la probabilidad de selección sin reemplazo.

¹⁰ Si bien el sorteo de las manzanas es sin reemplazos, a efectos descriptivos se muestran las fórmulas de sorteo con reemplazo dada la complejidad de representar la probabilidad de selección sin reemplazo.

¹¹ pp 221-234. Principles of Social Research Methodology, capítulo Sampling Techniques for Quantitative Research.

La probabilidad de selección de una vivienda v_j en la manzana i viene dada por:

$$P(v_j) = \frac{n_i}{N_i}$$

Donde:

N_i : número de viviendas en la manzana i .

n_i : número de viviendas seleccionadas en la manzana i (Sí $N_i \leq 4$ entonces $n_i = N_i$).

VI.1.4. Probabilidad total y factor de expansión directo

La probabilidad total de selección de una vivienda en la muestra refresco se obtiene como el producto de las probabilidades en cada etapa:

$$\pi_j = P(c_h)P(m_i)P(v_j)$$

El factor de expansión directo se define como:

$$d_j = \frac{1}{\pi_j}$$

VI.1.5. Ajuste por sub o sobre representación

Al igual que en rondas anteriores, el diseño contempló el uso de viviendas de reemplazo¹², lo que permitió realizar entrevistas en todas las comunas y manzanas seleccionadas. Sin embargo, no se logró la meta de diseño en 24 manzanas, mientras que en otras 40 se entrevistó a más hogares de los previstos. La muestra efectiva fue de 2.848 hogares (31 más que lo planificado), lo que hizo necesario aplicar un ajuste por sub o sobre respuesta a nivel de hogares por manzana, expresado como:

$$d_j^{NR} = d_j \frac{1}{\phi_i}$$

Donde: $\phi_i = n'_i/n_i$ (cociente de la muestra de viviendas efectiva y teórica de la manzana i)

¹² La no respuesta en la muestra de refresco se mitiga mediante el uso de viviendas de reemplazo. Cada vivienda originalmente seleccionada contó con un panel de entre 1 y 4 viviendas de reemplazo, seleccionadas dentro de la misma manzana (Unidad de Selección Muestral, USM). Excepcionalmente, en manzanas con 4 viviendas o menos, los reemplazos se seleccionaron en otras manzanas de la misma comuna y estrato, las cuales debían cumplir con dos condiciones: tener el mismo tamaño en términos del número de viviendas, y minimizar la diferencia entre la suma de los avalúos fiscales de sus viviendas respecto a la suma de avalúos de la manzana original.

Este ajuste garantiza que la suma de los d_j^{NR} reproduzca el valor total de viviendas en el marco muestral.

La siguiente tabla, muestra la suma de los factores de expansión de refresco ajustados por sub o sobre representación de viviendas en las manzanas, según el diseño muestral.

Tabla 10

Suma de los Factores de Expansión ajustados de la muestra de refresco de la EFH 2024 según macrozona y estrato

Zona	Estrato 1	Estrato 2	Estrato 3	Total
Norte	319.319	198.791	131.534	649.644
Centro	803.623	486.676	348.776	1.639.075
Sur	208.037	132.855	80.290	421.182
RM	1.165.688	679.680	438.049	2.283.417
Total	2.496.667	1.498.002	998.649	4.993.318

VI.2. Factores de expansión muestra panel

Como se ha mencionado previamente, en la Encuesta Financiera de Hogares coexisten dos diseños muestrales independientes. El segundo corresponde a la **muestra panel**, conformada por un subconjunto de hogares que participaron en la muestra refresco de la EFH 2021. En consecuencia, el factor de expansión inicial de esta muestra se corresponde con el factor ajustado por sub o sobre respuesta de dicha ronda:

$$d_j^{NR,21} = d_j^{21} * \frac{1}{\phi_j} \quad \text{donde} \quad d_j^{21} = \frac{1}{\pi_j^{21}}$$

Tal como se detalla en la sección IV.2, la muestra refresco de 2021 fue visitada nuevamente en 2023, en un proceso de *fidelización* orientado a actualizar los datos de contacto necesarios para la aplicación de la encuesta. Como resultado, se logró actualizar la información de 2.350 hogares (82% de la muestra refresco de 2021), de los cuales 2.155 (76%) manifestaron su disposición a participar en la nueva ronda. Finalmente, durante el trabajo de campo de la EFH 2024, se logró encuestar a 1.801 hogares (63% de la muestra refresco de 2021), conformando así la muestra efectiva de panel.

El fenómeno de **atrición** entre rondas consecutivas hace necesario aplicar una corrección por no respuesta, dado que los hogares que no participan pueden tener características distintas a los que sí lo hacen (por ejemplo, zona geográfica, nivel educativo, edad del informante principal, entre otras). Esta corrección busca minimizar el sesgo atribuible a causas no aleatorias de la no respuesta.

Al igual que en la ronda anterior, el ajuste por no respuesta se realiza mediante un modelo binario tipo **logit**, que estima las probabilidades de respuesta condicionales a una serie de variables que caracterizan la muestra refresco de 2021. La especificación general del modelo es:

$$\Pr(R = 1|x_j\beta) = \Lambda(x_j\beta) = \phi_j$$

A partir de este modelo se obtienen probabilidades de respuesta predichas $\hat{\phi}_i$, que se agrupan en veintiles ($s = 1, \dots, 20$). El factor de expansión corregido por no respuesta se expresa como:

$$d_j^{NR',21} = d_j^{NR,21} * 1/\bar{\phi}_s$$

El factor $d_j^{NR',21}$ constituye la base para la muestra panel. Sin embargo, dado que expande al marco muestral de la ronda anterior, es necesario aplicar un último ajuste para que represente adecuadamente las macrozonas del marco de la EFH 2024 (CBR2023):

$$d_j^{NR',21,24} = d_j^{NR',21} * \frac{N_{m,e}^{24}}{\sum_{j \in m,e} d_j^{NR',21}}$$

Donde:

$N_{m,e}^{24}$: número de viviendas de la macrozona m y el estrato e en el CBR2023.

$d_j^{NR',21,24}$: factor final de expansión para la muestra de panel, utilizado en la construcción del factor de corte transversal de la EFH 2024.

La siguiente tabla muestra la suma de los factores de expansión de la muestra de panel, corregidos por cobertura y no respuesta.

Tabla 11

Suma de los Factores de Expansión ajustados de la muestra de panel de la EFH 2024 según macrozona

Zona	Total
Norte	649.644
Centro	1.639.075
Sur	421.182
RM	2.283.417
Total	4.993.318

Y la siguiente tabla muestra una comparación de algunas estadísticas descriptivas de los factores de refresco de la EFH 2021 y los ajustes aplicados para la obtención de los factores de panel de la EFH 2024, donde se comprueba que la suma de los factores de expansión de la muestra panel, ajustados por no respuesta y por tamaño del marco muestral 2024, reproducen el total de viviendas urbanas del CBR2023.

Tabla 12
Comparación del efecto de los ajustes en los Factores de Expansión
para la muestra de panel de la EFH 2024

Estadísticos	Refresco 2021	Panel 2024 (Ajustados por NR)	Panel 2024 (Ajustados por NR y MM2024)
Sum	4.486.025	4.473.941	4.993.318
N	2.855	1.801	1.801
mean	1.571	2.484	2.773
sd	1.808	3.293	3.675
p5	216	276	316
p25	560	750	871
p50	1.255	1.716	1.929
p75	2.167	2.978	3.292
p95	3.503	7.072	7.905
max	21.921	49.260	54.477

VI.3. Factores de expansión de corte transversal EFH 2024

Al igual que en las dos rondas anteriores de la EFH, los factores corregidos de cada tipo de muestra se ensamblan en un único factor de corte transversal, definido de la siguiente forma:

$$w_j^{24} = \begin{cases} \alpha^R * d_j^{NR} & \text{si } j \in \text{Refresco} \\ (1 - \alpha^R) * d_j^{NR', 21, 24} & \text{si } j \in \text{Panel} \end{cases}$$

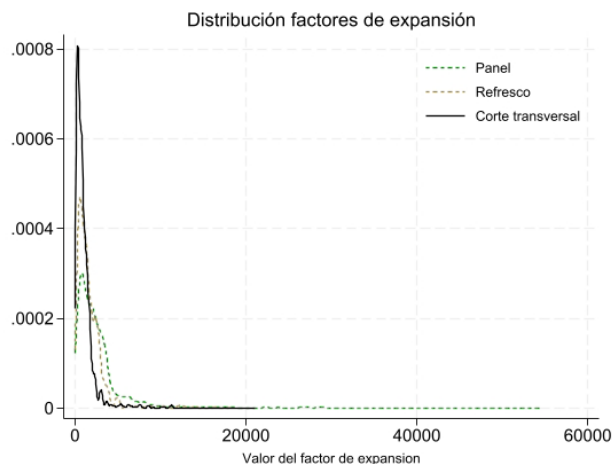
Donde:

α^R : representa la proporción de la muestra de refresco sobre la muestra total de la EFH 2024, calculada como $\alpha^R = n^R / n$, siendo $n = n^R + n^P$.

Esta metodología pondera la contribución de cada tipo de muestra según su participación en el tamaño total de la muestra. Este enfoque ha sido utilizado en la Encuesta Financiera de las Familias (EFF) del Banco de España (Bover et al., 2014) y en la Survey of Consumer Finances de Estados Unidos (Kennickell y Woodburn, 1998).

El siguiente gráfico muestra la densidad de los factores de expansión de panel y refresco, junto con los factores de corte transversal. Se observa que la combinación de ambos tipos de muestra reduce la dispersión de los factores individuales, logrando que cada vivienda entrevistada represente menos viviendas que las que originalmente representaba según su tipo muestral.

Gráfico 1
Distribución de los factores de expansión ensamblados (corte transversal)



VI.4. Calibración

El proceso de calibración consiste en ajustar los factores de expansión previamente corregidos por no respuesta y ensamblados en un único factor de corte transversal, con el fin de que las estimaciones reflejen mejor las características conocidas de la población objetivo.

Este ajuste es especialmente recomendable cuando se busca minimizar los efectos de la no respuesta y la subcobertura. Una vez aplicado, permite que los factores de expansión coincidan con los totales poblacionales de variables demográficas como grupos etarios, género y zonas geográficas.

Los totales poblacionales utilizados para calibrar la muestra de la EFH 2024 corresponden a las proyecciones de población urbana para el año 2024, publicadas por el INE. Al igual que en la ronda 2021, se emplea el método de **regresión lineal generalizada** para la calibración conjunta de ambos tipos de muestra. Los factores de corte transversal constituyen el insumo principal de este proceso.

Mediante el comando *calibrate* de Stata, los factores de expansión de corte transversal (a nivel de vivienda) se calibran para cada miembro P_i del hogar j , utilizando información de las proyecciones de población urbana por edad, género y macrozona, según la siguiente expresión:

$$w_{jP_i}^{24} = w_j^{24} * \delta_{jP_i}$$

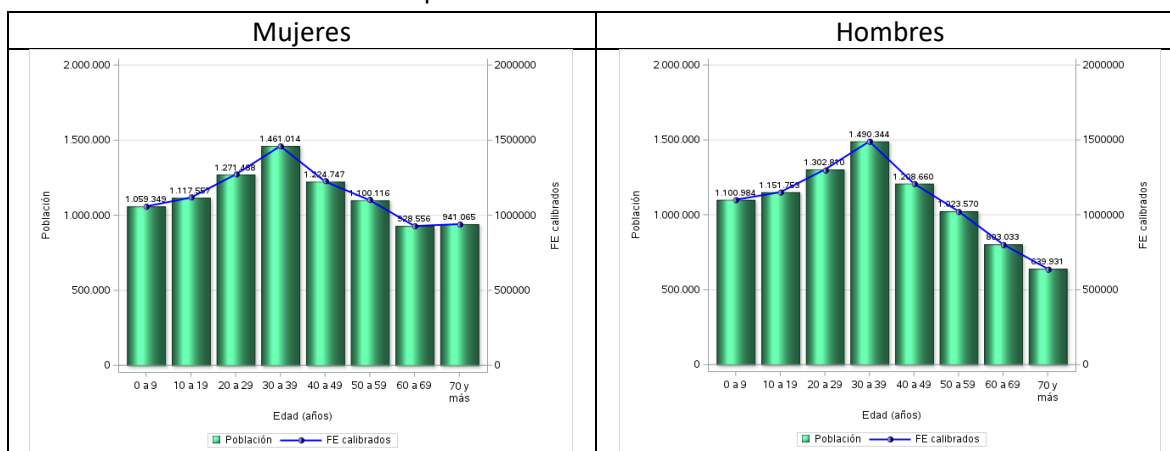
Donde:

δ_{jP_i} representa el ajuste por calibración para la persona i del hogar j .

Finalmente, se calcula el promedio de los $w_{jP_i}^{24}$ dentro de cada hogar, obteniendo así el factor de corte transversal calibrado EFH 2024 ($w_{jP_i}^{24}$).

Los siguientes gráficos muestran el ajuste de los factores de expansión calibrados. En ellos se observa que este proceso garantiza la reproducción exacta de los totales poblacionales proyectados para el año 2024, diferenciados por género:

Gráfico 2
Proyecciones de población del año 2024 y Factores
de Expansión Calibrados de la EFH 2024



VII. No respuesta parcial e imputaciones

La naturaleza sensible y detallada de la información solicitada en la Encuesta Financiera de Hogares (EFH)—como ingresos, activos, deudas y gastos—genera tasas de no respuesta parcial superiores a las observadas en encuestas de temáticas menos sensibles. En este contexto, los datos faltantes (*missing values*) representan un desafío relevante para el diseño y análisis de la encuesta. Si no se abordan adecuadamente, pueden introducir sesgos en los resultados y reducir la eficiencia de las estimaciones.

En encuestas como la *Survey of Consumer Finances* (SCF) en Estados Unidos y la *Encuesta Financiera de las Familias* (EFF) en España, los datos faltantes se abordan mediante imputación múltiple encadenada (MICE), un método que preserva la incertidumbre asociada a los valores ausentes para garantizar inferencias estadísticamente válidas, generando múltiples valores plausibles en lugar de una única estimación puntual.

Existen diversos métodos para la imputación de datos faltantes, entre ellos *Hot Deck*, *K-Nearest Neighbors* (KNN), modelos bayesianos jerárquicos y *Random Forest*. La elección de MICE responde a su formulación basada en modelos lineales, lo que facilita su estudio, implementación y trazabilidad en contextos estadísticos. El procedimiento se estructura mediante una secuencia encadenada de regresiones, diseñada para capturar relaciones multivariadas entre variables. Esta arquitectura implica una alta exigencia computacional, ya que no permite paralelización ni intervención durante la ejecución, de manera que la validación del proceso se realiza mediante la revisión previa de las regresiones utilizadas y la evaluación posterior de la coherencia y estabilidad de las imputaciones generadas. En el caso del Banco Central, este procedimiento se aplica de forma sistemática en cada ronda, generando 30 versiones imputadas por variable, consolidadas en una base única para análisis. Cada versión se obtiene tras 30 iteraciones encadenadas, necesarias para alcanzar convergencia, en un proceso que toma aproximadamente 30 horas de cómputo.

El tratamiento riguroso de los datos faltantes se ha consolidado como una práctica estándar en encuestas financieras internacionales. La EFH de Chile se alinea con estas buenas prácticas, adoptando procedimientos de imputación que permiten generar información representativa, precisa y comparable a nivel internacional.

VII.1. Estimación Imputación MICE en la EFH 2024

La implementación del método MICE en la EFH 2024 se estructura en los siguientes pasos:

1. **Identificación de los valores faltantes:** Se detectan las variables con datos ausentes y se clasifican según su tipo:
 - *Variables individuales:* por ejemplo, ingresos provenientes de la ocupación principal, consultados para cada miembro ocupado del hogar.

- *Variables a nivel de hogar*: como el valor de la deuda hipotecaria o el valor de la vivienda.
 - *Variables repetidas*: consultadas varias veces por hogar, como las tres deudas de consumo de mayor valor.
2. **Especificación del modelo de imputación**: Para cada variable con datos faltantes se define un modelo de regresión condicional apropiado.
 3. **Inicialización**: Se realiza una imputación inicial para completar el conjunto de datos y permitir el inicio del algoritmo.
 4. **Iteración encadenada**: Las variables faltantes se imputan secuencialmente hasta alcanzar estabilidad. En el caso de la EFH, se trabaja con tres bloques de variables definidos en el paso 1, lo que requiere un doble proceso de iteración encadenada: primero dentro de cada modelo y luego entre bloques.
 5. **Generación de múltiples conjuntos imputados**: Se repite el proceso completo para obtener varias versiones completas del conjunto de datos.
 6. **Análisis posterior (post-imputación)**: Se realizan análisis en cada conjunto imputado aplicando las reglas de Rubin (ver Anexo N°3), que permiten estimar parámetros y errores estándar válidos.

VII.2. Análisis de valores faltantes

Desde su primera aplicación en 2007, la EFH ha imputado exclusivamente variables continuas, como montos de ingresos, activos, deudas, gastos y horas trabajadas. En 2024, se identificaron 249 variables imputables, de las cuales 222 fueron efectivamente imputadas (ver Tabla 12).

Tabla 12
Variables imputables, imputadas y método de imputación

	Total	Imputar	Hotdeck	MICE	Máximo % Missing	Mínimo % Missing
Ingresos individuales	7	7	2	5	29,7%	2,8%
Ingresos del hogar	30	29	13	16	33%	1,4%
Activos del hogar	30	30	10	20	63,7%	4,8%
Deuda hipotecaria	53	41	24	17	71,4%	0,4%
Gasto del hogar	21	21	0	21	38,3%	4%
Deudas de consumo ¹³	108	94	4	90	44,7%	6,1%
Total	249	222	53	169	71,4%	0,4%

Las variables con menos de 55 observaciones se imputan por *Hotdeck*, mientras que el resto se imputan por MICE. Las 169 variables imputadas por MICE se agrupan en tres bloques temáticos (ver Anexo N°4).

¹³ Agregados de las 3 deudas de mayor valor. Por ejemplo, las variables cf_td1_1, cf_td1_2 y cf_td1_3, correspondientes a la carga financiera en deudas de tarjetas de créditos, para las 3 deudas de mayor valor, se agrupa en una sola variable para el proceso de imputación cf_td1_.

VII.3. Transformación de variables y valores iniciales

Para imputar un bloque utilizando variables de otro, se asignan valores iniciales a los datos faltantes de las variables externas mediante Hotdeck. Además, se aplican transformaciones estadísticas para mejorar el comportamiento de las variables, como la transformación de Box-Cox o logarítmica (ver Anexo N°5).

Pasos previos a la imputación múltiple encadenada:

- **Imputación inicial por Hotdeck:** Se imputan las 222 variables para obtener valores iniciales.
- **Validación con tramos de respuesta:** Se ajustan los valores imputados según los tramos declarados por los encuestados.
- **Transformación Box-Cox:** Se utiliza el comando `bcskew0` de Stata para reducir la asimetría de las variables.

VII.4. Evaluación y consistencia de las imputaciones

El proceso fue evaluado y ajustado iterativamente mediante:

1. **Evaluación del desempeño de los modelos MICE:**
 - *Métricas globales:* R^2 , R^2 ajustado, estadístico F, Root MSE y tamaño muestral.
 - *Estructura interna:* análisis de significancia de regresores, jerarquización por relevancia y coherencia conceptual.
2. **Evaluación de consistencia por bloques económicos:**
 - Comparación entre modelos de variables relacionadas.
 - Imitación metodológica entre modelos exitosos.
3. **Evaluación de los datos imputados:**
 - Comparación de distribuciones estadísticas entre datos originales e imputados.
 - Validación mediante histogramas, boxplots y análisis de percentiles.
 - Aplicación del marco metodológico de Rubin (1987) para consolidar estadísticas.

VIII. Estimación de la varianza

Tal como se aplica en otras encuestas financieras de hogares, como la *Household Finance and Consumption Survey* (HFCS) de la Comunidad Europea, la EFH pone a disposición del público réplicas de los factores de expansión o **pesos Bootstrap**. Este enfoque es altamente efectivo para el cálculo de la varianza de los estimadores, debido a su fácil implementación y a que permite obtener estimaciones consistentes del error, tanto para estimadores continuos como categóricos (Shao, 2003; Girard, 2009).

VIII.1. Métodos de estimación de varianza en encuestas complejas

Los métodos de estimación de varianza en encuestas complejas se dividen en dos grandes grupos:

- **Métodos paramétricos**, como la **linealización de Taylor (TLS)**.
- **Métodos no paramétricos o de réplicas**, entre los que se incluyen:
 - *Balanced Repeated Replication (BRR)*
 - *Jackknife Repeated Replication (JRR)*
 - *Bootstrap*

Independientemente del método utilizado, suelen aplicarse dos supuestos simplificadores:

1. La **Unidad Primaria de Muestreo (UPM)** se considera como la unidad última de información, lo que permite reducir cualquier diseño complejo a un muestreo estratificado bi-etápico (*Ultimate Cluster Sampling*).
2. Las UPM se seleccionan **con reemplazo** desde los estratos de primera etapa.

Estos supuestos tienden a generar una ligera sobreestimación de las varianzas (Heeringa et al., 2017).

VIII.1.1. Método de linealización de Taylor

Este método consiste en linealizar el estimador mediante una expansión de Taylor, y luego calcular su varianza utilizando métodos exactos. Para su implementación en software estadístico como Stata, se requiere disponer de las variables de agrupación (*cluster*), estrato y factor de expansión. Esta exigencia puede ser una limitación, ya que algunas encuestas no publican estas variables para proteger la confidencialidad de los entrevistados.

VIII.1.2. Métodos de réplicas

Los métodos BRR, JRR y Bootstrap utilizan réplicas de la muestra original para estimar la varianza de los estimadores. El procedimiento general incluye los siguientes pasos:

1. Definir $r=1, \dots, R$ réplicas de la muestra.
2. Calcular factores de expansión para cada réplica, reescalando los factores originales y aplicando ajustes por no respuesta y calibración.
3. Estimar el parámetro de interés con los factores originales y con los de cada réplica.
4. Calcular la varianza del estimador.

La diferencia entre los métodos radica en cómo se generan las réplicas:

- **BRR:** Diseñado para diseños con dos UPM por estrato. Cada réplica conserva una de las dos UPM por estrato, generando hasta $2H^2H$ réplicas. Se utiliza la matriz de Hadamard para seleccionar las UPM, asignando el doble del peso original a las seleccionadas y cero a las no seleccionadas.
- **JRR:** Cada réplica elimina una UPM de la muestra. Si hay 1.500 UPM, se generan 1.500 réplicas. Los pesos se reescalan por $n_h / (n_h - 1)$ para las UPM restantes.
- **Bootstrap:** Selecciona $n_h - 1$ UPM por estrato mediante muestreo aleatorio con reemplazo. El peso de cada UPM seleccionada se calcula como:

$$\omega_{hij}^* = \frac{n_h}{n_h - 1} \cdot m_{hij}^* \cdot \omega_{hij}$$

donde m_{hij}^* es la cantidad de veces que se selecciona la UPM en la réplica.

VIII.2. Estimación de varianza en la EFH 2024

En la EFH 2024 se opta por el método de **pesos Bootstrap**, ya que permite preservar la confidencialidad de los hogares. Las bases públicas no incluyen información sobre estratos ni conglomerados, lo que impide aplicar directamente métodos que requieren re-muestreo de la muestra original.

Los pesos Bootstrap se entregan como insumo público sin comprometer la privacidad de los datos, lo que los convierte en una opción óptima para el cálculo de varianzas por parte de usuarios externos.

El procedimiento aplicado en la EFH 2024 es el siguiente:

1. Para cada estrato h , se realiza un muestreo aleatorio con reemplazo para seleccionar una muestra de tamaño $n_h - 1$.
2. Se generan factores de expansión para cada réplica utilizando el mismo método que en la construcción de los factores originales. Estos se denotan como w_{hi}^* .
3. Se repiten los pasos anteriores 1.000 veces, obteniendo $\{w_{hi}^{*b}\}_{b=1}^{1000}$ para cada unidad (h, i) .

Dado el diseño muestral tri-etápico de la EFH 2024, este método asume que la variabilidad proviene principalmente de las dos primeras etapas (selección de comunas y manzanas), lo cual es razonable, ya que la variabilidad de la selección de hogares suele ser menor (HFCS, 2020).

Al igual que en las rondas 2017 y 2021, se incorporó una mejora metodológica: en lugar de aplicar directamente la fórmula de reescalamiento de Rao-Wu (Rao et al., 1992; Girard, 2009; HFCS, 2016), se replicó el proceso completo de generación de factores originales en cada réplica, incluyendo los ajustes por no respuesta y calibración.

IX. Referencias

- Albacete, N. (2014). "Multiple Imputation in the Austrian Household Survey on Housing Wealth". *Austrian Journal of Statistics*, 43(1), 5–28.
- Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research*, 20(1), 40–49.
- Barceló, C. (2006). "Imputation of the 2002 Wave of the Spanish Survey of Household Finances (EFF)". *Documentos Ocasionales N° 0603*, Banco de España.
- Ben-Gal, I. (2005). Outlier detection. En Maimon, O. & Rockach, L. (Eds.), *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*. Kluwer Academic Publishers.
- Bodner, T. (2008). What improves with increased missing data imputations? *Structural Equation Modeling*, 15(4), 651–675.
- Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211–252.
- Bover, O. (2004). "The Spanish Survey of Household Finances (EFF): description and methods of the 2002 wave". *Documentos Ocasionales N° 0409*, Banco de España.
- Bover, O. (2008). "The Spanish Survey of Household Finances (EFF): description and methods of the 2005 wave". *Documentos Ocasionales N° 0803*, Banco de España.
- Centro UC de Encuestas y Estudios Longitudinales (2016). "Diseño Muestral y Factores de Expansión de la VI Ronda EPS". Santiago de Chile.
- Deville, J., & Sarndal, C. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 87(418), 376–382.
- Heeringa, S. G., West, B. T., & Berglund, P. A. (2017). *Applied Survey Data Analysis* (2nd ed.). Chapman and Hall, London.
- HFCN (2016). "The Eurosystem Household Finance and Consumption Survey: methodological report for the second wave". *Statistics Paper Series N°17*.
- Instituto Nacional de Estadísticas – Chile (2025) "Fundamentos del Estándar para la evaluación de la calidad de las estimaciones en encuestas económicas". Documento de trabajo, Santiago de Chile.
- Kennickell, A., & Woodburn, R. (1997). "Consistent Weight Design for the 1989, 1992 and 1995 SCFs, and the Distribution of Wealth". *Working Paper*, Board of Governors of the Federal Reserve Board.

- Little, R., & Rubin, D. (2002). *Statistical Analysis with Missing Data* (2nd ed.). John Wiley & Sons, New York.
- Mander, A., & Clayton, D. (1999). HOTDECK: Stata module to impute missing values using the hotdeck method. *Statistical Software Components S366901*, Boston College Department of Economics.
- McCarthy, P. J. (1969). Pseudoreplication: Half Samples. *Review of the International Statistical Institute*, 37, 239–264.
- Observatorio Social (2016). “Metodología de Diseño Muestral Encuesta de Caracterización Socioeconómica Nacional, Casen 2015”. *Serie Documentos Metodológicos Casen N°3*.
- Portela, J., & Villeta, M. (2007). *Técnicas básicas de Muestreo con SAS*. 1ª Edición. ISBN 978846866133.
- Rao, J. N. K., Wu, C. F. J., & Yue, K. (1992). Some recent work on resampling methods for complex surveys. *Survey Methodology*, 18(2), 209–217.
- Royston, P. (2004). Multiple imputation of missing values. *The Stata Journal*, 4(3), 227–241.
- Rubin, D. B. (1976). Inference and Missing Data. *Biometrika*, 63(3), 581–592.
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, New York.
- Rubin, D. B. (2004). Multiple Imputation After 18+ Years. *Journal of the American Statistical Association*, 91(434), 473–489.
- Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. Chapman & Hall/CRC, New York.
- Singh, A. C., & Mohl, C. A. (1996). Understanding calibration estimators in survey sampling. *Survey Methodology*, 22, 107–115.
- StataCorp. (2019). *Stata 19 Multiple-Imputation Reference Manual*. College Station, TX: Stata Press.
- Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112–118.
- Van Buuren, S. (2012). *Flexible Imputation of Missing Data*. CRC Press.
- Van Buuren, S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, 18(6), 681–694.
- Vijayan, K. (1968). An Exact Sampling Scheme: Generalization of a Method of Hanurav. *Journal of the Royal Statistical Society*, 30(6), 499–513.

Valliant, R., Dever, J., & Kreuter, F. (2013). *Practical Tools for Designing and Weighting Survey Samples*. Springer, New York.

von Hippel, P. T. (2020). How many imputations do you need? A two-stage calculation using a quadratic rule. *Sociological Methods & Research*.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, 30(4), 377–399.

Wolter, K. M. (1985). *Introduction to Variance Estimation*. Springer, New York.

Anexo N°1: Error relativo de la muestra de refresco de la EFH 2024

Macrozona	Estrato	Proporción en 2021	Deff en 2021	Muestra EFH 2024	Error relativo en 2024
Norte	1	26,0%	1,94	132	40,1%
	2	44,9%	1,80	89	30,9%
	3	55,4%	0,68	150	11,9%
Centro	1	22,4%	2,18	337	29,3%
	2	34,8%	2,14	251	24,8%
	3	48,1%	1,63	362	13,7%
Sur	1	23,8%	1,22	184	28,6%
	2	42,2%	0,83	130	18,3%
	3	49,6%	0,59	207	10,5%
RM	1	27,6%	3,89	356	33,2%
	2	41,8%	2,68	238	24,5%
	3	53,4%	1,41	381	11,2%
TOTAL		34,9%	2,53	2817	8,0%

Anexo N°2: Lista de comunas EFH 2024

Macrozona	Comuna	Refresco 2024	Panel (posterior a recontacto)
1	ALTO HOSPICIO	x	x
1	ANDACOLLO		x
1	ANTOFAGASTA	x	x
1	ARICA	x	x
1	CALAMA	x	x
1	COPIAPÓ	x	x
1	COQUIMBO	x	x
1	IQUIQUE	x	x
1	LA SERENA	x	x
1	MEJILLONES	x	
1	OVALLE	x	x
2	ALGARROBO	x	x
2	CABRERO	x	
2	CALERA	x	
2	CHIGUAYANTE	x	x
2	CHILLÁN	x	x
2	CONCEPCIÓN	x	x
2	CONCÓN	x	x
2	CORONEL	x	x
2	CURANILAHUE		x
2	CURICÓ	x	x
2	EL QUISCO		x
2	EL TABO		x
2	FLORIDA		x
2	HUALPÉN	x	x
2	HUALQUI		x
2	LIMACHE	x	
2	LOS ÁNGELES	x	
2	MACHALÍ		x
2	MAULE		x
2	MULCHÉN	x	
2	NOGALES	x	x
2	PENCO	x	x
2	QUILPUÉ	x	x
2	RANCAGUA	x	x
2	SAN ANTONIO		x
2	SAN CARLOS		x
2	SAN FELIPE	x	
2	SAN PEDRO DE LA PAZ	x	x
2	TALCA	x	x
2	TALCAHUANO	x	x
2	TOMÉ		x
2	TUCAPEL		x
2	VALPARAÍSO	x	x
2	VILLA ALEMANA	x	x
2	VIÑA DEL MAR	x	x

Macrozona	Comuna	Refresco 2024	Panel (posterior a recontacto)
3	ANCUD		x
3	ANGOL		x
3	AYSÉN	x	x
3	COYHAIQUE	x	x
3	FREIRE		x
3	NACIMIENTO		x
3	NATALES		x
3	PADRE LAS CASAS		x
3	PORVENIR	x	
3	PUERTO MONTT	x	x
3	PUERTO VARAS		x
3	PUNTA ARENAS	x	x
3	TEMUCO	x	x
3	VALDIVIA	x	x
4	BUIN	x	x
4	CERRILLOS	x	x
4	CERRO NAVIA	x	x
4	COLINA		x
4	CONCHALÍ	x	x
4	EL BOSQUE	x	x
4	ESTACIÓN CENTRAL	x	x
4	HUECHURABA	x	x
4	INDEPENDENCIA	x	x
4	ISLA DE MAIPO		x
4	LA CISTERNA	x	x
4	LA FLORIDA	x	x
4	LA GRANJA	x	x
4	LA PINTANA	x	x
4	LA REINA	x	x
4	LAMPA		x
4	LAS CONDES	x	x
4	LO BARNECHEA	x	x
4	LO ESPEJO	x	x
4	LO PRADO	x	x
4	MACUL	x	x
4	MAIPÚ	x	x
4	ÑUÑO A	x	x
4	PADRE HURTADO	x	x
4	PEDRO AGUIRRE CERDA	x	x
4	PEÑAFLO R		x
4	PEÑALOLÉN	x	x
4	PROVIDENCIA	x	x
4	PUDAHUEL	x	x
4	PUENTE ALTO	x	x
4	QUILICURA	x	x
4	QUINTA NORMAL	x	x
4	RECOLETA	x	x
4	RENCA	x	x
4	SAN BERNARDO	x	x
4	SAN JOAQUÍN	x	x
4	SAN MIGUEL	x	x
4	SAN RAMÓN	x	x
4	SANTIAGO	x	x
4	TALAGANTE		x
4	VITACURA	x	x

Anexo N°3: Reglas de Rubin

Las reglas de Rubin son un conjunto de principios estadísticos propuestos por Donald Rubin para combinar los resultados de un proceso de imputación múltiple (Multiple Imputation, MI). Estas reglas permiten obtener estimaciones válidas y eficientes de parámetros y sus errores estándar después de haber imputado múltiples conjuntos de datos incompletos.

Media de una variable imputada

Suponga que se han generado m conjuntos de datos imputados. Entonces, un parámetro de interés Q (por ejemplo, una media, una regresión, etc.) con este conjunto de datos, se calcula de la siguiente forma:

$$\bar{Q} = \frac{1}{m} \sum_{i=1}^m Q_i$$

Donde Q_i es la estimación del parámetro en el conjunto de datos imputado i . Es decir, el estimador (media, mediana, beta de una regresión, etc.) de una variable imputada para la cual disponemos m conjuntos de datos, corresponde al promedio del estimador obtenido en cada uno de los conjuntos de datos.

Varianza de una variable imputada

La varianza de una variable imputada tiene dos componentes:

1. La varianza dentro de las imputaciones (within):

$$W = \frac{1}{m} \sum_{i=1}^m V_i$$

Donde V_i es la varianza de Q_i .

2. La varianza entre las imputaciones (between):

$$B = \frac{1}{m-1} \sum_{i=1}^m (Q_i - \bar{Q})^2$$

Entonces la varianza total de una variable imputada es:

$$T = W + \left(1 + \frac{1}{m}\right) \cdot B$$

Intervalos de confianza

El intervalo de confianza de una variable imputada, utilizando las reglas de Rubin, es de la siguiente forma:

$$\bar{Q} \pm t_{v, 1-\frac{\alpha}{2}} \cdot \sqrt{T}$$

Donde v corresponden a los grados de libertad ajustados:

$$v = \frac{(m-1) \left(1 + \frac{W}{\left(1 + \frac{1}{m}\right) B} \right)^2}{1 + \frac{1}{m+1} \left(\frac{W}{\left(1 + \frac{1}{m}\right) B} \right)^2}$$

Regresiones y significancia estadística

Para cada uno de los m conjuntos de datos se hace estima el modelo de regresión por MCO, rescatando en cada una de las regresiones los coeficientes estimados β_i y su correspondiente varianza V_i (cuadrado del error estándar).

Luego los coeficientes de regresión se obtienen promediando los coeficientes obtenidos en cada uno de los m conjuntos de datos:

$$\bar{\beta} = \frac{1}{m} \sum_{i=1}^m \beta_i$$

La varianza total de estos coeficientes corresponde a:

$$T = \underbrace{\frac{1}{m} \sum_{i=1}^m V_i}_W + \left(1 + \frac{1}{m}\right) \cdot \underbrace{\frac{1}{m-1} \sum_{i=1}^m (\beta_i - \bar{\beta})^2}_B$$

De esta forma, el error estándar es:

$$SE(\bar{\beta}) = \sqrt{T}$$

Así, el estadístico t es:

$$t = \frac{\bar{\beta}}{SE(\bar{\beta})}$$

El cual se debe comparar con una distribución t-student $t_{v, 1-\frac{\alpha}{2}}$, donde v son los grados de libertad corregidos, definidos en la sección A.3.

Anexo N°4: Análisis de valores faltantes EFH 2024

Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 1	yoprnm	Ingreso relacionado a la actividad principal del individuo	5841	4894	947	16%	MICE
	yosec	Ingreso relacionado otras ocupaciones distintas a la ocupación principal del individuo	787	553	234	30%	MICE
	yoprin_otro_1	OIAP: bonificaciones y gratificaciones	967	775	192	20%	MICE
	yoprin_otro_2	OIAP: remuneración en especies o regalías	107	84	23	21%	MICE
	yoprin_otro_3	OIAP: retiro de productos de su negocio (no agrícolas) para consumo propio	35	27	8	23%	HOTDECK
	yoprin_otro_4	OIAP: otros	27	24	3	11%	HOTDECK
	horas_trabajadas	Horas trabajadas	5841	5680	161	3%	MICE

Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 2	y_pen_1	Pensión de alimentos	408	390	18	4%	MICE
	y_pen_2	Dinero aportado por personas ajenas al hogar (residentes Chile)	185	177	8	4%	MICE
	y_pen_3	Dinero aportado por personas ajenas al hogar (no residentes Chile)	31	29	2	6%	HOTDECK
	y_pen_6	Pensión de vejez, jubilación o rentas vitalicias	792	732	60	8%	MICE
	y_pen_7	Pensión de invalidez	138	128	10	7%	MICE
	y_pen_8	Montepío o pensión de viudez	170	154	16	9%	MICE
	y_pen_9	Pensión de orfandad	18	15	3	17%	HOTDECK
	y_pen_10	Seguro de cesantía	65	51	14	22%	HOTDECK
	y_sub_4	Pensión básica solidaria de vejez o aporte previsional solidario	850	802	48	6%	MICE
	y_sub_5	Pensión básica solidaria de invalidez	159	145	14	9%	MICE
	y_sub_11	Subsidio familiar, duplo, discapacidad, cesantía	139	137	2	1%	MICE
	y_sub_12	Subsidio al agua potable	373	315	58	16%	MICE
	y_sub_13	Sistema de protección social"	20	16	4	20%	HOTDECK
	y_sub_14	Asignación familiar	273	260	13	5%	MICE
	y_sub_15	Otros subsidios del estado	153	143	10	7%	MICE
	y_otro_1	Intereses por ahorro	376	252	124	33%	MICE
	y_otro_2	Dividendos por acciones	63	46	17	27%	HOTDECK
	y_otro_3	Retiro de utilidades de su negocio	53	42	11	21%	HOTDECK
	y_otro_4	Retiro de productos de su negocio	20	16	4	20%	HOTDECK
	y_otro_5	Productos huerto familiar	7	5	2	29%	HOTDECK
	y_otro_6	Venta de productos caseros	50	41	9	18%	HOTDECK
	y_otro_7	Finiquito o indemnización laboral	144	135	9	6%	MICE
	y_otro_8	Devolución de impuestos	474	419	55	12%	MICE
	y_otro_9	Donaciones de instituciones o personas ajenas al hogar	12	10	2	17%	HOTDECK
	y_otro_14	Arriendo de propiedades	280	270	10	4%	MICE
	y_otro_15	Arriendo de maquinarias, herramientas o implementos	5	4	1	20%	HOTDECK
	y_otro_16	Arriendo de vehículos de transporte	16	13	3	19%	HOTDECK
	y_otro_18	Otros Ingresos	40	38	2	5%	HOTDECK

Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 2	act_acc	Valor de acciones transadas en bolsa	119	82	37	31%	MICE
	act_fm	Valor de fondos mutuos	386	308	78	20%	MICE
	act_rv	Valor de otros activos financieros de renta variable	28	21	7	25%	HOTDECK
	act_ps	Valor de participación en sociedades o en fondos de inversión	113	73	40	35%	HOTDECK
	act_cripto	Valor de criptoactivos	78	67	11	14%	HOTDECK
	act_ca	Valor de cuentas de ahorro	779	627	152	20%	MICE
	act_apv	Valor de ahorro previsional voluntario	368	256	112	30%	MICE
	act_rf	Valor de instrumentos de renta fija	433	348	85	20%	MICE
	act_seg	Valor de seguros con ahorro	206	131	75	36%	MICE
	act_acc_12	Valor de acciones transadas en bolsa, inversión en últimos 12 meses	119	48	71	60%	HOTDECK
	act_fm_12	Valor de otros activos financieros de renta variable, inversión en últimos 12 meses	386	222	164	42%	MICE
	act_rv_12	Valor de otros activos financieros de renta variable, inversión en últimos 12 meses	28	17	11	39%	HOTDECK
	act_ps_12	Valor de participación en sociedades o en fondos de inversión, inversión en últimos 12 meses	113	41	72	64%	HOTDECK
	act_cripto_12	Valor de criptoactivos, inversión en últimos 12 meses	78	45	33	42%	HOTDECK
	act_ca_12	Valor de cuentas de ahorro, inversión en últimos 12 meses	779	632	147	19%	MICE
	act_apv_12	Valor de ahorro previsional voluntario, inversión en últimos 12 meses	368	201	167	45%	MICE
	act_rf_12	Valor de instrumentos de renta fija, inversión en últimos 12 meses	433	327	106	24%	MICE
	act_seg_12	Valor de seguros con ahorro, inversión en últimos 12 meses	206	119	87	42%	MICE
	act_aut	Valor de autos o camionetas (incluye taxis)	2648	2514	134	5%	MICE
	act_mot	Valor de motos	208	198	10	5%	MICE
	act_utt	Valor de furgones y utilitarios	84	78	6	7%	MICE
	act_otrosvehic	Valor de otros activos automotrices (ej: embarcaciones, aviones, helicópteros)	16	15	1	6%	HOTDECK
	act_vp	Valor de vivienda propia	2607	2120	487	19%	MICE
	act_otp_1	Valor de otras propiedades	748	664	84	11%	MICE
	act_otp_2	Valor de otras propiedades	228	208	20	9%	MICE
	act_otp_3	Valor de otras propiedades	70	63	7	10%	HOTDECK
	act_ahcta	Valor de ahorro en cuenta corriente/cuenta vista	1181	968	213	18%	MICE
	habah_m	Monto ahorrado en los últimos 12 meses	1573	1344	229	15%	MICE
	cap_pen	Monto ahorrado en los últimos 12 meses	1313	1173	140	11%	MICE
	act_otros	Valor de otros activos reales (maquinaria industrial/agrícola, cabezas de ganado u otros animales, obras de arte, etc.)	53	45	8	15%	HOTDECK

Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 2	m_vp_compra	Valor de compra de vivienda principal	2607	1718	889	34%	MICE
	m_op_venta_1	Valor de venta de la otra propiedad 1	748	664	84	11%	MICE
	m_op_venta_2	Valor de venta de la otra propiedad 2	228	208	20	9%	MICE
	m_op_venta_3	Valor de venta de la otra propiedad 3	70	63	7	10%	HOTDECK
	m_vp_fr	Monto financiado de vivienda principal con recursos propios	1883	1384	499	27%	MICE
	m_vp_fs	Monto financiado de vivienda principal con subsidio	862	453	409	47%	MICE
	m_vp_fch	Monto financiado de vivienda principal con crédito hipotecario	1200	904	296	25%	MICE
	m_vp_foc	Monto financiado de vivienda principal con otras fuentes	129	96	33	26%	MICE
	m_vp_ftr	Monto financiado de vivienda principal transferencias	142	52	90	63%	HOTDECK
	m_vp_ffa	Monto financiado de vivienda principal a través de créditos de familiares o amigos	75	51	24	32%	HOTDECK
	m_otp_fr_1	Monto financiado de otra propiedad 1 con recursos propios	504	444	60	12%	MICE
	m_otp_fs_1	Monto financiado de otra propiedad 1 con subsidio	114	40	74	65%	HOTDECK
	m_otp_fch_1	Monto financiado de otra propiedad 1 con crédito hipotecario	344	307	37	11%	MICE
	m_otp_foc_1	Monto financiado de otra propiedad 1 con otros créditos de instituciones financieras	42	38	4	10%	HOTDECK
	m_otp_ftr_1	Monto financiado de otra propiedad 1 con transferencias	99	84	15	15%	MICE
	m_otp_ffa_1	Monto financiado de otra propiedad 1 con créditos de familiares o amigos	22	18	4	18%	HOTDECK
	m_otp_fr_2	Monto financiado de otra propiedad 2 con recursos propios	164	148	16	10%	MICE
	m_otp_fs_2	Monto financiado de otra propiedad 2 con subsidio	17	7	10	59%	HOTDECK
	m_otp_fch_2	Monto financiado de otra propiedad 2 con crédito hipotecario	92	79	13	14%	MICE
	m_otp_foc_2	Monto financiado de otra propiedad 2 con otros créditos de instituciones financieras	12	11	1	8%	HOTDECK
	m_otp_ftr_2	Monto financiado de otra propiedad 2 con transferencias	33	31	2	6%	HOTDECK
	m_otp_fr_3	Monto financiado de otra propiedad 3 con recursos propios	56	47	9	16%	HOTDECK
	m_otp_fs_3	Monto financiado de otra propiedad 3 con subsidio	7	2	5	71%	HOTDECK
	m_otp_fch_3	Monto financiado de otra propiedad 3 con crédito hipotecario	26	24	2	8%	HOTDECK
	m_otp_foc_3	Monto financiado de otra propiedad 3 con otros créditos de instituciones financieras	4	3	1	25%	HOTDECK
Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 2	div_chvp	Dividendo del crédito hipotecario para vivienda principal	686	662	24	3%	MICE
	div_ocvp	Dividendo del crédito complementario para vivienda principal	28	23	5	18%	HOTDECK
	div_favp	Dividendo del crédito con familiares y amigos para vivienda principal	13	8	5	38%	HOTDECK
	div_chotp_1	Dividendo del crédito hipotecario para otras propiedades	231	223	8	3%	MICE
	div_chotp_2	Dividendo del crédito hipotecario para otras propiedades	53	51	2	4%	HOTDECK
	div_chotp_3	Dividendo del crédito hipotecario para otras propiedades	16	15	1	6%	HOTDECK
	div_ocotp_2	Dividendo del crédito complementario para otra propiedad 2	4	3	1	25%	HOTDECK
	div_faotp_1	Dividendo del crédito con familiares y amigos para otra propiedad 1	5	3	2	40%	HOTDECK
	pp_m_chvp	Plazo por pagar en meses del crédito hipotecario para vivienda principal	686	683	3	0%	MICE
	pp_m_ocvp	Plazo por pagar en meses del crédito complementario para vivienda principal	28	26	2	7%	HOTDECK
	pp_m_favp	Plazo por pagar en meses del crédito con familiares y amigos para vivienda principal	9	7	2	22%	HOTDECK
	pp_m_chotp_1	Plazo por pagar en meses del crédito hipotecario para otras propiedades	231	230	1	0%	MICE
	pp_m_faotp_1	Plazo por pagar en meses del crédito con familiares y amigos para otra propiedad 1	5	3	2	40%	HOTDECK
	pp_m_chotp_2	Plazo por pagar en meses del crédito hipotecario para otras propiedades	53	52	1	2%	HOTDECK
	pp_m_faotp_2	Plazo por pagar en meses del crédito con familiares y amigos para otra propiedad 2	2	1	1	50%	HOTDECK

Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 2	gasto_m_super	Gasto mensual en Supermercado	4649	4467	182	4%	MICE
	gasto_m_almacen	Gasto mensual en Ferias libres, almacenes y negocios de barrio	4649	4393	256	6%	MICE
	gasto_m_agua	Gasto mensual en agua potable	4649	4467	182	4%	MICE
	gasto_m_luz	Gasto mensual en luz o electricidad	4649	4464	185	4%	MICE
	gasto_m_gas	Gasto mensual en gas (licuado o por cañería)	4649	4421	228	5%	MICE
	gasto_m_leña	Gasto mensual en leña	4649	4216	433	9%	MICE
	gasto_m_internet	Gasto mensual en servicios de internet y/o televisión pagada (cable, satelital o servicios de streaming)	4649	4323	326	7%	MICE
	gasto_m_fono	Gasto mensual en telefonía fija o móvil	4649	4313	336	7%	MICE
	gasto_m_gcomun	Gasto mensual en gastos comunes de condominio o edificio	4649	4147	502	11%	MICE
	gasto_m_publico	Gasto mensual en transporte público (buses, micros, metro, colectivos, trenes, taxis y taxis en aplicaciones)	4649	4327	322	7%	MICE
	gasto_m_bencina	Gasto mensual en bencina o combustibles para vehículos del hogar	4649	4268	381	8%	MICE
	gasto_m_tag	Gasto mensual en peajes o tag	4649	4070	579	12%	MICE
	gasto_m_cmedicas	Gastos en consultas médicas (incluya dentales, psicológicas, kinesiológicas, etc.)	4649	4277	372	8%	MICE
	gasto_m_exmedicos	Gastos en exámenes médicos y procedimientos médicos ambulatorios	4649	4236	413	9%	MICE
	gasto_m_medicamen	Gasto en medicamentos	4649	4289	360	8%	MICE
	gasto_m_electro	Gasto en electrodomésticos pequeños	4649	4558	91	2%	MICE
	gasto_m_mantenviv	Gasto en servicios y/o compra de materiales para conservación de la vivienda	4649	4566	83	2%	MICE
	gasto_m_vestuario	Gasto en vestuario y zapatos	4649	4555	94	2%	MICE
	gasto_m_celulares	Gasto en compra de equipos de telefonía móvil	4649	4572	77	2%	MICE
	gasto_m_elecgrandes	Gasto en electrodomésticos grandes	4649	4582	67	1%	MICE
	gasto_m_muebles	Gasto en muebles: mensualizado	4649	4588	61	1%	MICE
	m_arriendo_vp	Monto de arriendo de la vivienda	1233	1201	32	3%	MICE
	yai	Alquiler imputado vivienda propia	3408	3099	309	9%	MICE

Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 3	cf_td1_1	Carga financiera en deudas asociadas a tarjetas de crédito bancarias	796	718	78	10%	MICE
	cf_td1_2	Carga financiera en deudas asociadas a tarjetas de crédito bancarias	311	274	37	12%	MICE
	cf_td1_3	Carga financiera en deudas asociadas a tarjetas de crédito bancarias	91	70	21	23%	MICE
	cf_td2_1	Carga financiera en deudas asociadas a línea de crédito	282	229	53	19%	MICE
	cf_td2_2	Carga financiera en deudas asociadas a línea de crédito	76	55	21	28%	HOTDECK
	cf_td2_3	Carga financiera en deudas asociadas a línea de crédito	20	12	8	40%	HOTDECK
	cf_td3_1	Carga financiera en deudas asociadas a tarjetas de crédito de casas comerciales	914	811	103	11%	MICE
	cf_td3_2	Carga financiera en deudas asociadas a tarjetas de crédito de casas comerciales	342	286	56	16%	MICE
	cf_td3_3	Carga financiera en deudas asociadas a tarjetas de crédito de casas comerciales	113	83	30	27%	MICE
	cf_td4_1	Carga financiera en deudas asociadas a préstamos de consumo en casas comerciales	154	149	5	3%	MICE
	cf_td4_2	Carga financiera en deudas asociadas a préstamos de consumo en casas comerciales	28	26	2	7%	HOTDECK
	cf_td4_3	Carga financiera en deudas asociadas a préstamos de consumo en casas comerciales	8	5	3	38%	HOTDECK
	cf_td5_1	Carga financiera en deudas asociadas a préstamos de consumo bancarios o en financieras	480	455	25	5%	MICE
	cf_td5_2	Carga financiera en deudas asociadas a préstamos de consumo bancarios o en financieras	105	89	16	15%	MICE
	cf_td5_3	Carga financiera en deudas asociadas a préstamos de consumo bancarios o en financieras	21	18	3	14%	HOTDECK
	cf_td6_1	Carga financiera en deudas asociadas a préstamos de cajas de compensación, cooperativas u otros similares	383	353	30	8%	MICE
	cf_td6_3	Carga financiera en deudas asociadas a préstamos de cajas de compensación, cooperativas u otros similares	7	6	1	14%	HOTDECK
	cf_td7_1	Carga financiera en deudas asociadas a créditos automotrices	185	177	8	4%	MICE
	cf_td8_1	Carga financiera en deudas educacionales	337	271	66	20%	MICE
	cf_td8_2	Carga financiera en deudas educacionales	70	56	14	20%	MICE
	cf_td9_1	Carga Financiera Efectiva(monto facturado) de Préstamos parientes o amigos	92	90	2	2%	MICE
	cf_td12_1	Carga Financiera Efectiva(monto facturado) de Fiado	26	25	1	4%	HOTDECK
	cf_td13_1	Carga Financiera Efectiva(monto facturado) de Otras fuentes de crédito	41	39	2	5%	HOTDECK
	ct_td1_1	Cupo total de la tarjeta de crédito 1	813	740	73	9%	MICE
	ct_td1_2	Cupo total de la tarjeta de crédito 2	328	290	38	12%	MICE
	ct_td1_3	Cupo total de la tarjeta de crédito 3	108	76	32	30%	MICE
	ct_td2_1	Cupo total de la línea de crédito 1	305	250	55	18%	MICE
	ct_td2_2	Cupo total de la línea de crédito 2	99	59	40	40%	MICE
	ct_td2_3	Cupo total de la línea de crédito 3	43	14	29	67%	HOTDECK
	ct_td3_1	Cupo total de la tarjeta de casas comerciales 1	935	829	106	11%	MICE
	ct_td3_2	Cupo total de la tarjeta de casas comerciales 2	363	308	55	15%	MICE
	ct_td3_3	Cupo total de la tarjeta de casas comerciales 3	134	103	31	23%	MICE

Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 3	m_td1_1	Monto inicial de la deuda asociada a tarjetas de crédito bancarias	796	759	37	5%	MICE
	m_td1_2	Monto inicial de la deuda asociada a tarjetas de crédito bancarias	311	292	19	6%	MICE
	m_td1_3	Monto inicial de la deuda asociada a tarjetas de crédito bancarias	91	77	14	15%	MICE
	m_td2_1	Monto inicial de la deuda asociada a línea de crédito	282	240	42	15%	MICE
	m_td2_2	Monto inicial de la deuda asociada a línea de crédito	76	58	18	24%	MICE
	m_td2_3	Monto inicial de la deuda asociada a línea de crédito	20	13	7	35%	HOTDECK
	m_td3_1	Monto inicial de la deuda asociada a tarjetas de crédito de casas comerciales	914	849	65	7%	MICE
	m_td3_2	Monto inicial de la deuda asociada a tarjetas de crédito de casas comerciales	342	310	32	9%	MICE
	m_td3_3	Monto inicial de la deuda asociada a tarjetas de crédito de casas comerciales	113	99	14	12%	MICE
	m_td4_1	Monto inicial de la deuda asociada a préstamos de consumo en casas comerciales	154	146	8	5%	MICE
	m_td4_2	Monto inicial de la deuda asociada a préstamos de consumo en casas comerciales	28	27	1	4%	HOTDECK
	m_td4_3	Monto inicial de la deuda asociada a préstamos de consumo en casas comerciales	8	7	1	13%	HOTDECK
	m_td5_1	Monto inicial de la deuda asociada a préstamos de consumo bancarios o en financieras	480	452	28	6%	MICE
	m_td5_2	Monto inicial de la deuda asociada a préstamos de consumo bancarios o en financieras	105	92	13	12%	MICE
	m_td5_3	Monto inicial de la deuda asociada a préstamos de consumo bancarios o en financieras	21	18	3	14%	HOTDECK
	m_td6_1	Monto inicial de la deuda asociada a préstamos de cajas de compensación, cooperativas u otros similares	383	360	23	6%	MICE
	m_td6_2	Monto inicial de la deuda asociada a préstamos de cajas de compensación, cooperativas u otros similares	53	51	2	4%	HOTDECK
	m_td7_1	Monto inicial de la deuda asociada a créditos automotrices	185	169	16	9%	MICE
	m_td8_1	Monto inicial de la deuda asociada a créditos educacionales	460	350	110	24%	MICE
	m_td8_2	Monto inicial de la deuda asociada a créditos educacionales	93	61	32	34%	MICE
	m_td8_3	Monto inicial de la deuda asociada a créditos educacionales	8	7	1	13%	HOTDECK
	m_td9_1	Monto inicial de la deuda asociada a préstamos de parientes o amigos	154	144	10	6%	MICE
	m_td12_1	Monto inicial de la deuda asociada a pedir fiado	50	49	1	2%	HOTDECK
	m_td13_1	Monto inicial de la deuda asociada a créditos de otras fuentes	51	46	5	10%	HOTDECK

Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 3	pp_td1_1	Plazo por pagar en meses en deudas asociadas a tarjetas de crédito bancarias	796	756	40	5%	MICE
	pp_td1_2	Plazo por pagar en meses en deudas asociadas a tarjetas de crédito bancarias	311	295	16	5%	MICE
	pp_td1_3	Plazo por pagar en meses en deudas asociadas a tarjetas de crédito bancarias	91	78	13	14%	MICE
	pp_td2_1	Plazo por pagar en meses en deudas asociadas a líneas de crédito	282	256	26	9%	MICE
	pp_td2_2	Plazo por pagar en meses en deudas asociadas a líneas de crédito	76	62	14	18%	MICE
	pp_td2_3	Plazo por pagar en meses en deudas asociadas a líneas de crédito	20	14	6	30%	HOTDECK
	pp_td3_1	Plazo por pagar en meses en deudas asociadas a tarjetas de crédito de casas comerciales	914	841	73	8%	MICE
	pp_td3_2	Plazo por pagar en meses en deudas asociadas a tarjetas de crédito de casas comerciales	342	301	41	12%	MICE
	pp_td3_3	Plazo por pagar en meses en deudas asociadas a tarjetas de crédito de casas comerciales	113	95	18	16%	MICE
	pp_td4_1	Plazo por pagar en meses en deudas asociadas a préstamos de consumo en casas comerciales	154	141	13	8%	MICE
	pp_td4_2	Plazo por pagar en meses en deudas asociadas a préstamos de consumo en casas comerciales	28	26	2	7%	HOTDECK
	pp_td4_3	Plazo por pagar en meses en deudas asociadas a préstamos de consumo en casas comerciales	8	5	3	38%	HOTDECK
	pp_td5_1	Plazo por pagar en meses en deudas asociadas a préstamos de consumo bancarios o en financieras	480	449	31	6%	MICE
	pp_td5_2	Plazo por pagar en meses en deudas asociadas a préstamos de consumo bancarios o en financieras	105	91	14	13%	MICE
	pp_td5_3	Plazo por pagar en meses en deudas asociadas a préstamos de consumo bancarios o en financieras	21	18	3	14%	HOTDECK
	pp_td6_1	Plazo por pagar en meses en deudas asociadas a créditos automotrices	383	356	27	7%	MICE
	pp_td6_2	Plazo por pagar en meses en deudas asociadas a créditos automotrices	53	50	3	6%	HOTDECK
	pp_td7_1	Plazo por pagar en meses en deudas asociadas a créditos de cajas de compensación, cooperativas u otros similares	185	175	10	5%	MICE
	pp_td8_1	Plazo por pagar en meses en deudas asociadas a créditos educacionales	337	189	148	44%	MICE
	pp_td8_2	Plazo por pagar en meses en deudas asociadas a créditos educacionales	70	37	33	47%	HOTDECK
	pp_td8_3	Plazo por pagar en meses en deudas asociadas a créditos educacionales	4	2	2	50%	HOTDECK
Bloque	Variable	Descripción	Observaciones	No missings	Missings (N)	Missings (%)	Método
Bloque 3	pr_td1_1	Pago total realizado en deudas asociadas a tarjetas de crédito bancarias	796	709	87	11%	MICE
	pr_td1_2	Pago total realizado en deudas asociadas a tarjetas de crédito bancarias	311	273	38	12%	MICE
	pr_td1_3	Pago total realizado en deudas asociadas a tarjetas de crédito bancarias	91	69	22	24%	MICE
	pr_td3_1	Pago total realizado en deudas asociadas a tarjetas de crédito de casas comerciales	914	765	149	16%	MICE
	pr_td3_2	Pago total realizado en deudas asociadas a tarjetas de crédito de casas comerciales	342	260	82	24%	MICE
	pr_td3_3	Pago total realizado en deudas asociadas a tarjetas de crédito de casas comerciales	113	79	34	30%	MICE
	pr_td4_1	Pago total realizado en deudas asociadas a préstamos de consumo en casas comerciales	154	135	19	12%	MICE
	pr_td4_2	Pago total realizado en deudas asociadas a préstamos de consumo en casas comerciales	28	25	3	11%	HOTDECK
	pr_td4_3	Pago total realizado en deudas asociadas a préstamos de consumo en casas comerciales	8	6	2	25%	HOTDECK
	pr_td5_1	Pago total realizado en deudas asociadas a préstamos de consumo bancarios o en financieras	479	434	45	9%	MICE
	pr_td5_2	Pago total realizado en deudas asociadas a préstamos de consumo bancarios o en financieras	105	86	19	18%	MICE
	pr_td5_3	Pago total realizado en deudas asociadas a préstamos de consumo bancarios o en financieras	21	18	3	14%	HOTDECK
	pr_td6_1	Pago total realizado en deudas asociadas a créditos de cajas de compensación, cooperativas u otros similares	383	343	40	10%	MICE
	pr_td6_2	Pago total realizado en deudas asociadas a créditos de cajas de compensación, cooperativas u otros similares	53	51	2	4%	HOTDECK
	pr_td6_3	Pago total realizado en deudas asociadas a créditos de cajas de compensación, cooperativas u otros similares	7	6	1	14%	HOTDECK
	pr_td7_1	Pago total realizado en deudas asociadas a créditos automotrices	185	170	15	8%	MICE
	pr_td7_2	Pago total realizado en deudas asociadas a créditos automotrices	16	14	2	13%	HOTDECK

Anexo N°5: Transformación Box-Cox

La transformación Box-Cox es una técnica utilizada para estabilizar la varianza y aproximar la normalidad en la distribución de la variable. Esto es específicamente útil en el contexto de imputación múltiple, donde los modelos utilizados tienen un mejor comportamiento bajo el supuesto de normalidad de las variables.

Los principales objetivos de aplicar esta transformación son:

- Reducir la asimetría de la distribución
- Mejorar el ajuste a la normalidad
- Facilitar la aplicación del método de imputación múltiple

La transformación de Box-Cox se define como:

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{si } \lambda \neq 0 \\ \log(y), & \text{si } \lambda = 0 \end{cases}$$

Donde:

y es el valor original de la variable

λ es el parámetro de transformación que se estima a partir de los datos por máxima verosimilitud.

En el caso particular de la EFH, se utiliza el comando `bcskew0` en Stata que aplica transformaciones de Box-Cox con el objetivo específico de reducir la asimetría (*skewness*) de una variable. Así, este comando busca el valor del parámetro de transformación λ que minimiza la asimetría de la variable transformada. A diferencia de la transformación de Box-Cox tradicional, que suele estimar λ mediante máxima verosimilitud para aproximar la normalidad, `bcskew0` se enfoca en encontrar el valor de λ que haga que la asimetría sea cero (o lo más cercana posible a cero).