

Instituto Nacional de Estadísticas

Documento Metodológico Encuesta Suplementaria de Ingresos (ESI) 2022

Agosto 2023

INSTITUTO NACIONAL DE ESTADÍSTICAS

Subdirección Técnica

Departamento de Estadísticas del Trabajo

Subdepartamento de Estadísticas de Ingresos del Trabajo

Departamento de Metodologías e Innovación Estadística

Departamento de Operaciones de Estadísticas Sociales

DOCUMENTO METODOLÓGICO ENCUESTA SUPLEMENTARIA DE INGRESOS (ESI)

Agosto / 2023

Versión 1.0

Dirección: Morandé N° 801, piso 22, Santiago, Chile

Código postal: 8340148

Teléfono: 562 3246 1010

Sitio web: www.ine.gob.cl

Correo electrónico: ine@ine.gob.cl

Facebook: @ChileINE

Twitter: @INE_Chile

Índice

1	Introducción	6
2	Contexto metodológico	7
2.1	Antecedentes	7
2.1.1	Actualizaciones metodológicas y refinamiento del cuestionario de ingresos realizados en 2010.....	7
2.1.2	Hitos y modificaciones más importantes 2011-2022	11
2.1.3	Necesidades de información	14
2.2	Objetivos de la Encuesta Suplementaria de Ingresos	15
2.2.1	Objetivo general	15
2.2.2	Objetivos específicos	15
3	Marco de referencia	16
3.1	Marco conceptual	16
3.1.1	Relevancia y directrices internacionales para la elaboración de estadísticas de ingresos	16
3.1.2	Lineamientos conceptuales y metodológicos para la elaboración de estadísticas de ingresos	17
3.1.3	Referencias nacionales e internacionales	19
4	Diseño estadístico.....	26
4.1	Universo	26
4.2	Población objetivo.....	26
4.3	Marco estadístico	26
4.4	Cobertura geográfica.....	26
4.5	Unidades estadísticas.....	26
4.5.1	Unidad de muestreo.....	27
4.5.2	Unidad de información y análisis	27
4.6	Definición de variables	28
4.6.1	Variables analíticas	28
4.6.2	Variables de información muestral.	29
4.6.3	Variables de identificación de la persona	30
4.7	Nomenclatura utilizada	30
4.7.1	Clasificación Internacional Normalizada de la Educación del año 1997 (CINE 97).....	30
4.7.2	Clasificación Internacional Uniforme de Ocupaciones del año 2008, adaptación nacional (CIUO 08.CL).....	31

4.7.3	Clasificación Internacional Industrial Uniforme revisión 4.CL según la Clasificación de Actividades Económicas Nacional de Encuestas Sociodemográficas (CAENES).....	32
4.7.4	Clasificación Internacional de la Situación en el Empleo del año 1993 (CISE 93) 33	
4.8	Diseño según tipo de fuente: Muestra.....	33
4.8.1	Diseño y selección muestral	33
4.8.2	Cálculo del factor de expansión	33
4.9	Medidas de precisión	41
4.9.1	Error absoluto e intervalo de confianza	41
4.9.2	Error relativo y coeficiente de variación	42
4.9.3	Contrastes de hipótesis	43
4.9.4	Representatividad de las estimaciones.....	46
4.10	Medidas complementarias.....	48
4.10.1	Tasa de respuesta y tasa de logro objetivo de la recolección.....	48
4.10.2	Tasas de supervisión.....	50
4.10.3	Tasa de no respuesta parcial y total.....	51
4.10.4	Tasas de imputación.....	53
4.11	Estimación de parámetros	53
4.11.1	Estimador de totales	53
4.11.2	Estimador de razón.....	54
4.11.3	Estimador de promedio.....	54
4.11.4	Estimador de mediana.....	56
5	Metodología de Recolección de datos	57
5.1	Técnica y métodos de recolección de los datos.....	57
5.2	Diseño de instrumento de captura de información.....	57
5.3	Prueba y evaluación instrumento de captura de información	61
5.3.1	Prueba Piloto 2010	61
5.3.2	Prueba Cualitativa 2021	62
5.4	Método de capacitación para la recolección de datos	62
5.5	Métodos de Supervisión	64
5.5.1	Supervisión directa	64
5.5.2	Supervisión indirecta re-entrevista	64
5.5.3	Supervisión de gabinete y reporte de consultas	65
6	Metodología de procesamiento de la información.....	68
6.1	Procesamiento de la información	68

6.1.1	Integración de datos.....	68
6.1.2	Clasificación y codificación	68
6.1.3	Revisión y validación	69
6.1.4	Imputación	70
6.1.5	Derivar nuevas variables y unidades	72
7	Metodología de análisis.....	76
7.1	Procedimiento para la anonimización de los datos	76
7.2	Cálculo de estimaciones y validación de resultados	76
7.3	Finalización de resultados	77
8	Alcance de la presentación de resultados.....	78
8.1	Productos estadísticos a publicar.....	78
8.2	Difusión de los resultados	79
8.2.1	Microdatos	79
8.2.2	Metadatos	80
8.2.3	Otros.....	81
8.2.4	Consideraciones sobre el uso de la información.....	81
8.3	Accesibilidad a los datos y Marco legal.....	82
9	Glosario.....	83
10	Bibliografía.....	86
11	Anexos	89
11.1	Formulario ESI 2022	89
11.2	Sintaxis para el cálculo de medianas en R.....	91

1 Introducción

La Encuesta Suplementaria de Ingresos (ESI), realizada por el Instituto Nacional de Estadísticas (INE), es un módulo complementario aplicado al interior de la Encuesta Nacional de Empleo (ENE). Se recolecta cada año en el trimestre móvil octubre, noviembre y diciembre (en adelante “trimestre OND” u “OND”), con el objetivo de caracterizar los ingresos laborales de las personas que son clasificadas como ocupadas en la ENE, así como los ingresos provenientes de otras ocupaciones, distintas de la principal.

La ESI, al ser un módulo complementario de la ENE, comparte sus mismos dominios de estudio o niveles de estimación, en los que el diseño muestral de la ENE garantiza estimaciones representativas a nivel nacional, nacional urbano, nacional rural, regional, en el área urbana de las 16 regiones y en el área rural de las regiones de Coquimbo, Valparaíso, O’Higgins, Maule, Biobío, La Araucanía, Los Lagos, Metropolitana, Los Ríos y Ñuble¹.

El presente documento tiene como propósito describir los procesos y metodologías utilizadas en la producción de la ESI, apuntando a lograr que las personas que acceden a esta tengan un conocimiento global de los productos estadísticos que se desprenden, desde su diseño hasta la entrega de resultados finales, aumentando el entendimiento de la encuesta y sacando el máximo provecho de la información proporcionada.

El documento se encuentra organizado en capítulos, luego de la presente introducción, el segundo de ellos se refiere al contexto metodológico de la encuesta, abarcando antecedentes, objetivos y necesidad de información. El tercer capítulo aborda el marco de referencia y el contexto metodológico en el que se enmarca el desarrollo de la ESI. El cuarto capítulo desarrolla el diseño estadístico de la encuesta. El capítulo cinco se refiere a la metodología de recolección de datos. El sexto capítulo describe la metodología para el procesamiento de la información. El séptimo capítulo desarrolla la metodología de análisis. Finalmente, el octavo capítulo expone el alcance de la presentación de resultados.

¹ Para más información, consultar el documento metodológico de la Encuesta Nacional de Empleo (ENE) 2021 https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodologia-ene-2020.pdf?sfvrsn=793380e9_14

2 Contexto metodológico

2.1 Antecedentes

Las estadísticas sobre fuerza de trabajo tienen una larga tradición en el INE, siendo –junto al Censo de Población y Vivienda y el Índice de Precios al Consumidor (IPC)– las mediciones más antiguas que se realizan en la institución. La Encuesta Nacional de Empleo (ENE), en particular, tiene origen en 1966. Por su parte, la ESI como módulo suplementario de la ENE se registra a partir de 1990.

A contar de 2010, los resultados de la ESI provienen de un nuevo cuestionario, el que incluyó una serie de mejoras a las preguntas de ingresos que permitieron una mayor precisión en su captura. Adicionalmente, constituyeron la primera medición de ingresos con base en la vigente ENE y, en consecuencia, incorporaron los cambios conceptuales de esta última. Dado lo anterior, a partir de 2010 los resultados no son comparables con versiones anteriores.

Adicionalmente, a contar de la ESI 2021 se acotaron los objetivos de la encuesta enfocándose en la captura y medición de ingresos del trabajo². Lo anterior, implica que ya no se capturan ingresos de otras fuentes de los hogares, lo que deriva en que ya no se difundirán estimaciones de ingresos a nivel de hogar.

A continuación, se presentan las modificaciones y actualizaciones de la ESI desde 2010.

2.1.1 Actualizaciones metodológicas y refinamiento del cuestionario de ingresos realizados en 2010

2.1.1.1 Actualizaciones metodológicas de medición

Al ser un módulo de la vigente ENE, las actualizaciones metodológicas en la medición del empleo y el desempleo, incorporadas en 2010, afectaron indirectamente a la ESI. Estas modificaciones se sintetizan en los siguientes puntos:

1. Actualización del concepto de persona ocupada, en el que se incluye el criterio explícito de haber trabajado al menos 1 hora durante la semana de referencia de la encuesta y de percibir un ingreso por el trabajo realizado, a excepción del familiar no remunerado.
2. Refinamiento de los filtros para capturar la situación en el empleo que derivó en un mayor rigor en la composición de las sub-poblaciones de la CISE³ y que, a su vez, resultó de

² Para más información, consultar [https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/comites-y-notas-tecnicas/2021/comit%C3%A9-t%C3%A9cnico-de-expertos-estad%C3%ADsticas-laborales-\(junio-2021\).pdf?sfvrsn=355552fe_4](https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/comites-y-notas-tecnicas/2021/comit%C3%A9-t%C3%A9cnico-de-expertos-estad%C3%ADsticas-laborales-(junio-2021).pdf?sfvrsn=355552fe_4)

³ Clasificación Internacional de Situación en el Empleo (CISE).

fundamental importancia en el módulo de ingresos y la pregunta del mismo por la que deben comenzar los distintos integrantes del hogar.

3. Actualización de nomenclaturas asociadas a nivel de educación (CINE-97), rama de actividad económica (CIU-Rev.3) y grupo ocupacional (CIUO-88)⁴. Respecto de rama de actividad, además, hubo un cambio en la pregunta que pasó de referirse del establecimiento a la empresa en que trabaja la persona ocupada. Esto no es menor, puesto que la empresa o razón social puede involucrar a más de un establecimiento y ello tiene efecto sobre la actividad en que, finalmente, es clasificado un trabajador.

Al igual que en las versiones anteriores, los ingresos del trabajo siempre se refieren a los ingresos netos. De esta forma, si se trata de trabajo asalariado se deben deducir del ingreso bruto los descuentos legales (pagos previsionales, de salud, seguro de cesantía y el denominado “impuesto único”); si se trata de trabajo independiente, a los ingresos brutos se les deben restar los costos totales asociados a la actividad realizada de manera que se obtenga el beneficio neto⁵.

En estricto sentido, y de acuerdo con las recomendaciones internacionales relativas a encuestas de ingresos, el ingreso disponible que mide la ESI es el ingreso corriente que puede destinarse íntegramente al consumo o al ahorro/inversión (United Nations, 2011). Este concepto no se alteró con respecto a la antigua medición.

2.1.1.2 Refinamiento del cuestionario de ingresos

Las mejoras al cuestionario de ingresos consistieron en lo siguiente:

1. Garantizar que los ingresos de la ocupación fueran plenamente concordantes con la actividad laboral reportada por los informantes en el módulo central del cuestionario.

Esta fue la mejora más importante sobre el formulario, y tuvo como objetivo distinguir si el empleo actual de la persona, o sea, aquel que se reportó en la semana de referencia, fuera el mismo que tenía en el mes anterior al mes de levantamiento de la encuesta.

Este cambio significó establecer un vínculo entre la caracterización del empleo que brinda el formulario central de la encuesta y la información de ingresos de la ocupación que brinda la ESI. En las antiguas ediciones de la encuesta no era posible garantizar este vínculo, pues nada aseguraba que el trabajo reportado en la encuesta de empleo fuera el mismo cuyo ingreso se señalaba en la antigua ESI⁶.

⁴ Clasificación Internacional de Nivel Educativo (CINE), Clasificación Internacional Industrial Uniforme (CIU), Clasificación Internacional Uniforme de Ocupación (CIUO).

⁵ En el caso de las personas trabajadoras independientes que se autoasignan un sueldo, éste se suma a los beneficios netos que reporten ya que es parte de su ingreso disponible como personas naturales.

⁶ Hay que tener en consideración que los períodos de referencia de estas encuestas son distintos: la vigente ENE consulta sobre la semana anterior, la ESI sobre el mes calendario anterior a la semana de referencia de la encuesta nacional de empleo.

Gracias a este cambio, el análisis de la información del empleo se enriquece con lo relativo a ingresos laborales.

2. Medir con mayor precisión los componentes variables y las regalías de la remuneración de las personas ocupadas dependientes, estableciendo una clara distinción entre ambas fuentes de ingreso.
3. Contabilizar el tiempo efectivo trabajado de las personas ocupadas dependientes en el empleo actual durante el mes calendario anterior a la semana de referencia. Esta es una mejora importante, por cuanto permite vincular de manera directa las horas trabajadas con el ingreso percibido por dicho trabajo.
4. Respecto del ingreso de las personas ocupadas independientes, se continúa preguntando por el ingreso neto, pero aclarando que quienes se autoasignaran un sueldo debían agregarlo a dicho neto. Asimismo, se mejoró la pregunta relativa al autoconsumo y se incluyó una asociada específicamente a la utilización de la vivienda propia en el proceso productivo.
5. Incorporar una pregunta que permite conocer los ingresos de otro empleo distinto al actual, que tanto las personas ocupadas como las no ocupadas de la semana de referencia, percibieron durante el mes de referencia. En el caso de la población ocupada, estos ingresos laborales se refieren conceptualmente a los ingresos por una actividad secundaria o distinta a la actual, en cambio en la población no ocupada, inactiva o desocupada, estos se explican porque la condición de actividad en el mes anterior puede ser distinta a la que actualmente tiene la persona, por ejemplo, porque fue despedida el mes pasado y actualmente se encuentra buscando trabajo. En suma, con esta pregunta se pueden tener indicadores asociados a la no ocupación reciente, obteniendo la masa de ingresos provenientes de todos los trabajos, denominada como ingresos de la ocupación⁷.
6. Las preguntas de ingresos de otras fuentes también fueron refinadas. Se distinguió el arriendo entre bienes inmuebles y muebles, diferenciando si existe carácter agrícola en ellos. Asimismo, se depuró la pregunta de arriendo imputado permitiendo distinguir los distintos tipos de propiedad de la vivienda y capturando información relevante para el proceso posterior de imputación de esta variable. En cuanto a los instrumentos de ahorro, se actualizaron los posibles tramos de percepción de intereses y se mejoró el reporte de no respuesta.
7. Lo relativo a transferencias significó la desagregación o apertura de las preguntas que capturan la gama de posibles pensiones, subsidios, y donaciones de la que puede beneficiarse una persona u hogar (jubilaciones/montepío, pensiones, subsidios especiales/becas, transferencias entre hogares, y otros ingresos del exterior). Lo anterior,

⁷ Los ingresos de la ocupación se dividen en ingresos por sueldos y salarios, ingresos de empleadores y cuenta propia e ingresos por otros trabajos. Este último ítem incorpora los ingresos de empleos distintos al actual en el mes anterior -actividad secundaria- como ingresos laborales recibidos por inactivos o desocupados, en el mes anterior de la encuesta.

implicó que la captura de la información mejorara tanto en términos de calidad como de cobertura, lo que se tradujo en niveles mayores, en comparación a los registrados en la antigua medición.

8. Otra de las mejoras relevantes al cuestionario de ingresos fue el énfasis en distinguir los diferentes alcances de la no respuesta. En ese sentido, se diferenció el “no sabe” del “no responde” en las preguntas claves del módulo y, se espera que esta distinción se profundice en los cuestionarios futuros de la encuesta⁸.

2.1.1.3 Implicancias en la comparabilidad de los datos.

De acuerdo con lo explicado en el punto anterior, la ESI no es comparable con las ediciones de la encuesta anteriores a 2010. Ello se debe principalmente a lo siguiente:

1. El nuevo módulo de ingresos es una extensión del nuevo formulario de empleo, por tanto, el concepto de ocupación no es el mismo que en la antigua ESI⁹.
2. El refinamiento del cuestionario implicó un mayor nivel de especificidad y precisión en la captura de todas las partidas que conforman el ingreso de las personas y los hogares. Por tanto, a pesar de que el concepto de ingreso que se busca medir es el mismo, existe un efecto asociado a cambios en preguntas del cuestionario (inclusión, reformulación y precisiones) que hace que los datos no sean comparables.
3. Adicionalmente, se depuró la población ocupada de manera que solamente se cuentan las personas que efectivamente mantuvieron su empleo en los dos períodos de referencia involucrados (semana de referencia y mes de referencia).
4. Al actualizarse las nomenclaturas para rama de actividad (a CIU Rev3.), para grupo de ocupación (a CIUO 88), y para nivel de educación (a CINE 97) las agregaciones de información no coinciden con las de las encuestas anteriores.

Respecto al ingreso total de los hogares y su distribución, no se implementaron actualizaciones conceptuales respecto de años previos, sin embargo, no garantiza su comparabilidad, ya que los nuevos filtros de la vigente ENE derivaron en una reestructuración de la oferta laboral dentro de los hogares, en la cual la población ocupada ganó peso relativo dentro de la población total.

⁸ Más detalles ver el apartado “Hitos y modificaciones más importantes 2010-2021”.

⁹ Se pide, además de estar ocupado, percibir un ingreso por la ocupación (salvo en el caso del familiar no remunerado). Adicionalmente, se es más riguroso en exigir el haber trabajado al menos una hora en la semana de referencia de la encuesta.

2.1.2 Hitos y modificaciones más importantes 2011-2022

2.1.2.1 Mejoras al cuestionario

Durante la recolección de la ESI, el aprendizaje en terreno y las mejoras en los procesos operativos y técnicos han conllevado a la realización de algunas precisiones al cuestionario¹⁰, las que se sintetizan a continuación:

Tabla 1. Resumen de modificaciones aplicadas al cuestionario ESI 2011-2022.

Año de levantamiento	Modificaciones
2011	<ul style="list-style-type: none">- Para facilitar el llenado en terreno, se agregan a un set de preguntas las opciones “No Sabe” y “No Responde”.- Como mecanismo de control que permita asegurar la captura de las personas ocupadas de referencia en el mes anterior, se crean preguntas filtro para ingresos de ocupados dependientes e independientes (D1_opcion y D5_opcion).- Mejora en la completitud de las opciones de respuesta: “No recibe” en D2, y opción “Sala cuna/Jardín infantil” en D3, mientras que en D14 cambia “Pensiones por parte del Estado” a “Pensión Básica Solidaria del Estado”.- Reformulación de algunas preguntas de acuerdo con la experiencia en terreno y calidad de la información obtenida: D6.- Fusión de preguntas para facilitar la captura: D12 y D13.
2012	<ul style="list-style-type: none">- Para facilitar la secuencia de llenado en terreno, se traspuso la pregunta D2, agregando conectores para facilitar su levantamiento.- Se intensifica uso de opciones “No Sabe” y “No Responde” en relación con montos, porcentajes y tramos para las preguntas D2; D3, y D13.- Ampliación de opciones de respuesta de acuerdo con la condición de actividad en D9.- Se elimina opción de ingreso variable “Otro tipo de incentivos” en D2.
2015	<ul style="list-style-type: none">- Para mejorar la información capturada relativa a los ingresos variables de las personas ocupadas dependientes (pregunta D2), se agregó una pregunta que indaga por la periodicidad de los ingresos reportados.- Para mejorar la información capturada relativa al ingreso de las personas ocupadas independientes, se agregó una pregunta que indaga a cuánto tiempo de trabajo correspondió el ingreso reportado (pregunta D5).- Se mejoró la diagramación de la pregunta D12, con el objetivo de mejorar su captura.

¹⁰ El cuestionario 2021 se mantuvo vigente para 2022. Disponible para consulta en <https://www.ine.gob.cl/docs/default-source/encuesta-suplementaria-de-ingresos/formularios/cuestionarios/cuestionario-esi-2021.pdf>

	<ul style="list-style-type: none"> - Para mejorar la captura de los ingresos provenientes de activos financieros (pregunta D13) se incrementó el número de tramos de intereses. - Para mejorar la captura de los ingresos provenientes de becas de estudio, se dividió la pregunta D14_11 en 2 alternativas: becas de estudio monetarias y en especie.
2016	<ul style="list-style-type: none"> - Dada la baja frecuencia e incidencia de algunas alternativas de la pregunta D3, se eliminaron las preguntas sobre beneficios o regalías de ropa y/o calzado, teléfono personal, servicio de estacionamiento y viáticos. Sin embargo, estas siguen consignándose en “<i>otros (especifique)</i>”. - Se incorporaron 2 secciones de “cálculos” para facilitar el levantamiento de información compleja por parte del encuestador(a). - Con el objetivo de mejorar su captura se modificó la diagramación de la pregunta D12. En particular, se agregó un salto a la alternativa “propia, pero paga dividendo” para pasar directamente a la estimación del valor de la vivienda, a la vez que se reordenó para facilitar la lectura y los saltos de dicha pregunta.
2017	<ul style="list-style-type: none"> - Incorporación de mejoras en la redacción de preguntas e instrucciones: D1b, D2a, D2b, D2c, D3, D4, D5b, D6, D8, D9a, D11, D13 - Se elimina la palabra "viáticos" del ejemplo en la pregunta D3. Esto, en concordancia con la eliminación de la pregunta por este beneficio realizada el 2016 (baja prevalencia). - En D14 se incorporó "NS/NR" (no sabe, no responde), como opción para quienes declaran recibir una partida de ingreso, pero no el monto.
2018	<ul style="list-style-type: none"> - En la pregunta D3, se incorpora en la definición de beneficio como todo aquello que “NO es necesario para el desarrollo de su trabajo”. - En D12, opción 5 de “cedida por el empleador”, se añade una pregunta 5.1 de “necesaria para el desarrollo de su trabajo”, que vincula el tipo de vivienda (cedida o necesaria) con el beneficio que pueda percibir por parte del empleador (vínculo con la pregunta D3). - En reemplazo del clasificador de ocupaciones CIUO-88 se adopta el estándar CIUO-08.CL, el cual corresponde a una versión adaptada a la realidad chilena del clasificador del clasificador CIUO-08, manteniendo los principios básicos y la estructura del mismo.
2019	<ul style="list-style-type: none"> - Sin cambios.
2020	<ul style="list-style-type: none"> - Cambio de nombre en todas las variables y preguntas del cuestionario ENE-ESI. Las letras destinadas al módulo ESI son K, L, M, N y O.
2021	<ul style="list-style-type: none"> - Se eliminaron las preguntas D10 a la D14, es decir los ingresos de otras fuentes, las cuales capturan ingresos provenientes de rentas de la propiedad, activos financieros y transferencias, lo que implica que la encuesta cambia sus objetivos, enfocándose en la captura y medición de ingresos asociados al trabajo.
2022	<ul style="list-style-type: none"> - Sin cambios.

2.1.2.2 Revisión del módulo de ingresos

Hasta 2012 los stocks de personas ocupadas estimadas para el trimestre OND entre la ENE y la ESI sólo diferían en dos aspectos: por una parte, en que los primeros no estaban filtrados para aquellos que tenían al menos un mes de permanencia en el actual empleo y, por otra parte, en que los factores de expansión de ambas encuestas eran distintos. No obstante, considerando que los tiempos de análisis de la ESI son más amplios que para la ENE (la ESI publica sus indicadores de manera anual, mientras que la ENE lo hace mensualmente), existe la posibilidad de hacer una revisión en mayor profundidad del módulo de ingresos para encontrar posibles errores en las distintas clasificaciones que trae el módulo de empleo.

En respuesta a lo anterior, desde 2013 se determinó flexibilizar el criterio de mantener toda la información del módulo de empleo inalterada y se permite, a través de un extenso análisis al perfil de la ocupación, que puedan existir cambios en el módulo ENE. En consecuencia, pueden existir diferencias de stocks de personas ocupadas entre la ENE y la ESI, que afectan exclusivamente al proceso de producción de la ESI y que no implican una rectificación de cifras inmediata de la ENE.

Adicional a lo anterior, en línea con el Código de Buenas Prácticas de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), se anunció en febrero de 2020 la implementación de la política de revisión y actualización anual de cifras de la Encuesta Nacional de Empleo (ENE)¹¹, la que comenzó a implementarse a contar del año 2021, y que considera la evaluación constante de todas las áreas que intervienen en la generación de la información de la Encuesta Nacional de Empleo. En ese sentido, el punto 3 de la política de rectificación de cifras de la ENE considera también como insumo ediciones que se puedan realizar en el marco del procesamiento de la ESI.

2.1.2.3 Uso de una tabla de referencia y efecto en las transferencias en educación

La experiencia en el levantamiento de encuestas de ingresos muestra que los informantes muchas veces desconocen el monto de sus beneficios o subestiman su valor, esto se hace más recurrente cuando se contemplan transferencias en especies como becas de alimentación o la subvención escolar. La recomendación internacional sugiere la utilización de una tabla de referencia en el trabajo de campo como una de las formas para soslayar esta subestimación.

Hasta 2019 la tabla de referencia ESI contenía montos vigentes por (i) ingreso mínimo neto (ii) bonos al trabajo y aguinaldos, (iii) ingresos por asignación y subsidio familiar (SUF), (iv) ingresos y subsidios por programas solidarios, (v) subsidios de cesantía, (vi) pensiones básicas solidarias y por discapacidad y (vii) subvenciones escolares. A partir de 2021, considerando que la ESI dejó de consultar ingresos del hogar, la tabla de referencia se limitó sólo a contener el monto del ingreso mínimo neto, junto a bonos al trabajo y aguinaldos.

En las evaluaciones en terreno que se hicieron durante el proceso de levantamiento de la primera versión de la ESI (2010), se observó un proceso gradual de aprendizaje en el uso de la tabla de referencia por parte de los encuestadores, ya que varios de ellos olvidaban que tenían esta herramienta a disposición y dejaban a criterio del informante la estimación del valor de la transferencia. Esta situación implicó realizar un plan de capacitación a los equipos de

¹¹ https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/rectificacion-de-cifras/separata-t%C3%A9cnica-pol%C3%ADtica-de-revisi%C3%B3n-y-actualizaci%C3%B3n-de-cifras-ene-febrero-2020.pdf?sfvrsn=72b9a8c7_4

levantamiento y la elaboración de instructivos operativos sobre el uso de la tabla de referencia, que se reflejaron en un mayor aprendizaje en el uso del instrumento en el trabajo de campo de 2011.

De esta forma, se observó que el aprendizaje en el uso de la tabla de referencia tuvo efecto en el levantamiento de la información, ya que al comparar los resultados de 2011 con respecto al 2010, se identificó una variación en el número de personas que señalaban tener el beneficio, aumentando en 76,4% el número de casos muestrales.

Por lo tanto, si se incluye dentro del ingreso de los hogares la partida “becas de estudios”, la estructura de ingresos de los hogares de 2011 **no es comparable** con la de 2010. La medida que se tomó fue publicar tabulados incluyendo y excluyendo la partida de “becas de estudio” en el componente de transferencias de los hogares, con el fin de que los tabulados sin “becas de estudios” se puedan comparar adecuadamente entre el 2010 y 2011, ya que así se aislaría el efecto levantamiento/aprendizaje que incidió en la captura de “becas de estudios”.

Respecto de los resultados de la ESI 2012 a 2020, dado que son comparables con los de 2011, los análisis posibles permiten incluir la partida de transferencias en educación o aislarla, dependiendo de su objetivo, para lo cual se publican en la página web institucional las dos versiones de los tabulados. Sin embargo, los documentos con la descripción de los principales resultados que son producidos oficialmente con la ESI se refieren a la versión que no incluye la partida de becas de estudio.

2.1.2.4 Cambio en el clasificador de rama de actividad económica

A partir de la necesidad de incluir las nuevas realidades productivas y tecnológicas creadas, y mitigar los problemas de clasificación que se presentan actualmente en las encuestas sociodemográficas (encuestas de hogares), el INE ha transitado desde el Clasificador Industrial Internacional Uniforme de todas las actividades económicas (CIIU) rev.3, a CIIU rev.4, este último adaptado a la realidad chilena para el caso de encuestas de hogares (CAENES).

Para llevar a cabo una correcta transición, desde el 2013 al 2016 se publicaron los tabulados con ambos clasificadores, no obstante, desde la medición 2017 en adelante el clasificador oficial de rama de actividad económica corresponde a CAENES.

2.1.3 Necesidades de información

La ESI responde a la necesidad de conocer en detalle la realidad de los ingresos laborales e ingresos de otra(s) ocupación(es) distinta(s) a la principal de la población en el país. Constituye un complemento a las estadísticas sobre empleo obtenidas mediante la aplicación de la ENE, permitiendo caracterizar los ingresos de las personas ocupadas del país.

Busca responder a la necesidad de académicos, investigadores, estudiantes y público en general respecto a tener información oportuna, año a año, sobre las condiciones en las que se encuentra el país y sus regiones, en relación con los ingresos del trabajo. La información que proporciona la encuesta sirve de apoyo para distintos organismos del Estado puedan elaborar políticas en beneficio de toda la población.

2.2 Objetivos de la Encuesta Suplementaria de Ingresos

2.2.1 Objetivo general

- Caracterizar los ingresos laborales de la ocupación principal en el mes de referencia de las personas que son clasificadas como ocupadas en la ENE.
- Caracterizar los ingresos de otra(s) ocupación(es) distinta(s) a la principal para las personas que son clasificadas como ocupadas y no ocupadas en la ENE.

2.2.2 Objetivos específicos

- Caracterizar los ingresos del trabajo principal, en función del año, sexo, clasificación en el empleo, nivel educacional, formalidad, grupo ocupacional, rama de actividad económica, entre otros, a nivel nacional y regional.
- Caracterizar los ingresos del trabajo principal de las personas ocupadas por cuenta propia, en función del sexo, jornada laboral, lugar de trabajo, cotización previsional, entre otros, a nivel nacional y regional.
- Caracterizar los ingresos del trabajo principal de las personas asalariadas, en función del sexo, tamaño de empresa, tramo de ingreso mínimo neto, ingreso por hora, entre otros, a nivel nacional y regional.
- Caracterizar los ingresos de otras ocupaciones distintas a la principal.

3 Marco de referencia

3.1 Marco conceptual

Los aspectos teórico-conceptuales que constituyen el fundamento de la encuesta pueden ser descritos en torno a dos ejes: uno relacionado a su relevancia social y otro relacionado a sus directrices conceptuales y metodológicas.

3.1.1 Relevancia y directrices internacionales para la elaboración de estadísticas de ingresos

La ESI responde a la necesidad de conocer en detalle la realidad de los ingresos laborales e ingresos de otra(s) ocupación(es) distinta(s) a la principal de la población en el país. Constituye un complemento a las estadísticas sobre empleo obtenidas mediante la aplicación de la ENE, permitiendo caracterizar los ingresos de las personas ocupadas del país.

Su formulación toma en consideración un conjunto de recomendaciones y discusiones internacionales en torno a la deseabilidad y la relevancia social de la compilación y estandarización de la información sobre ingresos:

- Lineamientos internacionales sobre la materia, como los expresados en el Manual del Grupo de Canberra (United Nations, 2011), mencionan que la compilación de datos acerca de la distribución de ingresos es relevante en la medida que permite la comprensión de los patrones de distribución de estos y su relación con la forma de organización de las sociedades, permite evaluar el impacto de las políticas públicas en distintos ámbitos de interés, así como también analizar cómo diferentes patrones de distribución de ingresos influyen en el bienestar del hogar y en las habilidades de las personas para adquirir bienes y servicios que necesitan para satisfacer sus necesidades.

Complementariamente, menciona que el interés en la distribución del ingreso se puede justificar en la medida que permite observar cómo el producto nacional es distribuido entre la población e indirectamente, como medida de la distribución del bienestar económico (United Nations, 2011).

- Por su parte, la Oficina Internacional del Trabajo (OIT), entrega lineamientos a través del desarrollo de “Conferencias Internacionales de Estadísticos del Trabajo” (en adelante CIET). La resolución de la 16° CIET (OIT, 1998) menciona la importancia que los países complementen sus programas de estadísticas sobre el empleo, desempleo, subempleo y los salarios, con información sobre ingresos relacionados al empleo, con el fin de analizar la capacidad generadora de ingresos de las actividades económicas y analizar el bienestar económico de las personas basado en las oportunidades de empleo.
- Complementariamente, la 17° CIET enfatiza que las estadísticas de ingresos (y también de gastos) de los hogares pueden servir de base para la descripción de distintos aspectos económicos, sociales y otros, dentro de lo cual se puede destacar la utilidad de las estadísticas

para valorar el nivel, estructura y tendencias del bienestar económico de los hogares y los individuos, aspecto que enfatiza el hecho que las entradas contribuyen en la medida que aumentan el potencial de los individuos para consumir o ahorrar (OIT, 2003).

3.1.2 Lineamientos conceptuales y metodológicos para la elaboración de estadísticas de ingresos

3.1.2.1 Ingreso corriente

En particular, la ESI mide una fuente de ingreso que se circunscriben en el concepto de **ingreso corriente**: aquel proveniente de la ocupación principal e ingresos de otra(s) ocupación(es) distinta(s) a la principal, durante el periodo de referencia.

La definición de ingreso corriente considerada en la encuesta se enmarca conceptualmente en el Sistema de Cuentas Nacionales (SCN), donde se menciona que, desde el punto de vista teórico, el ingreso corresponde al “monto máximo que un hogar u otra unidad puede permitirse gastar en bienes o servicios de consumo durante el período contable, sin tener que financiar sus gastos reduciendo su dinero, disponiendo de otros activos financieros o no financieros o incrementando sus pasivos” (GITCN, 2009). Esta definición excluiría de la medición del ingreso eventos tales como transferencias de capital, las ganancias o pérdidas por tenencia reales y otras variaciones en el volumen de activos debidas a acontecimientos como los desastres naturales.

Si bien el enfoque responde a distintos intereses analíticos (macroeconómico en el caso del SCN y microeconómico en el caso de las estadísticas de distribución del ingreso de personas) (United Nations, 2011), es posible establecer lineamientos comunes que enmarcan el concepto de ingreso de interés:

- Refiere a las entradas regulares o periódicas y no a las extraordinarias.
- Está orientado al gasto corriente y, por lo tanto, **ingresos netos** provenientes del empleo (asalariado e independiente).
- Excluye en su definición transacciones que involucren cambios en el patrimonio de las unidades.

3.1.2.2 Ingresos procedentes del empleo u ocupación

A la base de la consideración del empleo y los ingresos que se generan se encuentra el **concepto de trabajo**. Si bien en la práctica ambos son utilizados como sinónimos, la acepción de trabajo es más amplia que la de empleo. Trabajo refiere al “(...) conjunto de actividades humanas, remuneradas o no, que producen bienes o servicios en una economía, o que satisfacen las necesidades de una comunidad o proveen los medios de sustento necesarios para los individuos” (OIT, 2004). Por su parte, el **concepto de empleo** se refiere a un trabajo realizado a cambio de pago, sin importar la relación de dependencia entre las partes (OIT, 2004).

En ese sentido, el concepto de **empleo u ocupación** es de especial relevancia para la medición de los ingresos de las personas, en torno al cual se establecen recomendaciones internacionales para su adecuado registro.

En la resolución de la 16° CIET, así como también es señalado por el Grupo de Canberra, los ingresos de la ocupación o procedentes del empleo comprenden las entradas por concepto de participación en actividades económicas en un contexto estrictamente relacionado con el empleo (OIT, 1998). Consiste en pagos en dinero o especies recibidos por individuos para sí o con respecto a sus familiares, como resultado de su involucramiento actual o futuro en trabajos remunerados o por cuenta propia (United Nations, 2011).

El marco de referencia para la medición de los ingresos está limitado, a nivel general, por la denominada **frontera de producción del SCN**, que en términos generales involucra a las actividades –procesos físicos- llevadas a cabo por una unidad institucional (dentro de las cuales se incluyen los hogares, personales o constituidos por grupos de personas), para transformar insumos de bienes y servicios en productos de otros bienes y servicios que están destinadas al mercado para su venta o trueque. Esta concepción de frontera de producción considera la posibilidad de transacciones no monetarias entre unidades distintas de la economía e incluso dentro de una misma unidad, como puede ser un hogar (a excepción de los servicios para autoconsumo, según se mencionará en un párrafo posterior).

De acuerdo con lo indicado en la resolución de la 19° CIET, el puesto de trabajo o actividad productiva “se define como el conjunto de tareas y obligaciones desempeñado por una persona, o que se prevé que una persona desempeñe, para una sola unidad económica” (OIT, 2013). Esto implica que una persona puede tener uno o más puestos de trabajo, y que en ese conjunto también puede definirse un **puesto de trabajo principal** como aquel que corresponde el mayor número de horas habitualmente trabajadas.

Según lo detallado por Grupo de Canberra y lo resuelto en las CIET, se puede mencionar que los ingresos procedentes del empleo o la ocupación incluyen:

- **Ingresos del empleo asalariado**, que pueden ser recibidos en dinero o especies (bienes o servicios), comprendiendo sueldos y salarios directos, además de un conjunto de otras entradas proporcionadas por el empleador, que se recomienda sean registrados en forma separada cuando correspondan. Cuando se trata de pagos en bienes o servicios, estos son valorizados y considerados solo en la medida que no se trata de condiciones impuestas, en cuyo caso se excluyen o se valoran en cero.
- **Ingresos del trabajo independiente**, que se vincula con las entradas recibidas como resultado de la ocupación en trabajos independientes y que, por lo tanto, depende directamente de los beneficios derivados de los bienes y servicios producidos. Los trabajadores independientes se definen así porque son, a diferencia de los trabajadores asalariados, directamente dependientes del bienestar de la empresa (OIT, 1993, p.2). Los ingresos del empleo independiente, en correspondencia con la definición del SCN serían de carácter mixto, en la medida que no puede separarse fácilmente la auto-remuneración de los beneficios obtenidos por el propietario en tanto empresario (GITCN, 2009).

3.1.3 Referencias nacionales e internacionales

Del apartado anterior, se explica las referencias internacionales en las cuales se basa la ESI. A continuación, la Tabla 2 resume las características principales de productos estadísticos similares a la ESI en cuanto a temática (ingresos en personas) y metodología (encuestas presenciales en hogares, de carácter probabilístico), desarrollados en Chile y otros países.

El criterio de inclusión de otros países fue un ingreso per cápita similar en el contexto geográfico de América del Sur¹², a los que se agrega México, dado que además de Chile, es el otro país de América Latina miembro de la OECD.

¹² Consultado en: <https://estadisticas.cepal.org/cepalstat/>

Tabla 2. Referencias nacionales e internacionales atinentes a la ESI.

País	Estudio y Periodicidad	Indicadores clave; Población de Estudio e informante	Diseño muestral y Tamaño de la muestra	Periodo de aplicación y Periodo de referencia	Instrumento	Observaciones
Argentina	Encuesta Permanente de Hogares (EPH) – INDEC ¹³	<ul style="list-style-type: none"> - Características del hogar, características demográficas de personas - Características ocupacionales y migraciones - Características habitacionales, educacionales y de ingresos <hr/> <ul style="list-style-type: none"> - Hogares y personas - Cuestionario individual entrega información de personas de 10 años y más 	<p>Muestras bietápicas estratificadas, unidades primarias (radios censales) agrupadas en estratos, selección proporcional a tamaño (cantidad de viviendas). Segunda etapa: viviendas</p> <hr/> <p>52.340 encuestas 1º trimestre 2018</p>	<p>Se entregan resultados 4 veces al año: trimestres</p> <hr/> <p>Periodo de referencia es semanal, salvo indicadores en base a mes</p>	<p>3 cuestionarios:</p> <ul style="list-style-type: none"> - Cuestionario Individual: características ocupacionales y de ingreso - Cuestionario de vivienda - Cuestionario de hogares 	<p>Cuestionario individual:</p> <ul style="list-style-type: none"> - Conjunto de preguntas destinadas a indagar las distintas fuentes de ingreso, tanto de las personas en forma individual como de los hogares en su conjunto. - Los ingresos relevados provienen tanto de fuentes laborales, ya sean de ocupados asalariados o independientes, como no laborales, principalmente jubilaciones y pensiones y subsidios. - Incluye preguntas de movimientos interurbanos en ocupación. <p>Cuestionario de hogares:</p> <ul style="list-style-type: none"> - Contiene preguntas de ingresos no laborales. - Incluye módulo de preguntas por labores domésticas. - Módulo de “estrategias de manutención del hogar”, perceptores de ingresos
Brasil	Encuesta Mensual continua de viviendas (PNAD Continua mensal) – IBGE ¹⁴	<ul style="list-style-type: none"> - Ocupación, inserción de personas en mercado del trabajo <hr/> <ul style="list-style-type: none"> - Personas - Cuestionario individual entrega información de personas de 10 años y más. - Son entrevistadas personas en edad de trabajar: 14 o más años 	<p>Muestra probabilística</p> <hr/> <p>211.344 domicilios particular permanentes por trimestre</p>	<p>Se entregan resultados 4 veces al año: trimestres</p> <hr/> <p>Semana de referencia para ocupación</p> <hr/> <p>Mes de referencia anterior a semana para ingresos</p>	<p>Cuestionario único</p>	<ul style="list-style-type: none"> - Cuestionario tiene módulo de caracterización de personas, además de módulo de caracterización de ocupación primaria, secundaria y otros. - Lógica del cuestionario integra preguntas por ocupación primaria, secundaria y otras con ingresos según tipo, así como también otros ingresos complementarios entregados por el empleador o dispuestos para autoconsumo. Después continúa con consulta de horas, subocupación, no ocupado. - Incluye actividades de autoconsumo (servicios). Incluye también módulo de trabajo voluntario y cuidado de personas. Incluye también módulo de labores domésticas - Consulta por ingresos brutos del trabajo

¹³ Consultado en: https://www.indec.gov.ar/uploads/informesdeprensa/eph_pobreza_01_18.pdf

¹⁴ Consultado en: <https://www.ibge.gov.br/estatisticas-novoportal/sociais/trabalho/9171-pesquisa-nacional-por-amostra-de-domicilios-continua-mensal.html>

						<ul style="list-style-type: none"> - Consulta por descuentos por atrasos, faltas, etc. - Consulta por cálculo de ingreso habiendo calculado esos descuentos - Incluye módulos muy básicos de ingresos de propiedad. - Incluye módulo más detallado de transferencias - Incluye consulta por alquiler imputado
Chile	Encuesta CASEN – Ministerio de Desarrollo Social ¹⁵	<p>Pobreza por ingresos de personas y hogares.</p> <p>Identificar carencias en áreas de salud, educación, vivienda, trabajo e ingresos.</p> <p>Evaluar brechas de ingresos.</p> <p>Estimar cobertura, focalización y distribución del gasto fiscal de los principales subsidios.</p> <p>Personas y hogares</p> <p>La unidad de información de Casen en Pandemia 2020 es el hogar y la persona. Por su parte, el informante idóneo es el jefe de hogar o en su ausencia, algún miembro del hogar de 18 años o más.</p>	<p>El diseño muestral de la encuesta Casen en Pandemia 2020 se define como probabilístico y estratificado, según área geográfica y por tamaño poblacional, tanto en el área urbana como rural. La selección de la muestra se realiza en dos etapas - muestreo bietápico en las áreas rurales y urbanas.</p> <hr/> <p>Muestra 2020:</p> <p>Viviendas: 62.540; hogares: 62.540; personas: 185.437.</p>	<p>Entrega de resultados cada 2 años</p> <hr/> <p>Referencia de ocupación a semana anterior, mes</p> <p>Referencia de ingresos mes anterior</p>	<p>Siete módulos: Registro de Residentes, Educación, Trabajo, Ingresos, Salud, Identidades y Vivienda. El cuestionario Casen en Pandemia 2020 considera una versión reducida respecto del cuestionario Casen 2017, con foco en las preguntas atinentes a la medición de pobreza por ingresos.</p>	<ul style="list-style-type: none"> - Cuestionario incluye pregunta por ingresos provenientes del trabajo, propiedad y transferencias - Dentro de la ocupación, incluye actividad principal, otras ocupaciones. - Además de consulta por mes de referencia anterior, incluye medición de ingresos a plazo de 12 meses.
México	Encuesta Nacional de Ingresos y Gastos de los Hogares – ENIGH ¹⁶	<ul style="list-style-type: none"> - Ingresos - Gastos - Actividades económicas desarrolladas por 	<p>Diseño probabilístico, estratificado y por conglomerados (bietápico). Unidad</p>	<p>Se aplica cada 2 años</p> <p>En el año de aplicación, el levantamiento se realiza en 3 meses,</p>	<p>Distintos cuestionarios:</p>	<ul style="list-style-type: none"> - Cuestionario para personas de 12 o más años incluye medición de aspectos ligados al empleo, así como también ingresos (ingreso corriente), tanto provenientes del trabajo como de otras fuentes.

¹⁵ Consultado en <https://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-en-pandemia-2020>.

¹⁶ Consultado en: <https://www.inegi.org.mx/programas/enigh/nc/2016/default.html>

		<p>quienes conforman los hogares</p> <hr/> <p>- Población objetivo: hogares</p> <hr/> <p>- Informante adecuado: jefe o jefa del hogar, ama de casa o algún integrante del hogar de 18 o más años que conozca la información del hogar en su conjunto.</p> <p>- Ideal que todas las personas de 12 años o más proporcionen información propia. Personas menores de 12 años: informa persona de 18 o más años que conozca información.</p>	<p>de selección última es la vivienda y de observación el hogar.</p> <hr/> <p>81.515 viviendas a nivel nacional, con un 90% de confianza y error relativo máximo de 7.20%</p>	<p>organizado en 9 periodos, cada uno de ellos de 10 días</p> <hr/> <p>b. Distintos</p> <p>Para módulo de ingresos del hogar en personas mayores de 12 años: mes anterior al de entrevista.</p> <p>También captura ingresos de meses anteriores al mes de referencia: hasta 6 meses.</p> <p>En niños menores 12 años también considera ingresos hasta 6 meses antes del mes de referencia</p>	<ul style="list-style-type: none"> - Cuestionario de hogares y vivienda - Cuadernillo de gastos diarios del hogar - Cuestionario de gastos del hogar - Cuestionario para personas de 12 o más años: ocupación, ingresos del trabajo principal, ingresos del trabajo secundario, ingresos de otros trabajos del mes pasado, ingresos ajenos al trabajo, redes sociales, uso del tiempo, salud. - Cuestionario para negocios del hogar: características e ingresos de los negocios del hogar, consistencia de la información - Cuestionario para personas menores de 12 años: ingresos, acceso a la salud 	<ul style="list-style-type: none"> - Anterior cuestionario incluye además la medición de entradas que no corresponden a la definición de ingreso corriente: percepciones financieras y de capital que modifican el valor neto del patrimonio del hogar y no satisfacen los criterios de regularidad y destino - Incluye cuestionario que reúne información de los negocios del hogar cuando corresponde - Además, un cuestionario para personas menores de 12 años que registra sus ingresos a partir de informante jefe de hogar o persona encargada de 18 años o más y que conozca la información. - Incluye cuestionarios y cuadernillos destinados a gastos <hr/> <ul style="list-style-type: none"> - Ingresos del trabajo integran la medición de indemnizaciones por accidentes de trabajo, indemnizaciones por despido y retiro voluntario. - Excluye de ingresos del trabajo independiente ingresos percibidos bajo la forma de sueldos o ganancias de trabajadores que son socios de alguna empresa constituida en sociedad, y también de trabajadores que son dueños de empresas y que, si bien no están constituidas en sociedad, realizan prácticas contables completas (cuasisociedades) - También considera la medición de “ingresos mixtos” en el trabajo independiente. - Considera expresamente la medición de regalías (derechos de autor, patentes) en las rentas de propiedad
México	Encuesta Nacional de Ocupación y Empleo (ENOE) - ENIGH ¹⁷	<ul style="list-style-type: none"> - Ocupación y empleo - Ingresos <hr/> <p>- Se publican datos de población de 15 años o más y hogares. Se captan datos económicos</p>	<p>El tipo de muestreo utilizado es probabilístico, bietápico, estratificado y por conglomerados.</p>	<p>Periodicidad trimestral. Corte mensual y trimestral.</p> <hr/> <p>Características sociodemográficas:</p>	<ul style="list-style-type: none"> - Cuestionario sociodemográfico. - Cuestionario de ocupación y empleo, básico y ampliado 	<p>Menciona que se basa en marco normativo de sistemas de contabilidad nacional.</p> <ul style="list-style-type: none"> - Considera en definición de actividad económica que, si un servicio no es solicitado, no es una auténtica transacción y por lo tanto no es una actividad económica puesto que no es objeto de demanda en la sociedad: en esta encuesta se distingue y no queda al arbitrio de encuestado que se defina o no como

¹⁷ Consultado en: <https://www.inegi.org.mx/programas/enoe/15ymas/>

		<p>sobre personas entre 12 y menos de 15 años.</p> <p>- Informante idóneo: jefe o jefa del hogar o algún otro integrante de éste que tenga 15 o más años. Ideal que todas las personas de 12 o más proporcionen propia información</p>	<p>Trimestralmente: 120.260 encuestas</p>	<p>momento de la entrevista.</p> <p>Características económicas: semana anterior (de lunes a domingo)</p>		<p>ocupación. Ejemplo: limpiaparabrisas o quienes piden limosna.</p> <ul style="list-style-type: none"> - Los “ocupados” se definen como tales bajo la premisa de que alguien demanda lo que ofrecen. - Captura ingresos de la ocupación principal. No captura ingresos por ocupación secundaria u otras, tampoco provenientes de la propiedad y tampoco transferencias.
Uruguay	Encuesta Continua de Hogares ¹⁸ - INE	<ul style="list-style-type: none"> - Indicadores oficiales del mercado laboral (actividad, empleo y desempleo) - Ingresos de los hogares y las personas con - Línea de pobreza y de indigencia de forma anual. - Base para estudios sociales de otras temáticas - Hogares y todas las personas de los hogares de viviendas seleccionadas. Ocupación e ingresos: personas de 14 o más años. 	<p>Diseño muestral complejo que incluye varias etapas de selección: estratificación de departamentos y localidades según tamaño; selección de zonas censales; selección de viviendas.</p> <p>53.580 casos anuales, 4.465 mensual.</p>	<p>Periodicidad mensual, trimestral, semestral y anual.</p> <p>El período de referencia es la semana anterior a la realización de la entrevista: ocupación</p>	<p>Cuestionario único:</p> <ul style="list-style-type: none"> - Módulo caracterización vivienda y hogar - Datos de la persona 	<ul style="list-style-type: none"> - Recabada con dispositivos electrónicos. - La investigación sobre las características de la mano de obra y los ingresos percibidos se aplica únicamente a las personas de 14 o más años. Este umbral coincide con la edad que define a la Población económicamente Activa (PEA) en el país. - Mide ingreso corriente de hogares, del trabajo, de la propiedad, transferencias - Considera expresamente la medición de regalías (derechos de autor, patentes) en las rentas de propiedad. - Caracteriza ingresos de trabajadores agropecuarios de forma especial. Considera ingresos posibles por “medianería o aparecería”, “pastoreo”, “ganado a capitalización”. - Para trabajadores independientes tiene una pregunta específica que distingue ingresos recibidos por distribución de utilidades, distinta a otra en que pregunta si se retiró dinero de negocios propios para gastos del hogar.

¹⁸ Consultado en: <https://www.gub.uy/instituto-nacional-estadistica/datos-y-estadisticas/encuestas/encuesta-continua-hogares>

		- Informante es miembro del hogar mayor de 18 años (excepto el servicio doméstico) mentalmente competente para brindar información acerca del resto de los integrantes del hogar.				
Uruguay	Encuesta de Gastos e Ingresos de los Hogares - ENGIH 2016/2017 ¹⁹ - INE	<ul style="list-style-type: none"> - Información de hogar y vivienda - Características de las personas: educación y trabajo, ingresos, gastos del hogar y sus miembros <hr/> <ul style="list-style-type: none"> - Análisis a nivel de Hogar particular, vivienda, miembros del hogar <hr/> <ul style="list-style-type: none"> - Se entrevista presencialmente a 	<p>Muestreo Aleatorio</p> <p>Estratificado con asignación óptima. Aleatorio, estratificado y por conglomerados, en dos o tres etapas.</p> <p>Subuniversos, estratos geográficos/área, viviendas con hogares particulares, hogares</p> <hr/> <p>7.500 viviendas, todos los miembros del hogar</p>	<p>Se realiza cada 10 años.</p> <p>La última versión fue aplicada durante todo el año 2017</p> <hr/> <p>Referencia semanal y mensual (ingresos mensuales)</p>	<p>4 instrumentos:</p> <p><u>Questionario 1:</u></p> <ul style="list-style-type: none"> - Características de la vivienda, del hogar, características sociodemográficas y otros sociales por persona - Ingresos, persona, hogar - Gastos del hogar - Gastos diarios del hogar <p><u>Questionario 2:</u> gastos diarios del hogar</p> <p><u>Questionario 3:</u> gastos diarios de las personas</p>	<p>Questionario 1:</p> <ul style="list-style-type: none"> - Mide ingreso corriente de hogares, del trabajo, de la propiedad, transferencias. - Considera expresamente la medición de regalías (derechos de autor, patentes) en las rentas de propiedad. - Caracteriza ingresos de trabajadores agropecuarios de forma especial. Considera ingresos posibles por “medianería o aparecería”, “pastoreo”, “ganado a capitalización”. - Para trabajadores independientes tiene una pregunta específica que distingue ingresos recibidos por distribución de utilidades, distinta a otra en que pregunta si se retiró dinero de negocios propios para gastos del hogar.

¹⁹ Consultado en: <https://www.gub.uy/instituto-nacional-estadistica/encuesta-gastos-ingresos-hogares>

		cada miembro del hogar			<u>Questionario 4:</u> gastos del hogar	
--	--	------------------------	--	--	--	--

4 Diseño estadístico

4.1 Universo

Anualmente, la ESI recaba información de ingresos de personas residentes en viviendas particulares ocupadas, presentes en el territorio nacional y que son cubiertas por el marco de muestreo definido para la encuesta en base al censo de población y vivienda 2017 (MMV 2017), con actualización al año 2020.

4.2 Población objetivo

La población objetivo, es la población en edad de trabajar, es decir, todas las personas de quince años y más, que residen habitualmente en viviendas particulares ocupadas en el país.

4.3 Marco estadístico

Un marco de muestreo se define como la lista o los procedimientos que permiten identificar a todos los elementos de una población objetivo (Groves, 2004, pág. 68). Los marcos constituyen el insumo fundamental para el diseño de encuestas por muestreo y, por consiguiente, resulta de suma importancia que su metodología de conformación sea debidamente documentada.

Particularmente la ESI, al ser un módulo suplementario de la ENE, comparte la totalidad de su marco estadístico. El detalle de este diseño se encuentra en el Documento Metodológico oficial de la ENE²⁰

4.4 Cobertura geográfica

Al igual que en la ENE, la cobertura geográfica de la ESI considera las 16 regiones del país, buscando presencia en todo el territorio nacional y abarcando 97% de las comunas del país. No obstante, se excluyen algunas áreas geográficas que, por presentar ciertas características especiales (insularidad, clima, altitud, tiempo de traslados, estado de los caminos, etc.), son excluidas del marco de muestreo de la ENE y, por tanto, de la ESI.

4.5 Unidades estadísticas

Al ser la ESI un módulo de la ENE y, en consecuencia, una encuesta por muestreo, las unidades estadísticas las conforman unidades de muestreo, unidades de información y unidades de análisis. Las unidades de muestreo son aquellas que son susceptibles de ser seleccionadas en las etapas de muestreo definidas en el diseño, pudiendo estar conformadas por una unidad elemental o por un

²⁰ Revisar apartado “4.3. Marco estadístico” en Documento metodológico Encuesta Nacional de Empleo (2021) en el enlace: https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodologia-ene-2020.pdf?sfvrsn=793380e9_14

conjunto de ellas (conglomerado). Por otra parte, las unidades de información son aquellas desde donde se obtiene la información de interés, pudiendo esta coincidir con la unidad de análisis que, es aquella unidad sobre la cual se puede realizar la medición del fenómeno objeto de estudio.

4.5.1 Unidad de muestreo

Dado el diseño bietápico existen dos unidades de muestreo para la ENE y, por tanto, para la ESI.

- a) Unidad primaria de muestreo (UPM): definida como un conglomerado homogéneo, en términos de la cantidad de viviendas particulares que los componen.
- b) Unidad secundaria de muestreo (USM): que corresponde a viviendas particulares ocupadas dentro de las UPM seleccionadas.

El tamaño promedio de las UPM en el área urbana es de 200 viviendas particulares, con un rango de 160 a 240 viviendas particulares, excluyendo viviendas de temporada. Mientras que, el tamaño promedio de las UPM rurales es de 90 viviendas particulares, con un rango de 70 a 110 viviendas particulares, excluyendo las de temporada.

4.5.2 Unidad de información y análisis

Para el cumplimiento del objetivo de la encuesta, se utiliza como unidades de información y análisis a las personas que cumplen con el requisito de ser parte de la población objetivo, esto es, que tengan quince años o más y tengan residencia habitual en hogares ubicados en viviendas particulares ocupadas dentro del territorio nacional. Sin embargo, se registran las características sociodemográficas de todas las personas residentes en la vivienda seleccionada mediante el cuestionario construido para estos fines.

4.6 Definición de variables

4.6.1 Variables analíticas

Una variable analítica es aquella que se genera en función de otras variables y que se incluye en el análisis estadístico. Las variables analíticas de la ESI corresponden a ingresos de origen laboral o de la ocupación. Específicamente, estos son:

- a) Variable **ing_t_d**, Total ingresos sueldos y salarios. Corresponde a la suma del total de ingresos provenientes de sueldos y salarios de las personas trabajadoras dependientes en su ocupación principal.
- b) Variable **ing_t_i**, Total ingresos trabajo independiente. Corresponde a la suma del total de ingresos proveniente del trabajo independiente.
- c) Variable **ing_t_p**, Ingresos del Trabajo Principal. Corresponde a la suma del total de ingresos provenientes de sueldos y salarios, y el total de ingresos del trabajo independiente. Es la variable principal de análisis de la ESI.
- d) Variable **ing_ot**, Ingresos de Otros Trabajos. Corresponde a los ingresos provenientes de otro(s) trabajos distintos de la ocupación principal.
- e) Variable **ing_t_t**, Ingresos Totales del Trabajo. Corresponde a la suma de los ingresos provenientes del trabajo principal y de los ingresos provenientes de otros trabajos.
- f) Variable **ocup_ref**, es la variable de filtro que permite distinguir personas ocupadas con al menos 1 mes en el empleo actual.
- g) Variable **fact_cal_esi**, es el Factor de expansión de la ESI, que permite expandir datos muestrales a poblacionales.

Respecto de los ingresos de la ocupación, es importante tomar en cuenta que la población ocupada se refiere a stocks demográficos del mes anterior a la semana de referencia de la ENE. En consecuencia, la población ocupada de acuerdo con la ESI hace referencia al conjunto de personas para las que su **empleo actual es el mismo que tenían en el mes de referencia** y, por lo tanto, todos los perfiles que caracterizan al empleo se extienden al análisis de los ingresos laborales de este grupo de personas.

Adicionalmente, la encuesta permite identificar a la población que tuvo el mes pasado un segundo empleo, o bien otro empleo distinto al que tenía en la semana de referencia de la encuesta (incluyendo a quienes no necesariamente hayan estado empleados en dicha semana, pero que sí trabajaron el mes previo), y que reportan información de ingresos laborales. Para este grupo de personas no es posible extender los perfiles de ocupación que se pudieran haber informado en el

cuestionario central de la ENE²¹, ya que el único empleo que se caracteriza en la encuesta de empleo es el trabajo principal realizado en la semana de referencia.

En síntesis, la información de ingresos de la ocupación de personas trabajadoras –dependientes e independientes– se refiere exclusivamente a aquellos cuyo empleo actual es el mismo que tenían en el mes de referencia.

Por otro lado, se destaca que la ESI mide los ingresos de las personas ocupadas independientes de acuerdo con dos categorías ocupacionales:

- “Empleadores”, que corresponde a aquellas personas que son propietarias de alguna empresa o negocio y que contratan a una o varias personas como asalariados a lo largo de un periodo continuo, que incluye el periodo de referencia (no se considera para este caso los familiares no remunerados).
- “Trabajadores por cuenta propia”, que corresponde a aquellas personas que trabajan por su cuenta, constituyendo el segundo grupo más numeroso dentro de la población ocupada y, a diferencia de los empleadores, no contratan de manera continua a otras personas asalariadas para que trabajen con ellos.

Para el caso de las personas ocupadas independientes, sean o no empleadoras, se busca separar los ingresos brutos de los gastos totales de manera que se obtenga, al igual que en el caso de las personas asalariadas, el ingreso neto percibido por la actividad realizada. Debido a que para las personas ocupadas independientes se mezclan los ingresos provenientes del capital y del trabajo, lo que se reporta en la encuesta es una combinación de utilidades y sueldos²², lo que se denomina en la literatura como ingreso “mixto”²³.

4.6.2 Variables de información muestral.

Además de las variables analíticas antes descritas, existe un set de variables de información muestral que se incluyen en las bases de datos. Estas variables son compartidas con la ENE, y corresponden a los datos que permiten identificar las características geográficas y del marco muestral utilizado en la encuesta.

²¹ El ingreso de trabajos secundarios se captura en un ítem específico del módulo ESI. Es una pregunta global que incorpora tanto ingresos en dinero como especies.

²² Por ejemplo, cuando el propietario de una empresa o negocio se asigna un sueldo a sí mismo y además se apropia de las utilidades de la empresa.

²³ La denominación de ingreso “mixto” es una forma de denotar las dificultades de distinción entre lo que corresponde a depreciación del capital y utilidades, y el pago de sueldos y salarios.

4.6.3 Variables de identificación de la persona

Dentro de las variables requeridas para la identificación individual de miembros de un hogar, se construyen dos variables para poder identificar a las personas entrevistadas, estas son:

- **nro_linea:** Corresponde al número de la línea con el cual la persona fue identificada en el cuestionario, o al orden correlativo que la persona toma dentro de la composición de un hogar.
- **idrp:** Identificador único de personas. Esta variable permite identificar a las personas encuestadas de forma longitudinal.

4.7 Nomenclatura utilizada

Las nomenclaturas utilizadas en la Encuesta Suplementaria de Ingresos provienen de los principales clasificadores utilizados para caracterizar a la población en edad de trabajar y, en la mayoría de los casos, a las personas ocupadas.

4.7.1 Clasificación Internacional Normalizada de la Educación del año 1997 (CINE 97)

Forma parte de la familia internacional de Clasificaciones Económicas y Sociales de las Naciones Unidas, las cuales son empleadas a nivel mundial en la elaboración de estadísticas con el objetivo de acopiar y analizar datos comparables a nivel internacional de manera consistente. Dentro de estas, la CINE representa una clasificación de referencia que permite ordenar los programas educativos y sus respectivas certificaciones por niveles de educación y campos de estudio según la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO).

Actualmente, a nivel nacional la categoría contempla el siguiente desglose que se aplica a todas las personas en edad en trabajar que conforman la población objetivo de la encuesta:

1. Nunca estudió
2. Educación preescolar
3. Educación primaria (nivel 1)
4. Educación primaria (nivel 2)
5. Educación secundaria
6. Educación Técnica (Educación Superior no Universitaria)
7. Educación universitaria
8. Post títulos y maestría
9. Doctorado
10. Nivel ignorado

En el caso de la categoría de personas ocupadas, existen varios clasificadores que contribuyen a caracterizar a las personas ocupadas, de acuerdo con sus actividades realizadas, a la empresa o lugar donde trabajan, entre otras dimensiones.

4.7.2 Clasificación Internacional Uniforme de Ocupaciones del año 2008, adaptación nacional (CIUO 08.CL)

El Clasificador Chileno de Ocupaciones, CIUO 08.CL, corresponde a la adaptación nacional de la Clasificación Internacional Uniforme de Ocupaciones, CIUO 08. El Clasificador Chileno de Ocupaciones, CIUO 08.CL es el estándar oficial para el procesamiento y análisis de las ocupaciones, que permite organizarlas de acuerdo con el tipo de tareas realizadas en un puesto de trabajo y las competencias requeridas para ello, dependiendo del nivel y la especialización de estas competencias.

1. Directores, gerentes y administradores
2. Profesionales, científicos e intelectuales
3. Técnicos y profesionales de nivel medio
4. Personal de apoyo administrativo
5. Trabajadores de los servicios y vendedores de comercios y mercados
6. Agricultores y trabajadores calificados agropecuarios, forestales y pesqueros
7. Artesanos y operarios de oficios
8. Operadores de instalaciones, máquinas y ensambladores
9. Ocupaciones elementales
10. No identificado

4.7.3 Clasificación Internacional Industrial Uniforme revisión 4.CL según la Clasificación de Actividades Económicas Nacional de Encuestas Sociodemográficas (CAENES)

El CAENES es una clasificación basada en el Clasificador Chileno de Actividades Económicas CIIU4.CL 2012. Por este motivo, se rige por los mismos conceptos y criterios del Clasificador Internacional CIIU Rev.4. En términos operativo corresponde al clasificador de la rama de actividad económica de la empresa o institución donde la persona ocupada trabajar o de la que es dueña. Posee 21 ramas de actividad.

- A. Agricultura, ganadería, silvicultura y pesca
- B. Explotación de minas y canteras
- C. Industrias Manufactureras
- D. Suministro de electricidad, gas, vapor y aire acondicionado
- E. Suministro de agua; evacuación de aguas residuales, gestión de desechos y descontaminación
- F. Construcción
- G. Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas
- H. Transporte y almacenamiento
- I. Actividades de alojamiento y servicio de comidas
- J. Información y comunicaciones
- K. Actividades financieras y de seguros
- L. Actividades inmobiliarias
- M. Actividades profesionales, científicas y técnicas
- N. Actividades de servicios administrativos y de apoyo
- O. Administración pública y defensa; planes de seguridad social de afiliación obligatoria
- P. Enseñanza
- Q. Actividades de atención de la salud humana y de asistencia social
- R. Actividades artísticas, de entretenimiento y recreativas
- S. Otras actividades de servicios
- T. Actividades de los hogares como empleadores; actividades no diferenciadas de los hogares como productores de bienes y servicios para uso propio
- U. Actividades de organizaciones y órganos extraterritoriales.

4.7.4 Clasificación Internacional de la Situación en el Empleo del año 1993 (CISE 93)

La CISE-93 clasifica los empleos según el tipo de riesgo económico, solidez del vínculo entre la persona y el empleo, y el tipo de autoridad, que tienen o tendrán los titulares sobre los establecimientos y sobre otros trabajadores. La actual CISE 93 utilizada en la ENE corresponde a la adaptación chilena y posee las siguientes categorías:

1. Empleador
2. Cuenta propia
3. Asalariado sector privado
4. Asalariado sector público
5. Personal de servicio doméstico puertas afuera
6. Personal de servicio doméstico puertas adentro
7. Familiar o personal no remunerado

4.8 Diseño según tipo de fuente: Muestra

4.8.1 Diseño y selección muestral

La explicación del diseño y selección muestral se encuentra en www.ine.gob.cl, apartado de Estadísticas, sección Mercado laboral, subsección Ocupación y desocupación, “Documento Metodológico Encuesta Nacional de Empleo (ENE) - 2020²⁴”.

Como ya se ha mencionado anteriormente, la ESI comparte todos los atributos del diseño muestral de la ENE, por lo tanto, también comparte la determinación del tamaño muestral teórico para los meses de octubre, noviembre y diciembre, que corresponden al trimestre de aplicación de la ESI.

4.8.2 Cálculo del factor de expansión

Los factores de expansión se definen como la cantidad de unidades en la población representadas por un elemento que ha sido seleccionado en la muestra.

Como el diseño de la ENE y, por ende, el de la ESI, es bietápico, el factor de expansión teórico o de selección de unidades se compone de:

- Ponderador de Unidades Primarias de Muestreo (UPM)
- Ponderador de Unidades Secundarias de Muestreo (USM)

El ponderador de UPM depende del número de UPM seleccionadas en cada estrato de muestreo h ; del número de viviendas totales en el estrato, y , del número de viviendas particulares de cada UPM seleccionada, según el marco de selección.

Por su parte, el ponderador de USM depende del número de viviendas seleccionadas en la UPM y del número de viviendas obtenido desde la actualización²⁵ de los directorios de cada UPM seleccionada.

El factor de expansión teórico es el mismo para todas las viviendas seleccionadas en una UPM y para todas las personas que conforman dichas viviendas.

Sin embargo, luego del trabajo de campo, se observan edificaciones que no forman parte de la población objetivo, es decir, ya no son viviendas particulares o bien, no se pudo dilucidar su estado de ocupación; así como también, edificaciones que corresponden a viviendas particulares, pero que no fue posible obtener respuesta, por varios motivos (rechazo, moradores ausentes o ausencia de informante idóneo al momento de la entrevista).

En el primer caso, la edificación queda catalogada como no elegible; en el segundo, como elegibilidad desconocida y ; en el tercero, como vivienda sin respuesta. Para cada uno de estos casos, el factor de expansión incorpora ajustes en forma secuencial. Estos son:

1. Ajuste por elegibilidad desconocida
2. Ajuste por no elegibilidad
3. Ajuste por no respuesta

Estos ajustes permiten, por un lado, distribuir los pesos originales (o de selección) de unidades en que se desconoce su estado de elegibilidad entre las unidades donde sí se conoce (elegibles y no elegibles); y , por otro lado, desechar los pesos de las unidades no elegibles ya que, por definición, estas unidades no contienen a la población objetivo por lo que, hacer inferencias sobre ellas sería incorrecto.

Finalmente, el último ajuste al factor de selección de viviendas permite asignar los pesos de viviendas que no respondieron a viviendas que sí lo hicieron dentro de “casillas” o “veintiles de respuesta²⁶” donde las probabilidades de responder la encuesta son similares.

²⁵ Dicha actualización corresponde a la enumeración en terreno o verificación en gabinete del número de viviendas particulares ocupadas que contiene cada UPM seleccionada en la muestra.

²⁶ A partir del modelo de no respuesta se obtiene la distribución de probabilidad de que una vivienda responda, a partir de la cual se obtienen veintiles, formando 20 grupos de viviendas que internamente tienen probabilidades similares de responder.

A continuación, se describen cada una de las etapas del factor:

4.8.2.1 Ponderador de selección de UPM

Al ser el diseño muestral bietápico y estratificado, en una primera etapa se seleccionan las unidades primarias de muestreo (UPM) al interior de cada estrato h con probabilidad proporcional al tamaño, en función al número de viviendas particulares ocupadas que la UPM registra en el marco muestral de selección.

El ponderador de selección de las UPM corresponde al inverso de su probabilidad de selección y se expresa a partir de la ecuación (1):

$$w_{hi} = (\pi_{hi})^{-1} = \frac{M_h}{n_h \cdot M_{hi}} \quad (1)$$

Donde,

w_{hi} : Ponderador de selección de la UPM i en el estrato h .

π_{hi} : Probabilidad de selección de la UPM i en el estrato h .

M_h : Número de viviendas particulares en el estrato h al Censo 2017 o al año de actualización del marco muestral.

M_{hi} : Número de viviendas particulares en la UPM i en el estrato h al Censo 2017 o al año de actualización del marco muestral.

n_h : Número de UPM seleccionadas en la muestra del estrato h .

Este ponderador se interpreta como la cantidad de UPM en el marco muestral que representa una UPM seleccionada en la muestra.

4.8.2.2 Ponderador de selección de viviendas

El ponderador de selección de viviendas $whij$, indica la cantidad de viviendas en la población que representa la vivienda seleccionada en la muestra y es calculado como el inverso de la probabilidad de selección de cada vivienda.

La probabilidad de selección de una vivienda j está condicionada a la selección de la UPM i que la contiene y se expresa como el producto entre la probabilidad de selección de la UPM i en el estrato h y la probabilidad de selección de la vivienda j en la UPM i . Esto se expresa según la ecuación (2):

$$\pi_{hij} = \pi_{hi} \cdot \pi_{hj/i}, \text{ con } \pi_{hj/i} = \frac{M_{hi}}{M'_{hi}} \quad (2)$$

Luego, la probabilidad de selección de la vivienda j se expresa según la ecuación (3):

$$w_{hij} = (\pi_{hij})^{-1} = \frac{1}{\pi_{hi}} \cdot \frac{1}{\pi_{hj/i}} = \frac{M_h}{n_h \cdot M_{hi}} \cdot \frac{M'_{hi}}{m_{hi}} \quad (3)$$

Donde,

w_{hij} : Factor de expansión de selección de la vivienda j dentro de la UPM i en el estrato h .

π_{hij} : Probabilidad de selección de la vivienda j dentro de la UPM i en el estrato h .

π_{hi} : Probabilidad de selección de la UPM i en el estrato h .

$\pi_{hj/i}$: Probabilidad de selección de la vivienda j dado que fue seleccionada en la UPM i en el estrato h .

M_h : Número de viviendas particulares en el estrato h al Censo 2017 o al año de actualización del marco muestral.

n_h : Número de UPM seleccionada en la muestra en el estrato h .

M_{hi} : Número de viviendas particulares de la UPM i en el estrato h al Censo 2017 o al año de actualización del marco muestral.

m_{hi} : Número de viviendas seleccionadas en la UPM i en el estrato h .

M'_{hi} : Número de viviendas particulares ocupadas, según actualización de los listados de viviendas, en la UPM i en el estrato h .

4.8.2.3 Ajuste por elegibilidad desconocida

Al momento de selección de la muestra de viviendas, desde los directorios actualizados, ya sea en terreno o en gabinete, todas deben tener un estatus de elegibilidad conocida; más aún, deben ser elegibles, es decir, pertenecer a la población objeto de estudio. Sin embargo, en ocasiones ocurre que, durante el trabajo de campo de la encuesta no es posible comprobar el estatus de algunas viviendas, debido al impedimento de acceso (por clima, seguridad, entre otros) o simplemente la muestra no fue gestionada oportunamente, quedando clasificadas, en ambos casos, como “Elegibilidad desconocida”. Para mitigar esta situación, se realiza un ajuste al factor de expansión teórico o de selección de vivienda con la finalidad de redistribuir los pesos de las unidades que resultaron con elegibilidad desconocida, en forma proporcional entre todas las unidades que resultaron con elegibilidad conocida (elegibles y no elegibles).

Así, el ponderador de ajuste es definido mediante la ecuación (4):

$$\hat{R}_{h,known} = \frac{\sum_{i \in \Omega_h} \sum_{j \in \theta_i} w_{hij}}{\sum_{i \in \Omega_h} \sum_{j \in \theta_{i,known}} w_{hij}} \quad (4)$$

Donde,

θ_i : Conjunto de viviendas j seleccionadas en la UPM i .

$\theta_{i,known}$: Conjunto de viviendas j seleccionadas en la UPM i y clasificadas como elegibilidad conocida (elegibles o no elegibles).

Ω_h : Conjunto de UPM i seleccionados en el estrato h .

w_{hij} : Factor de expansión de selección de la vivienda j dentro de la UPM i en el estrato h .

La expresión $\hat{R}_{h,known}$ es calculada al interior de cada estrato de muestreo a partir del cociente entre la estimación de viviendas totales (elegibles, no elegibles y elegibilidad desconocida) y la estimación del total de viviendas de elegibilidad conocida (elegibles y no elegibles).

Luego, se obtiene el factor de expansión de selección de vivienda ajustado por elegibilidad desconocida (w'_{hij}), como se muestra en la ecuación (5):

$$w'_{hij} = \hat{R}_{h,known} \cdot w_{hij} \quad (5)$$

4.8.2.4 Ajuste por no elegibilidad

Las unidades seleccionadas en la muestra, conforme a la población objetivo de la ENE (ESI) corresponden a viviendas particulares ocupadas donde residen personas de quince años o más. Sin embargo, pese a los esfuerzos por mantener el marco de muestreo lo más actualizado posible, siempre existe la posibilidad de incluir en la muestra viviendas que no contienen a la población objetivo, debido a errores en el proceso de actualización o por cambios en el uso de la vivienda.

Por estos motivos se realiza un ajuste al factor de expansión que implica la asignación de un valor nulo (*missing value*) en factores cuyas viviendas hayan resultado como “no elegibles”.

La implementación de este ajuste se realiza a partir de la ecuación (6):

$$R_{hij,eleg} = \begin{cases} 1, & \text{si la vivienda } j \text{ del conglomerado } i \text{ en el estrato } h \text{ es elegible} \\ ., & \text{si la vivienda } j \text{ del conglomerado } i \text{ en el estrato } h \text{ no es elegible} \end{cases} \quad (6)$$

Luego, el factor de expansión de viviendas, ajustado por elegibilidad desconocida y por no elegibilidad, queda expresado según la ecuación (7):

$$w''_{hij} = w'_{hij} \cdot R_{hij,eleg} \quad (5)$$

4.8.2.5 Ajuste por no respuesta

A partir de esta etapa, los ajustes se realizan sobre viviendas elegibles, es decir, aquellas viviendas particulares ocupadas donde, se obtuvo o no, respuesta de los integrantes de los hogares que las conforman. Este ajuste considera la probabilidad que tienen ciertas viviendas de no responder la encuesta, ya sea porque las personas informantes rechazaron responderla o bien, donde no fue posible establecer contacto o no se pudo llevar a cabo la entrevista por otros motivos.

El ajuste por no respuesta busca corregir los pesos para que la encuesta, a partir de las unidades que responden, sea representativa de toda la población objeto de estudio. Como resultado, los pesos de las unidades que respondieron a la encuesta se incrementan para compensar el peso de las unidades que no respondieron.

Para implementar este ajuste se desarrolla un modelo para determinar la probabilidad de respuesta de viviendas dadas ciertas características. Estas características se denominan covariables y se listan a continuación:

- Área
- Región
- Total de personas en la Fuerza de Trabajo
- Edad promedio de personas en edad de trabajar
- Edad promedio de personas en edad de trabajar, al cuadrado

- Cantidad de hombres por vivienda
- Cantidad de hombres por vivienda, al cuadrado
- Cantidad de mujeres por vivienda
- Cantidad de mujeres por vivienda, al cuadrado
- Promedio de ocupados por vivienda

Obtenido el modelo y sus probabilidades de respuesta, se crean 20 casillas o veintiles de respuesta, que corresponden al nivel donde se realizará el ajuste.

Luego, en cada uno de los veintiles, se obtiene el ponderador $\hat{R}_{h,R}$ definido como la razón entre la estimación del total de viviendas elegibles en la población y la estimación del total de viviendas de las que se obtiene respuesta de las personas informantes. Ambas estimaciones calculadas a partir del factor de expansión definido en la ecuación (7).

Así, el ajuste por no respuesta queda determinado por la ecuación (8):

$$\hat{R}_{h,R} = \begin{cases} \frac{\sum_{i \in \Omega_h} \sum_{j \in \theta_{i,elec}} W''_{hij}}{\sum_{i \in \Omega_h} \sum_{j \in \theta_{i,elec,R}} W''_{hij}} , & \text{si la vivienda } j \text{ es elegible y responde} \\ . & , \text{ si la vivienda } j \text{ es elegible y no responde} \end{cases} \quad (8)$$

Donde:

$\theta_{i,elec}$: Conjunto de viviendas j seleccionadas en la UPM i , clasificadas como elegibles.

$\theta_{i,elec,R}$: Conjunto de viviendas j logradas en la UPM i , clasificadas como elegibles y que responden la encuesta.

Ω_h : Conjunto de conglomerados i seleccionados en el estrato h .

W''_{hij} : Factor de expansión de selección de la vivienda j dentro de la UPM i en el estrato h , ajustado por no elegibilidad.

Finalmente, el factor de expansión ajustado por no respuesta se expresa según la ecuación (9):

$$w_{hij}^{NR} = \hat{R}_{h,R} \cdot w''_{hij} \quad (9)$$

Este factor de expansión es asignado a las k personas dentro de la vivienda j en la UPM i del estrato h .

4.8.2.6 Ajuste a stocks poblacionales

A partir del factor ajustado por no respuesta es posible obtener una estimación del total poblacional que puede no concordar con estimaciones obtenidas desde otras fuentes. Estas diferencias podrían explicarse, en principio, por desactualizaciones de marco, como también por los resultados propios del trabajo de campo.

Luego, como una medida de mitigar estos desbalances se realiza un ajuste de calibración a stocks poblacionales conocidos, como lo son las proyecciones de población, generadas por el INE, a partir de modelos demográficos. Dicho ajuste se realiza según variables como sexo y edad originando estimaciones aproximadamente insesgadas y con varianza más pequeña que la del estimador de Horvitz y Thompson.

Los métodos de calibración se basan en algoritmos iterativos que ajustan sucesivamente los factores de expansión para que alcancen totales poblacionales dados por las proyecciones, mientras se minimiza la distancia entre estos y los calibrados. Esto permite mantener, en los factores calibrados, las propiedades de los factores originales.

El método de calibración utilizado en la ESI es el método multiplicativo o Raking²⁷. Este método se aplica a partir de marginales de calibración definidas específicamente para la encuesta, cumpliendo que, en todas las categorías de cada una de ellas se alcancen los stocks poblacionales, dados por las proyecciones de población según el Censo de Población y Viviendas 2017.

Las marginales de calibración definidas son:

- A nivel nacional: según sexo y trece tramos de edad. Estos son: [0-14], [15-19], [20-24], [25-29], [30-34], [35-39], [40-44], [45-49], [50-54], [55-59], [60-64], [65-69] y [70+].
- A nivel nacional: estimación de la población extranjera.
- A nivel regional: Población total según sexo.
- A nivel regional: Población en Edad de Trabajar (PET) según sexo.
- A nivel regional: Población total en el área Urbana según sexo.
- Región Metropolitana: Población total para seis grupos de edad por sexo. Estos son: [15-24], [25-34], [35-44], [45-54], [55-64] y [65+].
- Regiones (sin RM): Población total para tres grupos de edad. Estos son: [15-34], [35-54] y [55+].

Adicionalmente, se incluyen marginales para ciertas provincias y grandes conurbaciones²⁸.

²⁷ Para mayor detalle revisar: Página 74 y Anexo N° 8. Construcción de insumo para la calibración por población extranjera en factores de expansión de la ENE. https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodologia-ene-2020.pdf?sfvrsn=793380e9_14.

²⁸ Para mayor detalle revisar: Página 75. https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodologia-ene-2020.pdf?sfvrsn=793380e9_14.

4.9 Medidas de precisión

Los errores de muestreo existen como consecuencia de la utilización de una parte de la población para estimar características de todo el conjunto poblacional. Sin embargo, existen distintas formas de medir estos errores para las variables de interés, como, por ejemplo, la varianza, la desviación estándar o algunas transformaciones de estas, como el error absoluto, el coeficiente de variación, el error relativo, entre otros. En la ESI se utilizan principalmente las tres últimas medidas.

4.9.1 Error absoluto e intervalo de confianza

Para el cálculo del intervalo de confianza de 95% para la estimación de la variable Y , se utiliza el error absoluto (error de estimación que está expresado en las mismas unidades de la variable en estudio²⁹).

El error absoluto permite observar cuanto es la diferencia máxima esperada entre el valor estimado y el verdadero parámetro poblacional, donde dicha estimación se puede realizar a distintos niveles de confianza, siendo la más común la del 95%. A partir de este error se construye el intervalo de confianza de la variable de interés Y .

El error absoluto se obtiene del producto entre la desviación estándar ($S(\hat{y})$) de la estimación y el valor crítico asociado al percentil de la distribución *t-student* con v grados de libertad ($t_{1-\frac{\alpha}{2},v}$), donde el primer argumento hace referencia al nivel de confiabilidad con el cual se desea hacer inferencia (95%), mientras que el segundo argumento corresponde a los v grados de libertad con los que se calcula el valor crítico³⁰. Su expresión algebraica está dada por:

$$E_{abs}(\hat{y}) = t_{1-\frac{\alpha}{2},v} \cdot S(\hat{y}) \quad (28)$$

Alternativamente, en términos probabilísticos, se puede expresar como:

$$\Pr(|\hat{Y} - Y| \leq E_{abs}(\hat{y})) = 1 - \alpha \quad (29)$$

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

Por su parte, el intervalo de confianza es el intervalo estimado que incluye el promedio de las estimaciones de todas las muestras posibles con una probabilidad del 95%. Este intervalo tiene un límite inferior y un límite superior dentro de los cuales se espera que se encuentre el valor real o poblacional con el nivel de confianza del 95%.

²⁹ Cuando la variable Y es un porcentaje o tasa, el error absoluto se expresa en puntos porcentuales.

³⁰ Los grados de libertad corresponden a la diferencia entre el número de conglomerados no vacíos (con presencia de la variable de interés) y el número de estratos no vacíos.

El intervalo de confianza es construido a través del error absoluto, considerando como el valor mínimo del parámetro poblacional el parámetro estimado menos el error absoluto, y como valor máximo el parámetro estimado más el error absoluto. Matemáticamente, el intervalo de confianza es igual a:

$$IC(\hat{y})_{95\%} = (\hat{y} - E_{abs}(\hat{y}); \hat{y} + E_{abs}(\hat{y})) = \left(\hat{y} - t_{1-\frac{\alpha}{2},v} \cdot S(\hat{y}); \hat{y} + t_{1-\frac{\alpha}{2},v} \cdot S(\hat{y}) \right) \quad (30)$$

4.9.2 Error relativo y coeficiente de variación

Adicionalmente, existen otros indicadores, que dan cuenta de los errores de muestreo, en términos relativos, el error relativo y el coeficiente de variación.

El error relativo, expresa la magnitud del error de muestreo en términos porcentuales, reflejando la máxima variación porcentual que se puede esperar entre el valor estimado y el valor poblacional con una probabilidad de $(1 - \alpha) \cdot 100\%$, y se obtiene de la siguiente forma:

$$E_{rel}(\hat{y}) = t_{1-\frac{\alpha}{2},v} \cdot \frac{\hat{\sigma}_{\hat{y}}}{\hat{y}} = t_{1-\frac{\alpha}{2},v} \cdot CV_{\hat{y}} \quad (31)$$

Donde, $\frac{\hat{\sigma}_{\hat{y}}}{\hat{y}}$ representa el coeficiente de variación $CV_{\hat{y}}$ asociado a la estimación de la variable \hat{y} .

La gran utilidad tanto del error relativo como del coeficiente de variación es que pueden ser expresados en porcentajes, lo que los distingue de otras mediciones del error, al permitir comparar entre diferentes estimaciones, soslayando las diferencias entre unidades de medida.

4.9.3 Contrastes de hipótesis

La realización de pruebas de contrastes de hipótesis permite comprobar, con cierto nivel de confianza, si la diferencia entre dos estimaciones es o no estadísticamente significativa³¹. Desde la ESI 2020, para realizar estas pruebas estadísticas³², se calcula el producto entre la desviación estándar de la estimación de la diferencia, obtenida a través del método delta; y el valor crítico asociado al percentil de la distribución T-student con los grados de libertad determinados por la estimación. En este apartado se ahonda en los conceptos estadísticos asociados a estas pruebas.

Un contraste de hipótesis o test de hipótesis, en forma amplia, es un procedimiento para contrastar dos hipótesis o afirmaciones sobre el valor de cierto parámetro θ en la población o de la diferencia de dos parámetros poblacionales $\theta_1 - \theta_2$ provenientes de dos subpoblaciones, respectivamente.

4.9.3.1 Planteamiento de hipótesis de contraste

Todo contraste de hipótesis se basa en la formulación de dos hipótesis exhaustivas y mutuamente excluyentes (hipótesis nula e hipótesis alternativa).

- **Hipótesis nula (H_0):** asigna un valor o rango de valores al parámetro en cuestión. Es la afirmación que podría ser rechazada por los hechos o evidencia empírica a través de un procedimiento muestral.
- **Hipótesis alternativa (H_1):** es la negación de la hipótesis nula, corresponde a la alternativa de interés en la cual se centra el investigador.

En el contexto de comparación de totales, proporciones, medias y ratios, y sin perder generalidad, se puede tomar como referencia al parámetro θ como si fuera la media poblacional, aunque puede contrastarse la diferencia de cualquiera de los otros estadísticos descriptivos anteriores.

$H_0: \theta_1 - \theta_2 = 0 \Leftrightarrow H_0: \theta_1 = \theta_2$	No existe diferencia entre las medias poblacionales
$H_1: \theta_1 - \theta_2 \neq 0 \Leftrightarrow H_0: \theta_1 \neq \theta_2$	Existe diferencia entre las medias poblacionales

4.9.3.2 Estadístico de contraste:

El estadístico de contraste constituye un resultado muestral que cumple con las condiciones de:

- Proporcionar información empírica relevante sobre la afirmación propuesta en la hipótesis nula (H_0). Diferencias estimadas (a través de las muestras) alejadas relativamente de cero entregan evidencia en contra de H_0 y a favor de H_1 .

³¹ Que la diferencia entre dos estimaciones sea estadísticamente significativa, da a entender que en las poblaciones subyacentes existe diferencia entre los parámetros.

³² Estas pruebas estadísticas se refieren específicamente a pruebas de diferencias de medias, totales, proporciones y ratios dentro de un mismo año. Se excluye la realización de pruebas de diferencia de medianas; ya que sus medidas de variabilidad se obtienen a partir de diseños con repetición sub-bootstrap cuyas implicancias metodológicas están siendo evaluadas para su próxima incorporación.

- Poseer distribución muestral conocida. Bajo la hipótesis nula (bajo el supuesto que las medias poblacionales son iguales) y muestras lo suficientemente grandes como para aplicar el Teorema Central del Límite, la diferencia muestral se distribuye aproximadamente Normal si la varianza poblacional de la diferencia es conocida. Cuando esta varianza poblacional es desconocida (casi siempre) y estimada a través de las mismas muestras, la distribución de la diferencia estandarizada es T-Student, con ν grados de libertad.

En el contexto de muestreo bietápico por conglomerado y estratificado, los grados de libertad ν de la T-Student corresponden al número de conglomerados no repetidos con muestra en la subpoblación evaluada, menos el número de estratos no repetidos con muestra en la subpoblación evaluada. Para la ESI 2020 y 2021 se estuvo en presencia de dos marcos muestrales (marco dual) en que los conglomerados del marco antiguo y nuevo no poseían la misma configuración geográfica ni tampoco los estratos (insumos con los que se calculan los grados de libertad). Aunque los marcos no poseían la misma configuración geográfica, existió un solapamiento parcial y total de los conglomerados y estratos.

Con el objetivo de establecer una estrategia semi-conservadora en las estimaciones, se utiliza grados de libertad $\nu = 60$ para estos años. Con lo anterior, el percentil de la T-Student correspondiente al test bilateral (dos colas) con un nivel de significación $\alpha = 5\%$, es $t_{1-\alpha/2}^{\nu} = t_{0,975}^{60} = 2.000297822 \dots \approx 2$.

El estadístico estandarizado de contraste es:

$$t = \frac{(\hat{\theta}_1 - \hat{\theta}_2) - (\theta_1 - \theta_2)_0}{\sqrt{\hat{v}(\hat{\theta}_1 - \hat{\theta}_2)}} = \frac{(\hat{\theta}_1 - \hat{\theta}_2)}{\sqrt{\hat{v}(\hat{\theta}_1) + \hat{v}(\hat{\theta}_2) - 2 \cdot \text{COV}(\hat{\theta}_1; \hat{\theta}_2)}} \sim t^{60}, \text{ bajo } H_0, \text{ es decir, } (\theta_1 - \theta_2)_0 = 0$$

Nótese que la varianza estimada de la diferencia incluye la covarianza estimada entre las medias de las poblaciones subyacentes debido a que, en diseños complejos por conglomerado, las varianzas se calculan por el método del conglomerado último y las muestras podrían compartir conglomerados y/o estratos o dominios comunes. Los softwares estadísticos como Stata, R³³ y otros (SAS y SPSS, por ejemplo), emplean el método Delta o linealización de Taylor que incluye el término de la covarianza.

Para la ESI 2022 y sus próximas ediciones, en presencia de un marco único, el cálculo de los grados de libertad se realiza de forma efectiva para cada estimación.

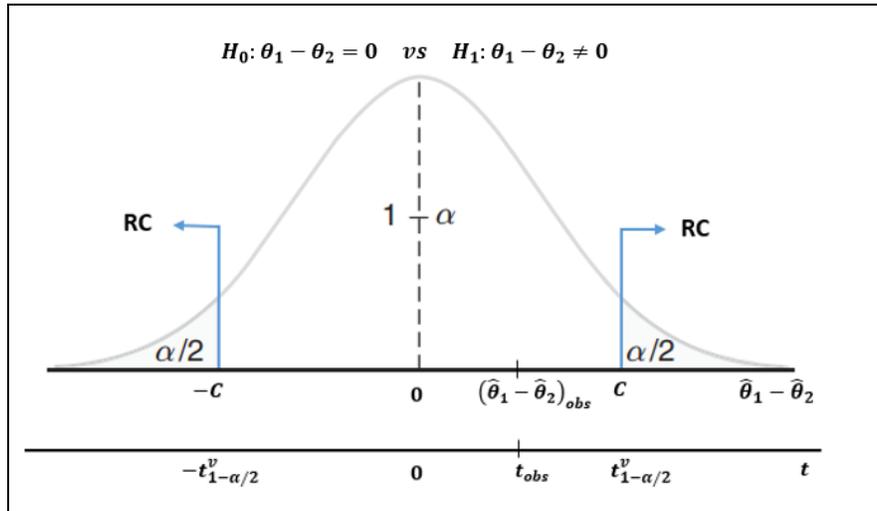
4.9.3.3 Región crítica o de rechazo de H_0 (RC)

Para un contraste de hipótesis de diferencias, se establece una región crítica o de rechazo de la hipótesis nula, de modo que, si la diferencia observada $(\hat{\theta}_1 - \hat{\theta}_2)_{obs}$ cae en esta RC, se toma la decisión de rechazar la hipótesis nula de igualdad, estableciéndose estadísticamente que existe diferencia o que la diferencia poblacional de los estadísticos es significativa. En el gráfico, la

³³ En el caso de R, específicamente a través de la función *svycontrast* del paquete *survey*.

diferencia observada no cae en la RC, sino que cae en la región de aceptación de la hipótesis nula, de tamaño $1 - \alpha$, concluyendo que no existe evidencia para rechazar H_0 o que la diferencia no es significativa o que no existe diferencia entre los estadísticos poblacionales o que los estadísticos poblacionales son iguales.

Figura 1. Región Crítica de nivel α para el contraste de hipótesis bilateral



Fuente: Instituto Nacional de Estadísticas.

La RC está formada por el conjunto de todas las diferencias absolutas observadas en el eje $\hat{\theta}_1 - \hat{\theta}_2$ del gráfico que son mayores que el valor crítico $C > 0$, o equivalentemente, formada por todas las diferencias absolutas estandarizadas (valores en el eje t del gráfico) que son mayores que el percentil de la T-Student con v grados de libertad, que, dada la simetría de esta distribución, corresponde al percentil de orden $1 - \alpha/2$, dejando un área bajo la curva de tamaño $\alpha/2$ para cada extremo de la distribución.

$$RC = \{ \hat{\theta}_1 - \hat{\theta}_2 ; |\hat{\theta}_1 - \hat{\theta}_2| > C \} \Leftrightarrow \{ \hat{\theta}_1 - \hat{\theta}_2 ; \hat{\theta}_1 - \hat{\theta}_2 < -C \text{ ó } \hat{\theta}_1 - \hat{\theta}_2 > C \} \Rightarrow Prob(RC) = \alpha$$

$$RC = \{ t ; |t| > t_{1-\alpha/2}^v \} \Leftrightarrow \{ t ; t < -t_{1-\alpha/2}^v \text{ ó } t > t_{1-\alpha/2}^v \} \Rightarrow Prob(RC) = \alpha$$

4.9.3.4 Nivel de significación a posteriori (*p-value*)

El verdadero nivel de significación de un contraste estadístico está dado por el *p-value* $= \alpha^*$ y corresponde al valor de significación mínimo a partir del cual se comienza a rechazar la hipótesis nula. Es decir, para cualquier nivel de significación fijado a priori $\alpha \geq \alpha^*$, existirá suficiente evidencia estadística para rechazar H_0 ($\theta_1 = \theta_2$) y equivalentemente aceptar H_1 ($\theta_1 \neq \theta_2$). En la ESI el valor de significación a priori está fijado en $\alpha = 0,05$, por lo que podemos rechazar H_0 cuando el *p-value* $< 0,05$. El *p-value* se puede calcular como:

$$\alpha^* = Prob\left(|\hat{\theta}_1 - \hat{\theta}_2| > |\hat{\theta}_1 - \hat{\theta}_2|_{obs}\right) = Prob\left(\hat{\theta}_1 - \hat{\theta}_2 < -|\hat{\theta}_1 - \hat{\theta}_2|_{obs} \text{ ó } \hat{\theta}_1 - \hat{\theta}_2 > |\hat{\theta}_1 - \hat{\theta}_2|_{obs}\right)$$

Dada la simetría de la distribución T-Student, el p-value también puede calcularse como:

$$\alpha^* = 2 \cdot Prob\left(|\hat{\theta}_1 - \hat{\theta}_2| > |\hat{\theta}_1 - \hat{\theta}_2|_{obs}\right)$$

En términos de la diferencia observada estandarizada, se tiene:

$$\alpha^* = Prob(|t| > |t_{obs}|) = 2 \cdot Prob(t > |t_{obs}|)$$

4.9.3.5 Contraste de hipótesis entre años

A partir de la ESI 2020 se utiliza un marco muestral diferente al marco muestral de la ENE-ESI de años anteriores (2010-2019). La variable que designaba a las unidades primarias de muestreo en el marco antiguo era "id_directorio"; mientras que en el marco nuevo esta es "conglomerado".

Para poder realizar contrastes de hipótesis entre años, las bases ESI 2010-2019 cuentan con la variable "conglomerado", la cual homologa la denominación de las unidades primarias de muestreo en dichas bases de datos a su nueva nomenclatura en el marco muestral de la ESI 2020 y 2021. De esta forma, se pueden unir las bases de datos (ESI 2020 o posterior + ESI de un año anterior) y declarar el diseño complejo en un software estadístico utilizando esta variable de conglomerados homologados a su nomenclatura en el marco 2020³¹; lo que permite mejorar la medición de la covarianza para la realización de pruebas de hipótesis.

4.9.4 Representatividad de las estimaciones

Basados en el Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares, el que una estimación sea o no fiable depende del enfoque planteado en el flujograma para publicar o suprimir estimaciones (INE, 2020b). El estándar fue adoptado desde la ESI 2019, el cual contempla los siguientes elementos:

4.9.4.1 Elementos relacionados con el tamaño muestral

a) Tamaño muestral, sin pérdida de generalidad, al señalar el término "tamaño muestral" se hará referencia al total de unidades de análisis consideradas para obtener las estimaciones. El umbral de aceptación considera un tamaño muestral mayor o igual que 60 observaciones.

b) Grados de libertad, los grados de libertad (gl) se definen como el número de observaciones independientes que son libres de variar (número de individuos en los datos) menos el número de parámetros estimados (número de relaciones impuestas a los datos), lo que en un diseño muestral complejo se obtiene como la diferencia entre el número de conglomerados no vacíos (con presencia de la variable de interés) y el número de estratos no vacíos. Además, según (Heeringa, West, & Berglund, 2010) estos también, dan cuenta de la precisión con la que se están estimando los parámetros a través de la muestra, por lo que, a mayores grados de libertad, las estimaciones serán

más precisa, y viceversa. El umbral de aceptación contempla que la estimación evaluada tenga nueve o más grados de libertad.

Si el nivel de desagregación evaluado no cumple con el criterio del tamaño muestral o con los grados de libertad, entonces se debe suprimir la estimación para dicho nivel.

4.9.4.2 Elementos relacionados con la precisión de la estimación

Si el estimador es proporción será evaluado por medio del error estándar (ee), los umbrales quedaran definidos por la siguiente función:

$$\text{Máximo ee tolerable} = \begin{cases} \sqrt[3]{p^2}/9, & \text{si } p \leq 0,5 \\ \sqrt[3]{(1-p)^2}/9, & \text{si } p > 0,5 \end{cases}$$

Si el nivel de desagregación evaluado cumple con los criterios del tamaño muestral y de los grados de libertad, pero no cumple los criterios anteriores, se publica la estimación advirtiendo que se trata de una estimación poco fiable.

Si el estimador no es proporción será evaluado a través del coeficiente de variación, y se considera:

$$\text{Máximo CV tolerable} = \begin{cases} CV \leq 0,15 \rightarrow \text{Aceptable} \\ 0,15 < CV \leq 0,30 \rightarrow \text{Estimación poco fiable} \end{cases}$$

Si el nivel de desagregación evaluado cumple con los criterios del tamaño muestral y de los grados de libertad, pero presenta un coeficiente de variación mayor que 0,30 se debe suprimir la estimación para dicho nivel.

Posteriormente, la publicación del tabulado depende del número de estimaciones que se han suprimido o generado con alerta de poca fiabilidad, en este caso, se recomienda que, si estas estimaciones suman más de 50% de las que forman el tabulado, este último no debe ser publicado.

4.10 Medidas complementarias

4.10.1 Tasa de respuesta y tasa de logro objetivo de la recolección

La tasa de respuesta describe el número de unidades donde se ha obtenido una entrevista completa o parcial, sobre el total de unidades muestrales a encuestar que son potencialmente elegibles (entrevistadas, rechazadas, no contactadas y no entrevistadas por otros motivos) y aquellas de elegibilidad desconocida, en un período de tiempo determinado.

$$Tasa\ de\ respuesta = \frac{E}{E + R + NC + O + ED}$$

Donde:

E: número de unidades entrevistadas

R: número de unidades rechazadas

NC: número de unidades no contactadas

O: número de unidades no entrevistada por otros motivos

ED: número de unidades de elegibilidad desconocida

La tasa de logro objetivo es un complemento de la tasa de respuesta, ya que permite evaluar la calidad del proceso de recolección en relación con los niveles de precisión establecidos en el diseño muestral de cada estudio.

Describe el número de unidades muestrales a encuestar en un período determinado y donde se obtiene una entrevista (completa o parcial), en relación con las unidades de la “muestra objetivo” definidas en la encuesta. A diferencia de la tasa de respuesta este indicador no distingue en su cálculo a aquellas unidades que no forman parte de la población objetivo del estudio (no elegibles).

Por lo general, este indicador se emplea para estudios donde el diseño muestral considera un número de unidades como sobremuestra con el propósito de mitigar la falta de respuesta.

$$Tasa\ de\ logro\ objetivo = \left[\frac{E + P}{Mo} \right] * 100$$

Donde,

E: número de unidades con entrevista completa

P: número de unidades con entrevista parcial que cumplen con los criterios de suficiencia ENE

Mo: Total de unidades muestra objetivo en el período t

Las viviendas de la muestra pueden estar en diferentes estados según su nivel de logro. A continuación, se describen las principales categorías:

- **Entrevistadas (E):** viviendas u hogares contenidos en la muestra en las que fue posible hacer contacto y aplicar la encuesta, obteniendo información suficiente para el estudio, de acuerdo a la evaluación de suficiencia de datos.
- **Rechazo (R):** viviendas u hogares contenidos en la muestra en los que fue posible hacer contacto, sin embargo, un informante idóneo no acepta participar en el estudio desde un inicio, no existiendo recolección de información de estas unidades.
- **Otras no entrevistadas (O):** viviendas u hogares contenidos en la muestra en los que fue posible interactuar presencialmente con un informante idóneo, pero no fue posible aplicar la encuesta por razones distintas al rechazo, por ejemplo, por problemas de comunicación con el informante debido al idioma; porque el informante estaba impedido física y/o mentalmente para contestar o por muerte del informante o de algún familiar, entre otros motivos.
- **Break-off:** *«Se refiere a unidades en las que fue posible hacer contacto y realizar la entrevista, sin embargo, no reportan la información suficiente requerida para ser consideradas en la categoría entrevistadas, por lo que no son incluidas en los resultados finales de la encuesta. Esta categoría constituye una clasificación generada post-recolección de datos, por lo que no es labor del encuestador(a) distinguir estos casos al momento de realizar el trabajo de recolección de datos. La categoría de clasificación “break-off”, se asigna en la fase de procesamiento, al momento de determinar la cantidad de información recogida en cada unidad y decidir si la encuesta cumple con la información suficiente para ser incluida en la base de datos final»* INE (2021). «Guía para la implementación de indicadores de rendimiento y calidad en el monitoreo del proceso de recolección de datos (documento interno)».

4.10.2 Tasas de supervisión

La etapa de supervisión, tanto en la ENE como en la ESI, cumple el rol de garantizar la correcta ejecución del trabajo de recolección de la encuesta, con el objetivo de confirmar que la información recolectada es fidedigna y además identificar problemáticas, así como errores u omisiones que afecten el cumplimiento de los objetivos del estudio.

Respecto a las medidas de precisión relacionadas con la supervisión de la recolección, se establecieron los siguientes indicadores:

- **Tasa de supervisión:** Es el porcentaje de unidades supervisadas respecto a la muestra teórica del periodo.

$$Tasa\ de\ supervisión = \frac{Total\ de\ unidades\ supervisadas}{Muestra\ teórica} * 100$$

- **Tasa de supervisión indirecta re entrevista:** Es el porcentaje de unidades con entrevista completa supervisada respecto al total de unidades entrevistadas para el periodo.

$$Tasa\ de\ supervisión\ re\ entrevista = \frac{Unidades\ con\ entrevista\ completa\ supervisadas}{Total\ de\ unidades\ entrevistadas} * 100$$

- **Tasa de corrección consultas recuperaciones:** Es el porcentaje entre el número de consultas corregidas más el número de consultas por región respecto al número de consultas recibidas a nivel nacional durante el período. Esta tasa se calcula en el contexto de la supervisión de gabinete.

$$Tasa\ de\ corrección = \left(\frac{N\ consulta\ corregida + n\ consulta\ consultada\ por\ región}{N^\circ\ de\ consultas\ recibidas\ en\ el\ periodo\ a\ nivel\ nacional} \right) \times 100$$

Se abordará más en detalle el trabajo de supervisión, en el próximo capítulo.

4.10.3 Tasa de no respuesta parcial y total

La falta de respuesta, tanto total³⁴ como parcial, es una situación recurrente en las encuestas de hogares, especialmente en encuestas de ingresos como la ESI. En la Tabla 3 se presentan las tasas de no respuesta en partidas de ingreso para 2022.

Tabla 3. Tasas de no respuesta partidas de ingreso, OND 2022³⁵

Región	Tasa de no respuesta		Tasa de no respuesta	
	No ponderada (%)		Ponderada (%)	
	Dependiente	Independiente	Dependiente	Independiente
Región de Arica y Parinacota	16,79	14,71	0,29	11,17
Región de Tarapacá	1,65	0,66	0,00	0,00
Región de Antofagasta	5,97	5,61	8,73	1,15
Región de Atacama	5,45	4,23	0,69	0,1
Región de Coquimbo	3,41	2,35	0,08	9,95
Región de Valparaíso	7,01	6,05	13,31	12,39
Región Metropolitana	15	15,74	7,92	7,71
Región del Libertador Gral. Bernardo O`Higgins	4,61	4,47	0,93	0,45
Región del Maule	8,42	6,73	4,5	0,55
Región del Ñuble	4,4	3,26	2,83	0,59
Región del Biobío	2,82	1,91	0,05	0,05
Región de La Araucanía	4,55	3,55	0,23	0,03

³⁴ Debido a que la ESI está asociada directamente a la ENE, no existen problemas de falta de respuesta TOTAL, ya que esta se corrige a través de *un ajuste a los factores de expansión teóricos*, que implica eliminar las viviendas sin respuesta operativa, de modo que las unidades que permanecen en la muestra estimen – sin sesgo– el total poblacional.

³⁵ Tasa de no respuesta calculada con la base de datos intermedia a la que se le aplicará la metodología de imputación.

Región de Los Ríos	8,98	8,23	1,09	21,26
Región de Los Lagos	13,89	15,54	6,08	44,81
Región de Aysén del Gral. Carlos Ibáñez del Campo	3,74	2,47	7,38	0,11
Región de Magallanes y Antártica Chilena	3,35	9,4	0,44	12,1
Total	8,05	7,81	5,87	6,22

Fuente: Instituto Nacional de Estadísticas.

La tasa de no respuesta ponderada se calcula como la proporción expandida de personas ocupadas con al menos un mes en la ocupación principal actual que responden las opciones “No Sabe o No Responde” a monto (d1_opcionb para el caso de las personas dependientes o d5_opcionb, en el caso de las personas independientes), sobre el total de personas ocupadas que tengan al menos un mes en la ocupación principal, diferenciado por personas ocupadas dependientes e independientes. En el caso, de la tasa de no respuesta no ponderada, se aplica la misma condición anterior sin considerar el factor de expansión.

Las tasas de no respuesta ponderada y no ponderada se encuentran dentro de rangos aceptables para este tipo de encuestas (por debajo del 10%), lo cual se explica tanto por factores operativos como técnicos, del control de calidad del levantamiento de la información y los instrumentos de supervisión que se implementan para minimizar los errores.

En cuanto a la no respuesta total, desde 2020 el cálculo del tamaño muestral de la ENE-ESI es sobredimensionado como alternativa al mecanismo de reemplazos que se utilizaba anteriormente y tiene el objetivo de contrarrestar la pérdida de muestra en las sucesivas rondas de la encuesta y la no respuesta por parte de los encuestados.

De esta forma, la ESI, no presenta errores por falta de respuesta total por unidad estadística, debido a que se realiza un proceso de ajuste por no respuesta, con la consecuente supresión de las viviendas no encuestadas³⁶.

³⁶ La no respuesta parcial y la metodología de imputación, se aborda en la sección 2.4.

4.10.4 Tasas de imputación

La ESI realiza imputación a la no respuesta parcial en las partidas de ingreso del trabajo neto (k1b_monto y l1_monto), junto con los días y horas trabajadas de los dependientes (k4a y k4b). Como se observa en la Tabla 4, en 2022 se imputó en 3.098 observaciones las variables de ingreso, correspondiente a 2.112 casos de monto de personas ocupadas dependientes y 732 de monto personas ocupadas independiente y 254 de días y horas efectivas de personas ocupadas dependientes. Esto significó una tasa de imputación no ponderada (TIs/pond)³⁷ de 8,05% para k1b_monto y de 7,81% para l1_monto. Por otra parte, la tasa de imputación ponderada (TIpond), que se construye utilizando el factor de expansión, corresponde a 9,79% en el caso de k1b_monto y a 9,70% para la variable de monto de las personas ocupadas independientes.

Tabla 4. Indicadores imputación

Variable	Total observaciones	Casos imputados	TIs/pond (%)	TIpond (%)
k1b_monto	26.238	2.112	8,05	9,79
l1_monto	9.371	732	7,81	9,70
k4a – k4b	26.238	254	0,97	1,20

Fuente: Instituto Nacional de Estadísticas.

4.11 Estimación de parámetros

4.11.1 Estimador de totales

El estimador de totales insesgado se obtiene multiplicando la variable de interés por el factor de expansión asociado a cada observación (Horvitz et al, 1952), de esta forma es posible llegar a los stocks poblacionales proyectados. Matemáticamente, se expresa como:

$$\hat{\pi}_{Fy} = \sum_{i=1}^n F_i \cdot y_i$$

Con:

F_i : Factor de expansión ESI asociado a la persona i .

³⁷ Correspondiente a la razón entre el número de valores imputados y el número total de valores para la variable dada.

y_i : Variable de interés asociado a la persona i . Ejemplo es: ingreso del trabajo principal. En caso de querer calcular el número de personas ocupados, la variable y_i tomará el valor 1 cuando la persona i pertenece a la población de análisis, y 0 en caso contrario.

4.11.2 Estimador de razón

El estimador de razón corresponde al cociente entre el estimador de totales de una variable de interés (x_i) sobre el estimador de totales de otra variable de interés (y_i). Matemáticamente, se expresa como:

$$\hat{R} = \frac{\sum_{i=1}^n F_i \cdot x_i}{\sum_{i=1}^n F_i \cdot y_i} = \frac{\hat{\pi}_{Fx}}{\hat{\pi}_{Fy}}$$

Con:

F_i : Factor de expansión ESI asociado a la persona i .

x_i : Variable de interés asociado a la persona i . Ejemplo es: ingreso del trabajo principal.

y_i : Variable de interés asociado a la persona i . Ejemplo es el número de horas trabajadas.

Algunos ejemplos son:

- Ingreso por hora: corresponde al total de ingresos de los asalariados provenientes de la ocupación principal ($\hat{\pi}_{Fx}$) dividido por el total de horas trabajadas mensualmente para el caso de los asalariados ($\hat{\pi}_{Fy}$).
- Esta expresión también se utiliza cuando la variable de interés (y_i) es binaria (0,1), esto corresponde a lo que llamamos proporciones. Por ejemplo, la proporción de personas asalariadas corresponde al total de personas asalariadas con al menos un mes en el trabajo ($\hat{\tau}_{Fx}$) dividido por el total de personas ocupadas con al menos un mes en el trabajo ($\hat{\tau}_{Fy}$).

$$\hat{R} = \frac{\hat{\tau}_{Fx}}{\hat{\tau}_{Fy}}$$

4.11.3 Estimador de promedio

El estimador de promedios se obtiene del cociente entre el estimador de totales de una variable de interés (x_i) sobre el estimador de totales de las unidades que presentan información en la variable de interés. Este estimador podría interpretarse como un caso puntual del estimador de razón. Matemáticamente, se expresa como:

$$\widehat{media} = \frac{\sum_{i=1}^n F_i \cdot x_i}{\sum_{i=1}^n F_i \cdot z_i} = \frac{\hat{\lambda}_{Fx}}{\hat{\lambda}_{Fz}}$$

Con:

F_i : Factor de expansión ESI asociado a la persona i .

x_i : Variable de interés asociado a la persona i . Ejemplo el ingreso del trabajo principal.

z_i : Variable que tomará el valor 1 cuando la persona i presenta información en la variable de interés (x_i), y 0 en caso contrario.

Ejemplo: el ingreso promedio de la ocupación principal corresponde al total de los ingresos declarados ($\hat{\lambda}_{Fx}$) dividido por el total de personas en la población ($\hat{\lambda}_{Fz}$).

4.11.4 Estimador de mediana

El estimador de mediana se obtiene ordenando la variable de interés (por ejemplo, alguna medida de ingreso) de forma ascendente y construyendo una variable p_k para el dominio de interés (subpoblación) en base a la definición N°5 propuesta en Hyndman & Fan (1996), adaptada al uso de encuestas de muestreo complejo. Esta regla de interpolación es utilizada debido a sus propiedades favorables para distribuciones continuas, tales como ingresos. La fórmula matemática para este método corresponde a:

$$p_k = (C_k - w_k/2)/C_n$$

Donde

p_k : Valor k-ésimo que determina el cuantil

C_k : Función de distribución acumulada para el factor de expansión del valor k

w_k : Factor de expansión para el valor k

C_n : Suma total del factor de expansión para el dominio de interés (subpoblación)

De tal forma que la mediana corresponde al valor de la variable de interés (x_k) donde se cumple que p_k es igual al cuantil buscado (en el caso de la mediana "0,5"). Cuando ningún valor de la variable de interés en la muestra cumple exactamente con lo anterior, la mediana se encontrará entre los puntos de la muestra $[(x_1, p_1)$ y $(x_2, p_2)]$ y estará determinada por la ecuación lineal:

$$mediana = x_1 + \left(\frac{x_2 - x_1}{p_2 - p_1} \right) * (0,5 - p_1)$$

Donde

x_1 : Valor de la variable de interés presente en la muestra cuyo p_k se encuentra justo por debajo de 0,5

x_2 : Valor de la variable de interés presente en la muestra cuyo p_k se encuentra justo por sobre 0,5

p_1 : Valor respectivo de p_k para el valor x_1

p_2 : Valor respectivo de p_k para el valor x_2

El cálculo de medianas oficial según el método presentado se realiza empleando el paquete "survey" del software R para muestras complejas. En el Anexo 11.2 del presente documento se detalla la sintaxis de esta programación.

5 Metodología de Recolección de datos

5.1 Técnica y métodos de recolección de los datos

El proceso de recolección de datos de la ESI se realiza anualmente, entre el trimestre OND de cada año. El periodo de referencia de los datos corresponde al mes anterior a la semana de referencia, pudiendo ser septiembre, octubre o noviembre de cada año.

La recolección de datos se realiza mediante método CAPI (Computer-Assisted Personal Interviewing), con la utilización de un dispositivo móvil de captura (Tablet). A su vez, la actividad de trabajo de campo que corresponde al levantamiento o recolección de la información propiamente tal era tradicionalmente presencial, sin embargo, a raíz de la pandemia se rotó a una metodología de recolección telefónica y en algunos casos presencial. A partir de esta experiencia, en el INE se determinó que para futuras recolecciones de la ENE y, por ende, de la ESI, se realizará de manera mixta, es decir, contemplará la manera presencial y la manera telefónica.

Al igual como ocurre con el diseño estadístico de la encuesta, las técnicas y metodologías de recolección de datos son compartidas con la ENE, ya que la ESI forma parte del mismo proceso de recolección de ese producto. Es decir, poseen la misma organización del trabajo de campo, distribución operativa de la muestra, estructura organizativa del equipo de recolección en terreno y modelo de supervisión. El desarrollo de estos tópicos se encuentra en el Documento Metodológico oficial de la ENE³⁸.

5.2 Diseño de instrumento de captura de información

El cuestionario ESI³⁹ tiene un diseño único que se aplica a cualquier unidad de información. Está compuesto por tres grandes secciones:

- Sección K "Ingresos de los ocupados dependientes"
- Sección L "Ingresos de los ocupados independientes"
- Sección M "Ingresos de otros trabajos"

El formulario ESI se anexa al cuestionario ENE en el trimestre octubre-diciembre de cada año, y es contestado por el informante idóneo que es aquella persona de quince años o más, con residencia habitual en el hogar encuestado y que se encuentra en condiciones de responder el cuestionario por él, parte o todos los miembros del hogar. La Tabla 5, a continuación, presenta las preguntas del formulario con su respectiva sección y objetivo

³⁸ Revisar apartado "5.1 Técnica y métodos de recolección de los datos" en Documento metodológico Encuesta Nacional de Empleo (2021) en el enlace: https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodolog%C3%ADa-encuesta-nacional-de-empleo-ene-2020.pdf?sfvrsn=793380e9_10

³⁹ El formulario, se encuentra disponible en el apartado Anexos del presente documento

Tabla 5. Preguntas que conforman el cuestionario ESI

Sección	Pregunta	Pregunta	Objetivo
Sección K	k1a	El empleo actual de [Nombre de persona], ¿es el mismo que tenía en el mes de [Mes de referencia]?	Indagar si el empleo actual, declarado en el módulo B, es el mismo que tenía en el mes de referencia.
Sección K	k1b	¿Cuál fue el monto en pesos de su SUELDO NETO (SALARIO NETO)?	Indagar el sueldo neto de la actividad principal del mes de referencia, incluyendo ingresos variables regulares y fijos.
Sección K	k2a	En el mes de [mes de referencia], ¿recibió alguno de los siguientes ingresos variables) 1. Bonos por productividad 2. Bonos de otro tipo 3. Comisiones por venta 4. Pago de horas extraordinarias 5. Otros ingresos variables	Indaga la existencia de algún ingreso variable regular.
Sección K	k2b	¿Cuánto fue el monto en pesos (\$) o porcentaje (%) que representa de su sueldo neto?	Consigna el monto (\$) o porcentaje (%) del sueldo neto que representa dicho ingreso variable.
Sección K	k2c	¿Incluyó este ingreso en el sueldo neto declarado anteriormente?	Indaga si dicho ingreso variable se incluyó o no, en lo declarado en k1b.
Sección K	k2d	¿Con qué periodicidad recibe este ingreso variable?	Indaga la periodicidad de cada ingreso variable declarado.
Sección K	k3	En el mes de [mes de referencia], ¿le proporcionaron beneficios o regalías de ... 1...vivienda? 2...alimentación? 3...becas de estudio? 4...transporte y/o bencina? 5...propinas? 6...sala cuna o jardín infantil? 7...otros? (especifique)	Identificar los beneficios y regalías del empleo actual, recibidos en especie o en dinero.
Sección K	k3_monto	Indique el monto	Valorar los beneficios y regalías del empleo actual, recibidos en especie o en dinero.

Sección K	k4a	Este sueldo neto, ¿a cuántos días trabajados correspondió?	Capturar los días efectivamente trabajado en el empleo actual, durante el mes de referencia, correspondiente al sueldo (salario) neto reportado en k1.
Sección K	k4b	En promedio, ¿cuántas horas diarias trabajó durante esos días?	Capturar las horas diarias efectivamente trabajado en el empleo actual, durante el mes de referencia, correspondiente al sueldo (salario) neto reportado en k1.
Sección L	l1a	La actual actividad principal de <i>[Nombre de persona]</i> , ¿es la misma que tenía en el mes de <i>[Mes de referencia]</i> ?	Indagar si la actividad actual, declarada en el módulo B, es la misma que tenía en el mes de referencia.
Sección L	l1b	Durante dicho mes, ¿cuál fue el monto en pesos del ingreso neto (sueldos asignados, ganancias y/o retiros) que obtuvo en su actividad principal?	Indagar sobre el ingreso neto en el mes de referencia, que obtuvo o retiró de la empresa, negocio o actividad por cuenta propia.
Sección L	l1c	¿A cuánto tiempo de trabajo correspondió dicho ingreso?	Indagar sobre el número de meses trabajados, por los cuales el ocupado percibió dicho ingreso.
Sección L	l2	En dicho mes, ¿cotizó en alguno de los siguientes sistemas previsionales? 1_Salud (ISAPRE o FONASA) 2_Pensiones (AFP o IPS ex INP)	Conocer si la persona cotiza en el sistema previsional de salud (Isapre/Fonasa) o de pensiones (AFP/IPS). Por su actividad principal como empleador o trabajador por cuenta propia.
Sección L	l3	En el mes de <i>[Mes de referencia]</i> , ¿usó productos, bienes o servicios de su empresa negocio o actividad por cuenta propia, para su consumo personal o para el hogar?	Indagar y valorizar el autoconsumo de productos, bienes o servicios de su actividad principal como ocupados independientes.
Sección L	l4	Su empresa, negocio o actividad por cuenta propia, ¿es desarrollada en esta vivienda?	Indagar si el ocupado independiente realiza su actividad o negocio por cuenta propia en su vivienda. De ser afirmativa la respuesta, indagar el valor estimado por el arriendo del espacio utilizado.
Sección M	m1	Independiente de lo anterior, durante el mes de <i>[Mes de referencia]</i> , <i>[Nombre de persona]</i> ¿percibió ingreso/s por otra/s ocupación/es, distinta/s a la actual?	Indagar los ingresos recibidos: Durante el mes de referencia, correspondientes a un trabajo distinto al declarado en el módulo B. Por actividades secundarias, distintas a la

			actividad principal, realizadas durante el mes de referencia.
Sección M	m2	A pesar de no haber trabajado la semana pasada, durante el mes de [Mes de referencia], [Nombre de persona] ¿percibió ingresos por algún trabajo u ocupación?	Indagar los ingresos de las personas no ocupadas, que durante el mes de referencia recibieron un pago correspondiente a un empleo o actividad anterior.

5.3 Prueba y evaluación instrumento de captura de información

5.3.1 Prueba Piloto 2010

En agosto de 2010 se realizó un ejercicio piloto dentro de la Nueva Encuesta Nacional de Empleo (NENE). En el contexto de la actualización de la antigua Encuesta de Empleo, se aprovechó de actualizar el módulo de ingresos incluido en el último trimestre del año en la vigente NENE. En este marco, la prueba piloto tenía por objetivo probar en terreno el cuestionario elaborado por el equipo técnico a cargo de la Nueva Encuesta Suplementaria de Ingreso (NESI), basándose en líneas teóricas provenientes de:

- Las disposiciones y sugerencias del Grupo de Canberra (The Canberra Group, 2001)
- La Decimosexta Conferencia Internacional de Estadísticos del Trabajo, reporte sobre los ingresos del trabajo (OIT, 1998)
- El Sistema de Cuentas Nacionales de la ONU (GITCN, 1993)

A su vez, esto fue realizado con el fin de que el nuevo cuestionario permitiera medir ingreso mixto, reforzar el enfoque macroeconómico de los ingresos de los hogares y salvaguardar el enfoque microeconómico sobre distribución del ingreso y análisis de bienestar de los hogares y personas.

En términos operativos, el levantamiento tuvo una duración de dos semanas, en el mes de agosto de 2010 y fue realizado en la Región Metropolitana. El cuestionario NESI fue anexado al formulario central NENE, siendo aplicada inmediatamente después de esta.

Los resultados obtenidos producto del levantamiento piloto permitieron establecer las siguientes recomendaciones generales para la aplicación del instrumento:

- Existió una buena acogida de la NESI y se estableció una preferencia por incluir esta como módulo dentro de la NENE y no como una segunda visita.
- Permitted capturar información en hogares con ingresos altos.
- Se estableció que no existían grandes divergencias en la estimación de arriendo imputado en hogares de la misma manzana o calle.
- Se propuso trabajar en un sistema integrado de encuestas de ingresos, con acento en Encuestas de Presupuestos Familiares, por sobre el levantamiento coyuntural, en vistas de evitar sesgos en la entrega de información.

A nivel específico, se desarrollaron mejoras en el cuestionario de ingresos que determinaron su no comparabilidad con versiones de la ESI anteriores al 2010, pero que permitieron mayor precisión en la medición de los ingresos laborales y del hogar, según es descrito también en este documento.

5.3.2 Prueba Cualitativa 2021

En el contexto de un Compromiso de Desempeño Colectivo institucional (CDC) se desarrolló durante el 2020 una prueba de carácter cualitativo para revisar el instrumento de recolección vigente de la ESI, propiciando generar una discusión técnica y metodológica que permitiese la consecución de los objetivos de la encuesta en línea con estándares internacionales.

Se realizaron entrevistas cognitivas a través de un cuestionario diseñado por el equipo técnico ESI y revisado por expertos en la materia. Específicamente, se aplicó el cuestionario a 16 personas con perfiles socioeconómicos y laborales diversos, grabando y posteriormente, transcribiendo la conversación desarrollada por analistas de la encuesta. En este cuestionario se aplicó una versión reducida de la ESI y posteriormente, se consultó a cada usuario sus opiniones sobre cada una de ellas, vinculadas a las etapas del proceso cognitivo de acuerdo las cuatro etapas cognitivas de la respuesta identificadas en Tourangeau (1984) y Tourangeau & Rasinski (1998).

Los hallazgos más relevantes del estudio son que en el formulario evaluado (i) existen conceptos que las personas informantes no están entendiendo adecuadamente (ingreso y sueldo neto, la distinción entre líquido y bruto, descuentos no-legales, becas en especie, entre otros), (ii) no se presentan mayores problemas en la estimación de montos ni en incomodidad por la sensibilidad de preguntas y (iii) es posible aplicar medidas para que el cuestionario no resulte tan agotador de contestar, en especial con lo que estaba relacionado con el módulo de los otros ingresos.

Este último punto, que vinculó el agotamiento de las personas informantes con la realización del módulo de otros ingresos, va en sintonía con el cambio de objetivos que la encuesta realizó para su versión 2021, donde efectivamente dejó de aplicar la sección de “otros ingresos” para centrarse en caracterizar los ingresos del trabajo. Respecto al primer hallazgo presentado, si bien no se contempla en el corto plazo realizar modificaciones al cuestionario, es de utilidad tener presente estos problemas de comprensión cuando se evalúe introducir mejoras o cambios al formulario.

5.4 Método de capacitación para la recolección de datos

La ESI contempla una capacitación anual, generalmente un mes antes de iniciar la recolección oficial. Las personas recolectoras que asisten a esta capacitación corresponden al mismo personal que recolecta la ENE el resto del año. El proceso de capacitación en la ESI toma un enfoque de reforzamiento de contenidos y/o introducción de nuevos conceptos. La realización del material de apoyo supone la creación de instructivos sobre las modificaciones a las preguntas, si es que fuera el caso, y la construcción de estudios de casos que simulan situaciones.

El formato de las capacitaciones tradicionalmente se ha realizado de forma presencial, pero en las últimas versiones de la encuesta esta conducción se ha flexibilizado. En 2021, las capacitaciones estuvieron dirigidas a todas las personas encuestadoras, supervisoras y coordinadoras, realizándose vía videoconferencia en la plataforma Microsoft Teams.

Para el 2022 se habilitó una capacitación en la plataforma de la academia virtual INE en formato e-learning, donde el equipo de recolección pudo acceder a los contenidos pregrabados a su propio ritmo, en un lapso de dos semanas. El curso estuvo disponible a nivel nacional, teniendo una duración aproximada de dos horas, más una evaluación de alternativas destinada a verificar el

manejo de contenidos. Adicionalmente, se realizó una capacitación vía videoconferencia para personas recolectoras con poca experiencia y para personal de recolección que prefiriese esta modalidad.

La estructura temática de las capacitaciones se estructura a partir de presentaciones en PowerPoint, de la siguiente forma:

1. Introducción general a la Encuesta Suplementaria de Ingresos
2. Instrucciones de llenado del módulo (tanto para el formulario en papel como en el DMC)
3. Presentación de la sección K: corresponde a la explicación de la sección de ingresos de los trabajadores dependientes, junto a la revisión de los objetivos de cada pregunta.
4. Presentación de la sección L: corresponde a la explicación de la sección de ingresos de los trabajadores independientes, junto a la revisión de los objetivos de cada pregunta.
5. Presentación de la sección M: corresponde a la explicación de la sección de ingresos de otros trabajos, junto a la revisión de los objetivos de cada pregunta
6. Presentación de Alertas y justificaciones: su objetivo es explicar los tipos de alertas y ejemplos de correctas justificaciones en el DMC.

Independiente la modalidad en el que se realicen las capacitaciones de la ESI, el material dispuesto para los equipos de recolección siempre contempla presentaciones en Powerpoint (que incluye casos de estudio), manual del encuestador y manual de casos específicos.

5.5 Métodos de Supervisión

El Subdepartamento de Supervisión y Verificación de Estadísticas Sociales (SSVES) es el responsable de la supervisión y verificación de la Encuesta Nacional de Empleo (ENE) y Encuesta Suplementaria de Ingresos (ESI), por lo tanto, debe garantizar la correcta ejecución del trabajo de recolección, asegurando que el/la recolector/a de datos aplique de manera correcta el instrumento con base a los procedimientos metodológicos definidos; esto con el objetivo de confirmar que la información recolectada es fidedigna. A su vez, la supervisión tiene la finalidad de identificar problemáticas periódicas o esporádicas, desviaciones, errores u omisiones que afecten el cumplimiento de los objetivos del estudio.

El proceso de supervisión permite identificar errores no muestrales, que se pueden clasificar en errores de observación y de medición. Los primeros están relacionados con errores de cobertura y la falta de respuesta, y los segundos identifican problemas de diseño del cuestionario que afectan la captura de datos y el actuar de la persona que realiza la entrevista.

El modelo de supervisión general de la ESI y sus procedimientos se enmarca en la recolección de la ENE, contando la ESI con un formulario diferenciado para la supervisión directa y re-entrevista. El detalle de las fases operativas de la supervisión, impacto de la pandemia y aspectos de la supervisión que se vinculan exclusivamente con la ENE se pueden consultar en el Documento Metodológico oficial de la encuesta⁴⁰.

Por su parte, los tipos de supervisión en el producto se presentan a continuación.

5.5.1 Supervisión directa

El objetivo de esta supervisión es evaluar y observar in situ el cumplimiento de los protocolos operativos por parte del equipo de recolectores/as, enfocándose en la preparación de la salida a terreno, el desplazamiento a las viviendas, los intentos de contacto, el desarrollo de la entrevista y la revisión de preguntas específicas que conforman la ESI. Para la versión 2020 y 2021 de la encuesta, producto de la pandemia COVID-19, se suspendió la recolección presencial y se activó una recolección telefónica, imposibilitando realizar la observación directa al desempeño del recolector/a. Desde el año 2022 se reactiva la supervisión directa.

5.5.2 Supervisión indirecta re-entrevista

Apunta a detectar inconsistencias en la entrevista, así como también eventuales falseamientos de información, implicando volver a realizar la encuesta en la vivienda entrevistada por el/la recolector/a. La información es contrastada en gabinete por la SDT a fin de evaluar eventuales medidas correctivas, la principal innovación es que se pone el foco en el informante idóneo.

⁴⁰ Revisar apartado “5.1 Técnica y métodos de recolección de los datos” en Documento metodológico Encuesta Nacional de Empleo (2021) en el enlace: https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodolog%C3%ADa-encuesta-nacional-de-empleo-ene-2020.pdf?sfvrsn=793380e9_10

Producto de la pandemia, para la versión 2020 de la ESI esta supervisión se realizó de manera telefónica. Posteriormente, durante el 2021 y 2022, la supervisión se ha realizado de manera mixta, es decir, de manera telefónica y presencial según el caso.

Paralelo a la supervisión de terreno se ejecuta la supervisión en gabinete, cuyo objetivo es garantizar la calidad y consistencia de los datos levantados el terreno.

5.5.3 Supervisión de gabinete y reporte de consultas

Corresponde a la revisión y monitoreo de la calidad del dato una vez realizada la entrevista. A partir de 2020 la supervisión de gabinete consiste en implementar un procedimiento de recuperaciones de datos, consistente en recepcionar y gestionar semanalmente un archivo Excel desde la base una tabla con datos alojada en el servidor de la SDT con las inconsistencias y/o valores a recuperar.

Los cambios derivados del proyecto de modernización de la Encuesta Nacional de Empleo y la pandemia Covid-19, han impactado en los procesos de aseguramiento de la calidad del dato, conllevando a la implementación del proceso operativo denominado “*Recuperación y edición de datos*” dentro de la supervisión de gabinete, que deriva de la etapa posterior a la recolección y que tiene por objetivo asegurar la calidad del dato recolectado en la encuesta mediante el procesamiento de datos, cumpliendo con el registro de datos demandados de las validaciones y criterios de suficiencia definidos por el Subdepartamento de Estadísticas Continuas del trabajo.

Durante los meses de recolección de la ESI (trimestre octubre-diciembre), este proceso de recuperación y edición de datos es realizado conjuntamente con la ENE, sumando validaciones adicionales específicas para el módulo de ingresos. De esta forma, se constituye semanalmente el “*Reporte de consultas*”, el cual se comparte con el Subdepartamento de Supervisión y Verificación de Estadísticas Sociales (SSVES) para resolver posibles inconsistencias y recuperar determinada información en terreno de forma oportuna.

Para identificar información potencialmente errónea es necesario definir un marco consistente en términos lógicos, matemáticos y estadísticos al que deben ajustarse los datos. Esta verificación se realiza mediante la implementación de validaciones, que corresponden a un conjunto de controles lógicos que permiten detectar inconsistencias o datos potencialmente erróneos.

Las validaciones ESI incluidas en los reportes de consultas se clasifican en 5 categorías según su objetivo:

1. **Campo:** Evalúa que los valores ingresados por la persona recolectora en el dispositivo móvil estén dentro del rango definido para cada variable y que en cada una de las variables de la base de datos existan sólo valores permitidos para dicha variable. En otras palabras, verifica que no haya respuestas fuera del conjunto de respuestas posibles definidas para cada pregunta.
2. **Secuencia:** Evalúa que se cumplan las secuencias lógicas del cuestionario, es decir, verifica que no existan violaciones a las reglas de secuencia entre preguntas definidas en el instrumento de recolección. Esta validación debe capturar errores asociados a saltos incorrectamente implementados, respuestas vacías (que, por secuencia, no deberían estarlo) y respuestas con información (cuando por secuencia no deberían contenerla).

3. **Incidencia:** Evalúa si una observación impacta de manera significativa el promedio de ingresos en una subpoblación.
4. **Perfiles:** Evalúa si la información de ingreso y horas trabajadas son acorde al perfil laboral del individuo revisado.
5. **Otros Validadores:** Esta categoría agrupa distintas validaciones, principalmente las orientadas a analizar otros ingresos que no sea el principal (en el caso de dependientes) y evaluar la categoría de ocupación.

El proceso de recuperación de información es iterativo, es decir, una consulta puede aparecer en el reporte de consultas en más de una semana, o también la corrección de información puede llevar a que se genere una nueva consulta sobre la misma persona caracterizada en la encuesta. En estos casos cobra importancia la respuesta entregada por el SSVES, que es considerada en la revisión del equipo técnico ESI.

Para este proceso, se definen los estados del resultado de las consultas y se crean guías para su justificación. La Tabla 6 resume los estados de resultado de las consultas y los tipos de justificación que se pueden ingresar.

Tabla 6. Opciones de cada Estado y Justificación

Variable	Opciones seleccionables
Estado	Corregido
	Consultado
	Sin Corrección
Justificación	Se corrobora información
	Informante no sabe información
	Rechazo (justificar en observaciones)
	No se logra contacto con informante
	Otro (justificar en observaciones)

A su vez, cada estado tiene un objetivo y se ha determinado en qué situación específica se debe utilizar, señalando, además, que cada justificación es de uso exclusivo en cada estado. La definición de los estados también permite minimizar posibles sesgos asociados a la revisión manual de las consultas.

Tabla 7. Descripción de estados

Estado	Descripción de estados
Corregido	Cuando se recupera, corrige o mejora el dato (mejora de glosa, recuperación de No sabe/No responde, corrección de inconsistencia).
Consultado	Cuando hay contacto con la persona informante, se realiza la consulta y el informante corrobora los datos ya registrados, por lo tanto, no existe una corrección, mejora o cambio de secuencia. Existen dos situaciones en particular en las que se debe usar esta opción: Cuando se corrobora la información y cuando la persona informante no sabe la información.
Sin Corrección	Cuando no se logra contacto con el informante, cuando este rechaza entregar información u otra situación cuando no se logra realizar la consulta. Existen tres situaciones en particular cuando se debe usar esta opción: Rechazo, no logro contacto y otra situación inusual.

6 Metodología de procesamiento de la información

6.1 Procesamiento de la información

El procesamiento de la información se entiende como la operación que convierte los datos brutos, es decir, la información que no ha sido organizada y/o preparada, en información adecuada para el análisis y posterior difusión.

Es así como, en cada periodo de recolección se obtiene información bruta de las entrevistas, provenientes de los instrumentos que se utilizan en este proceso, como la hoja de ruta, el cuestionario sociodemográfico y el cuestionario central de la encuesta, que entran en el procesamiento. A continuación, se describirán las etapas de esta secuencia.

6.1.1 Integración de datos

Los datos de la encuesta recolectados en terreno son íntegramente almacenados en servidores INE internos. Con estas bases de datos brutas se generan la base ESI trimestral OND con las que se calculan los factores de expansión y que posteriormente, se usa para los cálculos de las variables principales y la difusión de los indicadores. Las bases internas para integrar son:

1. **Base de Hoja de Ruta:** contiene la información muestral de las viviendas que se recolectan en un periodo, y los registros de visitas e intentos de contacto a una vivienda.
2. **Base de Personas:** contiene toda la información recolectada en las entrevistas a los hogares, a partir del cuestionario sociodemográfico y el cuestionario central de la ENE.

Las bases internas integradas son validadas a través de programaciones en softwares estadísticos. Estos códigos identifican y resuelven posibles inconsistencias en secuencias del cuestionario y errores de campo. No se realizan actividades adicionales de integración de fuentes externas.

6.1.2 Clasificación y codificación

La ESI no realiza clasificaciones o codificaciones en su base de datos. Todas sus codificaciones provienen de la ENE, los cuales son construidos a partir de una metodología de clasificación y codificación automática de textos para el proceso de codificación de las descripciones de actividad económica y ocupación de la población ocupada entrevistada.

Para mayor detalle sobre la metodología de clasificación y codificación de la ENE, se recomienda revisar en la página web institucional el documento “Sistema de Clasificación y Codificación Automática en la Encuesta Nacional de Empleo”⁴¹ (INE, 2019).

⁴¹ INE. (Mayo de 2019). Instituto Nacional de estadísticas. Obtenido de Sistema de Clasificación y Codificación Automática en la Encuesta Nacional de Empleo: <https://www.ine.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion>

6.1.3 Revisión y validación

Dentro del procesamiento de la ENE y ESI es de suma importancia que exista consistencia de los datos publicados. El primer filtro que se utiliza para que el dato entregado sea válido es la existencia de validadores internos dentro del dispositivo móvil (Tablet) con el que se realiza la encuesta. Estos validadores corroboran la secuencia de preguntas que se realizan, la omisión de las preguntas, posibles inconsistencias con el valor entregado o advierte algún número extraño, y buscan evitar el registro de la información que no corresponda.

Pese a la presencia de validadores incorporados en el DMC, que fueron descritos anteriormente, se realiza luego una revisión de consistencia durante la etapa de procesamiento, verificando la consistencia de las preguntas y flujos de la encuesta a través de programaciones en el *software* estadístico R-Studio. Es así que, para que exista consistencia en todos los datos dentro de la base de datos se revisa que los saltos de preguntas y secuencias estén correctos, y también se verifica que los valores de respuesta estén dentro de los campos programados.

Una vez que la base de datos ESI ha quedado consistente en sus secuencias y valores dentro de los campos admitidos, se realiza una metodología de revisión de perfiles, con el objetivo de identificar casos que potencialmente presenten errores. En esta revisión se toma en cuenta el perfil socioeconómico completo de la persona y su hogar, identificando situaciones atípicas que puedan ser revisadas por analistas en profundidad.

Posteriormente, se realiza una depuración en gabinete con el objetivo de pulir errores conceptuales fácilmente identificables mediante una programación. Se corrigen de forma automática los casos de personas ocupadas catalogadas como servicio doméstico puertas adentro y que declaran percibir ingresos por regalías de alimentación y vivienda: se eliminan los montos o porcentajes declarados en las partidas de regalías en vivienda (pregunta K3_1) y alimentación (pregunta K3_2). Esto ocurre porque para el caso particular del servicio doméstico puertas adentro, estas regalías corresponden a elementos necesarios para la ejecución de su empleo, dado que habitan en su lugar de trabajo.

Finalmente, al terminar esta revisión y validación de datos en algunas ocasiones se realizan instructivos para apoyar la recolección en función de situaciones encontradas en este proceso, que tienen como objetivo la mejora continua de los datos de la ESI. Estos instructivos tienen diferentes finalidades, algunas de estas son: hallazgos importantes en la revisión y se debe retroalimentar; casos especial en cómo se debe pesquisar alguna pregunta; o debido a alguna contingencia que amerita una nueva instrucción, como por ejemplo el impacto que tuvo la pandemia del COVID-19 en la recolección y aplicación de la encuesta, entre otras razones.

6.1.4 Imputación

La falta de respuesta, tanto total⁴² como parcial, es una situación recurrente en las encuestas de hogares, especialmente en encuestas de ingresos como la ESI, al punto que puede impactar significativamente las estimaciones y su posterior inferencia sobre la población. Esto puede ocurrir debido a diversas causas, siendo las más relevantes la fatiga o cansancio del informante, desconocimiento de la información solicitada y rechazo a responder la encuesta de ingresos.

El proceso de imputación requiere la evaluación de las tasas de no respuesta en las preguntas de ingreso. En general, las tasas brutas de no respuesta que había registrado la antigua ESI cada año, y que registra actualmente la ESI, se encuentran dentro de rangos aceptablemente buenos para este tipo de encuestas (por debajo del 10%), lo cual se explica tanto por factores operativos como técnicos, del control de calidad del levantamiento de la información y los instrumentos de supervisión que se implementan para minimizar los errores. Las tasas de no respuesta parcial calculadas incorporan los casos con rechazo (*No responde*) y los casos de no respuesta por desconocimiento de la información (*No sabe*).

Las partidas sujetas a omisiones de información que se consideran para efectos de aplicar la imputación se dividen en dos grandes grupos:

- a) **Ingresos laborales:** las personas que se clasifican como ocupadas durante el mes de referencia, en una categoría distinta a la de familiar no remunerado, y que no responden las preguntas sobre los montos de los ingresos laborales. Estas partidas pueden ser *ingresos por Sueldos y Salarios* (ingresos provenientes de trabajos dependientes), o *ingresos Independientes* (ingresos provenientes de emprendimientos por cuenta propia o empleadores).
- b) **Horas efectivamente trabajadas:** las personas que se clasifican como ocupadas dependientes y que no responden la cantidad de horas o días efectivamente trabajados en el mes de referencia.

Para efectuar las imputaciones correspondientes a estas partidas es necesario utilizar un procedimiento que permita mitigar los posibles sesgos de estimación. Esto es, utilizar toda la información disponible que nos aporta la misma encuesta, ya sea los atributos laborales de cada receptor que declaró sus ingresos como las variables sociodemográficas de las personas.

Para la imputación de los Ingresos laborales, el método utilizado en la ESI es una variante del método **de medianas condicionadas** (Medina et al, 2007; Keifman et al, 1998), especialmente diseñado para la encuesta, el cual se aplica a nivel de las personas receptoras de ingresos y se basa en la búsqueda de un grupo de individuos con información de ingresos *-donantes-* que tengan características

⁴² Debido a que la encuesta de ingresos está asociada directamente a la encuesta de empleo, no existen problemas de falta de respuesta TOTAL, ya que esta se corrige a través de *un ajuste a los factores de expansión teóricos*, que implica eliminar las viviendas sin respuesta operativa, de modo que las unidades que permanecen en la muestra estimen –sin sesgo– el total poblacional.

similares a las del perceptor, imputándole la *mediana* de los ingresos del grupo que posee información⁴³.

Algunas de las ventajas de este procedimiento es que minimiza los posibles problemas de sesgo por errores de medición, ya que la búsqueda de donantes se realiza procurando buscar donantes con características similares al perceptor. Junto con esto, se incorporan variables geográficas que dan cuenta de las diferencias en los trabajos e ingresos en los distintos lugares del país.

La búsqueda de los donantes se basa en dos criterios fundamentales:

- i. *La utilización de niveles de exigencia*, que están elaborados en función de un set de variables categóricas o atributos laborales y sociodemográficos de las personas ocupadas. Se cuenta con una serie de niveles de exigencia que flexibilizan las variables utilizadas de manera gradual.
- ii. *La utilización de áreas geográficas de estimación*, que van desde el nivel más desagregado posible –conglomerado–, hasta llegar al nivel más agregado, es decir, al nivel país.

Como resultado, el procedimiento combina tanto variables cualitativas (niveles de exigencia basado en características de los preceptores de ingresos) como geográficas (áreas de estimación). La aplicación se realiza de forma vertical, es decir, recorriendo todas las áreas de estimación antes de pasar a un siguiente nivel de exigencia, con el objetivo de privilegiar la búsqueda de donantes con las mismas características del receptor en desmedro de donantes con menos características en común.

Respecto a la metodología de imputación de las horas efectivamente trabajadas, el método de imputación corresponde al de imputación *por sustitución*. De acuerdo con éste, se utilizan datos provenientes de la misma observación para realizar la imputación. En el caso de las horas trabajadas se utilizan las horas habitualmente trabajadas durante la semana de referencia, respondidas en el módulo de empleo. La principal ventaja de esta metodología es que hace uso de la información de las mismas unidades de observación, por lo que la correspondencia entre el perfil laboral y sociodemográfico del “donante” y del perceptor es completa, y por ello es habitualmente recomendada por la literatura como una primera opción de imputación. Finalmente, cuando no hay información disponible en el módulo de empleo para completar las horas efectivamente trabajadas en la ESI, se acude al mismo método de *medianas condicionadas* empleado para recuperar los ingresos de los trabajadores dependientes.

⁴³ A diferencia de otras aplicaciones, aquí se utiliza la “mediana” en vez de la media, ya que es un mejor estimador en presencia de distribuciones asimétricas, debido a que no está afectado por los extremos de la distribución (colas).

6.1.5 Derivar nuevas variables y unidades⁴⁴

En esta subsección se explican las principales variables creadas que sirven de insumo para el análisis de la ESI. Una variable analítica es aquella que se genera en función de otras variables y que se incluye en el análisis estadístico.

Las variables analíticas que se producen en la ESI son principalmente sumatorias y agregaciones de partidas de ingresos del cuestionario. El detalle de estas variables se presenta en el Tabla 8.

Tabla 8. Variables analíticas ESI del tipo agregación de ingresos

Variable	Nombre	Construcción
ss_t	Ingresos por sueldos y salarios netos	$ss_t = d1_monto$
svar_t	Ingresos por sueldos y salarios variables no incluidos en D1	$svar_t = \sum_{i=1}^5 d2_i_monto + \sum_{i=1}^5 \frac{d2_i_porcentaje * d1_monto}{100}$
reg_t	Total ingresos por regalías (beneficios entregados por el empleador)	$reg_t = \sum_{i=1}^7 d3_i_monto + \sum_{i=1}^7 \frac{d3_i_porcentaje * d1_monto}{100}$
gan_t	Ingresos por ganancias del negocio independiente	$gan_t = \begin{cases} d5_monto \\ \text{ó} \\ d5_promedio \end{cases}$
autoconsumo_t	Ingresos por autoconsumo	$autoconsumo_t = \begin{cases} d7_monto \\ \text{ó} \\ d7_promedio \end{cases}$
ahorroimp_t	Gasto imputado por arriendo de local	$ahorroimp_t = d8_monto$

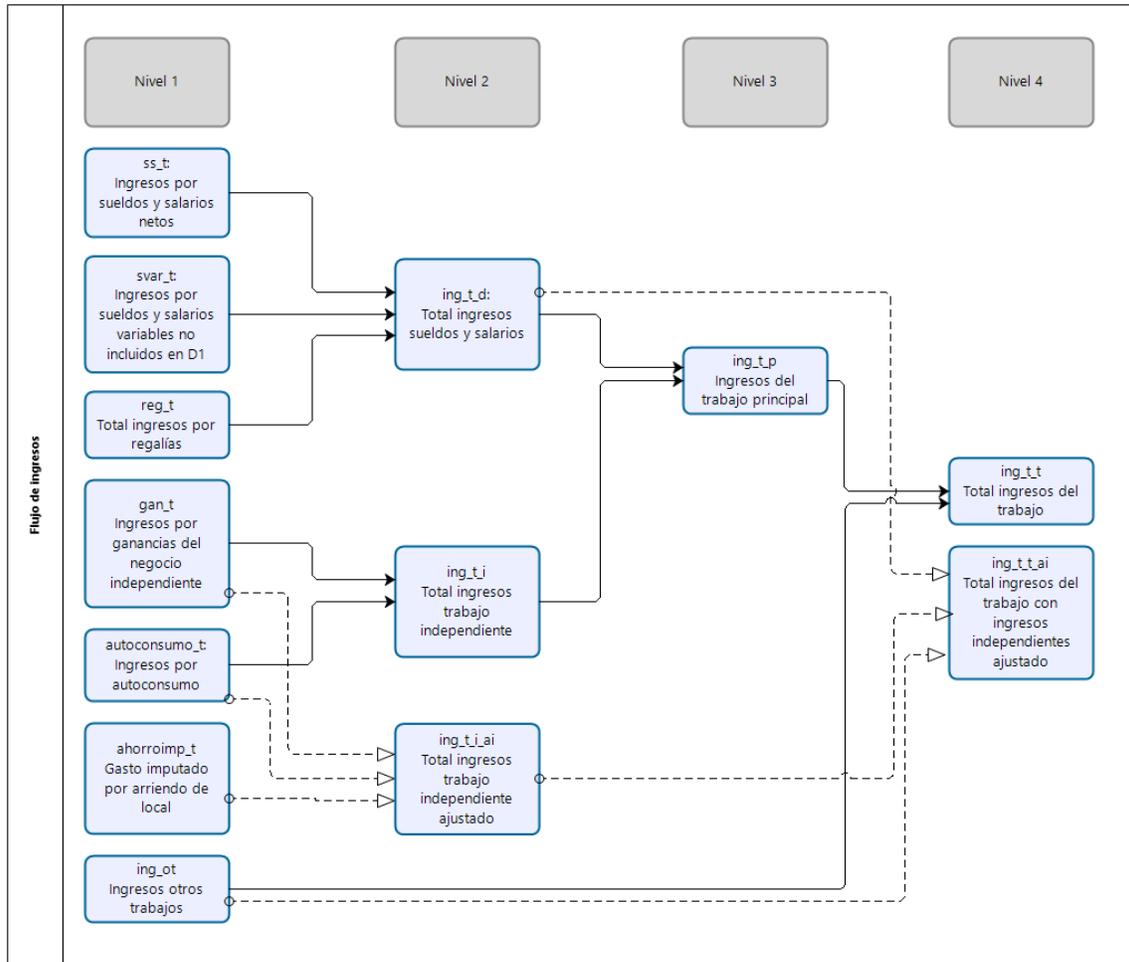
⁴⁴ La base de datos publicada en el microsítio oficial de la ESI utiliza la nomenclatura antigua del cuestionario. La equivalencia de variables se puede consultar en el presente documento <https://www.ine.gob.cl/docs/default-source/encuesta-suplementaria-de-ingresos/formularios/cuestionarios/correspondencia-cuestionario-esi-2021.pdf>

ing_ot	Ingresos otros trabajos	$ing_ot = \begin{cases} d9_monto \text{ ó } d9_promedio \\ \text{ó} \\ d9_montob \text{ ó } d9_promediob \end{cases}$
ing_t_d	Total ingresos sueldos y salarios	$ing_t_d = ss_t + svar_t + reg_t$
ing_t_i	Total ingresos trabajo independiente	$ing_t_i = gan_t + autoconsumo_t$
ing_t_i_ai	ingresos del trabajo independiente ajustado	$ing_t_i_ai = \begin{cases} gan_t + \\ autoconsumo_t - \\ ahorroimp_t \end{cases}$
ing_t_p	Ingresos del trabajo principal	$ing_t_p = \begin{cases} ing_t_d \\ \text{ó} \\ ing_t_i \end{cases}$
ing_t_t	Total ingresos del trabajo	$ing_t_t = \begin{cases} ing_t_d + ing_ot \\ \text{ó} \\ ing_t_i + ing_ot \end{cases}$
ing_t_t_ai	ingresos del con trabajo independientes ajustados	$ing_t_t_ai = \begin{cases} ing_t_d + ing_ot \\ \text{ó} \\ ing_t_i_ai + ing_ot \end{cases}$

Fuente: Instituto Nacional de Estadísticas

La agregación de estas mismas partidas de ingresos se sintetiza en la siguiente figura:

Figura 3. Diagrama de construcción de variables analíticas, por nivel



Fuente: Instituto Nacional de Estadísticas

Adicionalmente, se generan dos variables analíticas categóricas: “Tramos de ingresos mínimos netos” y “Binario ocupado de referencia tabulados de personas”.

Tabla 9. Variables analíticas ESI del tipo categóricas

Variable	Nombre	Valor	Etiqueta
tim	Tramos de ingresos mínimos netos ⁴⁵ (SM) según ingresos del trabajo principal	1	0 - 1SM
		2	1SM – 2SM
		3	2SM -4SM
		4	4SM – 6SM
		5	6SM – 8SM
		6	8SM – 10SM
		7	10SM y más
ocup_ref	Binario ocupados de referencia tabulados de personas	0	Ocupados con menos de 1 mes en el empleo actual
		1	Ocupados con más de 1 mes en el empleo actual

La variable “ocup_ref” es de especial importancia para las estimaciones que se realizan empleando la ESI, ya que es un filtro que permite identificar a aquellas personas ocupadas con al menos 1 mes de antigüedad en el empleo actual. Para todas las estimaciones oficiales de la encuesta se exige que la subpoblación cumpla con la variable “ocup_ref” igual a 1.

⁴⁵ Ingreso mínimo neto vigente al momento de recolectar la ESI en el trimestre octubre-diciembre.

7 Metodología de análisis

En esta sección se describirán los procesos que se realizan en la Encuesta Suplementaria de Ingresos (ESI) en lo referido al análisis de datos y los métodos utilizados para explicar e interpretar los resultados para su difusión y divulgación.

7.1 Procedimiento para la anonimización de los datos

Al formar la ESI parte de la ENE, su proceso de anonimización es homólogo. En el caso de la ENE este proceso viene dado por dos elementos centrales.

El primer elemento que se realiza para la indeterminación de las viviendas y personas es dentro de la base de datos, ya que no se considera ninguna variable que permita identificar a la persona entrevistada o a la vivienda seleccionada. Es decir, no se incluyen variables como el nombre, teléfono o RUN de la persona, como tampoco se incluyen variables que permitan identificar a las viviendas desde sus características muestrales, geográficas o de manera cruzada con otros productos estadísticos. En definitiva, la estructura de la base de datos permite indeterminar y anonimizar a los informantes y a las unidades muestrales (viviendas).

Además, para minimizar aún más el riesgo de determinación de las unidades entrevistadas, se enmascaran dos identificadores, estos son: la variable **idrph**, que es el identificador único de personas, y la variable **conglomerado**, que es la identificadora del conglomerado consolidado del marco nuevo/marco antiguo. Se genera un conglomerado y un idrph ficticio que no permita realizar ninguna conexión con la vivienda seleccionada.

En el próximo capítulo se profundiza sobre la accesibilidad a los datos y el marco legal que regula la protección y la difusión de la información por parte del INE.

7.2 Cálculo de estimaciones y validación de resultados

Posterior a la validación y anonimización de la base de datos, se realizan los cálculos de los tabulados que formarán parte de los productos de difusión ESI. Estos se ejecutan en una rutina automatizada en el software estadístico *R Studio* para la generación e interpretación de resultados. Además, se producen los gráficos y tablas correspondientes a la publicación. Los productos de difusión se abordarán en el apartado “8.1 Productos estadísticos a publicar”.

Una vez realizadas las estimaciones, análisis, gráficos, tablas para el boletín, se lleva a cabo una validación de los resultados mediante un contraste espejo de los cálculos generados, a partir de la base de datos consolidada y con el uso de un *software* estadístico distinto al que genera el boletín (*STATA*) y en caso de encontrar diferencias o tener comentarios que puedan complementar y mejorar el análisis, estos se aplican y se implementan las mejoras para obtener una versión final completamente revisada del producto a difundir.

Además de lo anterior, para cada estimación se aplica el estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares (INE, febrero de 2020). Esta metodología consiste en aplicar un flujo que incluye el número de observaciones o tamaño muestral, el coeficiente de

variación, los grados de libertad, el error estándar, los intervalos de confianza de la estimación. A partir de esta información, se determina si son publicables los tabulados asociados a los productos de difusión, y se incorporan las alertas respectivas cuando existan estimaciones con problemas de fiabilidad.

7.3 Finalización de resultados

Esta etapa tiene por objetivo realizar las revisiones pertinentes a los contenidos de los cálculos asociados a los productos de difusión en su última versión. Estas revisiones que tienen como fin garantizar la calidad de la información entregada a las personas usuarias comprende la revisión, implementación de mejoras y, finalmente, el visto bueno de las jefaturas correspondientes a nivel institucional para entregar el producto final, para así iniciar el proceso difusión de los resultados.

8 Alcance de la presentación de resultados

8.1 Productos estadísticos a publicar

Anualmente, con el fin de satisfacer las necesidades de información de los usuarios, se pone a disposición del público en la página web de la ESI los principales resultados de cada versión de la encuesta, así como también las bases de datos a partir de 2010 y su documentación asociada. A continuación, se detallan los productos de difusión de la encuesta:

- **Bases de datos:** se encuentra disponible la base personas para todos los años a partir de 2010, mientras que las bases de datos a nivel de hogar (Base de Hogares Con Becas de Estudio y Base de Hogares Sin Becas de Estudio⁴⁶) se encuentran hasta el 2020.
- **Síntesis de Resultados – Nacional:** se encuentra disponible de forma anual a partir de 2010. Corresponde a la presentación de los principales resultados obtenidos de la encuesta aplicada en un año en particular a nivel nacional. En ella se dan a conocer los principales indicadores de la encuesta (ingreso medio, ingreso mediano, ingreso por hora, masa de ingresos, brecha de género, entre otros), según distintas desagregaciones de relevancia.
- **Síntesis de Resultados – Regional:** corresponde a la presentación de los principales resultados obtenidos de la encuesta aplicada en un año en particular a nivel regional. En ella se dan a conocer los principales indicadores de la encuesta (ingreso medio, ingreso mediano, masa de ingresos, brecha de género, entre otros), según distintas desagregaciones de relevancia regional.
- **Nota complementaria:** este documento es publicado a partir de la ESI 2020 y tiene como objetivo realizar análisis complementarios que no se encuentren abordados en la síntesis de resultados, de acuerdo a las necesidades de información presentadas por el contexto nacional y el mercado laboral.
- **Comunicado de prensa – Nacional:** corresponde a un resumen de los principales resultados de la encuesta a nivel nacional. Se publican en el sitio de prensa del INE⁴⁷ a partir de 2014.

⁴⁶ La base de hogares con Becas de Estudio educación se construye a partir de las variables que incorporan en su cálculo las transferencias por concepto de educación (ya sean monetaria y/o no monetaria) de la base de personas, y posee una decilización construida en base al ingreso monetario del hogar contemplando estas becas. La base de hogares sin Becas de Estudio, contrariamente, excluye estas transferencias por educación y, en consecuencia, no las contempla dentro de su decilización. Para más detalles, revisar el Manual y Guía de variables de alguna de estas bases en el micrositio de la ESI <https://www.ine.gob.cl/estadisticas/sociales/ingresos-y-gastos/encuesta-suplementaria-de-ingresos>

⁴⁷ Presentes en [https://www.ine.gob.cl/prensa/detalle-prensa/2021/08/24/ingreso-laboral-promedio-mensual-en-chile-fue-de-\\$635.134-en-2020](https://www.ine.gob.cl/prensa/detalle-prensa/2021/08/24/ingreso-laboral-promedio-mensual-en-chile-fue-de-$635.134-en-2020)

- **Comunicado de prensa – Regional:** se encuentran disponibles en la sección “Comunicados de prensa” de cada región⁴⁸. Corresponden a un resumen de los principales resultados de la ESI para una región en particular y un resumen de los resultados a nivel nacional.
- **Cuadros estadísticos:** se encuentran disponibles desde 2010 al 2017. Corresponden a tabulados con la estimación de indicadores principales de la encuesta, así como tabulados con la estimación del coeficiente de variación. Desde la versión 2018 de la ESI, se reemplaza la publicación de cuadros estadísticos por indicadores de ingreso incluidos en la plataforma de INE.Stat, cuyo funcionamiento se detallará más adelante.
- **Enfoques, infografías y otros estudios:** elaborados a partir de una investigación en particular relacionada con los ingresos laborales.

Además de los productos indicados anteriormente, junto a cada publicación oficial se actualiza la información de indicadores en INE.Stat, con información desde 2010. En esta plataforma web se encuentran indicadores, presentando el ingreso medio, ingreso mediano y brecha de género, entre otros.

INE.Stat permite a las personas usuarias acceder a tabulados interactivos y modificables según sus preferencias, sin la necesidad de emplear softwares externos. La plataforma cuenta con la función de descarga de los tabulados en formato XLS y CSV, entregando metadatos de la encuesta (fuentes, características de los datos, población y alcance, conceptos claves y clasificaciones, y comentarios de calidad) y orienta a las personas usuarias a la documentación descargable en la página web de la ESI.

8.2 Difusión de los resultados

8.2.1 Microdatos

Como se mencionó anteriormente, la ESI difunde parte de sus datos a través de la plataforma INE.Stat, que es un repositorio de datos estadísticos flexible, interactivo y de uso gratuito, administrado por el Instituto Nacional de Estadísticas (INE). INE.Stat le brinda a la persona usuaria la posibilidad de buscar datos de indicadores estadísticos por fuente de información (por ejemplo, nombre de la encuesta), tema de interés, palabra clave o usando el árbol temático de la plataforma. La flexibilidad de la plataforma le permite a la persona usuaria personalizar la visualización de sus consultas y modificar la disposición de las filas, columnas y filtros en las tablas.

Junto con los cuadros estadísticos, se encuentran disponibles en la plataforma los metadatos respectivos de cada estimación. Estos corresponden a fichas metodológicas que describen a los

⁴⁸ Por ejemplo, en la sección de “Comunicados de prensa” de la Región de Valparaíso <https://regiones.ine.gob.cl/valparaiso/prensa/ingreso-laboral-promedio-en-la-regi%C3%B3n-de-valpara%C3%ADso-lleg%C3%B3-a-584.124-pesos-mensual-en-2020>

indicadores, información relevante acerca de las variables o series, y notas al pie respecto de datos específicos. Adicionalmente, la persona usuaria puede exportar cada consulta de los tabulados en formatos XLS y CSV.

Por otra parte, las bases de datos de la ESI se encuentran publicadas en el portal web del INE⁴⁹, desde el periodo octubre-diciembre de 2010, correspondiente a la serie vigente de la encuesta. Las bases de datos están disponibles en tres formatos para utilizar en los softwares: SPSS (.sav), Stata (.dta) y RStudio (.RData), además del formato comma-separated values (.csv), para trabajar con diferentes softwares. Adicionalmente, existe un libro de códigos actualizado para apoyar el uso de las bases de datos.

Es importante mencionar, que las bases de datos de la ENE-ESI no permiten identificar ni determinar geográfica ni individualmente a ningún informante, en sintonía con las buenas prácticas institucionales derivadas del Secreto estadístico, el cual establece que quien entrega información estadística debe resguardar el anonimato de las personas que entregaron dicha información. En el siguiente apartado se indagará más sobre la anonimización e indeterminación de la información.

8.2.2 Metadatos

Los metadatos engloban atributos de los datos procesados para un producto estadístico cualquiera, otorgando una caracterización de ellos. Esta caracterización abarcará desde el contenido, calidad, disponibilidad, metodología y otras particularidades en cuestión. Estos se encuentran disponibles en el micrositio de la encuesta.

Por lo anterior, para que la comprensión de la ESI sea más sencilla es que se listan los metadatos que la encuesta ofrece a las personas usuarias:

- **Manual y guía de variables:** se encuentran disponibles los documentos asociados a las bases de datos. Estos documentos sirven de guía para los usuarios, facilitando su uso y comprensión, entregando información acabada acerca de las variables contenidas en las bases correspondientes.
- **Cuestionarios:** están disponibles en su versión completa desde el 2010 hasta el año vigente de la encuesta.
- **Documento Metodológico:** este documento se actualiza todos los años, de manera acumulativa, junto a la publicación de la ESI vigente. Tiene como propósito describir los procesos y metodologías utilizadas en la producción de la ESI, otorgando un conocimiento global de la encuesta y sus productos estadísticos. En el documento metodológico se detalla información sobre el contexto de la encuesta, el marco de referencia, el diseño estadístico, la metodología de recolección y procesamiento de datos, información referente a la presentación de resultados y conceptos claves del producto.

⁴⁹ Se puede acceder al micrositio de la ESI en el siguiente enlace:

<https://www.ine.gob.cl/estadisticas/sociales/ingresos-y-gastos/encuesta-suplementaria-de-ingresos>

- **Documento metodológico de factores de expansión:** este documento tiene como objetivo describir la metodología utilizada para generar el factor de expansión.
- **Reporte de calidad:** disponible a partir de 2021. Describe las acciones que se desarrollan dentro del producto referidas a los cinco principios del pilar “Producción Estadística” del Código de buenas prácticas para la Estadísticas Chilenas (INE, 2015).
- **Ficha Técnica:** se pone a disposición la ficha técnica de la Encuesta Suplementaria de Ingresos del año vigente. Esta resume las principales características del producto y se puede encontrar información sobre objetivos, unidad de análisis, descripción del producto, fuente de información, cobertura geográfica, periodo de publicación, población de referencia, procesos referidos a la calidad de los datos, metodologías de imputación y declinación, nomenclatura utilizada, las principales variables del estudio, periodo de levantamiento y medios de publicación del producto.
- **Ficha Nesstar**

8.2.3 Otros

Junto con todos los documentos y productos de difusión, el INE pone a disposición de sus usuarios/as el portal del Sistema Integral de Información y Atención Ciudadana (SIAC), disponible en <https://atencionciudadana.ine.cl/>, para canalizar cualquier requerimiento adicional de información, el que será evaluado de acuerdo con dos criterios:

- Disponibilidad de la información solicitada.
- Protección de la confidencialidad de los datos, de acuerdo al secreto estadístico establecido en la Ley 17.374.

8.2.4 Consideraciones sobre el uso de la información

En relación con la información que se pone a disposición de las personas usuarias de los datos, es importante tener algunas consideraciones.

Primero, **la actual serie de datos de la ESI comienza en el 2010**, y se encuentra a disposición del público en la página institucional de INE para su descarga directa. Para información anterior a este periodo, se deben solicitar las bases de datos a través del Sistema Integral de Información y Atención Ciudadana (SIAC), disponibles desde 1990 en adelante.

Segundo, **los datos de la serie vigente no son comparables con series anteriores de la ESI**, debido a la incorporación de importantes cambios conceptuales y metodológicos a partir de 2010.

Finalmente, si bien se encuentra disponible información sobre la comuna de residencia habitual o de trabajo de las personas en las bases de datos, tanto **la ESI como la ENE no poseen representatividad comunal**. Su diseño garantiza errores de muestreo aceptables sólo para los dominios nacional, nacional urbano, nacional rural, regional urbano (para todas las regiones) y regional rural (sólo para regiones Metropolitana, O’Higgins, Maule, Ñuble, La Araucanía, Los Ríos y

Los Lagos) (INE, 2020a). El INE no se puede hacer cargo de estimaciones fuera de estos dominios oficiales de estudio, siendo responsabilidad de las personas investigadoras realizar las advertencias necesarias sobre esta situación.

8.3 Accesibilidad a los datos y Marco legal

De acuerdo con lineamientos institucionales, la información producida por la ENE y ESI se encuentra supeditada a los compromisos de confidencialidad y reserva de datos protegidos por secreto estadístico. Estos lineamientos pueden ser resumidos en el siguiente marco legal:

- Compromiso de confidencialidad y reserva de datos protegidos por el secreto estadístico e información sensible y/o sujeta a la Ley 19.628 sobre la protección de datos personales.
- Ley 19.628 de protección de vida privada y datos.
- Ley 20.285 regula el principio de transparencia de la función pública y el derecho de acceso a la información de los órganos de la Administración del Estado.
- Artículo 29° de la Ley 17.374- El Instituto Nacional de Estadísticas, los organismos fiscales, semifiscales y Empresas del Estado, y cada uno de sus respectivos funcionarios, no podrán divulgar los hechos que se refieren a personas o entidades determinadas de que hayan tomado conocimiento en el desempeño de sus actividades.

El estricto mantenimiento de estas reservas constituye el "Secreto Estadístico". Su infracción por cualquier persona sujeta a esta obligación, hará incurrir en el delito previsto y penado por el artículo 247, del Código Penal, debiendo en todo caso aplicarse pena corporal.

- Artículo 30° de la Ley 17.374- Los datos estadísticos no podrán ser publicados o difundidos con referencia expresa a las personas o entidades a quienes directa o indirectamente se refieran, si mediare prohibición del o los afectados." (Extractos de la Ley Orgánica 17.374 de creación del Instituto Nacional de Estadísticas, publicada el 10-12-1970.)
- Ley N° 17.336 sobre Propiedad Intelectual, en su inciso primero del Artículo 88: *"El Estado, los Municipios, las Corporaciones oficiales, las Instituciones semifiscales o autónomas y las demás personas jurídicas estatales serán titulares del derecho de autor respecto de las obras producidas por sus funcionarios en el desempeño de sus cargos"*.

9 Glosario

A continuación, se presentarán los conceptos claves que permiten entender la metodología expuesta anteriormente.⁵⁰

- Asalariado/a: es la persona que trabaja para un empleador, jefe o patrón y percibe una remuneración en forma de sueldo, salario, comisión, propinas, pagos a destajo o en especie.
- Asalariado/a privado: esta categoría incluye a todos los asalariados que no trabajan en instituciones o empresas del Estado, independiente del sector económico donde se desempeñan.
- Asalariado/a público: esta categoría incluye a todos los asalariados que trabajan en instituciones o empresas del Estado, ya sea dentro de la administración pública como en otros sectores económicos (salud, enseñanza, etc.).
- Brecha de género: es la diferencia que se manifiesta entre la situación, condición y/o posición de las mujeres y la de los hombres, es decir, la distancia que hace falta recorrer para alcanzar la igualdad en una situación determinada.
- Brecha de género en el ingreso: es la diferencia porcentual que existe en el ingreso medio, mediano o por hora, de las mujeres con respecto al de los hombres.
- Población desocupada: todas las personas en edad de trabajar que no estaban ocupadas durante la semana de referencia, que habían llevado a cabo actividades de búsqueda de un puesto de trabajo durante las últimas cuatro semanas (incluyendo la de referencia) y que estaban disponibles para trabajar en las próximas dos semanas (posteriores a la de referencia).
- Fuerza de trabajo: personas en edad de trabajar que durante la semana de referencia cumplen los requisitos para ser incluidas en la categoría de las personas ocupadas o desocupadas.
- Fuerza de trabajo potencial (Inactivos potencialmente activos): personas en edad de trabajar que, durante la semana de referencia, no estaban ocupadas ni desocupadas y que habían llevado a cabo «actividades de búsqueda» pero no estaban «actualmente disponibles para trabajar»; o que no llevaron a cabo «actividades de búsqueda», pero que si estaban «actualmente disponibles para trabajar».
- Fuerza de trabajo ampliada: se define como la suma de la fuerza de trabajo más los iniciadores disponibles y la fuerza de trabajo potencial (inactivos potencialmente activos).

⁵⁰ Este listado de términos es extraído de: Instituto Nacional de Estadísticas. (enero de 2019). Glosario ENE. Obtenido de Instituto Nacional de Estadísticas: https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/glosario-ene.pdf?sfvrsn=e8d50659_5

- Horas efectivas: cuantifican el tiempo dedicado por las personas al desempeño de actividades económicas que contribuyen a la producción de bienes y servicios durante un período de referencia corto (la semana de referencia de la encuesta).
- Horas habituales: corresponden a las horas comúnmente trabajadas en un período de referencia determinado (promedio de los últimos tres meses).
- Iniciadores/as disponibles: todas las personas fuera de la fuerza de trabajo que esgrimen no haber buscado trabajo en las últimas cuatro semanas debido a que iniciarán pronto una actividad laboral y que, al mismo tiempo, declaran disponibilidad. Este grupo poblacional se incluye dentro de los Inactivos, pero para ciertos efectos (de armonización estadística con la OCDE), se suma a la población desocupada para calcular una tasa de desocupación con iniciadores disponibles.
- Ingreso del trabajo principal: total de ingresos provenientes de la actividad principal actual, que se recibieron en el mes de referencia de la encuesta.
- Ingreso mediano: corresponde al ingreso de la persona situada en la posición central del conjunto de la población, una vez ordenados los ingresos de menor a mayor. Es decir, es el monto que divide a la población en dos partes iguales: una mitad recibe un ingreso igual o menor que la mediana, mientras que la mitad restante recibe un ingreso mayor.
- Ingreso medio o promedio: suma de todos los ingresos de la población dividido por la cantidad total de individuos que perciben ese ingreso.
- Masa de ingresos: corresponde a la suma de los ingresos obtenidos por la cantidad total de individuos que perciben ese ingreso.
- Población Ocupada: todas las personas en edad de trabajar, que durante la semana de referencia dedicaron al menos una hora a alguna actividad para producir bienes o servicios a cambio de una remuneración o beneficios.
- Ocupación informal: comprende a todos los trabajadores dependientes que carecen de acceso al sistema de seguridad social (pensión y salud) por concepto de su vínculo laboral. Para el caso de los trabajadores independientes, se considera que poseen una ocupación informal si la empresa, negocio o actividad que desarrollan pertenece al sector informal. Familiar no remunerado son todos clasificados como informales.
- Población en edad de trabajar (PET): población actualmente residente en el país de 15 años y más.
- Población fuera de la fuerza de trabajo o inactiva: todas las personas de la población en edad de trabajar que no son clasificadas como ocupadas ni desocupadas.
- Sector informal: conjunto de unidades económicas que carecen de registro en el Servicio de Impuestos Internos (SII) y no llevan una contabilidad a través de la cual puedan separar los gastos del negocio de los del hogar.

- Tasa de desocupación: número de personas desocupadas expresado como porcentaje de la fuerza de trabajo (no incluye a los iniciadores disponibles).
- Tasa de desocupación ajustada estacionalmente: es la tasa de desocupación que elimina los efectos de los factores exógenos estacionales de naturaleza no económica que influyen en su comportamiento coyuntural. Para el ajuste estacional de las series de empleo, se utiliza el programa X13 ARIMA SEATS. Se aplica este método independientemente en la serie de personas desocupadas y de ocupadas, para luego obtener la tasa de desocupación ajustada sin estacionalidad (método indirecto). De este modo es posible comparar de forma directa los valores de la serie entre distintos momentos del tiempo.
- Tasa de desocupación con iniciadores disponibles (SU1): número de personas desocupadas más personas iniciadores disponibles expresados como porcentaje de la fuerza de trabajo más iniciadores disponibles.
- Tasa combinada de desocupación y tiempo parcial involuntario (SU2): número de personas desocupadas más personas iniciadores disponibles y personas ocupadas a tiempo parcial involuntario expresado como porcentaje de la fuerza de trabajo más iniciadores disponibles.
- Tasa combinada de desocupación y fuerza de trabajo potencial (SU3): número de personas desocupadas más personas iniciadoras disponibles e inactivas potencialmente activas expresado como porcentaje de la fuerza de trabajo ampliada.
- Tasa de ocupación: número de personas ocupadas como porcentaje de la población en edad de trabajar.
- Tasa de ocupación en el sector informal (TOSI): número de personas ocupadas que trabajan en unidades económicas informales expresado como porcentaje del total de personas ocupadas.
- Tasa de ocupación informal (TOI): número de personas con ocupaciones informales expresado como porcentaje del total de personas ocupadas.
- Tasa de participación: número de personas en la fuerza de trabajo expresado como porcentaje de la población en edad de trabajar.
- Tasa de presión laboral: número de personas desocupadas, más personas ocupadas que buscan empleo, más iniciadores disponibles, expresado como porcentaje de la fuerza de trabajo más los iniciadores disponibles.

10 Bibliografía

- Alvarado, M., y Pizarro, M. (2019). *Fundamentos de la nueva metodología de calibración de los factores de expansión de la Encuesta Nacional de Empleo*. Documentos de Trabajo, Serie Documentos Metodológicos N°12, diciembre 2019. Instituto Nacional de Estadísticas, Santiago de Chile. https://www.ine.cl/docs/default-source/documentos-de-trabajo/documento-de-trabajo-fundamentos-de-la-nueva-calibraci%C3%B3n-de-los-factores-de-expansi%C3%B3n-en-la-ene.pdf?sfvrsn=3de3a0e1_4.
- Banco Central de Chile. (2005). *Indicadores económicos y sociales de Chile: 1969-2000*. Santiago. Recuperado de https://www.bcentral.cl/documents/33528/133439/bcch_archivo_098139_es.pdf
- Groves, R. (2004). *Survey Errors and Survey Cost*. Nueva York: John Wiley & Sons.
- Grupo Intersecretarial de Trabajo sobre Cuentas Nacionales (GITCN) (1993). Sistema de Cuentas Nacionales, 1993. Recuperado de: <http://comuna.cat/-/sctasnac93.pdf>
- Grupo Intersecretarial de Trabajo sobre Cuentas Nacionales (GITCN) (2009). Sistema de Cuentas Nacionales, 2008. Recuperado de: <https://unstats.un.org/unsd/nationalaccount/docs/SNA2008Spanish.pdf>
- Gutiérrez, A. (2016). *Estrategias de muestreo, diseño de encuestas y estimación de parámetros*. Ediciones de la U, Bogotá
- Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). *Applied Survey Data Analysis*. Obtenido de Boca Ratón, FL: Taylor and Francis Group, LLC.
- Horvitz, D. G., & Thompson, D. J. (1952). *A generalization of sampling without replacement from a finite universe*. Journal of the American statistical Association, 47(260), 663-685.
- Hyndman & Fan (1996) "Sample Quantiles in Statistical Packages" <https://www.istor.org/stable/2684934>
- INE. (2015). *Código de buenas prácticas Estadísticas*. Disponible en https://www.ine.gob.cl/docs/default-source/institucionalidad/buenas-pr%C3%A1cticas/codigo_buenas_practicas_2015.pdf
- INE (2016), *CAENES Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas*. Recuperado de: https://www.ine.gob.cl/docs/default-source/buenas-practicas/clasificaciones/caenes/clasificador/caenes.pdf?sfvrsn=16306d99_4
- INE (2018), *Manual conceptual y metodológico, diseño muestral, Encuesta Nacional de Empleo (ENE)*. Santiago. Recuperado de: https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodolog%C3%ADa-y-dise%C3%B1o-muestral.pdf?sfvrsn=94e53c33_11
- INE. (2019a). *Ciudades, pueblos, aldeas y caseríos 2019*. Obtenido de https://geoarchivos.ine.cl/File/pub/Cd_Pb_Al_Cs_2019.pdf

- INE. (2019). *SEPARATA Codificación Automática ENE*. Obtenido de https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/tematicas/codificacion-automatica-2019.pdf?sfvrsn=273dd63c_7
- INE. (2019b). *Unidades primarias de muestreo. Nuevo marco muestral de viviendas*. Documento interno. Instituto Nacional de Estadísticas (INE).
- INE. (2019c). *Estratificación socioeconómica del marco muestral de viviendas 2017 (MMV 2017)*. Documento de Trabajo. N°11. Obtenido de https://www.ine.cl/docs/default-source/documentos-de-trabajo/estratificaci%C3%B3n-socioecon%C3%B3mica-del-marco-muestral-de-viviendas-2017.pdf?sfvrsn=e4482f3a_2
- INE. (2020a). *Separata Técnica. Fortalecimiento de la Encuesta Nacional de Empleo 2018-2022*. Obtenido de https://ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/tem%C3%A1ticas/separata-t%C3%A9cnica-fortalecimiento-ene-febrero-2020.pdf?sfvrsn=c894aba2_4.
- INE. (2019b2020b). *Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares*. Obtenido de <https://www.ine.gob.cl/docs/default-source/institucionalidad/buenas-pr%C3%A1cticas/clasificaciones-y-estandares/est%C3%A1ndar-evaluaci%C3%B3n-de-calidad-de-estimaciones-publicaci%C3%B3n-27022020.pdf>
- INE (2021). *Documento Metodológico Encuesta Nacional de Empleo (ENE)*. Obtenido de https://www.ine.gob.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodologia-ene-2020.pdf?sfvrsn=793380e9_14
- Keifman, S. (1998). *Imputación de ingresos de hogares: la experiencia de la Encuesta Nacional de Gastos de los Hogares de la Argentina*. En: Segundo Taller Regional sobre Medición del Ingreso en las Encuestas de Hogares-LC/R. 1886-1998-p. 423-430.
- Lohr, S. L. (2000). *Muestreo: Diseño y Análisis*. México: International Thomson Editores.
- Lumley, T. (2020) *survey: Analysis of Complex Survey Samples*. R package version 3.37. <http://r-survey.r-forge.r-project.org/survey/>
- Medina, F., & Galván, M. (2007). *Estudios estadísticos y prospectivos. Imputación de datos: Teoría y práctica*. División Estadística y proyecciones Económicas Naciones Unidas. CEPAL.
- Naciones Unidas. (2009). *Diseño de muestras para encuestas de hogares. Directrices prácticas (Vol. Serie F N°98)*. Nueva York. Obtenido de https://unstats.un.org/unsd/publication/seriesf/Seriesf_98s.pdf
- OIT (1993), *Resolución sobre la Clasificación Internacional de la Situación en el Empleo (CISE), adoptada por la 15ª Conferencia Internacional de Estadísticos del Trabajo, Ginebra*. Recuperado de: <http://www.ilo.org/public/spanish/bureau/stat/download/res/icse.pdf>
- OIT (1998), *Resolución sobre la medición de los ingresos relacionados con el empleo, adoptada por la 16ª Conferencia Internacional de Estadísticos del Trabajo*. Ginebra.

Recuperado de: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087492.pdf

- OIT (2003), *Informe de la Conferencia, 17° Conferencia Internacional de Estadísticos del Trabajo. Ginebra*. Recuperado de: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms_087590.pdf
- OIT (2004), *¿Qué es el trabajo decente?*. Recuperado de: http://www.ilo.org/americas/sala-de-prensa/WCMS_LIM_653_SP/lang--es/index.htm
- OIT (2013), *Resolución I, Resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo de la 19° Conferencia Internacional de Estadísticos del Trabajo. Ginebra*. Recuperado de: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_234036.pdf
- Särndal, C.E. (2007) *The calibration approach in survey theory and practice. Survey Methodology*, Vol. 33, No. 2, (Dec. 2007), pp. 99-119.
- Silva, P. L. (2004) *Calibration Estimation: When and Why, How Much and How*. Textos para discussão, Diretoria de Pesquisas, ISSN 1518-675X, No. 15. Instituto Brasileiro de Geografia e Estatística (IBGE), Rio de Janeiro.
- The Canberra Group (2001), *Expert Group on Household Income Statistics, Final Report and Recommendations, Ottawa*. Recuperado de: http://www.lisdatacenter.org/wp-content/uploads/canberra_report.pdf
- Tourangeau, Roger. (1984) “*Cognitive Sciences and survey methods. Cognitive aspects of survey methodology: building a bridge between disciplines*”
- Tourangeau & Rasinski (1988) “*Cognitive processes underlying context effects in attitude measurement*” *Psychological Bulletin*, 103, 299-314
- United Nations (2011), *United Nations Economic Commission for Europe. Canberra Group, Handbook on Household Income Statistics, Second Edition. Geneva*.
- Vanderhoeft, C. (2001). *Generalised calibration at Statistics Belgium: SPSS® module g-CALIB-S and current practices*. Statistics Belgium.

11 Anexos

11.1 Formulario ESI 2022




■■■■■▶ Si b3 = 4 continuar en m1

SECCIÓN K INGRESOS DE LOS OCUPADOS DEPENDIENTES

k1a El empleo actual [nombre de persona], ¿es el mismo que tenía en el mes de [mes de referencia]?

1 Sí

2 No ▶ Pasa a m1

k1b ¿Cuál fue el monto en pesos de su SUELDO NETO (SALARIO NETO)?

! El SUELDO NETO INCLUYE: bonos regulares, tiempo extra regular y componentes variables. Se puede calcular de dos formas:
 1. Ingreso Líquido (+) Descuentos no Legales (bienestar, préstamo empresa, etc.).
 2. Ingreso Bruto (-) Descuentos Legales (AFP, ISAPRE/FONASA, Seguro de cesantía, Impuesto único a la remuneración).

1 Monto ▶ \$

88 No sabe

99 No responde ▶ Pasa a k3_1

k2a En el mes de [mes de referencia], ¿recibió alguno de los siguientes ingresos variables?

! - En caso de declarar que sí lo recibió, especifique inmediatamente en k2b, k2c y k2d, según corresponda.
 - En k2a_1, k2a_2, k2a_3 y k2a_4 si no lo recibió, no sabe (NS) o no responde (NR), pase al siguiente ingreso.
 - Si en k2a_5, no lo recibió, no sabe (NS) o no responde (NR) pase a k3_1.

1. Bonos por productividad	2. Bonos de otro tipo	3. Comisiones por venta	4. Pago de horas extraordinarias	5. Otros ingresos variables
<input type="checkbox"/> 1 Sí	<input type="checkbox"/> 1 Sí			
<input type="checkbox"/> 2 No	<input type="checkbox"/> 2 No			
<input type="checkbox"/> 88 NS	<input type="checkbox"/> 88 NS			
<input type="checkbox"/> 99 NR	<input type="checkbox"/> 99 NR			

k2b ¿Cuánto fue el monto en pesos (\$) o porcentaje (%) que representa de su sueldo neto?

\$	\$	\$	\$	\$
<input style="width: 100px;" type="text"/>				
%	%	%	%	%
<input style="width: 100px;" type="text"/>				
<input type="checkbox"/> 88 NS				
<input type="checkbox"/> 99 NR				

k2c ¿Incluyó este ingreso en el sueldo neto declarado anteriormente?

<input type="checkbox"/> 1 Sí				
<input type="checkbox"/> 2 No				
<input type="checkbox"/> 88 NS				
<input type="checkbox"/> 99 NR				

k2d ¿Con qué periodicidad recibe este ingreso variable?

<input type="checkbox"/> 1 Mensual				
<input type="checkbox"/> 2 Bimestral				
<input type="checkbox"/> 3 Trimestral				
<input type="checkbox"/> 4 Semestral				
<input type="checkbox"/> 5 Anual				
<input type="checkbox"/> 88 NS				
<input type="checkbox"/> 99 NR				

k3 En el mes de [mes de referencia], ¿le proporcionaron beneficios o regalías de ...

! BENEFICIO: es todo aquello que NO es necesario para el desarrollo de su trabajo. Ya sea, pagos en especie o dinero.

Fuente de ingreso	Sí	Monto		No	NSNR
		\$	% NSNR		
k3_1 ...vivienda?	<input type="checkbox"/>	<input style="width: 100px;" type="text"/>	<input style="width: 100px;" type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
k3_2 ...alimentación?	<input type="checkbox"/>	<input style="width: 100px;" type="text"/>	<input style="width: 100px;" type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
k3_3 ...becas de estudio?	<input type="checkbox"/>	<input style="width: 100px;" type="text"/>	<input style="width: 100px;" type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
k3_4 ...transporte y/o bencina?	<input type="checkbox"/>	<input style="width: 100px;" type="text"/>	<input style="width: 100px;" type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
k3_5 ...propinas?	<input type="checkbox"/>	<input style="width: 100px;" type="text"/>	<input style="width: 100px;" type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
k3_6 ...sala cuna o jardín infantil?	<input type="checkbox"/>	<input style="width: 100px;" type="text"/>	<input style="width: 100px;" type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>
k3_7 ...otros? (especifique)	<input type="checkbox"/>	<input style="width: 100px;" type="text"/>	<input style="width: 100px;" type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>

! Ejemplos k3_7: vestuario, teléfono personal, estacionamiento, pago en especie, etc.

k4a Este sueldo neto, ¿a cuántos días trabajados correspondió?

! Anote 888 en caso de "No sabe" y 999 en caso de "No responde".

k4b En promedio, ¿cuántas horas diarias trabajó durante esos días?

! Anote 888 en caso de "No sabe" y 999 en caso de "No responde".

Pasa a m1

Cálculos:

SECCIÓN L INGRESOS DE LOS OCUPADOS INDEPENDIENTES

l1a Su actual actividad principal de [nombre de persona], ¿es la misma que tenía en el mes de [mes de referencia]?

1 Sí
 2 No **→ Pasa a m1**

l1b Durante dicho mes, ¿cuál fue el monto en pesos del ingreso neto (sueldos asignados, ganancias y/o retiros) que obtuvo en dicha actividad principal?

! **INGRESO NETO** = (Ingresos por venta + Sueldos asignados + Retiros) - (Costos por venta + Gastos operacionales). Si no es posible indicar un monto, solicite al informante que estime un promedio mensual de los últimos doce meses.

1 Monto \$ **Pasa a l1c**
 2 Promedio \$
 88 No sabe **Pasa a l2**
 99 No responde

l1c ¿A cuánto tiempo de trabajo correspondió dicho ingreso?

1 Un mes o menos
 2 Más de un mes **¿Cuántos meses?**
 88 No sabe
 99 No responde

l2 En dicho mes, ¿cotizó en alguno de los siguientes sistemas previsionales?

		Sí	No	NS	NR
l2_1	Salud (Isapre o Fonasa)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
l2_2	Pensiones (AFP o IPS ex INP)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

l3 En el mes de [mes de referencia], ¿usó productos, bienes o servicios de su empresa, negocio o actividad por cuenta propia, para su consumo personal o para el hogar?

1 Sí **→ l3_1** Estime el monto que hubiese tenido que pagar por los productos o servicios utilizados.

! Si no es posible indicar un monto, solicite al informante que estime un promedio mensual de los últimos doce meses.

2 No
 88 No sabe
 99 No responde

Monto \$
Promedio \$

l4 Su empresa, negocio o actividad por cuenta propia, ¿es desarrollada en esta vivienda?

! Sólo considere aquellas cuyo uso PRINCIPAL es para el desarrollo de su actividad económica.

1 Sí **→ l4_1** Si tuviese que pagar un arriendo mensual por el espacio que utiliza en esta vivienda, ¿cuánto estima que le costaría?

Monto \$

2 No
 88 No sabe
 99 No responde

Cálculos:

SECCIÓN M INGRESOS DE OTROS TRABAJOS

■■■■■ **Pregunta solo para OCUPADOS (respondió sección B)**

m1 Independiente de lo anterior, durante el mes de [mes de referencia], [nombre de persona] ¿percibió ingreso(s) por otra(s) ocupación(es), distinta(s) a la actual?

! Si no es posible indicar un monto, solicite al informante que estime un promedio mensual de los últimos doce meses.

1 Sí **{ Monto** \$
 2 No **{ Promedio** \$
 88 No sabe
 99 No responde **Pasa a pre11**

■■■■■ **Pregunta solo para NO OCUPADOS (respondió sección J)**

m2 A pesar de no haber trabajado la semana pasada, durante el mes de [mes de referencia], [nombre de persona] ¿percibió ingresos por algún trabajo u ocupación?

! Si no es posible indicar un monto, solicite al informante que estime un promedio mensual de los últimos doce meses.

1 Sí **{ Monto** \$
 2 No **{ Promedio** \$
 88 No sabe
 99 No responde **Pasa a pre11**

11.2 Sintaxis para el cálculo de medianas en R

El método de interpolación N°5 de *Hydnaman & Fan* (1996) descrito en el apartado 4.11.4 de este documento está presente en la función *svyquantile* del paquete *survey* del software R y se aplica especificando en el parámetro *qrule* el valor "hf5". Esta función es levemente modificada en el cálculo oficial para que en los casos donde la subpoblación se constituya por un solo caso la mediana corresponda a ese mismo valor. Para facilitar la replicabilidad de los resultados, antes de estimar medianas se debe ordenar según el vector de orden oficial ESI (variable "id_mediana").

Por su parte, para calcular medidas de dispersión de las medianas en la ESI se utiliza un diseño repetido *subbootstrap* de 2.000 repeticiones. Se debe fijar una semilla para la aleatorización (la semilla oficial de la ESI es el número "234262762"), declarar el diseño muestral identificando las variables "conglomerado_correlativo", "estrato" y "fact_cal_esi" como unidad primaria muestral, estrato y factor de expansión de la base, respectivamente.

A continuación, se presenta un ejemplo para el cálculo de medianas por región y sexo, empleando el método de interpolación N°5 y calculando medidas de dispersión en la ESI 2022:

```
library(survey) #Para estimaciones a partir de diseño muestral complejo
library(magrittr) #Para usar operador %>%
library(dplyr) #Para funciones de manejo de datos

#a. Leer base y ordenar en base al vector oficial de orden ESI -----
esi_2022 <- [cargar base de datos utilizando la función correspondiente s
egún el formato de la base]

esi_2022 <- esi_2022 %>% arrange(id_mediana)

#b. Generación del diseño muestral: -----
diseno <- svydesign(id = ~conglomerado_correlativo, strata = ~estrato, ch
eck.strata = TRUE,
                  weights = ~fact_cal_esi, data = esi_2022)
options(survey.lonely.psu="remove")

#c. Se fijan las semillas
set.seed(234262762)

#d. Se realiza el diseño repetido de sub-bootstrap para 2.000 réplicas
diseno_rep <- as.svrepdesign(diseno, type = "subbootstrap", replicates =
2000)

#e. Se ejecuta la función modificada correspondiente a la regla de interp
olación del método N°5.
interpolacionMetodo5 <- function(x, w, p)
{
  if (any(zero <- w == 0)) {
    w <- w[!zero]
    x <- x[!zero]
```

```

}
n <- length(x)
ii <- order(x)
x <- x[ii]
w <- w[ii]
cumw <- cumsum(w)
pk <- (cumw - w/2)/(cumw[n])
#Modificación: si x es sólo 1 valor, muéstralo; en caso contrario; generar interpolación:
if(length(x) == 1){x}else{
  approx(pk, x, p, method = "linear", rule = 2)$y
}
}

```

#f. Se realiza la estimación especificando en el parámetro "qrule" la función definida en el paso anterior

```

estimacion<-svyby(~ing_t_p, by = ~region+sexo, data=esi_2022,
  subset(disenio_rep, ocup_ref==1),
  svyquantile, quantiles=c(0.5), interval.type="quantile",
  vartype=c("se"), ci=TRUE, na.rm.all=FALSE, qrule=interpolacionMetodo5)

```

#g. Mostrar el resultado de la estimación

```

estimacion
  region sexo  ing_t_p se.ing_t_p
1.1     1    1 598688.5 28286.538
2.1     2    1 792289.2 38072.838
3.1     3    1 663762.1 39272.557
4.1     4    1 500000.0 19587.492
5.1     5    1 502603.6 12526.733
6.1     6    1 500000.0  5764.639
7.1     7    1 462395.3 15010.599
8.1     8    1 522707.7 18912.837
9.1     9    1 450000.0 15871.569
10.1    10    1 500000.0  5167.151
11.1    11    1 693253.0 40769.021
12.1    12    1 713448.2 47115.173
13.1    13    1 618202.4 18437.099
14.1    14    1 510908.2 21180.841
15.1    15    1 488305.5 20996.426
16.1    16    1 451662.7 18133.648
1.2     1    2 436753.7 25050.340
2.2     2    2 495180.7 11313.680
3.2     3    2 450000.0 22019.773
4.2     4    2 400000.0  5532.318
5.2     5    2 404314.4  7203.963
6.2     6    2 400000.0  5713.409

```

7.2	7	2 420000.0	11629.719
8.2	8	2 416175.7	13362.725
9.2	9	2 400000.0	5592.067
10.2	10	2 420677.3	17782.800
11.2	11	2 495180.7	15611.542
12.2	12	2 502603.6	21294.115
13.2	13	2 502603.6	8882.812
14.2	14	2 432239.1	18802.811
15.2	15	2 371340.2	19161.105
16.2	16	2 400000.0	5700.427

En ocasiones, al ejecutar la estimación de la mediana utilizando el “método 5” de interpolación junto con el método de intervalo de confianza “quantile” (`interval.type="quantile"`) se puede desencadenar un error en la consola de R. Este problema ocurre cuando alguna subpoblación de la estimación posee un recuento no ponderado pequeño, teniendo como consecuencia de que la estimación no se realice. A continuación, se muestra un ejemplo:

```
#h. Se realiza la estimación especificando en el parámetro "qrule" la interpolación de método 5 y en el parámetro interval.type el método quantile, con subpoblación menores
region_cise <-svyby(~ing_t_p, by = ~region+~categoria_ocupacion, data=esi_2022,
  subset(diseño_rep, ocup_ref==1),
  svyquantile, quantiles=c(0.5), interval.type="quantile",
  vartype=c("se"), ci=TRUE, na.rm.all=FALSE, qrule=interpolacionMetodo5)

#i. Mostrar el resultado de la estimación
region_cise

Error in h(simpleError(msg, call)) :
error in evaluating the argument 'x' in selecting a method for function 't': need at least two non-NA values to interpolate
```

Para abordar este error, se recomienda realizar los siguientes dos pasos consecutivos:

1. Estimar la mediana utilizando el “método 5”, sin obtener el error estándar (SE), eliminando del código los parámetros: `interval.type="quantile"`, `vartype=c("se")`, `ci=TRUE`.
2. Estimar el error estándar (SE) solo para los casos que poseen un recuento no ponderado (RNP) mayor o igual a 60 casos (debido a que cuando el $RNP < 60$ la estimación pasa a ser “no

fiable” de forma inmediata (estándar de calidad de las estimaciones⁵¹), agregando los parámetros antes mencionados: (`interval.type="quantile"`, `vartype=c("se")`, `ci=TRUE`).

Lo anterior se visualiza de la siguiente forma en RStudio:

```
#j. Calcular el recuento no ponderado (RNP) del cruce a estimar
RNP <- esi_2022 %>% filter(ocup_ref == 1) %>%
  group_by(region, categoria_ocupacion) %>%
  dplyr::summarise(RNP = n())

#k. Pegar el vector RNP a la base de datos, y dejar lo NA con 0
esi_2022 <- left_join(esi_2022, RNP,
  by=c("region", "categoria_ocupacion"))%>%
  mutate(RNP=ifelse(is.na(RNP),0,as.numeric(RNP)))

#l. Dado que se modifica la base de datos, se vuelve a generar el diseño muestral
diseno <- svydesign(id = ~conglomerado_correlativo, strata = ~estrato, ch
eck.strata = TRUE,
  weights = ~fact_cal_esi, data = esi_2022)
options(survey.lonely.psu="remove")

set.seed(234262762)

diseno_rep <- as.svrepdesign(diseno, type = "subbootstrap", replicates =
2000)

#m. Se realiza la estimación de la mediana, sin los parámetros que afectan al error estándar
mediana <- svyby(~ing_t_p, by = ~region+~categoria_ocupacion, data=esi_2022,
  subset(diseno_rep, ocup_ref==1),
  svyquantile, quantiles=c(0.5), na.rm.all=FALSE, grule=interpolacionMetodo5) %>%
  select (-se.ing_t_p)

#n. Se realiza la estimación del error estándar, agregando el filtro del recuento no ponderado mayor o igual a 60 casos, con los parámetros necesarios
```

⁵¹ INE. (Febrero de 2020). Estándar para la evaluación de la calidad de las estimaciones en encuestas de hogares. Obtenido de <https://www.ine.gob.cl/docs/default-source/institucionalidad/buenas-pr%C3%A1cticas/clasificaciones-y-estandares/est%C3%A1ndar-evaluaci%C3%B3n-de-calidad-de-estimaciones-publicaci%C3%B3n-27022020.pdf>

```

error_estandar <- svyby(~ing_t_p, by = ~region+~categoria_ocupacion, data
=esi_2022,
  subset(disenio_rep, ocup_ref==1 & RNP>=60),
  svyquantile, quantiles=c(0.5), interval.type="quantile",
  vartype=c("se"), ci=TRUE, na.rm.all=FALSE, qrule=interpolacio
nMetodo5) %>%
  select (-ing_t_p)

```

#o. Se juntan los dos tabulados para tener toda la información en un mismo tabulado y así aplicar calidad a la mediana

```

est_se <- left_join(mediana, error_estandar,
  by=c("region", "categoria_ocupacion"))

```

#p. Se muestran los resultados

```

est_se

```

	region	categoria_ocupacion	ing_t_p	se.ing_t_p
1	1		1000000.0	NA
2	2		1200000.0	NA
3	3		721074.6	NA
4	4		808544.5	112575.002
5	5		798597.4	83622.497
6	6		800000.0	69152.516
7	7		566888.1	125981.567
8	8		804165.7	189091.944
9	9		840508.6	144468.557
10	10		1000000.0	89448.280
11	11		1202315.6	NA
12	12		1003262.3	NA
13	13		1010143.7	157530.911
14	14		990361.0	NA
15	15		793931.6	NA
16	16		517697.8	123471.170
17	1	2	316915.7	23853.460
18	2	2	300000.0	17070.195
19	3	2	250000.0	35880.419
20	4	2	277060.3	24533.771
21	5	2	297108.4	19874.593
22	6	2	300000.0	16809.754
23	7	2	301562.1	12035.042
24	8	2	281072.5	21509.156
25	9	2	220021.3	27387.500
26	10	2	351822.5	22667.896
27	11	2	303235.8	28490.584
28	12	2	400834.7	22458.563
29	13	2	321895.7	21087.548
30	14	2	297108.4	16204.723
31	15	2	292653.9	20233.601
32	16	2	203607.2	23303.514
33	1	3	541603.9	31201.767

34	2	3	694017.2	28099.788
35	3	3	600000.0	15052.601
36	4	3	500000.0	19665.767
37	5	3	500000.0	2932.222
38	6	3	452343.2	9919.321
39	7	3	452343.2	12266.004
40	8	3	500000.0	10815.944
41	9	3	452343.2	13784.191
42	10	3	495180.7	9900.719
43	11	3	512655.6	23974.205
44	12	3	583926.1	24256.981
45	13	3	603124.3	12803.578
46	14	3	495180.5	11120.704
47	15	3	450000.0	20545.059
48	16	3	450000.0	13964.065
49	1	4	859583.7	101615.609
50	2	4	987049.4	49834.359
51	3	4	824587.8	77414.787
52	4	4	544698.8	42410.814
53	5	4	792289.2	40253.110
54	6	4	747571.3	45232.306
55	7	4	652794.0	28373.327
56	8	4	683540.7	38699.176
57	9	4	678274.7	55287.363
58	10	4	693253.0	36056.384
59	11	4	1078664.9	87166.182
60	12	4	1100000.0	53820.745
61	13	4	848108.4	33307.382
62	14	4	884582.2	80825.629
63	15	4	800000.0	60487.136
64	16	4	694491.4	56817.041
65	1	5	322165.2	NA
66	2	5	245270.6	NA
67	3	5	220765.9	NA
68	4	5	238982.0	NA
69	5	5	273034.5	25063.541
70	6	5	297108.4	25415.096
71	7	5	250681.9	29755.026
72	8	5	205644.6	19384.839
73	9	5	247590.4	NA
74	10	5	251301.8	37772.846
75	11	5	326287.6	NA
76	12	5	247267.6	NA
77	13	5	372656.4	26571.199
78	14	5	198072.3	35714.823
79	15	5	118843.4	NA
80	16	5	215645.1	NA
81	2	6	603124.3	NA
82	3	6	415958.4	NA

83	4	6	302556.8	NA
84	5	6	498642.1	NA
85	6	6	467642.3	NA
86	7	6	523046.2	NA
87	8	6	504968.5	NA
88	9	6	392030.8	NA
89	10	6	421178.2	NA
90	11	6	504727.8	NA
91	13	6	527279.7	NA
92	14	6	423488.3	NA
93	15	6	346747.9	NA
94	16	6	378000.0	NA
95	1	7	0.0	NA
96	2	7	0.0	NA
97	3	7	0.0	NA
98	4	7	0.0	NA
99	5	7	0.0	NA
100	6	7	0.0	NA
101	7	7	0.0	NA
102	8	7	0.0	NA
103	9	7	0.0	NA
104	10	7	0.0	NA
105	11	7	0.0	NA
106	12	7	0.0	NA
107	13	7	0.0	NA
108	14	7	0.0	NA
109	15	7	0.0	NA
110	16	7	0.0	NA