Documento Metodológico de los Índices de Remuneraciones y Costos Laborales

Base promedio anual 2023=100

Instituto Nacional de Estadísticas

Mayo / 2026

Instituto Nacional de Estadísticas

Subdepartamento de Estadísticas de Ingresos del Trabajo

Subdepartamento de Diseño de Marcos y Muestras

Subdepartamento de Recolección de Estadísticas Económicas y de Precios

Subdepartamento de Supervisión y Verificación de Estadísticas Económicas

DOCUMENTO METODOLÓGICO DE LOS ÍNDICES DE REMUNERACIONES Y COSTOS LABORALES BASE PROMEDIO ANUAL 2023=100

Mayo / 2026

Versión 2.0

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

X: @INE_Chile

Índice

1. Introducción 6
2. Contexto metodológico 7
   2.1 Antecedentes 7
   2.2 Objetivos 11
3. Marco de referencia 12
   3.1 Marco teórico – conceptual 12
   3.2 Referencias nacionales e internacionales 13
   3.3 Marco normativo 15
4. Diseño estadístico 16
   4.1 Población objetivo 16
   4.2 Marco estadístico 16
   4.3 Cobertura geográfica 19
   4.4 Unidades estadísticas 19
   4.5 Definición de variables 19
   4.6 Clasificaciones estadísticas 22
   4.7 Diseño muestral 25
      4.7.1 Cálculo del tamaño muestral 25
      4.7.2 Selección muestral 28
      4.7.3 Distribución del tamaño muestral 28
   4.8 Factores de expansión 32
      4.8.1 Factor de expansión para las empresas del estrato aleatorio 32
      4.8.1.1 Factor de expansión teórico 32
      4.8.1.2 Ajuste por elegibilidad 33
      4.8.1.3 Ajuste por no respuesta 36
      4.8.1.2 Suavizamiento del factor de expansión 38
      4.8.1.3 Calibración del factor de expansión 39
      4.8.2 Factor de expansión para las empresas de inclusión forzosa 41
   4.9 Estimación de parámetros 42
5. Metodología de la recolección de datos 48
   5.1 Técnica y métodos de recolección de los datos 48
      5.1.1 Método de recolección de los datos 48
      5.1.2 Validación operativa 48
      5.1.3 Períodos de referencia y recolección 49
      5.1.4 Conformación del equipo de recolección y supervisión 49

5.1.5 Carga laboral del equipo de analistas 50
5.2 Estructura del instrumento de captura de la información 50
5.3 Método de capacitación para recolección de datos 51
5.4 Método de supervisión 52
5.4.1 Supervisión de gabinete 52
5.4.2 Herramienta para la ejecución de la supervisión 53
5.4.3 Actividades para la ejecución de la supervisión 54
5.4.4 Porcentaje de la muestra a supervisar 54
6. Procesamiento de datos 55
6.1 Integración y codificación de datos 55
6.2 Revisión y validación de datos 55
6.2.1 Auditorias técnicas continuas 56
6.2.2 Revisión final de datos para procesamiento 56
6.3 Edición e imputación de datos 56
6.3.1 Edición de datos 56
6.3.2 Imputación de datos 58
6.4 Asignación de códigos de disposición Final 65
6.5 Calcular ponderadores 66
6.5.1 Ponderadores para el cálculo de los índices 66
6.5.2 Ponderadores para el cálculo de las estimaciones 67
7. Análisis de resultados 69
7.1 Generación de resultados 69
7.1.1 Cálculo de los índices 69
7.1.2 Traslapes y política de rectificación de cifras 75
7.1.3 Empalme de las series 77
7.1.4 Desestacionalización de las series IR, ICL e IR real 78
7.1.5 Variación de los Índices 80
7.1.6 Incidencia de los índices por sección económica 81
7.1.7 Estimaciones analíticas 82
7.2 Validación e interpretación de resultados 90
7.3 Indicadores para la medición y aseguramiento de la calidad estadística 90
7.4 Tratamiento para la perdida de unidades muestrales en el tiempo 92
7.5 Método de control a la divulgación 92
7.5.1 Evaluación inicial del riesgo 92
7.5.2 Algoritmo de anonimización 93

7.5.3 Análisis de riesgo posterior 93
8. Alcance de la presentación de resultados 94
8.1 Productos estadísticos a publicar 94
8.1.1 Microdatos 94
8.1.2 Metadatos 94
8.2 Difusión de los resultados 95
8.3 Accesibilidad a los datos y marco legal 96
9. Glosario 97
10. Bibliografía 98
11. Anexo 100
11.1 Ponderaciones índices 100
11.2 Ponderaciones estimaciones 102
11.3 Modelo de desestacionalización 104

1. Introducción

La Encuesta Mensual de Remuneraciones y Costos Laborales (EMRCL), realizada por el Instituto Nacional de Estadísticas (INE) permite analizar a lo largo del tiempo el comportamiento de las remuneraciones y de los costos laborales de empresas formales de cinco o más personas trabajadoras en Chile.

Dada la misión del INE de entregar información confiable, oportuna, pertinente y comparable, y su compromiso de mejora continua de la producción estadística, se realiza una actualización metodológica enmarcada en el proyecto de la Encuesta Estructural de Remuneraciones (EERCL) 2021 y Cambio de Año Base (CAB) 2023=100[1], en línea con las recomendaciones internacionales, que permite que los actuales indicadores sigan siendo representativos del mercado laboral a través del tiempo.

Los principales indicadores de la encuesta son el Índice Nominal de Remuneraciones (IR) y el Índice Nominal de Costos Laborales[2] (ICL), y un conjunto de indicadores analíticos complementarios que permiten monitorear el comportamiento del mercado laboral. Adicionalmente, se deriva el Índice Real de Remuneraciones, que mide la evolución de las remuneraciones ajustadas por la variación mensual del Índice de Precios al Consumidor (IPC).

Los indicadores principales son representativos a nivel de sección económica, según el Clasificador Chileno de Actividades Económicas CIIU4.CL 2012, y a nivel de tamaño de empresas según número de personas trabajadoras. Tienen como año de referencia el promedio de 2023 y utilizan una estructura de ponderación provista por la EERCL 2021 a nivel de tamaño de empresas, sección económica, sexo y grupo ocupacional.

El presente documento describe la metodología asociada a la Encuesta Mensual de Remuneraciones y Costos Laborales, desde donde derivan los Índices de Remuneraciones y Costos Laborales, y en los principales aspectos metodológicos, técnicos y operativos del proyecto de Cambio de Año Base 2023=100, y forma parte de una serie de documentos de referencia y consulta que el INE pone a disposición de la ciudadanía, para quienes deseen conocer los procedimientos y las decisiones metodológicas del estudio, dando cumplimiento de este modo a las orientaciones generales para la documentación de metodologías de la institución. Tanto este documento como la ficha metodológica de la operación estadística se ponen a disposición del público en la página web institucional del INE.

---

[1] La metodología de los indicadores fue actualizada por última vez el 2017, luego del último cambio de año base 2016=100.

[2] A partir del Cambio de Año Base 2023=100, se reemplaza el término de Costo de la Mano de Obra por el de Costo Laboral, alineándose a las buenas prácticas internacionales en términos conceptuales y normativos. Lo anterior, implica que el anterior Índice el Costo de la Mano de Obra (ICMO) se denomina en adelante Índice de Costos Laborales (ICL).

El documento se divide en once capítulos y se estructura de la siguiente forma: el primer capítulo corresponde a la introducción, el segundo al contexto en el que se construyen los índices. El tercer capítulo explica el marco teórico y conceptual junto con las referencias internacionales, el cuarto profundiza en el diseño estadístico del muestreo y en la descripción de los indicadores. Desde el quinto hasta el octavo capítulo, se informa sobre el proceso de recolección, procesamiento y análisis de resultados, donde se aborda la construcción metodológica de los indicadores principales; por último, la difusión de estos resultados. Los capítulos nueve y diez, describen los conceptos claves que se tratan en este documento y la bibliografía. Por último, en el capítulo 11 se encuentran los anexos que incorporan información complementaria al documento.

2. Contexto metodológico

2.1 Antecedentes

- **Índices mensuales de remuneraciones y costos laborales**

Los índices de remuneraciones (IR) y de costos laborales (ICL) se calculan utilizando la información recolectada mediante la encuesta mensual de remuneraciones y costos laborales (EMRCL), que es aplicada a una muestra de seguimiento de empresas seleccionadas a partir de las unidades que contestaron la encuesta estructural de remuneraciones y costos laborales[3] 2021. La muestra responde a un diseño probabilístico, estratificado y bifásico que tiene representatividad a nivel nacional por sección económica según CIIU4.CL 2012 y por tamaño de empresas según número de personas trabajadoras. Su objetivo principal es conocer la evolución mensual de las remuneraciones y costos laborales en el país, de forma precisa, confiable y oportuna[4].

La necesidad de contar con estos índices se fundamenta en tener una medida de medición de las evoluciones de las remuneraciones y costos laborales en el país. Estos índices sirven entre otros usos, como instrumento de reajuste de salarios, de los topes imponibles utilizados por la Superintendencia de Pensiones, como instrumento de negociaciones por parte de gremios y sindicatos, de análisis y estudios por parte del público general o personas investigadoras, así como fuente complementaria de información para el Sistema de Cuentas Nacionales.

[3] Para más información dirigirse al Documento Metodológico de la Encuesta Estructural de Remuneraciones y Costos Laborales 2021, en el siguiente enlace.
[4] Los indicadores se publican con un rezago aproximado de 37 días.

Contexto histórico

Desde abril de 1959 el INE comenzó a publicar el Índice de Sueldos y Salarios (ISS), establecido por el DFL N°2 publicado en el diario oficial el 31 de julio de 1959 (Diario Oficial, 1959). Este indicador correspondió a una medida agregada de la evolución de las remuneraciones brutas pagadas a obreros y empleados (categorías ocupacionales) en empresas de tamaño mediano y grande, donde la medida de tamaño se definió con respecto al número de personas trabajadoras que tenían. Además, se medían cuatro sectores de la actividad económica: explotación de minas y canteras; industrias manufactureras; electricidad, gas y agua; y servicios comunales. El Índice general se desagregó en indicadores para cada una de las dos categorías ocupacionales y cuatro sectores.

El ISS, cuya base 100 se fijó en abril de 1959, se calculó para los meses de abril de cada año hasta 1962; para los meses de enero, abril, julio y octubre entre abril de 1963 y abril de 1976; y mensualmente entre mayo de 1976 y marzo de 1983.

Desde abril de 1963, bajo la ley N°15.163 del ministerio de obras públicas del 13 de febrero de 1963, se excluyó de la población objetivo del ISS a la gran minería del cobre, Compañía de Aceros del Pacífico, Empresa Nacional del Petróleo y a las compañías manufactureras de papel y celulosa (INE, 1963).

A comienzos de 1978, y con la necesidad de modernizar la cobertura, la estructura ocupacional y la muestra utilizada en el ISS (vigente desde 1959), demandó que el Departamento de Estadísticas Laborales del INE realizara estudios con el objetivo de reemplazar el ISS por un nuevo indicador denominado Índice de Remuneraciones (IR). Este indicador, de periodicidad mensual, fue establecido por la ley N°18.182 de noviembre de 1982[5], eligiéndose la base de diciembre de 1982 (INE, 1985). Sin variar sustancialmente las definiciones y la metodología de cálculo utilizadas en el ISS, en el nuevo indicador, IR, incrementó a ocho el número de ramas de actividad económica, conforme a la Clasificación Internacional Industrial Uniforme rev. 2, y se elevó a ocho el número de grupos ocupacionales, de acuerdo con la clasificación ocupacional COTA-1970 (INE, 1982).Sin variar sustancialmente las definiciones y la metodología de cálculo utilizadas en el ISS, en el nuevo indicador, IR, incrementó a ocho el número de ramas de actividad económica, conforme a la Clasificación Internacional Industrial Uniforme rev. 2, y se elevó a ocho el número de grupos ocupacionales, de acuerdo con la clasificación ocupacional COTA-1970 (INE, 1982).

El IR se definió como una medida agregada de la evolución de las remuneraciones brutas pagadas a las trabajadoras y trabajadores en empresas, establecimientos, instituciones u organizaciones con un mínimo número de trabajadores en todo el país.

[5] Ley N°18.182, modifica artículo 27 del DFL N°2, de 1959, sobre plan habitacional, fecha de publicación el 16-11-1982.

A partir de una encuesta mensual de entidades se calcularon estimaciones del índice general y de índices desagregados para cada una de las ocho ramas y ocho grupos ocupacionales.

En 1993 se actualizan los clasificadores utilizados, quedando como mes base abril de 1993=100, con una medición mensual que estuvo vigente hasta enero de 2006, donde se cambió el mes base a enero 2006=100, pero se siguió con la misma estructura utilizada con el fin enlazar las cifras (INE, 2017).

Durante 2009, el INE impulsó el ‘‘proyecto de actualización del IR’’[6], que implicó una actualización de la muestra y cobertura, así como la actualización de la metodología de cálculo. Dentro de los principales cambios metodológicos se contempló el cambio en la variable de ponderación del IR, que comenzó a ocupar las horas pagadas consistente con un índice de masa salarial; la utilización de ponderaciones para los niveles más agregados, dados por cuentas nacionales; la agregación en los niveles más bajos a través de media geométrica; una nueva forma de imputación; anualizar la base, año 2009; y la incorporación de la desagregación por sexo solo a nivel general (INE, 2010).

En virtud de lo expuesto y de las modificaciones metodológicas que experimentó el Índice de Remuneraciones, las variaciones que se registran a contar de 2010 no son empalmables estadísticamente hacia atrás[7].

El 7 de septiembre de 2017 se realizó una actualización metodológica, donde se incluye una nueva muestra seleccionada a partir de la encuesta estructural 2014, nuevos ponderadores, cambio de clasificador de actividades económicas a CIIU4.CL 2012, y se redefinió la medición de tamaño de empresas, pasando de tamaño por ventas anuales en U.F. a tamaño según número de personas trabajadoras, siendo esta estratificación de tamaño más acorde con las variables de interés de la encuesta, que son los puestos de trabajo y las remuneraciones y costos laborales que pagan las empresas. Esta nueva muestra seleccionada fue recolectada completamente el 2016, teniendo así un año base promedio 2016=100. Junto con ello, se recolectó la información coyuntural hasta julio de 2017, por lo que esta nueva versión de los índices y estimaciones hace referencia a julio de 2017 como primer dato oficial (INE, 2017).

El 7 de noviembre de 2018, y en línea a la Recomendación del Consejo sobre Buenas Prácticas Estadísticas de la Organización para la Cooperación y el Desarrollo Económicos (OCDE)[8], se estableció una política de rectificación de cifras del IR e ICMO, la cual implicaba la actualización de las series de datos para permitir un mejor análisis coyuntural de las variaciones mensuales, acumuladas y anuales. Las series con datos rectificados pertenecen a los siguientes indicadores: Índice de Remuneraciones (IR)

[6] Diario oficial de la república de chile, viernes 29 de enero de 2010.
[7] Decreto sobre indexación que establece la metodología de cálculo para la actualización de valores sobre la base del nuevo Índice de Precios al Consumidor (IPC) y del nuevo Índice de Remuneraciones elaborados por el Instituto Nacional de Estadísticas, publicado en el Diario Oficial el 29 de enero de 2010. Para más información sobre este decreto, ver aquí.
[8] Ver “Recomendación del consejo de la OCDE sobre buenas prácticas estadísticas” (OCDE, 2015), en http://www.oecd.org/statistics/good-practice-toolkit/

nominal y real, Índice del Costo de la Mano de Obra (ICMO) nominal, estimaciones de la remuneración promedio y costos medios por hora pagada y estimaciones de las horas medias pagadas. Por lo que, desde noviembre de 2018 (periodo de referencia de septiembre) las cifras publicadas en cada mes de análisis (período t) tienen un carácter de provisional[9] y son actualizadas en el mes siguiente.

Durante la pandemia de COVID-19, y derivado de la entrada en vigor de la Ley de Protección al Empleo (Ley N°21.227), se realizaron cambios en la metodología de imputación, estableciéndose una serie de criterios para determinar si correspondía o no imputar datos faltantes en ciertas empresas[10]. Estos cambios se realizaron a partir de la publicación del 7 de julio de 2020.

Durante el 2022 se recolectaron datos de empresas e instituciones con el fin de reunir los insumos necesarios para la actualización de los indicadores de remuneraciones y costos laborales, a través de la Encuesta Estructural de Remuneraciones y Costos Laborales 2021, cuyos datos de referencia corresponden al acumulado anual de 2021, obteniendo de este operativo estadístico, tanto los ponderadores como la muestra de seguimiento mensual. Los datos y metodología de la Encuesta Estructural de Remuneraciones y Costos Laborales se publicaron en diciembre de 2023[11].

Finalmente, este documento da cuenta de los cambios metodológicos implementados en esta nueva versión, transparentando la metodología de cálculo de los índices y de las estimaciones base promedio anual 2023=100, haciendo posible su réplica y su mejor disposición a las personas usuarias.

[9] Identificado en los productos de difusión oficial (boletín estadístico, cuadros estadísticos, entre otros) con la sigla "/p".
[10] Para más información sobre los cambios metodológicos asociados a la Ley 21.227, ver Separatas técnicas N°1 y N°2 de Contingencia Covid-19 en www.ine.gob.cl, sección mercado laboral, micrositio de Remuneraciones y Costos Laborales.
[11] Más información sobre la metodología, en el siguiente link.

2.2 Objetivos

Objetivo general

Medir la evolución de las remuneraciones y los costos laborales desembolsados por empresas de cinco o más personas trabajadoras a nivel nacional, a través de información mensual proporcionada por una muestra de empresas.

Objetivos específicos

- Obtener indicadores desagregados de remuneraciones y costos laborales según actividad económica y tamaño de empresas, y según grupo ocupacional y sexo de las personas trabajadoras.
- Obtener estimaciones analíticas de remuneraciones y costos medios por horas pagadas y horas pagadas medias por persona trabajadora.
- Obtener estimaciones de brechas salariales por sexo.

3. Marco de referencia

En esta sección se describen los elementos que dan soporte a la EMRCL y que configuran el marco conceptual de los índices de remuneraciones y costos laborales.

3.1 Marco teórico – conceptual

Las mediciones estadísticas sobre remuneraciones y costos laborales son, de acuerdo con la Organización Internacional del Trabajo (OIT), parte de los indicadores claves del mercado laboral, los cuales reflejan los dos aspectos principales de las medidas salariales existentes: la primera destinada a medir los ingresos de las personas asalariadas, y la segunda, destinada a medir los gastos de la unidad empleadora por contratar a dichas personas (OIT, 2015)

Desde hace décadas, OIT ha impulsado el desarrollo de estas estadísticas bajo un marco conceptual armonizado. En particular, desde la 11ª Conferencia Internacional de Estadígrafos del Trabajo (OIT, 1966) surge una resolución sobre las estadísticas de los costos laborales, y se emanan recomendaciones respecto a una clasificación internacional uniforme de los costos de la mano de obra, la cual sigue vigente hasta la fecha.

En Chile, se adoptan, en general, tales recomendaciones, y a través de la EMRCL se obtienen datos necesarios para la elaboración y cálculo de estadísticas mensuales de remuneraciones y de costos laborales por unidad de tiempo; en tanto que, los Índices de Remuneraciones y de Costos Laborales se construyen como una alternativa metodológica de seguimiento de la evolución en el tiempo de dichas estadísticas en base a números índices.

3.2 Referencias nacionales e internacionales

Referencias internacionales

La EMRCL recoge criterios y prácticas emanados de las Conferencias Internacionales de Estadísticos del Trabajo (CIET) de la Organización Internacional del Trabajo (OIT). En las CIET se formulan recomendaciones sobre estadísticas laborales en forma de resoluciones y directrices, las cuales se refieren a definiciones, métodos de compilación y clasificación, que promueven el desarrollo de tales estadísticas sobre una base equilibrada y contribuyen a mejorar su comparabilidad internacional.

En particular, la resolución de la 11.a CIET (OIT, 1966) y vigente hasta la fecha, establece y señala que los programas de estadísticas de costos laborales deberían elaborarse esencialmente con miras a proporcionar evaluaciones fidedignas del nivel, composición y evolución del costo que representa para el empleador la utilización de la mano de obra; la información compilada en las encuestas sobre los costos laborales también puede utilizarse para estudiar la parte de los ingresos de los trabajadores provenientes de sus empleos, como por ejemplo, los salarios directos, los elementos complementarios del salario no incluidos en las estadísticas ordinarias de salarios, el promedio de ganancias anuales, etc.

La incorporación de estándares y recomendaciones prácticos que se incorporan en la EMRCL se resumen en:

- Una adaptación de la Clasificación internacional uniforme de los costos de mano de obra de la resolución de la 11.a CIET (OIT, 1966), siguiendo las definiciones de la producción nacional reciente sobre remuneraciones y costos de la mano de obra, particularmente las definiciones de la Encuesta Estructural de Remuneraciones y Costo de la Mano de Obra 2014, la Encuesta Estructural de Remuneraciones y Costos Laborales 2021, y los indicadores de seguimiento mensual de remuneraciones (IR) y del costo de la mano de obra (ICMO) con base 2016=100.
- La adopción de recomendaciones de medición de estadística según ramas de actividad económica a través de la adaptación nacional de la cuarta versión del Clasificador Chileno de Actividades Económicas, CIIU4.CL 2012 (INE, 2012).
- La adopción de recomendaciones de medición de estadística según tamaño de empresas.
- La adopción de recomendaciones de medición de estadística según grupos de ocupación, de acuerdo con la adaptación nacional de la cuarta versión de Clasificador Chileno de Ocupaciones, CIUO 08.CL (INE, 2018).
- La adopción de recomendaciones de medición de estadística según sexo.
- La medición de horas de pagadas y horas de trabajo extraordinarias.

- La recomendación de mediciones sobre los doce meses del año de referencia no acumulados. En este caso, la EMRCL se recolecta con referencia mensual, tanto para el año base como para los meses de seguimiento coyuntural.

No se incorporan en la encuesta las siguientes recomendaciones:

- La medición por separado de las remuneraciones por tiempo no trabajado. En la EMRCL se incluyen, en general, dentro de los sueldos y salarios directos.
- La recomendación de medición de costo medio de la mano de obra por unidad de tiempo en base a horas efectivamente trabajadas. En la EMRCL se utilizan las horas pagadas.
- La recomendación de la unidad de análisis para la cual se obtiene información, que debería ser, en lo posible, el establecimiento, más bien que la empresa o la compañía. En la EMRCL, la unidad de análisis es la empresa o institución como unidad jurídica legal.
- La recomendación de mediciones, en lo posible, por áreas geográficas.
- La recomendación de mediciones, en lo posible, por grupos de edades.

Estos criterios adoptados en la EMRCL permiten determinados niveles de comparabilidad internacional en la medida en que pueda atenderse las especificaciones metodológicas incluidas en la producción de la estadística local. Actualmente, la producción de los indicadores de seguimiento mensual de IR e ICL son utilizados en comparaciones internacionales oficiales publicados por la OIT.

Referencias nacionales

Respecto a las referencias nacionales, la encuesta mensual de remuneraciones y costos laborales del Instituto Nacional de Estadísticas de Chile es la única que entrega información oficial respecto de la evolución de las remuneraciones y costos laborales. Para conocer mayor detalle sobre su contexto histórico ver apartado 2.1 de este documento y para conocer metodología, acceder al micrositio de la encuesta[12].

[12] Metodología disponible en el apartado de Metodologías del micrositio del producto, disponible en https://www.ine.gob.cl/estadisticas/sociales/mercado-laboral/remuneraciones-y-costos-laborales.

3.3 Marco normativo

- Ley N° 17.374 del 10 de diciembre 1970, del entonces Ministerio de Economía, Fomento y Reconstrucción, hoy Ministerio de Economía, Fomento y Turismo, que fija nuevo texto refundido, coordinado y actualizado del DFL. N° 313 de 1960, que aprueba la Ley Orgánica Dirección Estadística y Censos y crea el Instituto Nacional de Estadísticas.
- Decreto Supremo N° 322 del Ministerio de Economía, publicado en el Diario Oficial de 29 de enero de 2010, que establece la metodología de cálculo para la actualización de valores sobre la base del nuevo Índice de Precios al Consumidor (IPC) y del nuevo Índice de Remuneraciones (IR) elaborados por el Instituto Nacional de Estadísticas.
- Ley N° 19.653, de 14 de diciembre de 1999, Sobre probidad administrativa aplicable a los órganos de la Administración del Estado.
- Ley N° 20.285, de 20 de agosto de 2008, Sobre Acceso a la Información Pública.
- Decreto N° 1.062, de 17 de diciembre 1970, del entonces Ministerio de Economía, Fomento y Reconstrucción, hoy Ministerio de Economía, Fomento y Turismo, que aprueba reglamento del Instituto Nacional de Estadísticas.
- Decreto N° 305, de 17 de febrero de 2010, del entonces Ministerio de Economía Fomento y Reconstrucción, hoy Ministerio de Economía, Fomento y Turismo, que ordena la incorporación de la variable sexo en la producción de estadísticas y generación de registros administrativos[13].
- Ley N° 19.628, de 28 de agosto de 1999, Sobre protección de la vida privada.

Cabe señalar que, en vista de la normativa indicada, el Instituto Nacional de Estadísticas, se encuentra obligado a respetar las normas establecidas por la Ley N° 17.374, dentro de las cuales se contempla la estricta observancia del Secreto Estadístico, consagrado en los artículos 29° y 30° de la citada Ley N° 17.374.

Adicionalmente, de acuerdo a lo dispuesto en los artículos 20, 21 y 22 de la Ley N° 17.374, el INE podrá solicitar a los particulares, ya sean personas naturales o jurídicas, así como a los funcionarios encargados de organismos fiscales, semifiscales, empresas del Estado, municipalidades y demás instituciones públicas, la entrega de datos, antecedentes o informaciones de carácter estadístico, acerca de hechos que tengan relación con la formación de estadísticas oficiales; solicitud que será obligatoria para los informantes.

[13] Para más información revisar el decreto N°305 dispuesto en la biblioteca del Congreso Nacional de Chile disponible en el siguiente enlace: https://www.bcn.cl/leychile/navegar?idNorma=1011115.

4. Diseño estadístico

4.1 Población objetivo

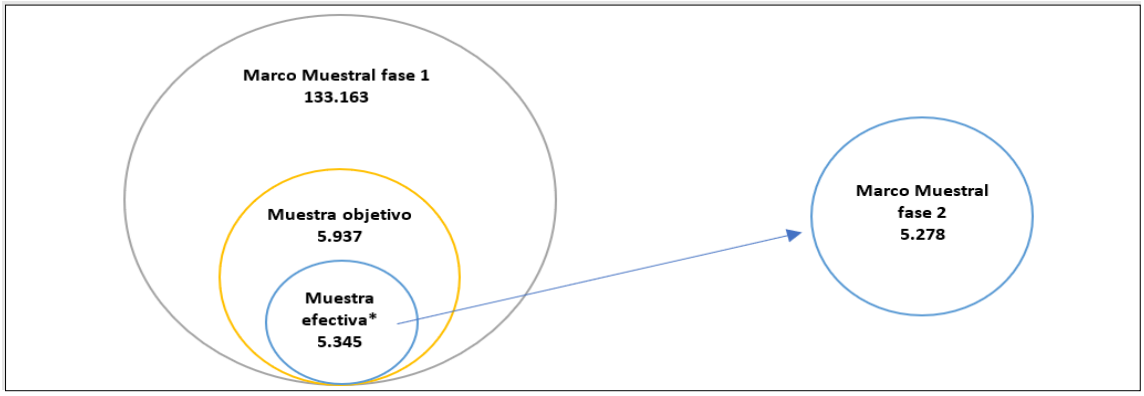
La población objetivo de la Encuesta Mensual de Remuneraciones y Costos Laborales (EMRCL), corresponde a empresas e instituciones formales con cinco o más personas trabajadoras contratadas directamente, que cuentan con iniciación de actividades en el Servicio de Impuestos Internos (SII), con RUT y contabilidad propia, y donde su actividad económica principal, es parte de las actividades de interés del estudio, según CIIU4.CL 2012, y que se desarrollan dentro de los límites geográficos del país.

4.2 Marco estadístico

El marco estadístico de la EMRCL se basa en un diseño muestral probabilístico, estratificado y bifásico, donde la primera fase se conforma por las empresas que provienen del Marco Maestro de Empresas[14] (MME), Formulario 1887[15] (F1887) y registros administrativos de la SUSESO[16]. El primer registro proporciona información sobre el Rut de las empresas y las actividades económicas que son partes del estudio, mientras que el número de personas trabajadoras, para el período de referencia, se obtiene del F1887 y de SUSESO, siendo este último registro la fuente principal para obtener la variable de interés que permite la estratificación del marco por tamaño de empresas. Respecto al marco muestral de la segunda fase, este se conforma a partir de la primera fase con las empresas vigentes en el período en estudio y que declaren cinco o más personas trabajadoras asalariadas en la Encuesta Estructural de Remuneraciones y Costos Laborales 2021, por lo que son aptas para procesar en la primera fase. Además, que realicen su actividad principal dentro del ámbito de estudio, en el territorio nacional.

[14] Registros administrativos que provienen del Servicio de Impuestos Internos (SII), relacionado a la información de la declaración mensual y pago simultáneo de impuestos (Formulario N° 29), impuestos anuales a la renta (Formulario N°22), entre otras fuentes y se complementa con información que se recoge en las encuestas continuas que realiza el INE.
[15] Registro administrativo que proporciona el SII en forma anual y corresponde a la declaración de rentas del Art. 42 N°1 (sueldos), otros componentes de la remuneración y retención del impuesto único de segunda categoría.
[16] Registro administrativo que recibe el INE desde la SUSESO, la cual proporciona mensualmente datos de las empresas.

Gráfico 1. Conformación del Marco Muestral EERCL (fase 1) y EMRCL (fase 2)



Fuente: Instituto Nacional de Estadísticas (INE).

* el marco muestral de segunda fase se conforma por encuestas vigentes al periodo de recolección y cuya respuesta fue apta para procesar en la primera fase. Así, se identificaron 67 empresas que no cumplieron dicho criterio, lo que origina diferencia entre la población objetivo que se identifica en la primera fase y el marco de selección de la segunda fase. Sin embargo, los ajustes que se realizan al factor de expansión de esta última fase mitigan el efecto de la pérdida muestral, a la vez que mantienen consistencia entre los stocks poblacionales representados tanto por la encuesta de primera como de segunda fase.

El marco muestral de la segunda fase queda conformado con la muestra efectiva de primera fase excluyendo algunas unidades por los ajustes ya mencionados, sobre el cual se realiza la post-estratificación a partir de la información proveniente de la recolección de primera fase. La configuración de los post-estratos permite reunir las empresas de la población objetivo en grupos homogéneos previo a la selección, con el propósito de mejorar las medidas de precisión de las estimaciones agrupando las unidades del marco muestral en clases homogéneas en su interior y que difieran de las características del resto, los cuales son mutuamente excluyentes y exhaustivos colectivamente.

Para la EMRCL, los estratos a considerar en el marco muestral se conforman a partir de las mismas variables utilizadas en la EERCL 2021, es decir, según el cruce de la sección económica, tamaño de empresa y estratos de inclusión forzosa y aleatoria.

1. **Sección económica:** este nivel se define a partir de la clasificación de las actividades económicas que se especifican en el apartado de Clasificaciones estadísticas, específicamente en Tabla 2: Secciones económicas consideradas en el estudio.

2. **Tamaño de empresas:** se define de acuerdo con el número de personas trabajadoras que declararon las empresas en la EERCL 2021, estableciendo cuatro categorías de clasificación para el EMRCL. Los límites se definen en la Tabla 1 que se presenta a continuación:

Tabla 1. Determinación del tamaño de las empresas según número de personas trabajadoras

Tamaño de empresas	Número de personas trabajadoras	
	Límite inferior	Límite superior
Mega	1.000	Más
Grande	200	999
Mediana	50	199
Pequeña	5	49

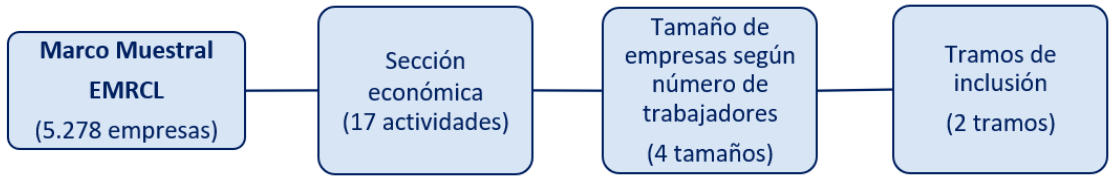
Fuente: Instituto Nacional de Estadísticas (INE).

3. **Estratos de Inclusión:** se conforma a partir de dos tipos de inclusión: 1) unidades de inclusión forzosa (IF), que se definen por su importancia en el sector y son consideradas como autorrepresentadas, es decir, que sólo se representan a sí mismas, y 2) las unidades de inclusión aleatoria (IA), que conforman el estrato aleatorio a partir del cual se obtiene una muestra probabilística de unidades.

Para la EMRCL, se mantienen las unidades pertenecientes al estrato de inclusión forzosa definidas en el marco de fase 1 y que respondieron la encuesta. El resto de las empresas conforman las unidades de inclusión aleatoria, del que finalmente se obtiene la muestra probabilística.

Finalmente, los niveles para la estratificación del marco muestral de la EMRCL quedan conformada de acuerdo con la interacción de las siguientes tres variables:

Gráfico 2. Estratificación del marco muestral para la EMRCL



Fuente: Instituto Nacional de Estadísticas (INE).

4.3 Cobertura geográfica

La cobertura geográfica de la EMRCL son las 16 regiones del país, donde las empresas formales que conforman la población objetivo desempeñan su actividad económica principal.

4.4 Unidades estadísticas

Unidad de muestreo

La unidad de muestreo es la empresa formal[17] que desarrolla su actividad económica principal en el ámbito de estudio y dentro de los límites territoriales del país, cumpliendo los requisitos para ser considerada parte de la población objetivo.

Unidad de observación y de análisis

La unidad de observación y análisis es la empresa formal que cumple con los requisitos para ser considerada parte de la población objetivo.

4.5 Definición de variables

Las variables más relevantes para la construcción de los indicadores son las siguientes:

- Remuneración ordinaria (RO): corresponde a la contraprestación en dinero y/o especies que las personas trabajadoras perciben de la empresa por las horas pactadas en el contrato de trabajo. Incluye los siguientes componentes: sueldo base, otros sueldos y salarios básicos y permanentes, sueldos por funciones ocasionales, incentivos y premios pagados, pagos por trato, comisiones por ventas, pagos en especies y/o gastos de vivienda, y pagos en dinero por beneficios de bienestar costeados por el empleador o empleadora.
- Remuneración extraordinaria (RE): contraprestación en dinero que percibe la persona trabajadora por las horas trabajadas que exceden la jornada laboral ordinaria.
- Costos laborales (CL): gastos en que incurre la empresa al mantener contratados a sus trabajadoras y trabajadores. Incluye los siguientes componentes: remuneraciones ordinarias, pagos por horas extraordinarias, reembolsos de gastos de la persona trabajadora por causa del trabajo, aportes patronales por seguros obligatorios, gastos del empleador o empleadora en servicios de bienestar, gastos de la persona empleadora en servicios de capacitación y otros costos laborales. Se

[17] Corresponde a la unidad que cuenta con iniciación de actividades en el Servicio de Impuestos Internos (SII), con RUT y contabilidad propia.

excluyen los pagos no mensuales y las indemnizaciones por término de la relación de trabajo.

- Número de puestos de trabajo (NT)[18]: 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. La noción de puesto de trabajo se utiliza en relación con el trabajo en la ocupación[19]. Las personas pueden tener uno o varios puestos de trabajo. En el caso de quienes poseen más de un puesto de trabajo, el puesto de trabajo principal se define como aquel al que corresponde el mayor número de horas habitualmente trabajadas, conforme a la definición establecida en las normas estadísticas internacionales sobre tiempo de trabajo.
- Horas ordinarias (HO): horas mensuales pagadas según la jornada laboral ordinaria, ya sea la pactada en el contrato de trabajo o la máxima legal.
- Horas extraordinarias (HE): horas pagadas que exceden la jornada laboral ordinaria.
- Horas totales (HT): corresponde a la suma de las horas ordinarias y las horas extraordinarias.
- Remuneración ordinaria por hora ordinaria (ROHO): corresponde al cociente entre las remuneraciones ordinarias (RO) y las horas ordinarias (HO), y es la medida de remuneración por unidad de tiempo objetivo del índice nominal de remuneraciones (IR).
- Costos laborales por hora toral (CLHT): corresponde al cociente entre el total de costos de costos laborales (CL) y las horas totales (HT), y es la medida de costo laboral por unidad de tiempo objetivo del índice nominal de costos laborales (ICL).
- Índice nominal de remuneraciones (IR): mide la evolución nominal de las remuneraciones ordinarias por hora ordinaria pagadas a las personas trabajadoras directamente contratadas por las empresas (base promedio año 2023=100). Indistintamente, se puede usar el nombre de este indicador como índice nominal de remuneraciones o solo índice de remuneraciones.
- Índice nominal de costos laborales (ICL): mide la evolución nominal del costo laboral por hora total en el que incurren las empresas para mantener a sus trabajadoras y trabajadores (base promedio año 2023=100). Indistintamente, se puede usar el nombre de este indicador como índice nominal de costos laborales o solo índice de costos laborales o índice del costo laboral.
- Índice real de remuneraciones: mide la evolución real de las remuneraciones por hora ordinaria en el país (base promedio año 2023=100), utilizando para ello el Índice de Precios al Consumidor (IPC).
- Índice desestacionalizado: mide la evolución del IR, ICL e IR Real, eliminando los factores y/o patrones exógenos, de naturaleza no económica, que influyen sobre la variable en estudio y que en ocasiones ocultan los aspectos netamente económicos, los que constituyen la esencia del análisis de una serie de tiempo económica.

[18] Definición según la 19ª Conferencia Internacional de Estadísticos del Trabajo (CIET).
[19] El trabajo en la ocupación comprende el trabajo realizado para terceros a cambio de remuneración o beneficios.

- Estimaciones o indicadores analíticos[20]: son cocientes que se construyen a partir de las horas ordinarias y extraordinarias, las remuneraciones ordinarias y extraordinarias y el costo laboral. Se utilizan como indicadores complementarios del comportamiento del mercado del trabajo y actualmente son seis: remuneraciones ordinarias por hora ordinaria, remuneraciones ordinarias y extraordinarias por hora total, costo laboral por hora total, horas ordinarias, extraordinarias y totales por persona trabajadora.
- Brecha por género: es la diferencia porcentual que existe entre la estimación de la remuneración ordinaria media por hora ordinaria[21] (o costos medios por hora total) de las mujeres con respecto a la de los hombres.
- Brecha con respecto al promedio según género: es la diferencia porcentual que existe entre la estimación de la remuneración ordinaria media por hora ordinaria (o costos medios por hora total) de las mujeres o de los hombres con respecto al promedio general en las estimaciones de la remuneración ordinaria por hora ordinaria (o del costo por hora promedio del total poblacional).
- Cadena larga: corresponde a la agregación de los datos de la encuesta a nivel de tamaño de empresas t, sección económica c, sexo s, grupo ocupacional g, y empresa e. Se expresa como: t, c, s, g, e.
- Cadena corta: corresponde a la agregación de los datos de la encuesta a nivel de tamaño de empresas t, sección económica c, sexo s y grupo ocupacional g. Se expresa como: t, c, s, g.
- Ponderación: determinación del peso o relevancia que tiene el atributo en la construcción del índice.
- Factor de expansión: en términos estadísticos se refiere a la inversa de la probabilidad de una empresa de ser seleccionada en la muestra. Se interpreta como la cantidad de empresas que representa una unidad estadística de la muestra en el marco muestral.

[20] Estas estimaciones se presentan a nivel general y según sexo (mujeres y hombres).
[21] O alguna otra variable de interés.

4.6 Clasificaciones estadísticas

En el siguiente apartado son presentados los clasificadores utilizados en las desagregaciones del índice a nivel de sección económica, tamaño de empresa, sexo y grupo ocupacional.

Clasificador Chileno de Actividades Económicas CIIU4.CL 2012

A continuación, se presentan dos tablas, en la primera se muestran las 17 actividades que son consideradas en el Clasificador Chileno de Actividades Económicas CIIU4.CL 2012 y en la segunda están las actividades no consideradas en la medición de los indicadores[22].

Tabla 2: Secciones económicas consideradas en el estudio.

Sección económica según CIIU4.CL 2012	
Sección	**Descripción**
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

Fuente: Instituto Nacional de Estadísticas (INE).

[22] Estos sectores fueron excluidos por criterios de cobertura y operativos, que restringen su incorporación en el estudio.

Tabla 3: Secciones económicas o clases especificas excluidas del estudio.

Sección económica según CIIU4.CL 2012	
Sección	**Descripción**
A	Agricultura, ganadería, silvicultura y pesca
O (8422)	Actividades de defensa
O (8423)	Actividades de mantenimiento del orden público y de seguridad
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

Fuente: Instituto Nacional de Estadísticas (INE).

Clasificación según tamaño de empresas[23]

El tamaño de una empresa se define en función del número de personas trabajadoras. Se establecen tres categorías distintas de ponderación para los indicadores: pequeñas empresas, que comprenden entre 5 y 49 personas; medianas empresas, que tienen entre 50 y 199 personas; y grandes empresas, que cuentan con 200 personas o más. En adelante, se empleará el término 'tamaño de empresa" para hacer referencia a la categoría correspondiente según el número de personas trabajadoras.

Tabla 4: Tamaño de empresas según número de personas trabajadoras consideradas en el estudio.

Tamaño	Descripción
Pequeña	De 5 a 49 personas trabajadoras
Mediana	De 50 a 199 personas trabajadoras
Grande	De 200 o más personas trabajadoras

Fuente: Instituto Nacional de Estadísticas (INE).

Clasificación según sexo

Tabla 5: Clasificación del sexo de las personas trabajadoras.

Sexo
1: Mujer
2: Hombres

Fuente: Instituto Nacional de Estadísticas (INE).

[23] Para la selección muestral existe una categoría adicional, Mega, que se construye con unidades que cuentan con 1.000 o más personas trabajadoras. Sin embargo, para fines de análisis la categoría Mega representa un subconjunto de la categoría Grande, por lo que para fines de este documento se considerarán las tres categorías descritas en la tabla 4.

Clasificación Chileno de Ocupaciones del año 2008 CIUO 08.CL

Uno de los principales cambios que se introdujo en el actual IR e ICL base promedio anual 2023=100 fue la actualización del clasificador de ocupaciones utilizado, pasando del Clasificador Internacional Uniforme de Ocupaciones CIUO-88 al Clasificador Chileno de Ocupaciones CIUO 08.CL.

Se utiliza el término grupo de ocupación para referirse al gran grupo (CIUO 08.CL) de ocupación, separando el gran grupo 5 de "Trabajadores de los servicios y vendedores de comercios y mercados" en 2 subgrupos principales, descritos en la tabla 6, y cuya agrupación se describe en la tabla 7.

Tabla 6: Clasificación según grupo ocupacional.

Grupo ocupacional según CIUO 08.CL
G1. Directores, gerentes y administradores
G2. Profesionales, científicos e intelectuales
G3. Técnicos y profesionales de nivel medio
G4. Personal de apoyo administrativo
G5.1. Trabajadores de los servicios a las personas, trabajadores de los cuidados personales, y personal de los servicios de protección y seguridad
G5.2. Vendedores
G7. Artesanos y operarios de oficios
G8. Operadores de instalaciones, máquinas y ensambladores
G9. Ocupaciones elementales

Fuente: Instituto Nacional de Estadísticas (INE).

Tabla 7: Subgrupos principales del gran grupo 5 de trabajadores de los servicios y vendedores de comercios y mercados (CIUO 08.CL).

Subgrupo CIUO 08.CL	Glosa	Subgrupo agrupado
5.1	Trabajadores de los servicios a las personas	5.1
5.2	Vendedores	5.2
5.3	Trabajadores de los cuidados personales	5.1
5.4	Personal de los servicios de protección y seguridad	5.1

Fuente: Instituto Nacional de Estadísticas (INE).

4.7 Diseño muestral

El diseño de la muestra para la EMRCL considera representatividad a nivel nacional, sección económica, tamaño de empresas y cruce entre sección y tamaño. Sustentado bajo un muestreo probabilístico, estratificado de dos fases; esta estrategia se basa en la necesidad de contar con información sobre remuneraciones, costos laborales y personas trabajadoras de las empresas, actualizada y alcanzando mayor desagregación que los niveles definidos en el diseño de la muestra. Así, además de la actividad económica y tamaño de empresas se requiere información de sexo y grupo ocupacional, que no está disponible en los marcos estadísticos de empresas del INE.

4.7.1 Cálculo del tamaño muestral

La estrategia del diseño de la muestra probabilística de la segunda fase considera los estratos que se definen por el cruce de la sección de actividad económica, tamaño de empresa y estratos de inclusión forzosa y aleatoria. Este diseño considera los siguientes elementos para calcular el tamaño de la muestra correspondiente a la EMRCL:

Tabla 8: Elementos considerados para calcular el tamaño de muestra

Parámetro	**Descripción**
Variable de diseño	Número de personas trabajadoras (y)
Parámetro asociado	Total de personas trabajadoras (Y)
Estimador asociado	Estimador del total de personas trabajadoras: $\hat{Y}$
Errores de muestreo	Los errores relativos para los distintos niveles de estimación no deben superar 20%
Tasa de sobremuestreo	Tasa promedio de no logro para el año 2022 de la encuesta mensual correspondiente al periodo base promedio anual 2016=100

Fuente: Instituto Nacional de Estadísticas (INE).

Según la estratificación definida en el marco muestral, se incorpora el tramo de inclusión, dividido en tipos de estratos de inclusión; inclusión forzosa (IF) corresponde a todas las unidades que se consideraron IF en la encuesta estructural y que respondieron la misma. Mientras que, el estrato de inclusión aleatoria, se extrae una muestra representativa y su metodología de cálculo se presenta en los siguientes pasos, además de las fórmulas empleadas para el cálculo del tamaño de muestra.

Paso 1. El tamaño de muestra objetivo por sección económica se calcula considerando un error relativo de 20%, mediante la siguiente formula:

$$n_h = \frac{N_h \cdot Z_{1-\alpha/2}^2 \cdot \sum_{v=1}^{V} N_{hv} \cdot \sigma_{hv}^2}{\left(e_r \cdot \sum_{v=1}^{V} \hat{Y}_{hv}\right)^2 + Z_{1-\alpha/2}^2 \cdot \left(\sum_{v=1}^{V} N_{hv} \cdot \sigma_{hv}^2\right)}$$

Donde:

n_h : Tamaño de muestra objetivo para EMRCL en la sección económica h.

N_h : Número de empresas en el marco muestral de la segunda fase para la sección económica h.

N_{hv} : Número de empresas en el marco muestral de la segunda fase para la sección económica h y tamaño v.

σ_{hv}^2 : Cuasi-Varianza de la variable número de personas trabajadoras en la sección económica h y tamaño v y se estima mediante la siguiente fórmula: $S_{hv}^2 = \sum_{v=1}^{V} \left[\hat{Y}_{hv} - \hat{\bar{Y}}_h\right]^2 / (n_h - 1)$.

e_r : Error relativo máximo admisible de 20%.

$\hat{Y}_{hv}$: Total de personas trabajadoras de las empresas en la sección económica h y tamaño v.

$Z_{1-\frac{\alpha}{2}}$: Cuantil de la distribución normal estándar correspondiente a un nivel de confianza[24].

Paso 2. Posteriormente, el tamaño de muestra obtenido en el paso 1, se reparte en cada uno de los estratos de tamaño v en forma proporcional al tamaño, utilizando la siguiente fórmula:

$$n_{hv} = n_h . \left(\frac{N_{hv}}{N_h}\right)$$

Donde:

n_{hv} : Tamaño de muestra objetivo en la sección económica h y tamaño v.

[24] Para el cálculo del tamaño muestral se utiliza un valor 2 y no 1,96, que es el valor del cuantil de la distribución normal estándar a un 95% de confianza. Esto debido a que, una vez realizado la recolección, para el cálculo de los errores de estimación se utiliza el cuantil de la distribución t student, el cual aumenta en la medida que disminuye el tamaño de muestral. Al utilizar un valor 1,96, podría subestimarse el tamaño de muestra necesario.

n_h : Tamaño de muestra objetivo para EMRCL en la sección económica h.

N_h : Tamaño del marco muestral de segunda fase para la sección económica h.

N_{hv} : Número de empresas en el marco muestral en la sección económica h y tamaño v.

Paso 3. Una vez distribuido (afijado) el tamaño muestral en los estratos de tamaño al interior del estrato, con la información de la variable número de personas trabajadoras contenida en el marco muestral de segunda fase, se calculan las varianzas y errores relativos esperados para los estratos y si se observa que estos superan el 20%, se ajustan, aumentando el tamaño de muestra en estos estratos y disminuyéndolo en aquellos en que el error relativo es inferior, tratando de mantener fijo el tamaño calculado inicial a nivel de sección.

Paso 4. Con la finalidad de cubrir posibles incidencias de terreno (como empresas cerradas, fuera de ámbito, fuera de marco, no respuesta, entre otros), se calcula el tamaño con sobremuestreo para cada estrato. Sin embargo, para la EMRCL donde su muestra es de segunda fase y considerando el mínimo desfase de tiempo que existe entre el término de la recolección de la primera fase y el inicio de la segunda, se decide utilizar un promedio de las tasas de no logro del año 2022, referida a la encuesta mensual correspondiente al periodo base promedio anual 2016=100, minimizando la perdida de unidades debido a sus características. Se aplica la siguiente fórmula:

$$n_{hv}^{s} = \frac{n_{hv}}{1 - tnl_{hv_{2022}}}$$

Donde:

n_{hv}^{s} : Tamaño de muestra con sobremuestreo en la sección actividad h y tamaño v.

n_{hv} : Tamaño de muestra objetivo en la sección económica h y tamaño v.

$tnl_{hv_{2022}}$: Tasa promedio de no logro para el año 2022 de la encuesta mensual correspondiente al periodo base promedio anual 2016=100, según sección económica h y tamaño v.

4.7.2 Selección muestral

Posterior al cálculo del tamaño muestral, se realiza la selección de unidades de acuerdo con el siguiente procedimiento:

Paso 1. Se incluyen las empresas del estrato de inclusión forzosa en la muestra que provienen de la EERCL, las que se encuentran vigentes al periodo en estudio y respondieron la encuesta. Estas empresas son autorepresentadas y su probabilidad de selección es igual a 1.

Paso 2. Para las empresas del estrato de selección aleatoria dentro de cada sección y tamaño, se ordena el marco muestral en forma ascendente para la variable de la sección económica (B, C, D, …, Q, R) y descendente para el tamaño de empresa (mega[25], grande, mediana, pequeña)

Paso 3. Una vez ordenadas las empresas se realiza un muestreo sistemático aleatorio simple en cada estrato, con igual probabilidad de selección para todas las unidades pertenecientes al mismo estrato. Esta forma de ordenar y seleccionar en forma sistemática, se denomina muestreo aleatorio sistemático, el que asegura que en cada estrato se representen las empresas de todos los tamaños y que la probabilidad de selección sea la misma para todas las unidades de muestreo.

4.7.3 Distribución del tamaño muestral

En las tablas siguientes, se presenta el número de empresas en el marco muestral, el tamaño de muestra objetivo con su respectiva sobremuestra y la muestra efectivamente lograda en la recolección del Cambio de Año Base 2023=100. Donde se determina un tamaño de muestra objetivo de 1.943 empresas, de las cuales 88 son de estrato de inclusión forzosa (EIF) y 1.855 empresas del estrato de selección aleatoria (ESA). Finalmente, para lograr la muestra objetivo se debe gestionar en trabajo de campo 2.141 empresas, de los cual se logra una muestra de 87 empresas EIF y 1.851 unidades ESA, sumando una muestra efectiva de 1.938 unidades.

[25] Para la selección muestral existe una categoría adicional, Mega, que se construye con unidades que cuentan con 1.000 o más personas trabajadoras. Sin embargo, para fines de análisis la categoría Mega representa un subconjunto de la categoría Grande, por lo que para fines de este documento se considerarán las tres categorías descritas en la tabla 4.

Tabla 9. Tamaño muestral objetivo y efectiva según sección económica

Sección Económica	Marco Muestral	Muestra Objetivo			Muestra objetivo con sobremuestreo	Muestra Efectiva		
		Total	EIF	ESA		Total	EIF	ESA
Total	**5.278**	**1.943**	**88**	**1.855**	**2.141**	**1.938**	**87**	**1.851**
B	198	111	10	101	120	106	10	96
C	335	120	9	111	129	119	8	111
D	137	78	1	77	84	78	0	78
E	164	85	1	84	88	78	1	77
F	320	151	9	142	172	147	8	139
G	462	120	15	105	134	131	19	112
H	315	115	7	108	124	117	7	110
I	284	112	9	103	128	106	8	98
J	272	108	2	106	121	110	3	107
K	232	120	7	113	131	121	6	115
L	355	69	1	68	87	69	0	69
M	400	133	3	130	143	123	2	121
N	461	120	6	114	128	112	7	105
O	250	185	2	183	202	193	2	191
P	396	120	3	117	132	125	3	122
Q	455	120	2	118	136	124	2	122
R	242	76	1	75	82	79	1	78

Fuente: Instituto Nacional de Estadísticas (INE).

Tabla 10. Tamaño muestra objetivo y efectiva según tamaño de empresas

Tamaño de empresas	Marco Muestral	Muestra Objetivo			Muestra objetivo con Sobremuestreo	Muestra Efectiva		
		Total	EIF	ESA		Total	EIF	ESA
Total	**5.278**	**1.943**	**88**	**1.855**	**2.141**	**1.938**	**87**	**1.851**
Pequeña	2.629	503	-	503	560	483	-	483
Mediana	1.068	503	-	503	562	517	1	516
Grande	966	505	-	505	554	537	-	537
Mega	615	432	88	344	465	401	86	315

Fuente: Instituto Nacional de Estadísticas (INE).

Tabla 11. Tamaño muestral objetivo y efectiva según sección económica y tamaño de empresas

Sección Económica	Tamaño de empresas	Marco Muestral	Muestra Objetivo			Muestra objetivo con Sobremuestreo	Muestra Efectiva		
			Total	EIF	ESA		Total	EIF	ESA
Total		**5.278**	**1.943**	**88**	**1.855**	**2.141**	**1.938**	**87**	**1.851**
B	**Total**	**198**	**111**	**10**	**101**	**120**	**106**	**10**	**96**
	Pequeña	82	30	-	30	33	29	-	29
	Mediana	56	30	-	30	33	27	-	27
	Grande	39	30	-	30	33	28	-	28
	Mega	21	21	10	11	21	22	10	12
C	**Total**	**335**	**120**	**9**	**111**	**129**	**119**	**8**	**111**
	Pequeña	133	30	-	30	31	30	-	30
	Mediana	68	30	-	30	31	28	-	28
	Grande	78	30	-	30	34	28	-	28
	Mega	56	30	9	21	33	33	8	25
D	**Total**	**137**	**78**	**1**	**77**	**84**	**78**	**-**	**78**
	Pequeña	83	30	-	30	33	27	-	27
	Mediana	36	30	-	30	33	32	-	32
	Grande	17	17	-	17	17	16	-	16
	Mega	1	1	1	-	1	3	-	3
E	**Total**	**164**	**85**	**1**	**84**	**88**	**78**	**1**	**77**
	Pequeña	109	30	-	30	33	29	-	29
	Mediana	23	23	-	23	23	19	-	19
	Grande	24	24	-	24	24	22	-	22
	Mega	8	8	1	7	8	8	1	7
F	**Total**	**320**	**151**	**9**	**142**	**172**	**147**	**8**	**139**
	Pequeña	126	30	-	30	35	34	-	34
	Mediana	64	30	-	30	41	37	-	37
	Grande	80	45	-	45	48	45	-	45
	Mega	50	46	9	37	48	31	8	23
G	**Total**	**462**	**120**	**15**	**105**	**134**	**131**	**19**	**112**
	Pequeña	221	30	-	30	34	30	-	30
	Mediana	88	30	-	30	33	31	1	30
	Grande	100	30	-	30	31	34	-	34
	Mega	53	30	15	15	36	36	18	18
H	**Total**	**315**	**115**	**7**	**108**	**124**	**117**	**7**	**110**
	Pequeña	163	30	-	30	33	31	-	31
	Mediana	60	30	-	30	33	27	-	27
	Grande	67	30	-	30	33	34	-	34
	Mega	25	25	7	18	25	25	7	18
I	**Total**	**284**	**112**	**9**	**103**	**128**	**106**	**8**	**98**
	Pequeña	149	30	-	30	33	22	-	22
	Mediana	72	30	-	30	33	31	-	31
	Grande	41	30	-	30	40	30	-	30
	Mega	22	22	9	13	22	23	8	15
J	**Total**	**272**	**108**	**2**	**106**	**121**	**110**	**3**	**107**
	Pequeña	156	30	-	30	32	26	-	26

	Mediana	61	30	-	30	34	32	-	32
	Grande	45	38	-	38	45	37	-	37
	Mega	10	10	2	8	10	15	3	12
K	**Total**	**232**	**120**	**7**	**113**	**131**	**121**	**6**	**115**
	Pequeña	102	30	-	30	33	30	-	30
	Mediana	53	30	-	30	35	31	-	31
	Grande	47	30	-	30	33	32	-	32
	Mega	30	30	7	23	30	28	6	22
L	**Total**	**355**	**69**	**1**	**68**	**87**	**69**	**-**	**69**
	Pequeña	288	30	-	30	45	36	-	36
	Mediana	58	30	-	30	33	26	-	26
	Grande	8	8	-	8	8	7	-	7
	Mega	1	1	1	-	1	0	-	-
M	**Total**	**400**	**133**	**3**	**130**	**143**	**123**	**2**	**121**
	Pequeña	242	30	-	30	32	28	-	28
	Mediana	81	30	-	30	37	30	-	30
	Grande	62	58	-	58	59	53	-	53
	Mega	15	15	3	12	15	12	2	10
N	**Total**	**461**	**120**	**6**	**114**	**128**	**112**	**7**	**105**
	Pequeña	174	30	-	30	33	29	-	29
	Mediana	91	30	-	30	31	25	-	25
	Grande	117	30	-	30	32	30	-	30
	Mega	79	30	6	24	32	28	7	21
O	**Total**	**250**	**185**	**2**	**183**	**202**	**193**	**2**	**191**
	Pequeña	23	23	-	23	23	18	-	18
	Mediana	40	30	-	30	33	49	-	49
	Grande	55	30	-	30	33	57	-	57
	Mega	132	102	2	100	113	69	2	67
P	**Total**	**396**	**120**	**3**	**117**	**132**	**125**	**3**	**122**
	Pequeña	169	30	-	30	33	32	-	32
	Mediana	84	30	-	30	33	30	-	30
	Grande	89	30	-	30	33	32	-	32
	Mega	54	30	3	27	33	31	3	28
Q	**Total**	**455**	**120**	**2**	**118**	**136**	**124**	**2**	**122**
	Pequeña	249	30	-	30	31	26	-	26
	Mediana	67	30	-	30	33	30	-	30
	Grande	82	30	-	30	36	32	-	32
	Mega	57	30	2	28	36	36	2	34
R	**Total**	**242**	**76**	**1**	**75**	**82**	**79**	**1**	**78**
	Pequeña	160	30	-	30	33	26	-	26
	Mediana	66	30	-	30	33	32	-	32
	Grande	15	15	-	15	15	20	-	20
	Mega	1	1	1	-	1	1	1	-

Fuente: Instituto Nacional de Estadísticas (INE).

4.8 Factores de expansión

Los factores de expansión, en una muestra probabilística, afectan las magnitudes recabadas en una encuesta, de manera que puedan ser representativas de la población objetivo. Este cálculo es, técnicamente, el inverso de la probabilidad de selección de cada unidad. Sin embargo, luego de haber calculado dicha probabilidad de selección, se deben realizar ajustes para incorporar información proveniente del trabajo de campo, donde se reflejan las incidencias que impiden la recolección total de la muestra seleccionada.

4.8.1 Factor de expansión para las empresas del estrato aleatorio

4.8.1.1 Factor de expansión teórico

El factor de expansión teórico corresponde al inverso de la probabilidad de selección de la empresa i perteneciente a la actividad económica h y tamaño v, la cual se calcula mediante la siguiente fórmula:

$$P_{hvi} = \frac{n_{hv}^{s}}{N_{hv}}$$

Donde:

P_{hvi} : Probabilidad de selección de la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

n_{hv}^{s} : Tamaño de muestra con sobremuestreo de la EMRCL en la sección económica h y tamaño v.

N_{hv} : Número de empresas en el marco muestral de la EMRCL en la sección económica h y tamaño v.

Por lo tanto, el factor de expansión teórico se obtiene mediante la siguiente fórmula:

$$w_{hvi} = \frac{1}{P_{hvi}}$$

Donde:

w_{hvi} : Factor de expansión teórico de la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

P_{hvi} : Probabilidad de selección de la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

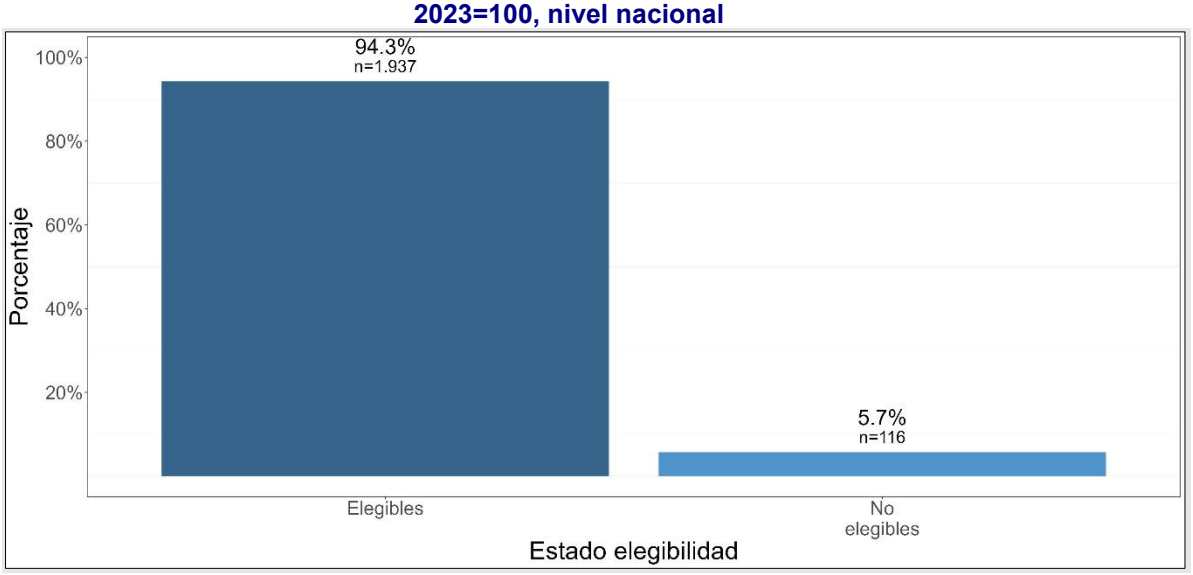
4.8.1.2 Ajuste por elegibilidad

Ajuste por elegibilidad desconocida

Durante el trabajo de campo, se presentan empresas que se clasifican según su estado de elegibilidad, diferenciando entre unidades “elegibles”, “no elegibles” o “elegibilidad desconocida”. A nivel nacional, como se observa en gráfico 3, las empresas elegibles corresponden a 1.937 empresas, cubriendo 94,3% de la muestra total. Mientras que, las empresas no elegibles alcanzan el 5,7%.

En este caso, no se presentan empresas con estado de elegibilidad desconocida, por lo cual no fue necesario realizar este ajuste.

Gráfico 3: Distribución estado de elegibilidad de las empresas recolectadas en el CAB 2023=100, nivel nacional

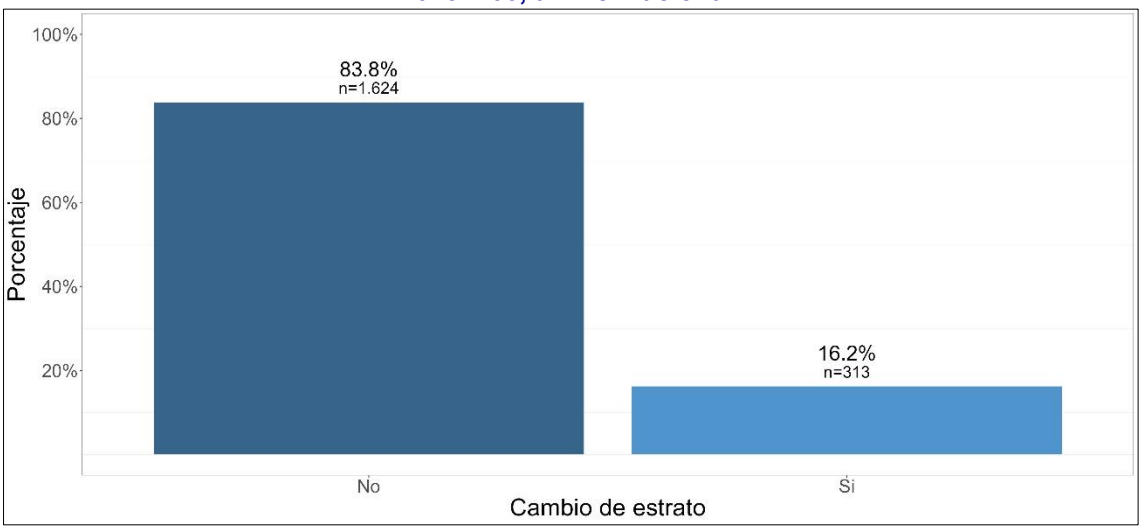


Fuente: Instituto Nacional de Estadísticas (INE).

Ajuste por cambio de estrato y no elegibilidad

En las encuestas se observa que la información de las variables de estratificación reportada en el marco muestral difiere del dato recolectado, de modo que la empresa puede no pertenecer al estrato en el que fue seleccionada inicialmente. En este caso, se presentan empresas con cambio de estrato por sección económica y tamaño de empresas, según el número de personas trabajadoras que ahora reporta la persona informante. Con el objetivo de analizar este ajuste, se determina que un cambio de estrato ocurre cuando al menos una de las dos variables del estrato registra algún cambio. Siendo así, en el gráfico 4, se observa que a nivel nacional 16,2% de las empresas elegibles cambiaron su estrato inicial.

Gráfico 4: Proporción de empresas elegibles que cambiaron de estrato en CAB 2023=100, a nivel nacional



Fuente: Instituto Nacional de Estadísticas (INE).

A su vez, al analizar los datos de la encuesta se identifican unidades clasificadas como “No elegibles” (ver gráfico 3). Ambos casos se deben principalmente al desfase entre la información contenida en el marco muestral y el trabajo de campo, lo cual se corrige realizando un ajuste por no elegibilidad y cambio de estratos en los factores de expansión. Por lo tanto, para realizar este ajuste se identifican los siguientes tres casos:

1. Si la empresa i es elegible, fue seleccionada en el estrato k, según la encuesta pertenece al estrato z y el estrato z es aleatorio, su factor de expansión no cambia.
2. Si la empresa i es elegible y sigue perteneciendo al estrato donde fue seleccionada, el factor de expansión es ajustado por cambio de estrato y no elegibilidad. Este ajuste consiste en descontar la proporción de empresas no elegibles entre las de elegibilidad conocida y descontar la suma de los factores de expansión de las empresas que fueron seleccionadas en el estrato k y según la encuesta pertenecen al estrato z.
3. Si la empresa i es no elegible, se asigna una ponderación nula.

Este ajuste se expresa en la siguiente fórmula:

$$w_{hv_i} = \begin{cases} w_{hv_i}^{ed} & , si\ i \in \Omega_{hv}^{el} \wedge i \in \Omega_{hv}^{k \to z} \\ \dfrac{N_{hv}\left(1-\frac{n_{hv}^{ne}}{n_{hv}^{ec}}\right) - \left(\sum_{k \neq z} \sum_{i=1}^{n_{hv}^{el,k \to z}} w_{hv_i}^{ed}\right)}{n_{hv}^{el,k \to k}} & , si\ \ i \in \Omega_{hv}^{el} \wedge i \in \Omega_{hv}^{k \to k} \\ . & , si\ i \in \Omega_{hv}^{ne} \end{cases}$$

Donde:

w_{hv_i} : Factor de expansión ajustado por cambio de estrato y no elegibilidad para la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

$w_{hv_i}^{ed}$: Factor de expansión ajustado por elegibilidad desconocida para la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

N_{hv} : Número de empresas en el marco muestral, pertenecientes a la sección económica h y tamaño v.

n_{hv}^{ne} : Número de empresas en la muestra, pertenecientes a la sección económica h y tamaño v, clasificada como no elegibles.

n_{hv}^{ec} : Número de empresas en la muestra, pertenecientes a la sección económica h y tamaño v, clasificada con elegibilidad conocida (elegibles y no elegibles).

$n_{hv}^{el,k \to k}$: Número de empresas en la muestra, pertenecientes a la actividad económica h y tamaño v, clasificada como elegibles y que han sido seleccionadas en el estrato k y después del trabajo de campo se mantienen en el mismo estrato.

Ω_{hv}^{el} : Conjunto de empresas en la muestra pertenecientes a la sección económica h y tamaño v, clasificadas como elegibles.

$\Omega_{hv}^{k \to z}$: Conjunto de empresas en la muestra que fueron seleccionadas en el estrato k y según la encuesta pertenecen al estrato z (cambio de estrato).

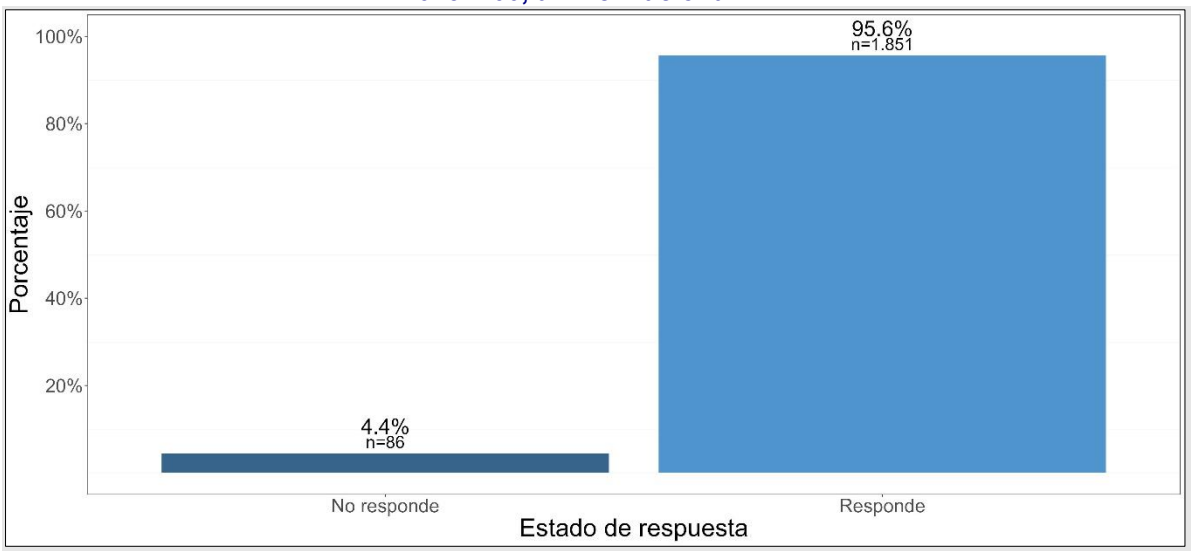
$\Omega_{hv}^{k \to k}$: Conjunto de empresas en la muestra que fueron seleccionadas en el estrato k y según la encuesta siguen perteneciendo al estrato k.

Ω_{hv}^{ne} : Conjunto de empresas en la muestra pertenecientes a la sección económica h y tamaño v, clasificadas como no elegibles.

4.8.1.3 Ajuste por no respuesta

Existen algunas empresas que, aun siendo elegibles, no proporcionan la información solicitada o lo hacen de forma incompleta, lo que se denomina “no respuesta”. Esto tiene como consecuencia que, al momento de realizar los análisis con el factor de expansión definido hasta el paso anterior, no se estaría dando cuenta de la totalidad de las empresas elegibles de la población. Se observa en el gráfico 5 que, de las empresas elegibles a nivel nacional un 95,6% responde la encuesta.

Gráfico 5: Proporción de empresas elegibles que respondió o no respondió en CAB 2023=100, a nivel nacional



Fuente: Instituto Nacional de Estadísticas (INE).

Una vez abordada las estrategias operativas necesarias para conseguir la información sin mayor éxito, se deben aplicar metodologías para corregir o mitigar los efectos en los estimadores. Esto implica realizar un ajuste por no respuesta en los factores de expansión. Se opta por la estrategia de ajuste por celda, que consiste en distribuir los pesos de las unidades que no responden entre las unidades que sí responden de forma independiente al interior de cada estrato. Para realizar este ajuste, primero se calcula una razón de ajuste por estrato, correspondiente a la suma de los factores de expansión de las unidades elegibles (que responden (R) y no responden (NR) la encuesta) sobre la suma de los factores de expansión de las unidades que respondieron la encuesta, tal como se observa en la siguiente fórmula:

$$\hat{R}_{hv}^{nr} = \frac{\sum_{i \in \Omega_{hv}^{el}} w_{hv_i}}{\sum_{i \in \Omega_{hv}^{el,R}} w_{hv_i}}$$

Donde:

$\hat{R}_{hv}^{nr}$: Razón de ajuste por no respuesta para las empresas pertenecientes a la sección económica h y tamaño v.

w_{hv_i} : Factor de expansión ajustado por cambio de estrato y no elegibilidad para la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

Ω_{hv}^{el} : Conjunto de empresas en la muestra pertenecientes a la sección económica h y tamaño v, clasificadas como elegibles.

$\Omega_{hv}^{el,R}$: Conjunto de empresas en la muestra pertenecientes a la sección económica h y tamaño v, clasificadas como elegibles y que respondieron la encuesta.

Posteriormente, para las unidades que respondieron la encuesta se multiplica el factor de expansión ajustado por cambio de estrato y no elegibilidad por la razón de ajuste por no respuesta, mientras que para el resto de las unidades se deja un valor perdido o *missing*.

$$w_{hv_i}^{nr} = \begin{cases} \hat{R}_{hv}^{nr} \cdot w_{hv_i} & , si \quad i \in \Omega_{hv}^{el,R} \\ . & , si \quad i \in \Omega_{hv}^{el,NR} \end{cases}$$

Donde:

$w_{hv_i}^{nr}$: Factor de expansión ajustado por no respuesta para la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

$\hat{R}_{hv}^{nr}$: Razón de ajuste por no respuesta para las empresas pertenecientes a la sección económica h y tamaño v.

w_{hv_i} : Factor de expansión ajustado por cambio de estrato y no elegibilidad para la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

$\Omega_{hv}^{el,R}$: Conjunto de empresas en la muestra pertenecientes a la sección económica h y tamaño v, clasificadas como elegibles y que respondieron la encuesta.

$\Omega_{hv}^{el,NR}$: Conjunto de empresas en la muestra pertenecientes a la sección económica h y tamaño v, clasificadas como elegibles y que no respondieron la encuesta.

4.8.1.2 Suavizamiento del factor de expansión

Tras los factores obtenidos en el paso anterior, se observa la presencia de valores atípicos, de modo que surge la necesidad de realizar un suavizamiento.
Antes de realizar el suavizamiento, se debe establecer o definir algún criterio o mecanismo para identificar valores atípicos en los factores de expansión. Para ello, se establece como regla que, al interior de cada estrato, todo aquel factor de expansión que se encuentre a dos o más desviaciones estándar respecto al factor de expansión promedio del estrato es considerado atípico. Es decir, el factor de expansión $w_{hv_i}^{nr}$ se considera atípico si está a una distancia superior a 2 desviaciones estándar respecto al promedio de su estrato. Si denotamos el factor atípico como A_{hv_i}, entonces A_{hv_i} no pertenece al intervalo I_{hv}, tal como se expresa en la siguiente fórmula:

$$A_{hv_i} \notin \mathrm{I}_{hv} = \left[\bar{w}_{hv}^{nr} - 2 \cdot \sigma_{w_{hv}}^{nr} \;, \bar{w}_{hv}^{nr} + 2 \cdot \sigma_{w_{hv}}^{nr}\right]$$

Donde:

A_{hv_i} : Factor de expansión de la empresa i en la muestra, considerado atípico, perteneciente a la sección económica h y tamaño v.
$\bar{w}_{hv}^{nr}$: Promedio de los factores de expansión ajustados por no respuesta del conjunto de empresas pertenecientes a la sección económica h y tamaño v.
$\sigma_{w_{hv}}^{nr}$: Desviación estándar de los factores de expansión del conjunto de empresas, pertenecientes a la sección económica h y tamaño v.

Una vez identificados los estratos con presencia de valores atípicos, se suavizan todos los factores de expansión al interior de estos, mientras que en aquellos estratos donde no hay presencia de valores atípicos, se mantiene el factor de expansión anterior. La fórmula para realizar el suavizamiento es la siguiente[26]:

$$w_{hv_i}^{s} = \begin{cases} \dfrac{\left(w_{hv_i}^{nr} + \bar{w}_{hv}^{nr}\right)}{2} & , si \quad i \in \left\{\Omega_{hv}^{el,R} \ni A_{hv_i}\right\} \\ w_{hv_i}^{nr} & , si \quad i \in \left\{\Omega_{hv}^{el,R} \not\ni A_{hv_i}\right\} \end{cases}$$

Donde:

$w_{hv_i}^{s}$: Factor de expansión suavizado de la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

[26] La fórmula utilizada para realizar el suavizamiento del factor de expansión consiste en obtener un promedio simple entre el factor de expansión y el promedio de los factores de expansión de su estrato. Esta fórmula puede ser modificada dependiendo si se desea dar más ponderación al promedio o al valor del factor de expansión de la empresa. En este caso, se entrega igual ponderación al promedio y al factor de expansión.

$w_{hv_i}^{nr}$: Factor de expansión ajustado por no respuesta para la empresa i en la muestra, perteneciente a la sección económica h y tamaño v.

$\overline{w}_{hv}^{nr}$: Promedio de los factores de expansión ajustados por no respuesta del conjunto de empresas de la sección económica h y tamaño v.

A_{hv_i} : Factor de expansión de la empresa i en la muestra, considerado atípico, perteneciente a la sección económica h y tamaño v.

$\Omega_{hv}^{el,R}$: Conjunto de empresas en la muestra pertenecientes a la actividad económica h y tamaño v, clasificadas como elegibles y respondieron la encuesta.

La fórmula de suavizamiento utilizada garantiza que el factor de expansión suavizado expanda a la misma población que el factor de expansión ajustado por no respuesta. Asimismo, garantiza mantener el orden de los factores de expansión previo y posterior al suavizamiento, es decir, si antes de realizar el suavizamiento $w_{hv_{i=1}}^{nr} > w_{hv_{i=2}}^{nr}$, una vez realizado el suavizamiento $w_{hv_{i=1}}^{s} > w_{hv_{i=2}}^{s}$.

4.8.1.3 Calibración del factor de expansión

La calibración de los factores de expansión implica ajustar los pesos de la muestra para que reflejen de manera más precisa las características de la población que se está estudiando. Esto se logra mediante la comparación de las características de la muestra y la población en diferentes variables.

Los métodos de calibración utilizan algoritmos iterativos que ajustan los factores de expansión con la finalidad de que alcancen totales poblacionales de variables auxiliares, minimizando la diferencia entre los factores de expansión originales y los calibrados. Esto permite asociar, por aproximación, cualquier propiedad óptima de los ponderadores iniciales a los calibrados. Asimismo, la calibración garantiza que las estimaciones de la encuesta sean coherentes con las estimaciones que ya han sido publicadas.

Tal como se mencionó anteriormente, no todas las empresas respondieron a la encuesta, por lo tanto, la suma de los factores de expansión de las unidades que sí respondieron no alcanza el mismo stock que el tamaño estimado de la población objetivo de la encuesta. Para corregir esto, se realiza un ajuste por no respuesta mediante un método de calibración por *raking*[27].

El método de calibración por *raking*, el cual se utiliza para este ajuste, asegura que los factores de expansión calibrados sean positivos, debido a que la función de distancia no se encuentra definida por valores negativos.
La función de distancia $G(w^c, w^s)$, se define mediante la siguiente fórmula:

[27] Para más información, revisar: Calibration Estimators in Survey Sampling (Deville & Särndal, 1992).

$$G(w^c, w^s) = w^c \log(w^c/w^s) - w^c + w^s$$

Donde:

$G(w^c, w^s)$	:	Función de distancia entre el factor de expansión calibrado y el factor de expansión suavizado.
w^c	:	Factor de expansión calibrado.
w^s	:	Factor de expansión suavizado.

Lo anterior permite obtener un factor de expansión calibrado a nivel de la marginal correspondiente, el cual se obtiene mediante la siguiente fórmula:

$$w^c = w^s \cdot F(w^c, w^s)$$

Donde:

w^c	:	Factor de expansión calibrado.
w^s	:	Factor de expansión suavizado.
$F(w^c, w^s)$	:	Función inversa de $\partial G(w^c, w^s)/\partial w^c$, también denotada por $g^{-1}(w^c, w^s)$.

Las marginales definidas para la calibración son las siguientes:

- Total de empresas por estrato (sección de actividad económica y tamaño de empresa), según estimación del total número de empresas obtenidos en la Encuesta Estructural de Remuneraciones y Costos Laborales (EERCL) 2021.
- Total de número de personas trabajadoras por estrato (tamaño de empresa), según estimación del total número de personas trabajadoras obtenidos en la EERCL 2021.

En línea con lo anterior y al igual que en la calibración de la EERCL, se realiza un tratamiento particular para la sección económica de “Administración Pública”, debido a que es una actividad que históricamente se comporta diferente al resto, por lo que se realiza una inspección exhaustiva del registro administrativo y de la muestra. En base a este análisis, se decide adicionar una variable auxiliar que permite alcanzar los stocks esperados. Por lo tanto, se agrega un nuevo nivel para definir las marginales, según pertenencia a municipalidad, donde “Administración Pública” puede tomar dos valores “municipal” y “no municipal”; mientras que el resto de las actividades económicas queda clasificado en “no aplica”. De esta manera, se crean las marginales definitivas para el caso de la administración pública:

- Total de empresas según sección de actividad económica, tamaño de empresa y pertenencia a municipalidad, de acuerdo con la estimación del total número de empresas obtenidos en la EERCL 2021.
- Total de número de personas trabajadoras, tamaño de empresa y pertenencia a municipalidad, de acuerdo con la estimación del marco muestral reconstruido con el año de referencia 2021.

En consecuencia, si bien las marginales se definieron para algunos de los dominios (actividad económica, tamaño de empresas y pertenencia a una unidad municipal), implícitamente quedan calibrados los totales nacionales para ambas variables al total de empresas y al total de número de trabajadores en los mismos dominios indicados.

Para realizar la calibración por *raking*, se utilizó la función *calibrate* en el software estadístico R.

4.8.2 Factor de expansión para las empresas de inclusión forzosa

Las unidades que pertenecen al estrato de inclusión forzosa ($j = 1$, inclusión forzosa; $j = 2$, inclusión aleatoria) tienen un factor de expansión igual a 1, ya que por definición se considera que su comportamiento es único, por lo que no pueden representar ni ser representadas por otras unidades.

4.9 Estimación de parámetros

La estimación estadística corresponde al procedimiento mediante el cual se calculan parámetros de la población a partir de los datos muestrales obtenidos. De este modo, para la EMRCL se realiza la estimación del total del número de personas trabajadoras, que es la variable de estudio de esta encuesta. En el siguiente apartado se describe el estimador y sus respectivas medidas de precisión, para posteriormente, presentar los resultados finales de las estimaciones y su error muestral.

La estimación del total del número de personas trabajadoras a nivel del cruce de sección económica y tamaño de empresas se obtiene mediante la siguiente fórmula:

$$\hat{Y}_{hv} = \sum_{j=1}^{2} \sum_{i=1}^{n_{hvj}} y_{hvji} \cdot w_{hvji}$$

Donde:

$\hat{Y}_{hv}$: Estimación del total del número de personas trabajadoras de la EMRCL para la sección económica h y tamaño v.

y_{hvji} : Número de personas trabajadoras de la EMRCL para la empresa i en la muestra, perteneciente a la sección económica h, tamaño v y tramo j.

w_{hvji}^{c} : Factor de expansión[28] de la empresa i en la muestra, perteneciente a la sección económica h, tamaño v y tramo j.

n_{hvj} : Número de empresas en la muestra, pertenecientes al tramo j de del grupo de tamaño v de la sección económica h.

A nivel de sección económica, la estimación del total del número de personas trabajadoras se obtiene mediante la siguiente fórmula:

$$\hat{Y}_{h} = \sum_{v=1}^{V} \hat{Y}_{hv}$$

Donde:

$\hat{Y}_{h}$: Estimación del total del número de personas trabajadoras de la EMRCL para la sección económica h.

$\hat{Y}_{hv}$: Estimación del total del número de personas trabajadoras de la EMRCL para la sección económica h y tamaño v.

[28] Representa al factor de expansión calibrado para las empresas de inclusión aleatoria y al factor de expansión igual a 1 para las empresas de inclusión forzosa.

A nivel de tamaño de empresas,

$$\hat{Y}_v = \sum_{h=1}^{H} \hat{Y}_{hv}$$

Donde:

$\hat{Y}_v$: Estimación del total del número de personas trabajadoras de la EMRCL para el tamaño de empresas v.

$\hat{Y}_{hv}$: Estimación del total del número de personas trabajadoras de la EMRCL para la sección económica h y tamaño v.

Asimismo, a nivel total nacional la estimación del número de personas trabajadoras se obtiene a partir de la siguiente fórmula:

$$\hat{Y} = \sum_{h=1}^{H} \sum_{v=1}^{V} \hat{Y}_{hv}$$

Donde:

$\hat{Y}$: Estimación del total del número de personas trabajadoras de la EMRCL a nivel total nacional.

$\hat{Y}_{hv}$: Estimación del total del número de personas trabajadoras de la EMRCL para la sección económica h y tamaño v.

Luego de calculadas las estimaciones, se obtienen las varianzas y errores muestrales de la estimación del indicador de interés, las cuales se estiman a nivel de sección económica y tamaño de empresas mediante la siguiente fórmula:

$$\hat{V}(\hat{Y}_{hv}) = \hat{N}_{hv} \,.\, (\hat{N}_{hv} - n_{hv}) \,.\, \frac{S^2(y)_{hv}}{n_{hv}}$$

Donde:

$\hat{V}(\hat{Y}_{hv})$: Estimación de la varianza de la estimación del total del número de personas trabajadoras en la sección económica h y tamaño v.

$\hat{N}_{hv}$: Estimación del número de empresas en la población, pertenecientes a la sección económica h y tamaño v.

n_{hv} : Número de empresas en la muestra efectiva, pertenecientes a la sección económica h y tamaño v.

$S^2(y)_{hv}$: Cuasi-varianza muestral de la variable del número de personas trabajadoras en la sección económica h y tamaño v.

Para obtener la varianza a nivel de la sección económica, se utilizan las siguientes fórmulas.

$$\hat{V}(\hat{Y}_h) = \sum_{v=1}^{V} \hat{V}(\hat{Y}_{hv})$$

Donde:

$\hat{V}(\hat{Y}_h)$: Estimación de la varianza de la estimación del total del número de personas trabajadoras en la sección económica h.

$\hat{V}(\hat{Y}_{hv})$: Estimación de la varianza de la estimación del total del número de personas trabajadoras en la sección económica h y tamaño v.

La estimación de la varianza a nivel del tamaño de empresa se determina de acuerdo con la siguiente fórmula:

$$\hat{V}(\hat{Y}_v) = \sum_{h=1}^{H} \hat{V}(\hat{Y}_{hv})$$

Donde:

$\hat{V}(\hat{Y}_v)$: Estimación de la varianza de la estimación del total del número de personas trabajadoras en el tamaño de empresa v.

$\hat{V}(\hat{Y}_{hv})$: Estimación de la varianza de la estimación del total del número de personas trabajadoras en la sección económica h y tamaño v.

A nivel total nacional,

$$\hat{V}(\hat{Y}) = \sum_{h=1}^{H} \sum_{v=1}^{V} \hat{V}(\hat{Y}_{hv})$$

Donde:

$\hat{V}(\hat{Y})$: Estimación de la varianza de la estimación del total del número de personas trabajadoras a nivel total nacional.

$\hat{V}(\hat{Y}_{hv})$: Estimación de la varianza de la estimación del total del número de personas trabajadoras en la sección económica h y tamaño v.

Por otro lado, para calcular el coeficiente de variación de la estimación del total del número de personas trabajadoras en cada estrato se utiliza la siguiente fórmula:

$$CV(\hat{Y}_{hv}) = \frac{\sqrt{\hat{V}(\hat{Y}_{hv})}}{\hat{Y}_{hv}}$$

Donde:

$CV(\hat{Y}_{hv})$: Coeficiente de variación de la estimación del total del número de personas trabajadoras para la sección económica h y tamaño v.

$\hat{V}(\hat{Y}_{hv})$: Estimación de la varianza de la estimación del total del número de personas trabajadoras en la sección económica h y tamaño v.

$\hat{Y}_{hv}$: Estimación del total del número de personas trabajadoras de la EMRCL para la sección económica h y tamaño v.

A nivel de sección económica:

$$CV(\hat{Y}_h) = \frac{\sqrt{\hat{V}(\hat{Y}_h)}}{\hat{Y}_h}$$

Donde:

$CV(\hat{Y}_h)$	:	Coeficiente de variación de la estimación del total del número de personas trabajadoras para la sección económica h.
$\hat{V}(\hat{Y}_h)$	:	Estimación de la varianza de la estimación del total del número de personas trabajadoras en la sección económica h.
$\hat{Y}_h$	:	Estimación del total del número de personas trabajadoras de la EMRCL para la sección económica h.

A nivel de tamaño de empresas:

$$CV(\hat{Y}_v) = \frac{\sqrt{\hat{V}(\hat{Y}_v)}}{\hat{Y}_v}$$

Donde:

$CV(\hat{Y}_v)$	:	Coeficiente de variación de la estimación del total del número de personas trabajadoras para el tamaño de empresas v.
$\hat{V}(\hat{Y}_v)$	:	Estimación de la varianza de la estimación del total del número de personas trabajadoras en el tamaño de empresa v.
$\hat{Y}_v$	:	Estimación del total del número de personas trabajadoras de la EMRCL para el tamaño de empresas v.

Finalmente, para el total nacional, se obtiene el coeficiente de variación a partir de la siguiente fórmula:

$$CV(\hat{Y}) = \frac{\sqrt{\hat{V}(\hat{Y})}}{\hat{Y}}$$

Donde:

$CV(\hat{Y})$	:	Coeficiente de variación de la estimación del total del número de personas trabajadoras a nivel del total nacional.
$\hat{V}(\hat{Y})$	:	Estimación de la varianza de la estimación del total del número de personas trabajadoras a nivel del total nacional.
$\hat{Y}$	:	Estimación del total del número de personas trabajadoras de la EMRCL a nivel del total nacional.

Asimismo, de acuerdo con la siguiente fórmula se calcula el error relativo de la estimación del total del número de personas trabajadoras para la EMRCL:

$$e(\hat{Y}_{hv}) = CV(\hat{Y}_{hv}) \cdot Z_{1-\frac{\alpha}{2}}$$

Donde:

$e(\hat{Y}_{hv})$: Error relativo de la estimación del total del número de personas trabajadoras en la EMRCL para la sección económica h y tamaño v.

$CV(\hat{Y}_{hv})$: Coeficiente de variación de la estimación del total del número de personas trabajadoras para la sección económica h y tamaño v.

$Z_{1-\frac{\alpha}{2}}$: Cuantil de la distribución normal estándar correspondiente a un nivel de confianza $(1 - \alpha)$.

A nivel de sección económica:

$$e(\hat{Y}_{h}) = CV(\hat{Y}_{h}) \cdot Z_{1-\frac{\alpha}{2}}$$

Donde:

$e(\hat{Y}_{h})$: Error relativo de la estimación del total del número de personas trabajadoras en la EMRCL para la sección económica h.

$CV(\hat{Y}_{h})$: Coeficiente de variación de la estimación del total del número de personas trabajadoras para la sección económica h.

$Z_{1-\frac{\alpha}{2}}$: Cuantil de la distribución normal estándar correspondiente a un nivel de confianza $(1 - \alpha)$.

A nivel de tamaño de empresas:

$$e(\hat{Y}_{v}) = CV(\hat{Y}_{v}) \cdot Z_{1-\frac{\alpha}{2}}$$

Donde:

$e(\hat{Y}_{v})$: Error relativo de la estimación del total del número de personas trabajadoras en la EMRCL para el tamaño de empresas v.

$CV(\hat{Y}_{v})$: Coeficiente de variación de la estimación del total del número de personas trabajadoras para el tamaño de empresas v.

$Z_{1-\frac{\alpha}{2}}$: Cuantil de la distribución normal estándar correspondiente a un nivel de confianza $(1 - \alpha)$.

A nivel de total nacional:

$$e(\hat{Y}) = CV(\hat{Y}) \cdot Z_{1-\frac{\alpha}{2}}$$

Donde:

$e(\hat{Y})$	:	Error relativo de la estimación del total del número de personas trabajadoras en la EMRCL para el nivel total nacional
$CV(\hat{Y})$	:	Coeficiente de variación de la estimación del total del número de personas trabajadoras a nivel del total nacional.
$Z_{1-\frac{\alpha}{2}}$	:	Cuantil de la distribución normal estándar correspondiente a un nivel de confianza $(1 - \alpha)$.

5. Metodología de la recolección de datos

5.1 Técnica y métodos de recolección de los datos

5.1.1 Método de recolección de los datos

El proceso de recolección de datos de la EMRCL se realizó a través del modo web de captura (*CAWI*[29]), para lo cual se habilitó una plataforma web con el formulario electrónico de la encuesta en un sistema dispuesto para ello. Esta plataforma web permite a cada unidad estadística que participa del estudio ingresar los datos de la encuesta en su respectivo formulario electrónico, o bien importarlos al sistema por medio de una réplica del formulario en formato Microsoft Excel. El acceso a la plataforma web se disponiblilizó a las personas informantes a través del sitio web institucional del INE.

El sistema web disponible para la recolección de datos incorporó un conjunto de validaciones básicas de completitud de información, formato y de cuadraturas de variables totalizadas, las cuales generan alertas y/o señales de error si estas no se cumplen al momento de intentar el envío de la encuesta. El conjunto de estas validaciones se denominan validaciones de primer nivel. Para la EMRCL, se estableció que un formulario con algún error de validación de primer nivel no podría enviarse exitosamente, mientras que, en caso contrario, la encuesta puede ser enviada y la información contenida en ella ingresa a las bases de datos de la operación estadística, para continuar con las siguientes etapas del proceso de recolección.

5.1.2 Validación operativa

Después que la encuesta del periodo de referencia cumple con todas las validaciones de primer nivel la persona informante puede hacer envío de la información para que sea recepcionada por la persona analista asignada. Desde este punto, el equipo de analistas de datos realiza un análisis de la información reportada mediante una planilla de análisis en formato Excel, que principalmente se centra en validadores de segundo nivel y variaciones. Se espera que tanto los indicadores principales de la encuesta como sus respectivas variaciones mensuales se encuentren dentro de un rango de parámetros esperados; si la persona analista detecta alertas en dicha planilla, se procede a consultar a la persona informante para corroborar la información o hacer las correcciones pertinentes, quedando respaldo y trazabilidad en el sistema de los cambios realizados.

[29] *CAWI, Computer Assisted Web Interviewing*. Entrevista web asistida por computador.

5.1.3 Períodos de referencia y recolección

La recolección del Cambio de Año Base (CAB) 2023=100 se inició en junio de 2023 finalizando en enero de 2024. Los períodos de referencia recolectados abarcaron todos los meses de 2023, entre enero y diciembre, con una periodicidad mensual con el siguiente detalle:

Tabla 12. Períodos de recolección y referencia CAB 2023

Períodos de recolección	Períodos de referencia
Junio 2023	Enero 2023
Julio 2023	Febrero 2023, marzo 2023, abril 2023
Agosto 2023	Mayo 2023, junio 2023, julio 2023
Septiembre 2023	Agosto 2023
Octubre 2023	Septiembre 2023
Noviembre 2023	Octubre 2023
Diciembre 2023	Noviembre 2023
Enero 2024	Diciembre 2023

Fuente: Instituto Nacional de Estadísticas (INE).

Adicionalmente, el Cambio de Año Base considera el levantamiento continuo de la nueva muestra de seguimiento, por tanto, a partir de febrero de 2024 comenzó la recolección mensual de los nuevos periodos de referencia, iniciándose esta en enero de 2024.

La recolección de cada período de referencia comienza el primer día hábil del periodo de recolección, mediante el envío de una carta de apertura a las unidades de la muestra. La persona informante a través de un usuario y contraseña pudo acceder a la aplicación informática para completar y enviar la encuesta en el plazo establecido para ello.

5.1.4 Conformación del equipo de recolección y supervisión

El proceso de recolección y supervisión del Cambio de Año Base 2023, se llevó a cabo con un equipo responsable de la recolección, el cual estuvo compuesto por los siguientes perfiles de cargo:

Tabla 13. Perfiles de cargo y cantidad de personas recolección

Perfil de cargo	Cantidad de personas
Coordinador(a) de recolección	1
Supervisor(a) de analistas de datos	2
Analistas de datos	16
Asistentes operativos	4
Analista de calidad	1

Fuente: Instituto Nacional de Estadísticas (INE).

Respecto al equipo de supervisión responsables de la supervisión operativa, este estuvo compuesto por los siguientes perfiles de cargo:

Tabla 14. Perfiles de cargo y cantidad de personas supervisión

Perfil de cargo	Cantidad de personas
Coordinador(a) de supervisión	1
Supervisor(a)	2

Fuente: Instituto Nacional de Estadísticas (INE).

El equipo operativo antes señalado estuvo a cargo de la recolección del año base y de la recolección continua hasta el periodo de referencia de febrero de 2024, abordando las unidades informantes exclusivas de la nueva muestra de seguimiento. A partir del periodo de referencia de marzo de 2024, se incorporó a las actividades de recolección de la nueva muestra el equipo operativo continuo del producto coyuntural vigente, siendo este último equipo el responsable de la recolección continua de la EMRCL a partir del periodo de referencia de junio 2024.

5.1.5 Carga laboral del equipo de analistas

La carga promedio asignada a cada persona analista de datos fue de 100 unidades aproximadamente durante todo el Cambio de Año Base 2023. Adicionalmente, se contó con la participación del equipo coyuntural de la encuesta mensual, quienes abordaron las unidades seleccionadas tanto en la muestra coyuntural actual como en la muestra del cambio de año base. El total de unidades gestionadas por el equipo operativo CAB fue de 1.600 unidades informantes (unidades nuevas), mientras que el equipo operativo de la encuesta mensual abordó un total de 541 unidades que ya estaban comprendidas en la muestra vigente del producto.

5.2 Estructura del instrumento de captura de la información

El cuestionario aplicado es único para todas las unidades de información y consta de un total de siete secciones, las que se detallan a continuación:

- Sección I: datos de la empresa o institución y de la persona informante.
- Sección II: mujeres con contrato directo. Esta sección se refiere solo a información de mujeres que recibieron pago de remuneraciones directamente de la empresa o institución en el mes de referencia.
- Sección III: hombres con contrato directo. Esta sección se refiere solo a información de hombres que recibieron pago de remuneraciones directamente de la empresa o institución en el mes de referencia.
- Sección IV: personal nacional y extranjero. Esta sección se recolectó para los periodos de referencia de junio de 2023 y diciembre de 2023. En ella se solicita

información de las remuneraciones de las personas trabajadores nacionales y personal extranjero.

- Sección V: personal contratado a honorarios. Esta sección se recolecta únicamente por las instituciones de la administración pública, y en ella se reportan las personas trabajadoras con contrato a honorarios.
- Sección VI: flujos laborales. Esta sección se refiere a la información de las nuevas contrataciones y desvinculaciones del personal del período de referencia.
- Observaciones: en esta sección la persona informante debe dejar toda información complementaria relevante para el estudio, ocurrida en el período de referencia.

El instrumento de captura de información descrito anteriormente se encuentra inserto íntegramente dentro del sistema de información web del producto, y en este se incorpora, adicionalmente, un instrumento complementario de uso interno, de trazabilidad de las gestiones realizadas por el equipo operativo respecto de los resultados obtenidos de cada intento de contacto con la unidad informante. Este instrumento se llamó Módulo de Registro de Intentos de Contacto, y contiene una estructura de códigos estandarizada y jerárquica para registrar el resultado de un intento de contacto.

5.3 Método de capacitación para recolección de datos

La estructura temática de las actividades de capacitación al equipo de recolección se efectuó según el flujo de proceso de la encuesta, en la que incorporó una capacitación para los equipos internos como también a las personas informantes.

Para las capacitaciones de las personas internas se contó con la asistencia de los equipos de recolección y supervisión, en modalidad presencial en dependencias institucionales del INE. Impartiéndose inicialmente a los siguientes perfiles de cargo: analistas de datos, supervisores, asistentes operativos y encargados de grupo, la que luego se impartió también a toda persona que se fue incorporando al equipo. La capacitación a todo el equipo es continua y se usa para reforzar aspectos en los que se identifican necesidades específicas como consecuencia de la implementación de nuevas directrices técnicas y/u operativas, con el fin de robustecer la información recolectada.

Los temas abordados en las capacitaciones de los equipos operativos fueron los siguientes:

- Gestión operativa en verificación
- Gestión operativa en recolección
- Gestión operativa en validación y análisis
- Revisión plantilla análisis de datos
- Gestión de observaciones de procesamiento y supervisión

Las capacitaciones fueron impartidas con apoyo de material audiovisual y desarrollo de casos prácticos.

La capacitación a las personas informantes se realizó de forma telemática, la que permitió aclarar dudas a través del chat de “Preguntas y Respuestas”. Las preguntas más frecuentes fueron compartidas con todas las unidades a través de correo electrónico, para facilitar la disposición de las aclaraciones. Como material de apoyo la capacitación efectuada quedó disponible en el acceso a personas informantes de la encuesta en página web de la institución[30].

5.4 Método de supervisión

La supervisión en el contexto del proyecto de cambio de año base tuvo como propósito garantizar la calidad de los datos recopilados durante todo el proceso de recolección.

El objetivo de la supervisión es asegurar la calidad del dato recolectado, minimizando errores no muestrales y garantizando que la información obtenida cumpla con los protocolos operativos establecidos en el proceso de recolección de información de la operación estadística.

5.4.1 Supervisión de gabinete

La supervisión aplicada en la operación estadística fue una supervisión de gabinete, la cual se ejecuta posterior a la aplicación del instrumento de recolección. La tarea de supervisión recae en las personas encargadas de grupo, no obstante, es la persona supervisora quien se encarga del proceso de registrar las observaciones o hallazgos en el sistema.

La supervisión de gabinete se realizó mediante un formulario de supervisión que permitió monitorear la información recabada para minimizar los potenciales errores en los datos informados o incidencias durante el trabajo de campo y, propiciar una

[30] Para acceder a la web entrar al enlace: https://www.ine.gob.cl/acceso-informantes/informacion-para-empresas-encuestadas

retroalimentación constante entre los equipos de recolección y supervisión durante el proceso de recolección.

5.4.2 Herramienta para la ejecución de la supervisión

El formulario de supervisión se diseñó para enfocar la supervisión en el cumplimiento de protocolos establecidos en la planificación de la operación en las distintas etapas de la recolección de datos. El detalle de estas y las principales actividades realizadas se presentan en la tabla 15.

Tabla 15: Supervisión de las etapas de recolección

Etapa	Actividades monitoreadas
Verificación	• Gestiones de primer contacto con la unidad (telefónico y por correo electrónico). • Gestiones de recuperación de información (búsqueda web). • Respaldo de información para unidades no elegibles.
Recolección	• Inducción a la encuesta por modo de contacto (telefónico y por correo electrónico). • Respaldo de correo electrónico de envío de carta de apertura. • Asistencia y soporte a la persona informante.
Análisis y validación	• Indicadores de recepción. • Gestiones de contacto (telefónico y correo electrónico). • Respaldo de correo electrónico con consultas de recepción y análisis. • Gestiones de seguimiento de la unidad. • Registro de observaciones por campo y en sistema. • Asistencia y soporte a la persona informante.
Procesamiento	• Consultas del equipo técnico. • Consultas del equipo de supervisión. • Gestiones de contacto (telefónico y correo electrónico). • Respaldo de correo electrónico para envío de consultas y ediciones de datos. • Asistencia y soporte al informante. • Gestión de observaciones en sistema.

Fuente: Instituto Nacional de Estadísticas (INE).

5.4.3 Actividades para la ejecución de la supervisión

A continuación, en la tabla 16, se describen las actividades de supervisión abordadas durante el proceso de recolección.

Tabla 16: Actividades desarrolladas en la supervisión

Actividad	Etapa	Descripción
Asignar carga de supervisión	Preparar la recolección	La persona encargada de supervisión realizó la asignación de carga de trabajo, utilizando un *script* de programación para identificar las unidades a supervisar. La metodología de recolección utilizada es CAWI y la modalidad de supervisión es de gabinete. La asignación de carga de supervisión desde la persona encargada a la supervisora o supervisor se realizó mediante la plataforma de Survey Solutions (SuSo).
Supervisión de gabinete	Ejecutar la supervisión	El supervisor(a) lleva a cabo la ejecución de la supervisión de gabinete, comenzando con la revisión y verificación de los protocolos establecidos por el equipo de recolección. Si los procedimientos no se cumplen, se gestiona con el equipo de recolección la recuperación de la información de manera oportuna.
Recuperación de información	Finalizar la supervisión	A partir del resultado de supervisión, la encargada o encargado de supervisión sube a sistema las observaciones (hallazgos detectados por el equipo de supervisión) que deben ser gestionadas por el equipo de recolección. En estas unidades se debió gestionar solicitudes de recuperación de información.
Cierre de la supervisión	Finalizar la supervisión	El procedimiento finaliza con la entrega de informes o bases de datos de supervisión al equipo de recolección, los cuales se enfocan en tratar específicamente los aspectos a mejorar. La reunión de retroalimentación se efectuó de manera mensual al equipo de recolección y se realizó una evaluación de hallazgos y mejoras.

Fuente: Instituto Nacional de Estadísticas (INE).

5.4.4 Porcentaje de la muestra a supervisar

Para la supervisión de la EMRCL, se selecciona aleatoriamente una muestra de encuestas que finalizaron la etapa de recolección considerando un mínimo de unidades a supervisar de un 10% según el avance diario o semanal del equipo de personas analistas de datos. A la muestra seleccionada se le aplica una supervisión de gabinete.

6. Procesamiento de datos

En esta sección se describen los procedimientos empleados en el procesamiento de los datos recopilados para los indicadores IR e ICL base 2023=100.

En el procesamiento de datos se consideran distintos subprocesos los cuales corresponden a la integración, validación, edición e imputación de datos. Cada nivel tiene un propósito estadístico específico que busca asegurar la calidad de los datos recolectados, para su posterior análisis y difusión. En este apartado se describen los niveles de procesamiento y sus alcances.

6.1 Integración y codificación de datos

Esta etapa consiste en la unión de los datos de las secciones principales de la encuesta disponibles en el sistema de información de la EMRCL, generando una base de datos (BBDD) para una posterior revisión y validación de datos.

Adicionalmente, se incorporan a la BBDD variables provenientes de RRAA de la Superintendencia de Seguridad Social y de Transparencia Activa para fines de ajustes en la Administración Pública (ver 6.3.1.).

6.2 Revisión y validación de datos

La revisión y validación de datos por parte del equipo técnico corresponde a una actividad que pretende asegurar la coherencia y consistencia económica de los datos recolectados mes a mes, a través de la definición de criterios de validación técnica, los cuales incorporan la detección de valores atípicos y la validación de variaciones inter-temporales de los indicadores y variables de interés. Los rangos establecidos para la detección de valores atípicos provienen del análisis de la información de la encuesta de remuneraciones y costo de la mano de obra base 2016=100,de los períodos de 2022, para cada perfil de empresa y grupo ocupacional.

También se incorpora en esta etapa una revisión de las acciones operativas realizadas en recolección respecto a los recontactos con la persona informante que permitieron confirmar o justificar el dato efectivo entregado.

La actividad de revisión y validación de datos se realiza en dos instancias con distintos objetivos: i) auditorías técnicas continuas durante la recolección de datos y ii) revisión final de datos para procesamiento.

6.2.1 Auditorias técnicas continuas

En paralelo con la etapa de recolección se realizó una revisión de calidad de la información a cargo del equipo técnico, en función del avance de la recolección, considerando una acumulación de periodos a revisar. Para ello se verificó y validó la consistencia económica de los datos, los cambios considerables en los totales de personas informadas, la revisión de variaciones de las variables principales derivadas de la encuesta, entre otros. En esta etapa, también se complementa la revisión técnica con el registro de observaciones de campo provenientes de la primera etapa de análisis de datos realizada en recolección. Los hallazgos del equipo técnico son derivados a la persona informante, a través de la gestión operativa, con el fin de corregir o justificar la información proporcionada por la unidad.

6.2.2 Revisión final de datos para procesamiento

En esta instancia el análisis técnico realiza la evaluación de datos finales de recolección luego de todos los intentos de corrección y/o justificación de datos realizados en el trabajo de campo. En esta etapa, la evaluación técnica tiene por objetivo determinar los casos que entran a procesamiento con datos efectivos y el conjunto de casos afectos a imputación.

6.3 Edición e imputación de datos

En este apartado se revisa el conjunto de acciones implementadas para corregir errores e inconsistencias en los datos recolectados a través de mecanismos de edición e imputación de datos. Se describirán los métodos implementados tanto para los meses correspondientes al año base como para los meses posteriores.

6.3.1 Edición de datos

Ajuste de pagos no mensuales en instituciones públicas en el año base

Este ajuste se implementó en los meses del año base en instituciones de la administración pública con anomalías en las remuneraciones ordinarias de marzo, junio, septiembre y diciembre a nivel de grupos ocupacionales. En la Administración pública, en los meses mencionados, se otorgan bonificaciones trimestrales que deben registrarse en la encuesta en la componente 9 de pagos no mensuales, la cual no es parte de la remuneración ordinaria. Sin embargo, se detectó que algunas instituciones registran erróneamente estas bonificaciones en las componentes de la remuneración ordinaria, lo que genera un aumento artificial durante estos meses de dicho concepto.

Para abordar esta problemática, se identificaron los grupos de ocupación de instituciones que presentan un aumento en la remuneración durante los meses mencionados. Posteriormente, se analizaron los componentes que más inciden en este incremento, determinando que se debe a la inclusión errónea de las bonificaciones.

A continuación, se calculó un promedio por hora ordinaria del componente con mayor incidencia, considerando los meses anterior y posterior al mes con bono (excepto para diciembre, donde solo se utiliza el mes anterior). Este valor promedio por hora se multiplicó por las horas ordinarias del mes con bono, obteniendo el nivel ajustado del componente problemático, como si no hubiera bonificación. Finalmente, la diferencia entre este valor ajustado y el valor original registrado se agrega a la componente 9 de pagos no mensuales, donde correspondía ser reportado.

Ajustes adicionales a la administración pública en el año base y adecuación para los meses posteriores

Este ajuste se realizó en determinadas instituciones de la administración pública, específicamente en municipalidades donde se detectó un subreporte de personal debido a que no se incluyó en los datos proporcionados en la encuesta el personal correspondiente a los departamentos de salud y/o educación que es contratado directamente por la municipalidad.

Este ajuste es similar al realizado en la Encuesta Estructural de Remuneraciones y Costos Laborales (EERCL) 2021, con algunas modificaciones para adaptarlo a las características de una encuesta mensual, y se basa en la incorporación de personas, horas, remuneraciones y costos laborales no reportados mediante el tratamiento de datos provenientes de Transparencia Activa y un mecanismo de integración.

El tratamiento de datos de Transparencia Activa incluye la codificación de sexo y grupo ocupacional del personal de salud y educación no reportado en los meses del año base siguiendo la metodología establecida en la EERCL[31]. Como en esta instancia la codificación y ajuste se debe realizar en cada mes del año base, se incorpora un criterio adicional para asignar un único grupo ocupacional por cada persona presente en el año base, seleccionando el grupo con mayor frecuencia, esto debido a que es posible que una misma persona tenga más de una clasificación de grupo ocupacional en meses distintos.

Los datos de transparencia activa, previamente codificados por sexo y grupo ocupacional, se utilizaron para ajustar los valores recolectados mensualmente, siempre que la institución cumpla con los siguientes requisitos en al menos tres períodos:

[31] Ver metodología en el siguiente link.

- El personal informado entre enero y diciembre de 2023 debe corresponder solo a personal de la administración municipal.
- El número de puestos de trabajo registrados en SUSESO debe ser similar al total de puestos de trabajo reportados en transparencia activa.[32]
- La municipalidad no debe tener departamentos o corporaciones de salud o educación con Rut propio.

Adicionalmente, solo se consideró para el ajuste aquellas instituciones a las que se les haya aplicado el mismo ajuste en la EERCL 2021.

Finalmente, la integración del personal faltante en los meses del año base sigue la misma metodología empleada en la EERCL 2021 la cual es descrita en el Documento Metodológico de la Encuesta Estructural de Remuneraciones y Costos Laborales 2021[33]; en tanto, el ajuste para los meses posteriores considerará factores de ajustes fijos determinados como el promedio de los factores de ajuste del año base, los cuales contemplan una revisión y actualización periódica.

6.3.2 Imputación de datos

Tratamiento de cadena larga no imputable en el año base según metodología de imputación de variables agregadas

Algunas unidades presentan cadenas (grupos ocupacionales) con datos con problemas de validación en todos los períodos enviados, imposibilitando, por este hecho, la imputación de variables agregadas definida para el año base y generando pérdida de información. Para abordar este problema particular, se implementó un tratamiento especial de imputación, permitiendo imputar valores problemáticos y conservar las cadenas correspondientes.

Se imputan dos variables que presentan problemas en todos sus períodos:

- **Error en las horas ordinarias (HO):** Se identificaron dos variables que presentan inconsistencias en todos sus períodos: Inconsistencias en el indicador de HONT y el indicador ROHO. Estas inconsistencias sugieren errores en el cálculo de las horas ordinarias. En el caso del indicador HONT, se detectó que en algunos casos las horas ordinarias reportadas correspondían a un período semanal en lugar de mensual. Esto generaba un valor de HONT artificialmente bajo y un valor de ROHO artificialmente alto.
- **Mal reporte de aportes patronales por remuneración imponible (APRI):** En el marco del análisis de datos, se detectaron inconsistencias en el reporte de aportes patronales por remuneración imponible (APRI). Se observó que en

[32] Para esto se calcula un ratio de la división entre el total en el portal de transparencia y Suseso, donde este valor debe estar entre 0.7 y 1.4.
[33] Publicada en el siguiente link.

algunos casos los valores de APRI superan el límite máximo permitido del 10,81%, establecido en las auditorías técnicas. Estas inconsistencias no se encuentran respaldadas por evidencia que justifique el aumento en los pagos de aportes, como aumentos en el pago de seguro social contra riesgo de accidentes del trabajo y enfermedades profesionales.

El tratamiento para estas se estableció según los siguientes pasos:

1. **Identificación de cadenas con problemas:** se identifican las cadenas que presenten datos incorrectos en todas las observaciones de las variables HONT o APRI.
2. **Búsqueda de donantes:** se buscan empresas con características similares a las cadenas con problemas (período de referencia, sexo, tamaño de empresa, sector económico, grupo ocupacional) que puedan proporcionar información válida sobre las variables HONT y APRI.
3. **Selección de donantes:** se seleccionan los donantes que cumplan con los criterios establecidos, considerando la posibilidad de flexibilizar estos criterios en caso de no encontrar un donante adecuado.
4. **Cálculo de la media del ratio**: para cada cadena con problemas, se calcula la media del ratio entre las variables HONT y APRI utilizando la información proporcionada por los *n* donantes seleccionados.
5. **Aplicación:** la media del ratio calculada en el paso anterior se aplica a las cadenas con problemas para imputar los valores de las variables HONT y APRI.
6. **Verificación:** se verifica que los valores imputados sean coherentes y se encuentren dentro de los rangos esperados. Para la primera edición, se comprueba que el nuevo valor de ROHO sea correcto, mientras que, para el segundo caso, se verifica que el indicador Costo Laboral por Hora Total (CLHT) esté dentro de los rangos.

Imputación de variables agregadas en el año base y adecuación para los meses posteriores

En este apartado se aborda el método de imputación para las variables agregadas de la EMRCL, esto es, *nt*, *ho*, *he*, *ro*, *re* y *og* (otros gastos laborales) implementado en los meses del año base y su adecuación para los meses posteriores. En términos generales, el método de imputación se basa en dos algoritmos de imputación elementales, arrastre y *hot-deck*, los cuales se explican a continuación[34]:

[34] En términos estrictos, la metodología de imputación no utiliza hot-deck en su sentido más clásico, sino una variante que toma de ideas de la imputación por razones. Para más detalles sobre el hot-deck clásico, la imputación por razones y otros métodos de imputación elementales, ver (De Waal, Pannekoek, & Scholtus, 2011).

Imputación por arrastre

En la imputación por arrastre, el valor faltante se imputa usando su valor en algún período de referencia (*v.gr.*, el período inmediatamente anterior).

Imputación por *hot-deck*

Existen varias variantes de la imputación por *hot-deck*. En este trabajo se utiliza una variante que explota tanto la información de empresas similares como la información de un período de referencia predefinido (*v.gr.*, el período inmediatamente anterior). Más concretamente, suponga que se desea imputar un dato perdido de una variable x (*v.gr.*, nt), usando una muestra de vecinos $\{(x_{i0}, x_{i1}): i = 1, \dots, N\}$, donde x_{i1} y x_{i0} representan el valor de x para el vecino i en los meses de interés (en lo que sigue, MI) y de referencia (en lo que sigue, MR), respectivamente. Entonces, si denotamos por x_{01} al dato a imputar y por x_{00} a su valor en el mes de referencia, el método de imputación por hot-deck utilizado en el CAB imputa x_{01} con

$$\hat{x}_{01} = \begin{cases} x_{00} \dfrac{\sum_{i \in I_1} x_{i1} / \#I_1}{\sum_{i \in I_0} x_{i0} / \#I_0} & \text{si } (\#I_1 > m_1) \wedge (\#I_0 > m_0) \wedge (\#I_1 / \#I_0 > \alpha), \\ \text{NA} & \text{en otro caso,} \end{cases}$$

donde $I_m := \{i > 0: x_{im} > 0\}$, los datos perdidos en la muestra de vecinos son ignorados, y (m_1, m_0, α) es una tupla de parámetros conocidos, la cual fue fijada en $(m_1, m_0, \alpha) = (4,\ 4,\ 0{,}5)$. Intuitivamente:

- m_1 es el mínimo # de vecinos con $x_{i1} > 0$ en el MI necesario para imputar.
- m_0 es el mínimo # de vecinos con $x_{i0} > 0$ en el MR necesario para imputar.
- α es la mínima fracción de vecinos (MI relativo a MR) necesaria para imputar.

En otras palabras, este método imputa el dato perdido multiplicando su valor en el período de referencia por un estimador del crecimiento de la variable entre el período de referencia y el de interés, pero solo si hay suficientes datos para hacerlo.

Note que los dos métodos de imputación recién expuestos asumen que tanto los vecinos del dato perdido como el mes referencia están dados. Sin embargo, en una muestra como la del CAB, no existe una sola manera de fijarlos. Respecto al conjunto de vecinos, se ha decidido siempre optar por el primer conjunto de variables que genere una imputación entre los siguientes conjuntos:

1. El conjunto de observaciones en la misma (categoría, grupo, tamaño de empresas, sexo).
2. El conjunto de observaciones en la misma (categoría, grupo, tamaño de empresas).

3. El conjunto de observaciones en la misma (categoría, grupo).

Respecto al mes de referencia, sin embargo, se ha optado por utilizar múltiples variantes. El siguiente ejemplo ayuda a entender por qué esto puede ser útil.

Ejemplo 1. Suponga que deseamos imputar el nt de un dato en febrero. ¿Cuál sería un mejor mes de referencia: enero o marzo?

El ejemplo anterior sugiere que, dependiendo del contexto, puede haber al menos dos meses de referencia útiles para imputar un dato faltante. La siguiente definición formaliza esta idea.

Definición 1. Si un método de imputación fija el mes de referencia como el mes anterior, se dice que el algoritmo imputa hacia adelante. Por otra parte, si es el mes posterior, se dice que imputa hacia atrás. Por conveniencia, estas variantes se abrevian como fw y bw, respectivamente.

Note que, incluso si decidimos cuáles serán los vecinos, sigue sin ser obvio si debiésemos usar o no las imputaciones previas como vecinos. El siguiente ejemplo ilustra este punto.

Ejemplo 2. Suponga que deseamos imputar el nt de un dato en marzo usando febrero como mes de referencia. ¿Debiésemos considerar las imputaciones de febrero como potenciales insumos para las imputaciones de marzo?

Así pues, además de distinguir la dirección de un método de imputación, debiésemos también distinguir si usa (o no) las imputaciones previas como insumos. La siguiente definición formaliza esta idea.

Definición 2. Si un algoritmo de imputación incluye datos previamente imputados entre los vecinos, se dice que el algoritmo es *recursivo.* En caso contrario, se dice que es *no-recursivo*. Por conveniencia, estas variantes se abrevian como rc y nrc, respectivamente.

Ya habiendo explicado los métodos básicos de imputación considerados para el año base, se procederá a explicar cómo se imputan concretamente las variables objetivo.

En términos generales, el proceso de imputación en el año base es el siguiente:

4. La muestra se divide entre roles de comportamiento común (CC) y único (CU), entendiendo como roles CU a aquellos que promediaron $nt > 1000$ en el año base (ignorando los meses sin datos).
5. Los datos de los roles CU son imputados por arrastre (rc/fw, luego rc/bw).
6. Los datos de los roles CC son imputados de un modo más complejo:
    1. La BBDD se imputa por hot-deck nrc/fw y nrc/bw (en paralelo).
    2. Los datos faltantes en la 1° BBDD creada en 3.1. se rellenan usando la 2°.

3. El resultado de 3.2. se imputa por hot-deck rc/fw.
4. El resultado de 3.3. se imputa por hot-deck rc/bw.
5. El resultado de 3.4. se imputa por arrastre rc/fw.
6. El resultado de 3.5. se imputa por arrastre rc/bw.

7. Los resultados de 2 y 3.6 se combinan en una única BBDD.

La idea de este algoritmo es imputar por hot-deck cuando sea posible, idealmente en su variante nrc, dejando espacio al arrastre y las variantes recursivas para casos en donde no haya suficientes datos para el hot-deck nrc. El algoritmo también incluye variantes bw a fin de (a) poder imputar los datos en enero y (b) posibilitar la imputación en situaciones no hay muchos datos en el período anterior pero sí en el período siguiente. Los datos de los roles CU siempre son imputados por arrastre pues, históricamente, no ha habido buenos vecinos para ellos.

Note que el pseudo-algoritmo anterior explica cómo se imputa la muestra *en general*, pero no explica cómo se imputan las variables objetivo *en particular*. Además, ignora que, tras versiones piloto del algoritmo de imputación, se decidió que algunas variables se imputarían siempre por arrastre. Así pues, se concluirá esta sección explicando específicamente en qué orden (y bajo qué transformaciones) se imputan las variables objetivo.

El siguiente diagrama explicita el proceso de imputación de las variables nt, og, ho, he, ro y re dentro de cada ítem del esquema general descrito en los párrafos anteriores:

Gráfico 6: Esquema de flujo de imputaciones

$$\overset{??}{\rightarrow} nt \begin{cases} \overset{??}{\rightarrow} hont \overset{r3}{\rightarrow} ho \overset{??}{\rightarrow} roho \overset{r3}{\rightarrow} ro \\ \overset{ar}{\rightarrow} hent \overset{r3}{\rightarrow} he \overset{ar}{\rightarrow} rehe \overset{r3}{\rightarrow} re \\ \overset{??}{\rightarrow} ognt \overset{r3}{\rightarrow} og \end{cases}$$

Fuente: Instituto Nacional de Estadísticas (INE).

donde el símbolo $x \overset{f}{\rightarrow} y$ denota que la variable y es imputada con el método f *después* de haber imputado la variable x, $r3$ denota regla de tres, ar denota arrastre, y el símbolo $??$ denota el método de imputación dominante en un paso concreto (hot-deck rc/fw para el paso 3.3 arrastre rc/bw para el paso 3.6). Así, por ejemplo, en el paso 3.3, la variable ro se imputaría a través de la siguiente cadena de imputaciones:

$$\overset{hrf}{\rightarrow} nt \overset{hrf}{\rightarrow} hont \overset{r3}{\rightarrow} ho \overset{hrf}{\rightarrow} roho \overset{r3}{\rightarrow} ro,$$

donde hrf es una abreviación del método *hot-deck* rc/fw. En otras palabras, primero se imputa nt usando hrf, luego $hont$ usando hrf, luego ho usando $r3$, luego $roho$ usando hrf, y finalmente ro usando $r3$.

Truncamiento de las imputaciones

Una vez imputadas las variables principales, se procede a truncarlas para que estas satisfagan todos los validadores de segundo nivel (*v.gr.*, que ho/nt esté dentro de los rangos pre-establecidos para cada dato). Más concretamente, el procedimiento de truncado opera introduciendo la variable auxiliar $aa := 1$ y actualizando las variables (secuencialmente) de la siguiente manera:

$$\begin{aligned} nt &\leftarrow trunc(nt, aa, lb_{ntaa}, ub_{ntaa}), \\ og &\leftarrow trunc(og, nt, lb_{ognt}, ub_{ognt}), \\ ho &\leftarrow trunc(ho, nt, lb_{hont}, ub_{hont}), \\ he &\leftarrow trunc(he, nt, lb_{hent}, ub_{hent}), \\ ro &\leftarrow trunc(ro, ho, lb_{roho}, ub_{roho}), \\ re &\leftarrow trunc(re, he, lb_{rehe}, ub_{rehe}), \end{aligned}$$

donde, dadas dos variables x e y, lb_{xy} y ub_{xy} son las cotas inferiores y superiores impuestas a la razón x/y, y la función $trunc(\cdot)$ de define como sigue:

$$trunc(x, y, lb, ub) := \begin{cases} y \times ub & \text{si } x/y > ub \\ y \times lb & \text{si } x/y < lb \\ x & \text{en otro caso} \end{cases}$$

Redondeo de las imputaciones

Una vez truncadas todas las variables imputadas, se procede a redondearlas para que arrojen valores enteros. Más concretamente, el procedimiento de redondeo actualiza las variables (secuencialmente) de la siguiente manera:

$$\begin{aligned} nt &\leftarrow 1 \vee [\lceil 10^6 nt \rceil / 10^6], \\ og &\leftarrow [\lceil 10^6 \times ognt \times nt \rceil / 10^6], \\ ho &\leftarrow [\lceil 10^6 \times hont \times nt \rceil / 10^6], \\ he &\leftarrow [\lceil 10^6 \times hent \times nt \rceil / 10^6], \\ ro &\leftarrow [\lceil 10^6 \times roho \times ho \rceil / 10^6], \\ re &\leftarrow [\lceil 10^6 \times rehe \times he \rceil / 10^6], \end{aligned}$$

donde todas las razones se calculan antes de actualizar las variables; $[\cdot]$ denota redondeo al entero más cercano, $\lceil \cdot \rceil$ denota redondeo hacia abajo, y $a \vee b$ denota el máximo entre a y b.

Porcentaje de imputación según variables y mes

En el marco del cambio de año base, se implementaron diversos métodos de imputación de datos para abordar la presencia de valores faltantes en la información recolectada.

El objetivo de este apartado es presentar los porcentajes de imputación por variable y mes, con respecto al total de valores no expandidos (muestrales) afectados por este proceso. Es importante destacar que los resultados de imputación provistos se calculan sobre la base de todos los cuestionarios que, tras la aplicación de los procedimientos de asignación de códigos de disposición final, fueron clasificados como "Elegibles, aptos para procesar".

En la tabla se observa que las variables con mayor porcentaje de imputación son Remuneración Extraordinaria y Horas Extraordinarias, presentando valores promedio de 10,4% y 10,2%, respectivamente. Por otro lado, las variables Puestos de trabajo y Otros Gastos exhiben los menores porcentajes de imputación, con promedios de 5,6% y 5,0%, respectivamente.

Tabla 17: Porcentaje de imputación según variable y mes para año base 2023

Mes	Puestos de trabajo	Horas Ordinarias	Remuneración Ordinaria	Costos Laborales	Remuneración Extraordinaria	Horas Extraordinarias	Otros Gastos
Enero	4,8%	5,2%	6,0%	6,4%	11,0%	8,6%	4,3%
Febrero	5,6%	6,0%	7,0%	7,5%	9,6%	8,6%	6,0%
Marzo	5,2%	6,1%	6,2%	6,7%	8,6%	8,3%	4,7%
Abril	4,1%	4,7%	5,3%	6,1%	11,2%	9,4%	4,1%
Mayo	5,9%	6,5%	6,8%	7,3%	9,8%	11,4%	7,2%
Junio	5,2%	5,7%	6,3%	7,1%	11,1%	12,3%	4,7%
Julio	5,8%	6,2%	6,9%	7,5%	10,1%	9,8%	5,7%
Agosto	6,6%	7,1%	6,4%	6,8%	11,5%	11,7%	4,5%
Septiembre	5,9%	6,5%	6,7%	7,2%	11,7%	11,1%	4,8%
Octubre	5,7%	6,2%	6,5%	7,1%	10,5%	10,3%	5,2%
Noviembre	5,7%	6,1%	6,3%	6,6%	9,0%	10,5%	4,2%
Diciembre	6,3%	6,8%	7,0%	7,5%	10,3%	10,7%	4,9%
Promedio	**5,6%**	**6,1%**	**6,5%**	**7,0%**	**10,4%**	**10,2%**	**5,0%**

Fuente: Instituto Nacional de Estadísticas (INE).

Imputación en período coyuntural

Para la imputación de los meses posteriores al año base, la metodología de imputación considera solo la implementación del algoritmo descrito anteriormente en su variante *hot-deck* fw (imputación hacia adelante) y rc (recursivo).

6.4 Asignación de códigos de disposición Final

Tras el procesamiento de la encuesta, se asignó un código de disposición final a cada unidad. Siendo las asignadas con el código “Encuesta apta para procesar” las que pasan a ser parte de la nueva muestra de seguimiento. Este código abarca las unidades con ‘Información completa’ como las con ‘Información parcial’. Las primeras son todas aquellas unidades que no fueron afectadas por ningún mecanismo de edición o imputación en ningún periodo del año base; en tanto las segundas corresponden a unidades que si fueron afectadas por algún mecanismo de edición o imputación en al menos un periodo del año base y en al menos una variable principal.

Otros códigos utilizados corresponden a “Cuestionario con información inconsistente” donde se incluyen encuestas que no cumplan con la calidad mínima establecida en la etapa de validación. Además, se incluyen los siguientes códigos:

- Rechazo de informante: asignado en casos que existió contacto con la persona informante a través de contacto telefónico, respuesta vía email o contacto presencial.
- Interrumpida, cuestionario con información insuficiente: corresponde a casos en donde la cantidad de periodos respondidos es menor al mínimo establecido de siete meses durante el año base o donde no existen respuestas en los periodos de octubre-noviembre-diciembre del 2023.
- Otra razón de no contacto: casos no contactados.
- Cese definitivo de actividades: se asignó este código cuando en el MRIC estaba presente el código de insolvencia.
- Paralizada temporalmente: se incluyeron los casos cuando hay evidencia que permita constatar que la empresa está paralizada o sin movimiento.
- Otros cambios en la unidad: cuanto el número de trabajadores promedio del año base no está dentro de la muestra objetivo, es decir, es menor a 5 personas.
- Otra razón para ser considerada no elegible: se asigna este código para otras situaciones de no elegibilidad no clasificadas previamente.

6.5 Calcular ponderadores

Los ponderadores de los indicadores de remuneraciones y costos laborales se calculan en base a la encuesta estructural de remuneraciones y costos laborales de 2021 junto con el cambio de año base 2023.

Se utilizarán N ponderadores asociados a una del total de cadenas[35] consideradas relevantes para el cálculo. Los criterios aplicados para considerar una cadena relevante para el cálculo son los siguientes:

1. La cadena debe tener información de la encuesta estructural 2021.
2. La cadena debe tener una participación mayor (o igual) a 0,002% en remuneraciones ordinarias y costos laborales en la encuesta estructural 2021.
3. La cadena debe tener información en la encuesta mensual del cambio de año base 2023.
4. La cadena debe tener 4 o más dadores en la encuesta mensual del año base 2023, se entiende como un dador a un rol de una empresa que tiene datos asociados a una determinada cadena corta.

6.5.1 Ponderadores para el cálculo de los índices[36]

1. Ponderador para el Índice de Remuneraciones (IR)

$$W\ IR_z = \frac{RO_z}{\sum_{i=z}^{Z} RO_z}$$

2. Ponderador para el Índice de Costos Laborales (ICL)

$$W\ ICL_z = \frac{CL_z}{\sum_{i=z}^{Z} CL_z}$$

Donde:

- z: corresponde a una combinatoria entre tamaño de empresas, actividad económica, sexo de las personas trabajadoras y grupo ocupacional.
- RO_z: es la sumatoria de las remuneraciones ordinarias expandidas para la cadena z, según la encuesta estructural 2021. Obtenidas como $RO_z = \sum_{z \in Z} ro_z * FE_z^Z$.
- CL_z: es la sumatoria de los costos laborales expandidos para la cadena z, según la encuesta estructural 2021. Obtenidas como $CL_z = \sum_{z \in Z} cl_z * FE_z^Z$.

[35] Una cadena corresponde a la agregación de los datos de una encuesta según algún atributo definido. Para más información ver apartado de conceptos claves de este documento.
[36] Los resultados del cálculo de ponderadores se encuentran incluidos en el apartado 11.2. Ponderadores Índices.

6.5.2 Ponderadores para el cálculo de las estimaciones[37]

Los ponderadores de las estimaciones de remuneración ordinaria por hora ordinaria, remuneración total por hora total, costo laboral por hora total se calculan en base a la encuesta estructural de remuneraciones y costos laborales de 2021 junto con la encuesta mensual de remuneraciones y costos laborales base 2023.

1. Ponderador para Remuneración Ordinaria por Hora Ordinaria (ROHO)

$$W\ horo_z = \frac{\left(\frac{HO}{RO}\right)_z}{\sum_{i=z}^{Z}\left(\frac{HO}{RO}\right)_z}$$

2. Ponderador para Remuneración Total por Hora Total (RTHT)

$$W\ htrt_z = \frac{\left(\frac{HT}{RT}\right)_z}{\sum_{i=z}^{Z}\left(\frac{HT}{RT}\right)_z}$$

3. Ponderador para Costo Laboral por Hora Total (CLHT)

$$W\ htc_z = \frac{\left(\frac{HT}{CL}\right)_z}{\sum_{i=z}^{Z}\left(\frac{HT}{CL}\right)_Z}$$

4. Ponderador para puestos de trabajo

$$W\ nt_z = \frac{NT_z}{\sum_{i=z}^{Z} NT_z}$$

Donde:

- HO_z: es la sumatoria de las horas ordinarias expandidas para la cadena z, según la encuesta estructural 2021. $HO_z = \sum_{z\in Z} ho_z * FE_z^Z$.
- HT_z: es la sumatoria de las horas totales expandidas para la cadena z, según la encuesta estructural 2021. $HT_z = \sum_{z\in Z} ht_z * FE_z^Z$.
- RO_z: es la sumatoria de las remuneraciones ordinarias expandidos para la cadena z, según la encuesta estructural 2021. $RO_z = \sum_{z\in Z} ro_z * FE_z^Z$.

[37] Los resultados del cálculo de ponderadores se encuentran incluidos en el apartado 11.3. Ponderadores Estimaciones.

- RT_z: es la sumatoria de las remuneraciones totales expandidas para la cadena z, según la encuesta estructural 2021. $RT_z = \sum_{z \in Z} rt_z * FE_z^Z$.

- CL_z: es la sumatoria del costo laboral expandido para la cadena z, según la encuesta estructural 2021. $CL_z = \sum_{z \in Z} cl_z * FE_z^Z$.

- NT_z: es la sumatoria de los puestos de trabajo expandidos para la cadena z, según la encuesta estructural 2021. $NT_z = \sum_{z \in Z} nt_z * FE_z^Z$.

7. Análisis de resultados

7.1 Generación de resultados

El principal análisis de los resultados de los indicadores está contenido en el boletín mensual. En el boletín se incluyen las descripciones de las evoluciones en doce meses y sus respectivos resultados de manera agregada y desagregada.

Las cifras de los indicadores principales y estimaciones se entregan junto con las variaciones mensuales, acumuladas y en doce meses. También se entrega la información de sus desagregaciones las que contienen las respectivas incidencias.

Actualmente los principales indicadores: IR nominal, ICL nominal e IR real se les realiza el proceso de ajuste estacional, método que descompone la serie principal en una serie desestacionalizada y otra de tendencia-ciclo, permitiendo aislar el efecto de factores estacionales que ocurren al comparar mensualmente los resultados.

7.1.1 Cálculo de los índices

El cálculo de los indicadores se realiza utilizando los datos que provee la encuesta mensual y los ponderadores[38] obtenidos de la encuesta estructural 2021. El IR e ICL tienen base promedio 2023=100 y utilizan ponderadores fijos que hacen referencia al año 2021.

El tipo de índice utilizado para el cálculo del IR e ICL corresponde a “Índices Laspeyres de precios modificado”. En este contexto, las empresas se identifican como los demandantes de personas trabajadoras; el precio, como el valor de los salarios y costos laborales por hora; y las cantidades consumidas se expresan en términos de las horas pagadas. En el concepto canasta, las funciones realizadas por las personas trabajadoras según grupo ocupacional y sexo, en todas las actividades económicas incluidas en los índices y los tres tamaños empresa.

El período de referencia de los índices comprende desde enero de 2023 hasta el mes anterior de la recolección de datos. Luego de la primera publicación, el período de referencia de los datos de la encuesta es el mes anterior a aquel que se lleva a cabo el levantamiento.

A continuación, se describirán en términos generales el cálculo de los índices, mientras que en el apartado siguiente se hará hincapié en la diferencia entre el cálculo provisorio y rectificado.

[38] Ver ponderadores desagregados en anexo 1.

7.1.1.1 Índice nominal de remuneraciones

El índice nominal de remuneraciones mide la evolución nominal de las remuneraciones ordinarias por hora ordinaria pagadas a las personas trabajadoras directamente contratados por las empresas.

i. Cálculo de remuneraciones ordinarias por hora ordinaria

$roho_k^m = \frac{RO_k^m}{HO_k^m}$,

Donde:

k: Corresponde a la característica tamaño de empresas t, sección económica c, sexo s, grupo ocupacional g, y empresa e

RO_k^m: Es la sumatoria de las remuneraciones ordinarias para el mes m a nivel de tamaño de empresas t, sección económica c, sexo s, grupo ocupacional g, y empresa e

HO_k^m: Horas ordinarias para el mes m a nivel de tamaño de empresas t, sección económica c, sexo s, grupo ocupacional g, y empresa e

ii. Cálculo índice básico

$$IBGE_{k(m-1)}^m = \frac{roho_k^m}{roho_k^{m-1}} * 100$$

Donde:

$roho_k^m$: Son las remuneraciones ordinarias por hora ordinaria para el mes m a nivel de tamaño de empresas t, sección económica c, sexo s, grupo ocupacional g, y empresa e.

$roho_k^{m-1}$: Son las remuneraciones ordinarias por hora ordinaria para el mes $m-1$ a nivel de tamaño de empresas t, sección económica c, sexo s, grupo ocupacional g, y empresa e.

iii. Agregación a nivel de grupo ocupacional

Se calcula la media geométrica de la agregación a nivel de empresa, nivel de tamaño de empresas t, sección c, sexo s y grupo ocupacional g.

-Índice por grupo ocupacional (IGO)

$$IGO_{t,c,s,g}^m = \prod_{t,c,s,g,e \in t,c,s,g} \left(IBGE_{t,c,s,g,e(m-1)}^m\right)^{\frac{1}{n(t,c,s,g)}}$$

Donde:

$n(t, c, s, g)$: Es el número agregaciones a nivel de empresa que pueden ser agregadas a nivel de tamaño de empresas t, sección económica c, sexo s y grupo ocupacional g

-Micro Índice en el nivel Grupo Ocupacional (MIGO)[39]

$$MIGO^{m}_{t,c,s,g} = IGO^{m}_{t,c,s,g} * \frac{MIGO^{(m-1)}_{t,c,s,g}}{100}$$

-Promedio anual MIGO, año base 2023[40]

$$P_MIGO^{2023}_{t,c,s,g} = \frac{\sum_{t=1}^{12} MIGO^{m,2023}_{t,c,s,g}}{12}$$

-MIGO recalculado[41]

$$R_MIGO^{m}_{t,c,s,g} = \frac{MIGO^{m}_{t,c,s,g} * 100}{P_MIGO^{2023}_{t,c,s,g}}$$

iv. <u>IR general</u>

$$\boldsymbol{IR^{m}} = \sum_{z=1}^{Z} (R_MIGO^{m}_{z} * W_IR_{z})$$

v. <u>IR a nivel de sección económica</u>

$$IR_sección^{m}_{c} = \sum_{z \in c} \left(R_MIGO^{m}_{z} * W_IR_{z,c}\right)$$

vi. <u>IR a nivel de tamaño de empresas</u>

$$IR_tamaño^{m}_{t} = \sum_{z \in t} \left(R_MIGO^{m}_{z} * W_IR_{z,t}\right)$$

vii. <u>IR a nivel de grupo ocupacional</u>

[39] Se usa a partir del cálculo de enero de 2024, debido a que el año base se utilizó el MIGO recalculado (descrito en el siguiente apartado)
[40] Solo se utiliza en el cálculo del año base 2023.
[41] Solo se utiliza en el cálculo del año base 2023.

$$IR_grupo_g^m = \sum_{z \in g} \left(R_MIGO_z^m * W_IR_{z,g}\right)$$

viii. IR según sexo

$$IR_sexo_s^m = \sum_{z \in s} \left(R_MIGO_z^m * W_IR_{z,s}\right)$$

7.1.1.2 Índice real de remuneraciones

El índice real de remuneraciones mide la evolución real de las remuneraciones por hora ordinaria en el país. Matemáticamente se puede expresar como:

$$\boldsymbol{IR_real^m} = IR_real^{m-1} * \frac{(1 + \Delta IR^m)}{(1 + \Delta IPC^m)}$$

Donde:

IR_real^m: Es el índice de remuneraciones real en el mes m.

IR_real^{m-1}: Es el índice de remuneraciones real en el mes anterior $m - 1$.

ΔIR^m: Es la variación mensual del índice nominal de remuneraciones en el mes m.

ΔIPC^m: Es la variación mensual del índice de precios al consumidor en el mes m.

Cabe destacar, que esta fórmula se utiliza desde el cálculo del IR real correspondiente a enero de 2019, debido a que durante esa fecha se realizó el cambio de año base en el IPC (IPC 2018=100). El objetivo de usar esta fórmula es evitar los empalmes derivados de los cambios de año base de los indicadores.

7.1.1.3 Índice nominal de costos laborales

El índice nominal de costos laborales mide la evolución nominal de los costos laborales por hora total en que incurren las empresas para mantener a sus empleados

i. Cálculo del costo por hora total

$$cht_k^m = \frac{C_k^m}{HT_k^m}$$

Donde:

k: Corresponde a la característica tamaño de empresas t, sección económica c, sexo s, grupo ocupacional g, y empresa e.

C_k^m: Es el costo laboral para el mes m a nivel de tamaño de empresas t, sección económica c, sexo s, grupo ocupacional g y empresa e.

HT_k^m: Corresponde a la suma de las horas ordinarias y las extraordinarias, para el mes m a nivel de tamaño de empresas t, sección económica c, sexo s, sexo s, grupo ocupacional g y empresa e.

ii. Cálculo índice básico

$$IBGE_k^m = \frac{cht_k^m}{cht_k^{m-1}} * 100$$

iii. Agregación a nivel de grupo ocupacional

Se calcula la media geométrica de la agregación a nivel de empresa, a nivel de tamaño de empresas t, sección c, sexo s y grupo ocupacional g.

-Índice por Grupo Ocupacional (IGO)

$$IGO_{t,c,s,g}^m = \prod_{t,c,s,g,e \in t,c,s,g} \left(IBGE_{t,c,s,g,e(m-1)}^m\right)^{\frac{1}{n(t,c,s,g)}}$$

Donde:

$n(t,c,s,g)$: Es el número agregaciones a nivel de empresa que pueden ser agregadas a nivel de tamaño de empresas t, sección económica c, sexo s y grupo ocupacional g.

-Micro Índice en el nivel Grupo Ocupacional (MIGO) en el mes m[42]

$$MIGO_{t,c,s,g}^{m} = IGO_{t,c,s,g}^{m} * \frac{MIGO_{t,c,s,g}^{(m-1)}}{100}$$

-Promedio anual 2023 MIGO[43]

$$P_MIGO_{t,c,s,g}^{2023} = \frac{\sum_{t=1}^{12} MIGO_{t,c,s,g}^{m}}{12}$$

-MIGO recalculado[44]

$$R_MIGO_{t,c,s,g}^{m} = \frac{MIGO_{t,c,s,g}^{m} * 100}{P_MIGO_{t,c,s,g}^{2023}}$$

iv. <u>ICL general</u>

$$ICL^{m} = \sum_{z=1}^{Z} \left(R_MIGO_{z}^{m} * W_{ICL_z}\right)$$

v. <u>ICL a nivel de sección económica</u>

$$ICL_sección_{c}^{m} = \sum_{z \in c} \left(R_MIGO_{z}^{m} * W_ICL_{z,c}\right)$$

vi. <u>ICL a nivel de tamaño de empresas</u>

$$ICL_tamaño_{t}^{m} = \sum_{z \in t} \left(R_MIGO_{z}^{m} * W_ICL_{z,t}\right)$$

vii. <u>ICL a nivel de grupo ocupacional</u>

$$ICL_grupo_{g}^{m} = \sum_{z \in g} \left(R_MIGO_{z}^{m} * W_ICL_{z,g}\right)$$

viii. <u>ICL según sexo</u>

$$ICL_sexo_{s}^{m} = \sum_{z \in s} \left(R_MIGO_{z}^{m} * W_ICL_{z,s}\right)$$

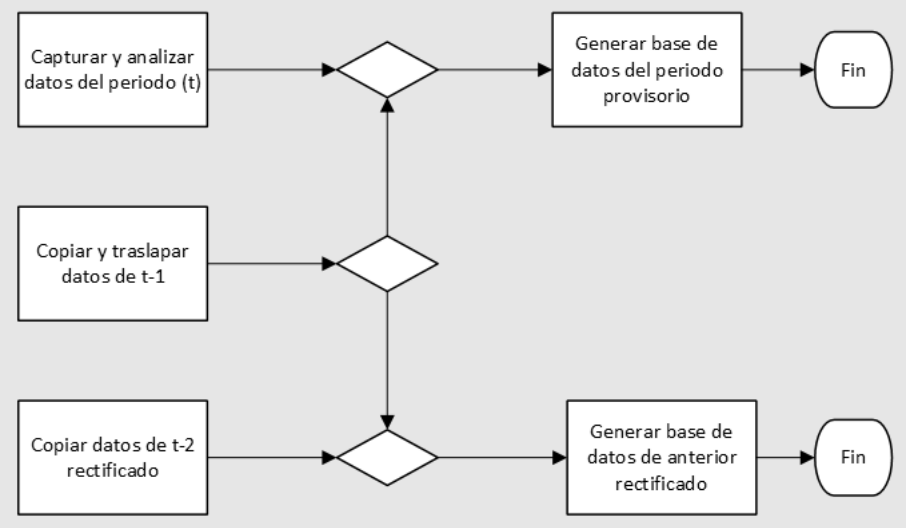
[42] Se usa a partir del cálculo de enero de 2024, debido a que el año base se utilizó el MIGO recalculado (descrito en el siguiente apartado)
[43] Solo se utiliza en el cálculo del año base 2023.
[44] Solo se utiliza en el cálculo del año base 2023.

7.1.2 Traslapes y política de rectificación de cifras

La política de rectificación de cifras del IR e ICL consiste en actualizar, por única vez en el mes de análisis (período t), la cifra del índice (nivel) del mes anterior (período t-1). Consecuentemente se actualizan las variaciones mensuales, acumuladas e interanuales, que fueron provisorias en la publicación del mes anterior, por su resultado final. La cifra del mes de análisis (período t) junto con sus variaciones mantienen el carácter de "provisionales" hasta la siguiente publicación de los indicadores.

La rectificación de datos se realiza por medio del método de traslape. Sin embargo, el procedimiento de generación de las bases de cálculo se resume en el gráfico 7.

Gráfico 7: Esquema de generación de bases de cálculo



Fuente: Instituto Nacional de Estadísticas (INE).

Al incorporar los resultados de una encuesta reciente en el índice del periodo actual, que aún se considera provisorio, se requiere modificar la información correspondiente al período anterior. Esta revisión implica ajustes tanto en el índice rectificado como en el índice provisorio para que ambos reflejen las últimas evidencias y correcciones.

Los motivos principales que conducen a la rectificación de cifras cada mes son:

- Rectificaciones por parte de las personas informantes: las entidades que suministran datos pueden corregir la información previamente entregada, afectando variables de la encuesta como, por ejemplo, los componentes de las remuneraciones para un grupo ocupacional específico.
- Cambios en las unidades estadísticas: se realizan actualizaciones debido a alteraciones estructurales como fusiones, absorciones o divisiones de empresas, siempre siguiendo los criterios de calidad y representatividad del Instituto Nacional de Estadísticas.

- Integración de respuestas tardías: las respuestas de la encuesta que no llegan a tiempo para ser incluidas en el mes correspondiente se procesan en el siguiente periodo de análisis. Como resultado, se ajusta el índice del mes previo para incluir estos datos recién recibidos.

A pesar de seguir una metodología establecida en la construcción de índices, los tipos de información sujetos a rectificación pueden variar. Los pasos para el cálculo del índice, tanto provisional como rectificado, se describen a continuación:

Tabla 18: Construcción IR provisorio versus rectificado

índice provisorio	**índice rectificado**
ROHO	
$ROHO_{CL}^{P} = \frac{RO^{P}}{HO^{P}}$	$ROHO_{CL}^{R} = \frac{RO^{R}}{HO^{R}}$
IBGE	
$IBGE_{CL}^{P} = \frac{ROHO^{P}}{ROHO^{R}} * 100$	$IBGE_{CL}^{R} = \frac{ROHO^{R}}{ROHO^{T-2}} * 100$
IGO	
$IGO_{cc}^{P} = \prod_{k \in cl} (IBGE_{k(R)}^{P})^{\frac{1}{n(cl)}}$	$IGO_{cc}^{R} = \prod_{k \in cl} (IBGE_{k(T-2)}^{R})^{\frac{1}{n(cl)}}$
MIGO	
$MIGO_{CC}^{P} = IGO_{CC}^{P} * \frac{MIGO_{CC}^{R}}{100}$	$MIGO_{CC}^{R} = IGO_{CC}^{R} * \frac{MIGO_{CC}^{T-2}}{100}$
IR General	
$IR^{P} = \sum_{Z=1} (MIGO_{Z}^{P} * W_{IR_Z})$	$IR^{R} = \sum_{Z=1} (MIGO_{Z}^{R} * W_{IR_Z})$

Tabla 19: Construcción ICL provisorio versus rectificado

índice provisorio	índice rectificado
CHT	
$CHT_{CL}^{P} = \frac{C^{P}}{HT^{P}}$	$CHT_{CL}^{R} = \frac{C^{R}}{HT^{R}}$
IBGE	
$IBGE_{CL}^{P} = \frac{CHT^{P}}{CHT^{R}} * 100$	$IBGE_{CL}^{R} = \frac{CHT^{R}}{CHT^{T-2}} * 100$
IGO	
$IGO_{cc}^{P} = \prod_{k \in cl} (IBGE_{k(R)}^{P})^{\frac{1}{n(cl)}}$	$IGO_{cc}^{R} = \prod_{k \in cl} (IBGE_{k(T-2)}^{R})^{\frac{1}{n(cl)}}$
MIGO	
$MIGO_{CC}^{P} = IGO_{CC}^{P} * \frac{MIGO_{CC}^{R}}{100}$	$MIGO_{CC}^{R} = IGO_{CC}^{R} * \frac{MIGO_{CC}^{T-2}}{100}$
ICL General	
$ICL^{P} = \sum_{Z=1} (MIGO_{Z}^{P} * W_{IR_Z})$	$ICL^{R} = \sum_{Z=1} (MIGO_{Z}^{R} * W_{IR_Z})$

Donde:

^{P}: Provisorio

^{R}: Rectificado

$^{T-2}$: Índice publicado (y rectificado) del periodo T-2

CC: Cadena corta (t,c,s,g)

CL: Cadena larga (t,c,s,g,e)

7.1.3 Empalme de las series

El cambio de base de los índices implica que las series publicadas con la base previa quedan descontinuadas. Sin embargo, estas rupturas temporales del índice se resuelven empalmando (enlazando o uniendo) los niveles de la base anual 2009=100 y 2016=100 al nivel de la nueva base anual 2023=100, conservando las variaciones mensuales publicadas.

Series empalmadas generales

Con respecto a la serie empalmada de los índices generales, se aplica la misma metodología a cada uno de los índices: IR, ICL e IR real.

Construcción del factor de enlace:

Para obtener las series empalmadas de los índices generales IR e ICL, se multiplica cada uno de los niveles mensuales del índice empalmado[45] respectivo con base anual 2016=100 (entre diciembre de 2009 y marzo de 2024) por el factor de enlace correspondiente, obteniendo como resultado la serie empalmada del IR e ICL con base anual 2023=100. En el caso particular del IR real, el empalme utiliza la misma técnica, pero la serie empalmada abarca desde enero 2010 hasta abril de 2024.

Series empalmadas desagregadas:

Los indicadores IR e ICL tienen desagregaciones según tamaño de empresas, sector económico, sexo y grupo ocupacional, no obstante, dados los cambios metodológicos introducidos durante la actualización realizada el 2017, es decir, el cambio de año base 2016=100, no se pueden realizar enlaces desagregados con la base anterior (2009=100). Por esto, se realizan los empalmes de las series base anual 2016=100 y base anual 2023=100 desagregados desde junio de 2017 hasta abril de 2024 por medio del factor de enlace obtenido en marzo de 2024, esto según tamaño de empresas, sector económico, sexo y grupo ocupacional.

$$\text{Factores de enlace (IR e ICL)}_{t,c,s,g} = \frac{(\text{Valor índice marzo 2024, base anual 2023} = 100)_{t,c,s,g}}{(\text{Valor índice marzo 2024, base anual 2016} = 100)_{t,c,s,g}}$$

Siendo así, cada tamaño de empresas, sector económico, sexo y grupo ocupacional tendrá un factor de enlace para unir las series en bases promedio anual 2016=100 y 2023=100.

7.1.4 Desestacionalización de las series IR, ICL e IR real

El ajuste estacional[46] de una serie de tiempo busca eliminar los factores y/o patrones exógenos, de naturaleza no económica, que influyen sobre la variable en estudio y que en ocasiones ocultan los aspectos netamente económicos, los que constituyen la esencia del análisis de una serie de tiempo económica. Los componentes que se usan para caracterizar una serie son:

- **La tendencia:** movimiento de largo plazo que refleja la dirección persistente de la serie, asociada a factores estructurales.
- **El ciclo:** movimiento periódico, en torno de la tendencia, que pone en evidencia una sucesión de etapas de crecimiento o de recesión.
- **La estacionalidad:** fluctuaciones mensuales o trimestrales que se repiten cada año de manera más o menos regular.

[45] Para más antecedentes sobre los índices empalmados con base anual 2016=100, revisar separata técnica sobre actualización de los indicadores coyunturales y empalme de las series IR-ICMO publicado en el link.
[46] Los modelos de desestacionalización se actualizaron en la publicación del 7 de mayo de 2026, para más información sobre el modelo anterior que estuvo vigente hasta la fecha mencionada, véase anexo de este documento.

- **El comportamiento irregular:** mide todas las fluctuaciones que no son incluidas en los componentes anteriores.

La aplicación del ajuste estacional al IR, IR real e ICL, permite comparar directamente las variaciones entre períodos consecutivos sin la distorsión de los patrones estacionales.

Es importante recalcar que, por la naturaleza misma del ajuste estacional, los índices y por ende la variación ajustada, cambian cada vez que se procesan los datos, ya que el modelo utiliza la información más actualizada posible para eliminar los componentes estacionales (modelo estacional autoregresivo integrado con media móvil). Por lo tanto, es posible que, para un mismo mes, con un número mayor de observaciones de las series, los valores estimados de los índices difieran.

Para el correcto cálculo de la serie desestacionalizada, tanto para el IR, ICL e IR real, se utiliza la serie empalmada, con el fin de tener una mayor cantidad de datos.

La siguiente tabla resume los modelos aplicados en la última actualización para cada uno de los índices.

Tabla 20: Modelos actualizados utilizados en las series del IR, ICL e IR Real para la aplicación del ajuste estacional[47]

Índice	Especificaciones	Modelos
IR	Ajuste estacional	Directo
	Modelo calendario	ls2010.01, ls2016.01, ls2020.04, ls2020.05, ls2021.05, ls2023.01, ao2023.08, ls2024.06, ls2024.07
	Estacionalidad	Estocástica
	SARIMA	(0 2 1) (0 1 1)
ICL	Ajuste estacional	Directo
	Modelo calendario	ls2010.03, ls2016.01, ls2020.04, ao2023.08, ls2024.06, ls2024.07
	Estacionalidad	Estocástica
	SARIMA	(0 2 1) (0 1 1)
IR real	Ajuste estacional	Directo
	Modelo calendario	ls2020.04
	Estacionalidad	Estocástica
	SARIMA	(0 1 0) (1 1 1)

Fuente: Instituto Nacional de Estadísticas (INE).

Nota 1: ls<año>.<mes> denota un cambio permanente de nivel.

Nota 2: ao<año>.<mes> denota un *outlier* aditivo.

[47] La metodología X13 ARIMA SEATS ajusta un modelo de regresión a la serie de tiempo, cuyos errores (Y_t) siguen un proceso SARIMA: $Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_k X_{kt} + Z_t$; e $Y_t - \beta' X_t = Z_t$. Donde Z_t es una serie de tiempo a la que se le aplica un modelo SARIMA y filtros de medias móviles para la obtención de los componentes estacional, componente de tendencia-ciclo y componente irregular.

7.1.5 Variación de los Índices

i. Variación mensual

$$vm_I_{m_y} = \left(\frac{I_{m_y}}{I_{m-1_y}} - 1\right) * 100$$

Donde:

I_{m_y}: Es el indicador en el mes m y en el año y

I_{m_y}: Es el indicador en el mes $m - 1$ y en el año y

ii. Variación a doce meses

$$v12m_I_{m_y} = \left(\frac{I_{m_y}}{I_{m_y-1}} - 1\right) * 100$$

Donde:

I_{m_y}: Es el indicador en el mes m y en el año y

I_{m_y}: Es el indicador en el mes m y en el año y-1

iii. Variación acumulada

$$vac_I_{m_y} = \left(\frac{I_{m_y}}{I_{dic_y-1}} - 1\right) * 100$$

Donde:

I_{m_y}: Es el indicador en el mes m y en el año y

I_{dic_y-1}: Es el indicador en el mes de diciembre y en el año $y - 1$

7.1.6 Incidencia de los índices por sección económica

i. Incidencia mensual

$$im_{m_y} = \frac{I_{c,m-1_y} * vm_I_{c,m_y} * W_c}{I_{m-1_y}}$$

Donde:

$I_{c,m-1_y}$: Es el indicador para la sección c en el mes $m-1$ y en el año y

vm_I_{c,m_y}: Es la variación mensual del indicador para la sección c en el mes m y en el año y

W_c: Ponderación para la sección c

I_{m-1} : Es el indicador a nivel general en el mes $m-1$

ii. Incidencia a doce meses

$$i12m_{m_y} = \frac{I_{c,m_y-1} * v12m_I_{c,m_y} * W_c}{I_{m_y-1}}$$

Donde:

I_{c,m_y-1}: Es el indicador para la sección c en el mes m y en el año $y-1$

$v12m_I_{c,m_y}$: Es la variación del indicador a doce meses para la sección c en el mes m del año y

W_c: Ponderación para la sección c

I_{m_y-1} : Indicador para el mes m y en el año $y-1$

iii. Incidencia acumulada

$$iac_{m_y} = \frac{I_{c,dic_y-1} * vac_I_{c,m_y} * W_c}{I_{dic_y-1}}$$

Donde:

I_{c,dic_y-1}: Es el indicador para la sección c en el mes de diciembre del año $y-1$

vac_I_{c,m_y}: Es la variación acumulada del indicador para la sección c en el mes m del año y

W_c: Ponderación para la sección c

I_{dic_y-1}: Indicador para el mes de diciembre del año $y-1$

7.1.7 Estimaciones analíticas

Las estimaciones que se presentan a continuación se utilizan como indicadores del comportamiento del mercado del trabajo y son cocientes que se construyen a partir de las horas ordinarias y extraordinarias, las remuneraciones ordinarias y extraordinarias, costos laborales y el número de puestos trabajo.

Estimaciones a nivel general

Las estimaciones a nivel general son seis y consideran en su cálculo los datos para hombres y mujeres.

i. <u>Remuneraciones ordinarias por hora ordinaria</u>

$$\left(\frac{\overline{RO}}{HO}\right)^m = \sum_{z=1}^{Z}\left(\frac{\sum_{e\in z} ro_e^m}{\sum_{e\in z} ho_e^m} * \frac{{HO_z}/{RO_z}}{\sum_{z=1}^{Z}\left({HO_z}/{RO_z}\right)}\right) = \sum_{z=1}^{Z}\left(\frac{RO_z^m}{HO_z^m} * W_horo_z\right)$$

Donde:

RO_z^m: Remuneraciones ordinarias en el mes m para la cadena z

HO_z^m: Horas ordinarias en el mes m para la cadena z

W_roho_z: Ponderación para la estimación y para la cadena z obtenida como $W_roho_z = \frac{{HO_z}/{RO_z}}{\sum_{z=1}^{Z}\left({HO_z}/{RO_z}\right)}$

ro_e^m: Remuneración ordinaria de la empresa e en el mes m

ho_e^m: Horas ordinarias de la empresa e en el mes m

RO_z: Sumatoria de las remuneraciones ordinarias expandidas para la cadena z, según la encuesta estructural 2021.

Obtenidas como $RO_z = \sum_{e\in z}(ro_e * FE_e^E)$

HO_z: Sumatoria de las horas ordinarias expandidas para la cadena z, según la encuesta estructural 2021.

Obtenidas como $HO_z = \sum_{e\in z}(ho_e * FE_e^E)$

ii. Remuneraciones totales por hora total

$$\left(\overline{\frac{RT}{HT}}\right)^m = \sum_{z=1}^{Z}\left(\frac{\sum_{e\in z} rt_e^m}{\sum_{e\in z} ht_e^m} * \frac{{HT_z}/{(RO+RE)_z}}{\sum_{z=1}^{Z}({HT_z}/{(RO+RE)_z})}\right) = \sum_{z=1}^{Z}\left(\frac{RT_z^m}{HT_z^m} * W_htrore_z\right)$$

Donde:

RT_z^m: Remuneraciones totales en el mes m para la cadena z

HT_z^m: Horas totales en el mes m para la cadena z

W_ht_z: Ponderación para la estimación y para la cadena z, obtenida como $W_ht_z = \frac{{HT_z}/{(RO+RE)_z}}{\sum_{z=1}^{Z}({HT_z}/{(RO+RE)_z})}$

rt_e^m: Remuneración total de la empresa e en el mes m

ht_e^m: Horas totales de la empresa e en el mes m

HT_z: Sumatoria de las horas totales expandidas para la cadena z, según la encuesta estructural 2021.

Obtenidas como $HT_z = \sum_{e\in z}(ht_e * FE_e^E)$

RT_z: Sumatoria de las remuneraciones totales expandidas para la cadena z, según la encuesta estructural 2021.

Obtenidas como $RT_z = \sum_{e\in z}(rt_e * FE_e^E)$

iii. Costo por hora total

$$\left(\overline{\frac{C}{HT}}\right)^m = \sum_{z=1}^{Z}\left(\frac{\sum_{e\in z} c_e^m}{\sum_{e\in z} ht_e^m} * \frac{{HT_z}/{C_z}}{\sum_{z=1}^{Z}({HT_z}/{C_z})}\right) = \sum_{z=1}^{Z}\left(\frac{C_z^m}{HT_z^m} * W_htc_z\right)$$

Donde:

C_z^m: Costo laboral en el mes m para la cadena z

HT_z^m: Horas totales expandidas en el mes m para la cadena z

W_ht_z: Ponderación para la estimación y para la cadena z, obtenida como $W_ht_z = \frac{{HT_z}/{C_z}}{\sum_{z=1}^{Z}({HT_z}/{C_z})}$

c_e^m: Costo laboral de la empresa e en el mes m

ht_e^m: Horas totales de la empresa e en el mes m

HT_z: Sumatoria de las horas totales expandidas para la cadena z, según la encuesta *estructural 2021.*

Obtenidas como $HT_z = \sum_{e \in z}(ht_e * FE_e^E)$

C_z: Sumatoria del costo laboral expandido para la cadena z, según la encuesta estructural 2021.

Obtenidas como $C_z = \sum_{e \in z}(c_e * FE_e^E)$

iv. Horas ordinarias por puesto de trabajo

$$\left(\overline{\frac{HO}{NT}}\right)^m = \sum_{z=1}^{Z}\left(\frac{\sum_{e \in z} ho_e^m}{\sum_{e \in z} nt_e^m} * \frac{NT_z}{\sum_{z=1}^{Z} NT_z}\right) = \sum_{z=1}^{Z}\left(\frac{HO_z^m}{NT_z^m} * W_nt_z\right)$$

Donde:

HO_z^m: Horas ordinarias en el mes m para la cadena z

NT_z^m: Número de puestos de trabajo en el mes m para la cadena z

W_nt_z: Ponderación para la estimación y para la cadena z, obtenida como $W_nt_z = \frac{NT_z}{\sum_{z=1}^{Z} NT_z} = \frac{NT_z}{\sum_{z=1}^{Z} NT_z}$

ho_e^m: Horas ordinarias de la empresa e en el mes m

nt_e^m: Número de puestos de trabajo de la empresa e en el mes m

NT_z: Sumatoria de los puestos de trabajo expandidos para la cadena z, según la encuesta estructural 2021.

Obtenidas como $NT_z = \sum_{e \in z}(nt_e * FE_e^E)$

v. Horas extraordinarias por puesto de trabajo

$$\left(\overline{\frac{HE}{NT}}\right)^m = \sum_{z=1}^{Z}\left(\frac{\sum_{e \in z} he_e^m}{\sum_{e \in z} nt_e^m} * \frac{NT_z}{\sum_{z=1}^{Z} NT_z}\right) = \sum_{z=1}^{Z}\left(\frac{HE_z^m}{NT_z^m} * W_nt_z\right)$$

Donde:

HE_z^m: Horas extraordinarias en el mes m para la cadena z

NT_z^m: Número de puestos de trabajo en el mes m para la cadena z

W_nt_z: Ponderación para la estimación y para la cadena z, obtenida como $W_nt_z = \frac{NT_z}{\sum_{z=1}^{Z} NT_z}$

he_e^m: Horas extraordinarias de la empresa e en el mes m

nt_e^m: Número de puestos de trabajo de la empresa e en el mes m

NT_z: Sumatoria de los puestos de trabajo expandidos para la cadena z, según la encuesta estructural 2021.

Obtenidas como $NT_z = \sum_{e \in z}(nt_e * FE_e^E)$

vi. Horas totales por puesto de trabajo

$$\left(\overline{\frac{HT}{NT}}\right)^m = \sum_{z=1}^{Z}\left(\frac{\sum_{e \in z} ht_e^m}{\sum_{e \in z} nt_e^m} * \frac{NT_z}{\sum_{z=1}^{Z} NT_z}\right) = \sum_{z=1}^{Z}\left(\frac{HT_z^m}{NT_z^m} * W_nt_z\right)$$

Dónde:

HT_z^m: Horas totales en el mes m para la cadena z

NT_z^m: Número de puestos de trabajo en el mes m para la cadena z

W_nt_z: Ponderación para la estimación y para la cadena z, obtenida como $W_nt_z = \frac{NT_z}{\sum_{z=1}^{Z} NT_z}$

ht_e^m: Horas totales de la empresa e en el mes m

nt_e^m: Número de puestos de trabajo de la empresa e en el mes m

NT_z: Sumatoria de los puestos de trabajo expandidos para la cadena z, según la encuesta estructural 2021.

Obtenidas como $NT_z = \sum_{e \in z}(nt_e * FE_e^E)$

Estimaciones diferenciadas según desagregación

- Estimaciones por tamaño de empresas
- Estimaciones por sector económico
- Estimaciones por grupo ocupacional
- Estimaciones según por sexo:

Las estimaciones diferenciadas por sexo se obtienen a partir de los mismos datos que las estimaciones a nivel general, pero utilizando nuevos ponderadores obtenidos a partir de los ponderadores de las estimaciones a nivel general, a los que se denominará como "ponderadores originales".

Paso N°1: Calcular nuevos ponderadores que consideren la apertura por sexo, para cada "ponderador original":

Obtener sumatoria de cada ponderador para mujeres y hombres:

Existen tres ponderadores distintos, para mujeres, hombres y el general.

Mujeres:

$$\sum_{t,c,s,g=1} W_{t,c,s,g} \quad Para\ s = 1$$

Hombres:

$$\sum_{t,c,s,g=1} W_{t,c,s,g} \quad Para\ s = 2$$

Donde:

$W_{t,c,s,g}$: Ponderador a nivel de tamaño de empresas t, sección económica c, sexo s y grupo ocupacional g.

Obtención de los "nuevos ponderadores" (NW_i)

Construir cociente a nivel de tamaño de empresas t, sección económica c, sexo s y grupo ocupacional g, entre el ponderador original y su sumatoria por sexo, y multiplicarlo por 100 para expresarlo en porcentaje. Estos ponderadores, a diferencia de los originales, permitirán realizar estimaciones diferenciadas por sexo.

Si es mujer:

$$NW_{t,c,s,g}^{mujer} = \frac{W_{t,c,s,g}}{\sum_{t,c,s,g=1} W_{t,c,s,g}} * 100 \;\; Para\; s = 1$$

Si es hombre:

$$NW_{t,c,s,g}^{hombre} = \frac{W_{t,c,s,g}}{\sum_{t,c,s,g=1} W_{t,c,s,g}} * 100 \;\; Para\; s = 2$$

Paso N°2: Obtener las estimaciones originales ($EO_{t,c,s,g}^{m}$) a nivel de cadena corta para cada mes m:

Calcular la sumatoria a nivel de t, c, s, g de las variables RO, RE, HO, HE, C y NT.

Calcular los cocientes respectivos a nivel de t, c, s, g, para cada mes con el que se cuentan datos:

- Remuneraciones ordinarias por hora ordinaria: $\left(\frac{RO}{HO}\right)_{t,c,s,g}^{m}$
- Remuneraciones ordinarias y extraordinarias por hora total: $\left(\frac{RO+RE}{HT}\right)_{t,c,s,g}^{m}$
- Costo laboral por horas totales: $\left(\frac{C}{HT}\right)_{t,c,s,g}^{m}$
- Horas ordinarias por número de puestos de trabajo: $\left(\frac{HO}{NT}\right)_{t,c,s,g}^{m}$
- Horas extraordinarias por número de puestos de trabajo: $\left(\frac{HE}{NT}\right)_{t,c,s,g}^{m}$
- Horas totales por número de puestos de trabajo: $\left(\frac{HT}{NT}\right)_{t,c,s,g}^{m}$

Paso N°3: Calcular las nuevas estimaciones por sexo para el mes m:

Obtener la estimación ponderada ($Est_pond_{t,c,s,g}^{m}$) a nivel de t, c, s, g

Calcular el producto entre las estimaciones y sus nuevos ponderadores respectivos a nivel de t, c, s, g y dividirlas por 100 (cada estimación tiene su ponderador previamente asignado, el cual se calculó en el paso N°1).

Ejemplo para las remuneraciones ordinarias por hora ordinaria:

$$Est_pond^{m}_{t,c,s,g} = \frac{EO^{m}_{t,c,s,g} * \ NW_{t,c,s,g}}{100} = \frac{\left(\frac{RO}{HO}\right)^{m}_{t,c,s,g} * \ NW_{t,c,s,g}}{100}$$

Donde:

$Est_pond^{m}_{t,c,s,g}$: Es la estimación ponderada para el mes m y cadena t, c, s, g

$EO_{t,c,s,g}$: Es la estimación original a nivel de cadena t, c, s, g para el mes m, descrita en el "paso N°2"

Obtener la sumatoria de la estimación ponderada para hombres y mujeres, para cada una de las estimaciones.

Mujeres:

$$\sum_{t,c,s,g=1} Est_pond^{m}_{t,c,s,g} = NE^{m} \ Para \ s = 1$$

Hombres:

$$\sum_{t,c,s,g=1} Est_pond^{m}_{t,c,s,g} = NE^{m} \ \ Para \ s = 2$$

- Calcular el nivel general de cada estimación: multiplicar estimaciones originales a nivel de cadena t, c, s, g para el mes m por sus ponderadores originales correspondientes, y obtener sumatoria de todas las cadenas.

$$E^{m} = \sum_{t,c,s,g=1} \frac{W_{t,c,s,g} * EO^{m}_{t,c,s,g}}{100}$$

- Obtener la brecha entre hombres y mujeres para cada una de las estimaciones en el mes m:

$$Brecha^{m} = \left(\frac{Estimación \ mujeres^{m}}{Estimación \ hombres^{m}} - 1\right) * 100$$

- Obtener la brecha de hombres o mujeres para cada una de las estimaciones contra el promedio general en el mes m:

$$Brecha\ con\ respecto\ al\ promedio\ mujeres^m = \left(\frac{Estimación\ mujeres^m}{Estimación\ promedio\ general^m} - 1\right)$$

$$Brecha\ con\ respecto\ al\ promedio\ hombres^m = \left(\frac{Estimación\ hombres^m}{Estimación\ promedio\ general^m} - 1\right)$$

7.2 Validación e interpretación de resultados

La metodología de construcción del Índice de Remuneraciones y de Costos Laborales entrega como resultado un índice general que, por sí solo, no posee interpretación. Para poder interpretarlo se requiere realizar una comparación con otros períodos, es decir, se debe tomar en consideración la variación del índice general con respecto al mes anterior, la variación acumulada en lo que va del año o la variación al comparar el mismo mes del año anterior (a doce meses). También es relevante considerar las incidencias de las distintas desagregaciones de los indicadores en cada una de estas variaciones. Generalmente, se analizan tanto las variaciones más importantes como las desagregaciones con mayores incidencias en la variación del índice general, lo que permite entender las fluctuaciones de las remuneraciones y costos laborales. Una vez que se procesan los datos de la muestra de seguimiento mensual y se obtienen los índices a los distintos niveles de desagregación, se realizan una serie de revisiones de los resultados agregados obtenidos (éstas son adicionales a las que se van realizado durante el mes).

7.3 Indicadores para la medición y aseguramiento de la calidad estadística

La calidad estadística se define como el grado o nivel de cumplimiento que deben tener los procesos y productos estadísticos, orientados a satisfacer las necesidades de información de instituciones, organizaciones y personas usuarias. En el ámbito de las estadísticas, la definición de calidad se materializa especificando un conjunto de dimensiones que caracterizan la calidad de los procesos y productos, en particular, en este apartado se presentan las fórmulas de cálculo de la tasa de elegibilidad, tasa de respuesta y tasa de imputación, asociadas a la dimensión de precisión y confiabilidad.

Tasa de elegibilidad:

La tasa de elegibilidad corresponde al porcentaje de unidades muestrales clasificadas como elegibles respecto al total de unidades de la muestra total a gestionar. La situación de elegibilidad se obtiene luego de realizar la asignación de códigos de disposición final a cada unidad de la muestra a gestionar.

La tasa de elegibilidad se calcula como:

$$te = 100 \cdot \frac{Total\ unidades\ elegibles}{Total\ muestra\ a\ gestionar}$$

La tasa de elegibilidad es un indicador que informa sobre el porcentaje de unidades que, luego del trabajo de campo y procesamiento de datos, se verifica que forma parte de la población de interés del estudio. El complemento de la tasa de elegibilidad corresponde

a la tasa de no elegibilidad la cual es informativa del porcentaje de unidades que no pertenecen a la población de interés luego del trabajo de campo y el procesamiento de datos. Los casos de no elegibilidad son situaciones en las que se verifica en la unidad informante incidencias como cese definitivo de actividades, unidades fusionadas con otras unidades legales, cambios en la unidad estadística tal que estas quedan fuera del ámbito del estudio, entre otras.

Tasa de respuesta:

La tasa de respuesta corresponde al porcentaje de unidades elegibles aptas para procesar, sobre el total de unidades que son elegibles. Estas clasificaciones y desagregación del tipo de elegibilidad de la muestra también derivan de la asignación de códigos de disposición final.

La tasa de respuesta se calcula como:

$$te = 100 \cdot \frac{Total\ unidades\ aptas\ para\ procesar}{Total\ muestra\ elegible}$$

La tasa de respuesta es un indicador de calidad del proceso de trabajo de campo, ya que indica el éxito de los esfuerzos operativos en la obtención de una respuesta. El complemento de la tasa de respuesta es la tasa de no respuesta, la cual es informativa del porcentaje de unidades elegibles que no entran al cálculo de los indicadores IR-ICL por motivos tales como rechazo, no contacto, encuestas no aptas para procesar, entre otras.

Tasa de imputación

La tasa de imputación que aborda este apartado corresponde al porcentaje muestral imputado de una variable respecto al total muestral de dicha variable, y se calcula para cada variable afecta a imputación. Estos porcentajes se obtienen sobre los datos muestrales del conjunto de unidades clasificadas como aptas para procesar.

La tasa de imputación se calcula como:

$$ti_v = 100 \cdot \frac{Total\ muestral\ imputado\ de\ la\ variable\ v}{Total\ muestral\ de\ la\ variable\ v}$$

La tasa de imputación es un indicador que se utiliza como medida para evaluar la calidad y completitud de un conjunto de datos, y la validez de los resultados realizados a partir de dicho conjunto de datos. El indicador complementario de la tasa de imputación se corresponde con el porcentaje de datos efectivos recolectados que entran directamente a los cálculos del IR e ICL.

7.4 Tratamiento para la perdida de unidades muestrales en el tiempo

Una vez incorporada la muestra del CAB a la recolección continua de la EMRCL, esta se mide durante todo el periodo de vigencia del CAB 2023, es decir, hasta un próximo cambio de año base. Es probable que durante este periodo de recolección se verifiquen incidencias de campo circunscritas dentro de las situaciones de no elegibilidad, o bien, de no respuesta o rechazo permanente por parte de una unidad informante, lo cual deriva en la pérdida de unidades de seguimiento y, por ende, la pérdida de información efectiva y el aumento de la tasa de imputación.

Dentro de las posibles metodologías que se evalúan para mitigar el impacto que presenta la perdida de empresas en el transcurso del seguimiento de la muestra coyuntural, se establecen varias líneas de trabajo como estrategia a implementar, entre las cuales se mencionan: 1) la incorporación de nuevas unidades en la medición, que permitan el refresco de la muestra en forma anual y a partir de un determinado período, 2) ajuste de los factores de expansión para las incidencias especiales y particulares registrada en terreno, tales como; fusiones, absorciones, divisiones, entre otros, 3) identificación de empresas que en un periodo corto de tiempo presentan alguna incidencia de cierre, quiebras, términos de giros u otros, que generen una perdida considerable de unidades en sus estratos, pudiendo establecer como método para su tratamiento el de reemplazo de empresas.

Los métodos que se evalúan aplicar dependerán de la identificación y clasificación de las incidencias en el trabajo de campo de modo que permita incorporar el tratamiento adecuado, los cuales se pueden definir desde métodos de rotación de muestra hasta reemplazos. Por otro lado, es relevante definir el período de tiempo con la que realizará la actualización de alguna porción definida de la muestra o el seguimiento de la unidad una vez validada su incidencia, para realizar los posibles análisis de impacto que implica la incorporar de nuevas unidades o ajustes al índice.

7.5 Método de control a la divulgación

7.5.1 Evaluación inicial del riesgo

Antes de generar el proceso de anonimización, se llevó a cabo un análisis preliminar del riesgo. Este análisis se enfocó en evaluar la posibilidad de re-identificación utilizando la base de datos en términos de variables comunes, teniendo en cuenta el factor de expansión. El propósito principal fue establecer un punto de referencia del riesgo asociado con la identificación de las entidades antes de cualquier intervención de anonimización.

7.5.2 Algoritmo de anonimización

Tras completar el análisis inicial del riesgo, se implementó el proceso de anonimización de los datos. En primer lugar, se eliminaron los identificadores directos, como el rol y el RUT, los cuales se consideraron de alto riesgo para la identificación de las empresas. Posteriormente, se genera un identificador único y no reversible para cada empresa, con el fin de permitir análisis agregados sin comprometer la confidencialidad de estas.

7.5.3 Análisis de riesgo posterior

Una vez completado el proceso de anonimización, se realizó una segunda evaluación del riesgo para verificar la efectividad de las medidas implementadas y confirmar que el nivel de riesgo de re-identificación fuera aceptablemente bajo. Específicamente, se aplicó el concepto de k-anonimato (Samarati & Sweeney, 1998), utilizando un valor de k = 3. Esto garantizó que cada combinación de variables cuasi-identificadoras no fuera distintiva para menos de tres empresas en la base de datos.

8. Alcance de la presentación de resultados

8.1 Productos estadísticos a publicar

8.1.1 Microdatos

Se publicarán dos bases de datos mensuales, una con los datos del período de referencia (t), y otra con los datos del período anterior (t-1), las variables que se incluyen son: remuneraciones ordinarias, extraordinarias, costos laborales, horas ordinarias, horas extraordinarias y totales, junto con el número de puestos de trabajo ocupados. Estas bases se ponen a disposición de las personas usuarias a nivel de empresa, tamaño, sector económico, grupo ocupacional y sexo de las personas trabajadoras. Dado lo anterior y con el fin de resguardar la información sometida a Secreto Estadístico, se lleva a cabo un proceso de anonimización de los datos antes de su divulgación. Este proceso consiste en no incluir identificadores expresos de la fuente de información. Estos archivos corresponden a un libro por mes, en formato MS Excel, que incluye 4 pestañas una con los datos del período, otra con los ponderadores, los reemplazos, y el diccionario de variables, respectivamente.

8.1.2 Metadatos

Los metadatos publicados[48] son:

- Documento metodológico de los Índices de Remuneraciones y Costos Laborales. Base promedio anual 2023=100: el propósito de este es documentar el proceso estadístico de manera integral, de acuerdo con las principales fases que lo componen (diseño y planificación, construcción, recolección de datos, procesamiento, análisis y difusión de resultados).
- Separata de actualización, cambios metodológicos y empalmes: este documento resume los principales cambios metodológicos realizados en el cambio de año base 2023, además de entregar un resumen del proceso de empalme de las series y los nuevos modelos de desestacionalización de IR-ICL.
- Separata con la política de rectificación de cifras: este documento entrega a las personas usuarias la actualización de la política de rectificación de cifras de IR-ICL.
- Separata de desestacionalización de las series: en esta se encuentra explicitado el modelo de desestacionalización seleccionado y aplicado en el IR-ICL e IR real, utilizando las series empalmadas de los 3 indicadores.

[48] Descarga directa en https://www.ine.gob.cl/estadisticas/sociales/mercado-laboral/remuneraciones-y-costos-laborales

- Ficha metodológica: corresponde a un resumen de la operación estadística, donde se incluye la descripción de la operación, descripción del diseño estadístico y un listado de los productos estadísticos a difundir.
- Diccionario de variables: cada uno de los archivos de cuadros estadísticos y bases de datos contiene un descriptor de las variables utilizadas, incluyendo el nombre de la variable o campo, la descripción de la misma, el tipo de variable y el rango de esta.

8.2 Difusión de los resultados

En el sitio web institucional está a disposición de todo público el índice nominal de remuneraciones, el índice nominal de costos laborales, el índice real de remuneraciones (estos últimos tanto en términos desestacionalizados como no), y las estimaciones de las remuneraciones ordinarias por hora ordinaria, remuneraciones extraordinarias por hora extraordinaria y de costos laborales por hora total, además de las estimaciones de horas ordinarias, extraordinarias y horas totales medias por persona trabajadora a niveles generales y desagregados. Esta información se presenta en cuadros estadísticos (índices y estimaciones rectificadas y provisorias). Por otro lado, en los cuadros estadísticos empalmados y empalmados desestacionalizados[49] se presenta toda la información histórica que se ha podido enlazar, pero haciendo hincapié en que la última cifra es provisoria (P).

Esta información se presenta y actualiza mensualmente a través de tabulados o cuadros estadísticos. Además, se publica un boletín en el que se exponen los aspectos más relevantes del comportamiento de cada indicador y estimación.

Todos los cuadros estadísticos se ponen a disposición de los usuarios en formato MS Excel incluyendo notas explicativas y diccionario de variables. A su vez, se pondrán a disposición de los usuarios cuadros estadísticos de indicadores y variaciones interanuales en Sistema de Información del Mercado Laboral (SIMEL) (https://www.simel.gob.cl/). Los indicadores incluidos en este portal son el IR e ICL nominal general, y a niveles desagregados por sector económico, tamaño de empresas, grupo ocupacional y sexo, además del IR real y las estimaciones de remuneraciones y horas medias generales.

[49] Los cuadros estadísticos desestacionalizados solo se presentan a nivel de índices generales.

8.3 Accesibilidad a los datos y marco legal

Las personas usuarias podrán encontrar las bases de datos en la página web del INE (www.ine.gob.cl). En caso de requerir mayor especificidad, estas deben ser solicitadas por transparencia, de manera de facilitar bases innominadas e indeterminadas.

De acuerdo con la Declaración de confidencialidad de los datos

- Artículo 29° - 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 hará incurrir en el delito previsto y penado por el artículo 247, del código penal, debiendo en todo caso aplicarse pena corporal.
- Artículo 30° - 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).

9. Glosario

Los principales conceptos incorporados dentro de este documento fueron extraídos del Glosario de conceptos estadísticos del INE (INE, 2022) y del Manual Metodológico del Índice de Precios al Consumidor (INE, 2023):

- **Cadena corta:** corresponde a la agregación de los datos de la encuesta a nivel de tamaño de empresas, sección económica, sexo y grupo ocupacional.
- **Cadena larga:** corresponde a la agregación de los datos de la encuesta a nivel de tamaño de empresas, sección económica, grupo ocupacional, sexo y empresa.
- **Empalmar:** conectar o unir dos series de índices de precios que se superponen en uno o más períodos.
- **Incidencia:** es el aporte a la variación que tiene una determinada unidad de análisis sobre el total. La suma de todas las incidencias da la variación entre dos periodos.
- **Imputación:** procedimiento por medio del cual se le asigna a una variable no reportada (o faltante) la variación de otra variable similar o de otras más o menos similares.
- **Número índice:** es una medida estadística que permite estudiar las fluctuaciones o variaciones de una (o más) magnitud(es) en relación con el tiempo o el espacio.
- **Período base:** suele denominarse a aquel período con el cual todos los demás se comparan. Sin embargo, el término tiene otros significados según el contexto. Se distinguen tres elementos:
  - Período de referencia de ponderaciones: período en el cual se obtuvo la información necesaria para construir los ponderadores de los productos de la canasta.
  - Período de referencia del índice: período para el cual el valor del índice se fija igual a 100.
  - Período de referencia de precios: período de duración definida cuyos precios se comparan con los precios de los demás períodos. El período de referencia de precios suele denotarse como período 0.
- **Variación acumulada:** es la suma de las variaciones da cada periodo entre un periodo inicial y el actual.
- **Variación anual:** describe la relación entre el valor de t-12 y el presente.
- **Variación mensual:** describe la relación entre el valor del mes anterior y el presente.
- **Variación porcentual:** describe la relación entre un valor pasado y uno presente. De manera específica, la variación porcentual representa la diferencia entre un valor pasado y uno presente en términos de un porcentaje del valor pasado.

10. Bibliografía

De Waal, T., Pannekoek, J., & Scholtus, S. (2011). *Handbook of statistical data editing and imputation* (Vol. 563). John Wiley \& Sons.

Deville, & Särndal. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 376-382.

Diario Oficial. (1959). Decreto con Fuerza de Ley N°2, artículo 27.

INE. (1963). *Metodología del Índice de Remuneraciones Base abril 1960=100.* Santiago: Biblioteca INE.

INE. (1982). *Índices de Remuneraciones (publicado desde 1982 a 1991).* Santiago: Biblioteca INE.

INE. (16 de noviembre de 1985). *Metodología del Índice de Remuneraciones (IR).* Santiago: Biblioteca INE.

INE. (2007). *Metodología de los Índices de Remuneraciones y Costo de la Mano de Obra (Base: enero 2006=100).* Santiago. Obtenido de https://www.ine.gob.cl/docs/default-source/sueldos-y-salarios/metodologias/base-enero-2006/manual-metodol%C3%B3gico-%C3%ADndice-remuneraciones-y-costo-de-la-mano-de-obra-base-enero-2006-=-100.pdf?sfvrsn=dd039af0_4

INE. (2010). *Manual metodológico del índice de remuneraciones (IR) índice del costo de mano de obra (ICMO), base anual 2009 = 100.* Santiago. Obtenido de https://www.ine.gob.cl/docs/default-source/sueldos-y-salarios/metodologias/base-anual-2009/manual-metodol%C3%B3gico-%C3%ADndice-remuneraciones-y-costo-de-la-mano-de-obra-base-2009-=-100.pdf?sfvrsn=e2fe71a6_4

INE. (2012). *Clasificador Chileno de Actividades Económicas CIIU4.CL 2012.* Obtenido de https://www.ine.gob.cl/docs/default-source/buenas-practicas/clasificaciones/ciiu/clasificador/ciiu4-cl-2012.pdf?sfvrsn=509b707c_3

INE. (2017). *Manual Metodológico de los Índices de Remuneraciones, del Costo de la Mano de Obra y de Puestos de Trabajo Base promedio anual 2016=100.* Santiago. Obtenido de https://www.ine.gob.cl/docs/default-source/sueldos-y-salarios/metodologias/base-anual-2016/manual-metodol%C3%B3gico-%C3%ADndice-remuneraciones-y-costo-de-la-mano-de-obra-base-2016-=-100.pdf?sfvrsn=c326269b_4

INE. (2018). *Clasificador Chileno de Ocupaciones CIUO 08.CL.* Obtenido de https://www.ine.gob.cl/docs/default-source/buenas-practicas/clasificaciones/ciuo/clasificador/ciuo-08-cl.pdf

INE. (2022). *Glosario de conceptos estadísticos.* Obtenido de https://www.ine.gob.cl/docs/default-source/buenas-practicas/estandares/estandar/glosario-de-conceptos-estad%C3%ADsticos/glosario-de-conceptos-ine-v_1-0.pdf?sfvrsn=43ff6ab6_2

INE. (2023). *Manual Metodológico Índice de Precios al Consumidor Base Anual 2023.* Obtenido de https://www.ine.gob.cl/docs/default-source/%C3%ADndice-de-precios-al-consumidor/metodologias/base-anual-2023-100/metodolog%C3%ADa-base-anual-2023-=-100.pdf?sfvrsn=b1cf9a46_8

OCDE. (2015). *Recomendación del consejo de la OCDE sobre buenas prácticas estadísticas.* OCDE. Obtenido de https://www.oecd.org/statistics/good-practice-toolkit/Brochure-Recommendation%20of%20the%20OECD%20Council%20on%20Good%20Statistical%20Practice_SPANISH_March2019.pdf

OIT. (1966). *Resolución sobre las estadísticas del costo de la mano de obra.* Conferencia Internacional de Estadísticos del Trabajo (CIET). Obtenido de https://ilostat.ilo.org/es/about/standards/icls/icls-documents/

OIT. (2015). *Indicadores Clave del Mercado del Trabajo (KILM) (9.a ed.).* Obtenido de http://www.inefop.org.uy/docs/Indicadores%20clave%20del%20mercado%20de%20trabajo_OIT%20Ginebra.pdf

Ralph, J. O. (2018). *A Practical Introduction to Index Numbers.*

Samarati, & Sweeney. (1998). *Protecting Privacy when Disclosing Information: k-Anonymity and its Enforcement through Generalization and Suppression.* Obtenido de http://www.csl.sri.com/papers/sritr-98-04/

11. Anexo

11.1 Ponderaciones índices

Tabla 21: Ponderaciones índices a nivel de sección económica

Sección económica (CIIU4.CL 2012)	Glosa	Ponderadores IR	Ponderadores ICL
B	Minería	3,23	3,57
C	Industria Manufacturera	11,18	11,60
D	Suministro de Electricidad, Gas, Vapor y Aire Acondicionado	0,78	0,81
E	Suministro de Agua, Evacuación de Aguas Residuales, Gestión de Desechos y Descontaminación	0,72	0,76
F	Construcción	9,70	9,98
G	Comercio	14,95	14,98
H	Transporte y Almacenamiento	5,36	5,65
I	Actividades de Alojamiento y Servicio de Comidas	2,05	2,07
J	Información y Comunicaciones	3,74	3,63
K	Actividades Financieras y de Seguros	6,90	6,69
L	Actividades Inmobiliarias	1,24	1,19
M	Actividades Profesionales, Científicas y Técnicas	6,76	6,57
N	Actividades de Servicios Administrativos y de Apoyo	5,57	5,88
O	Administración Pública	12,41	11,76
P	Enseñanza	9,28	8,78
Q	Actividades de Atención de la Salud Humana y de Asistencia Social	5,64	5,58
R	Actividades Artísticas, de Entretenimiento y Recreativas	0,51	0,51
Total		**100**	**100**

Fuente: Instituto Nacional de Estadísticas (INE).

Tabla 22: Ponderaciones índices a nivel de tamaño de empresas

Tamaño de empresas (según número de personas trabajadoras)	Descripción	Ponderadores IR	Ponderadores ICL
Pequeña	Entre 5 y 49 personas	22,26	21,89
Mediana	Entre 50 y 199 personas	19,04	19,00
Grande	200 o más personas	58,70	59,11
Total		**100**	**100**

Fuente: Instituto Nacional de Estadísticas (INE).

Tabla 23: Ponderaciones índices a nivel de sexo

Sexo	Ponderadores IR	Ponderadores ICL
Mujeres	36,24	35,66
Hombres	63,76	64,34
Total	**100**	**100**

Fuente: Instituto Nacional de Estadísticas (INE).

Tabla 24: Ponderaciones índices a nivel de grupo ocupacional CIUO 08.CL

Grupo ocupacional (CIUO-08CL)	Glosa	Ponderadores IR	Ponderadores ICL
1	Directores, gerentes y administradores	14,42	13,52
2	Profesionales, científicos e intelectuales	37,83	36,43
3	Técnicos y profesionales de nivel medio	13,08	13,39
4	Personal de apoyo administrativo	8,78	8,95
5.1	Trabajadores de los servicios a las personas, trabajadores de los cuidados personales, y personal de los servicios de protección y seguridad	1,82	2,01
5.2	Vendedores	4,18	4,12
7	Artesanos y operarios de oficios	7,56	8,13
8	Operadores de instalaciones, máquinas y ensambladores	5,52	6,25
9	Ocupaciones elementales	6,80	7,19
Total		**100**	**100**

Fuente: Instituto Nacional de Estadísticas (INE).

11.2 Ponderaciones estimaciones

Tabla 25: Ponderaciones estimaciones a nivel de sección económica, tamaño de empresas, sexo y grupo ocupacional

Sección económica (CIIU4.CL 2012)	Glosa	Ponderadores ROHO	Ponderadores CLHT	Ponderadores HONT, HENT, HTNT
B	Minería	3,46	3,28	1,67
C	Industria Manufacturera	7,88	7,76	11,99
D	Suministro de Electricidad, Gas, Vapor y Aire Acondicionado	2,89	2,80	0,40
E	Suministro de Agua, Evacuación de Aguas Residuales, Gestión de Desechos y Descontaminación	5,18	5,07	0,80
F	Construcción	6,18	6,13	11,73
G	Comercio	7,94	7,83	16,41
H	Transporte y Almacenamiento	6,92	6,74	6,34
I	Actividades de Alojamiento y Servicio de Comidas	9,49	9,61	4,21
J	Información y Comunicaciones	4,00	4,00	2,34
K	Actividades Financieras y de Seguros	3,38	3,41	3,53
L	Actividades Inmobiliarias	5,38	5,49	0,95
M	Actividades Profesionales, Científicas y Técnicas	5,55	5,58	4,25
N	Actividades de Servicios Administrativos y de Apoyo	8,29	8,12	10,40
O	Administración Pública	4,48	4,85	10,12
P	Enseñanza	6,68	6,99	9,02
Q	Actividades de Atención de la Salud Humana y de Asistencia Social	6,32	6,38	5,29
R	Actividades Artísticas, de Entretenimiento y Recreativas	5,99	5,97	0,54
Total		**100**	**100**	**100**

Fuente: Encuesta estructural de remuneraciones y costos laborales 2021, Instituto Nacional de Estadísticas (INE).

Tamaño de empresas (según NT)	Ponderadores ROHO	Ponderadores CLHT	Ponderadores HONT, HENT, HTNT
Pequeña	28,73	29,31	24,93
Mediana	32,71	32,72	18,54
Grande	38,55	37,97	56,53
Total	**100**	**100**	**100**

Fuente: Encuesta estructural de remuneraciones y costos laborales 2021, Instituto Nacional de Estadísticas (INE).

Sexo	Ponderadores ROHO	Ponderadores CLHT	Ponderadores HONT, HENT, HTNT
Mujeres	47,97	47,96	40,82
Hombres	52,03	52,04	59,18
Total	**100**	**100**	**100**

Fuente: Encuesta estructural de remuneraciones y costos laborales 2021, Instituto Nacional de Estadísticas (INE).

Grupo ocupacional (CIUO-08CL)	Glosa	Ponderadores ROHO	Ponderadores CLHT	Ponderadores HONT, HENT, HTNT
1	Directores, gerentes y administradores	3,54	3,80	3,86
2	Profesionales, científicos e intelectuales	7,66	7,98	22,53
3	Técnicos y profesionales de nivel medio	13,41	13,61	14,85
4	Personal de apoyo administrativo	17,15	17,25	12,91
5.1	Trabajadores de los servicios a las personas, trabajadores de los cuidados personales, y personal de los servicios de protección y seguridad	11,25	11,09	4,20
5.2	Vendedores	5,86	6,05	5,78
7	Artesanos y operarios de oficios	12,58	12,29	13,40
8	Operadores de instalaciones, máquinas y ensambladores	7,35	6,97	6,97
9	Ocupaciones elementales	21,21	20,97	15,49
Total		**100**	**100**	**100**

Fuente: Encuesta estructural de remuneraciones y costos laborales 2021, Instituto Nacional de Estadísticas (INE).

11.3 Modelo de desestacionalización descontinuado

La siguiente tabla corresponde a los modelos de desestacionalización del IR, ICL e IR Real utilizados desde la publicación del 6 de junio de 2024, y que fueron actualizados el 7 de mayo de 2026.

Tabla 26: Modelos utilizados en las series del IR, ICL e IR Real para la aplicación del ajuste estacional

Índice	Especificaciones	Modelos
IR	Ajuste estacional	Directo
	Modelo calendario	ls2010.01, ls2010.03, ls2010.10, ls2011.05, ls2013.07, ls2016.01, ls2020.04, ls2020.05, ls2021.05, ls2023.01, ao2023.08
	Estacionalidad	Estocástica
	SARIMA	(0 2 1)(0 1 1)
ICL	Ajuste estacional	Directo
	Modelo calendario	ls2010.01, ls2010.03, ls2010.10, ls2011.05, ls2013.07, ls2016.01, ls2020.04, ls2020.05, ls2022.08
	Estacionalidad	Estocástica
	SARIMA	(1 1 0)(0 1 1)
IR real	Ajuste estacional	Directo
	Modelo calendario	ls2010.01, ls2010.03, ls2010.10, ls2011.05, ls2013.07, ls2016.01, ls2020.04, ls2020.05, ao2012.02, ao2016.02, ao2020.02, ao2021.05
	Estacionalidad	Estocástica
	SARIMA	(0 1 0)(0 1 1)

Fuente: Instituto Nacional de Estadísticas (INE).

Nota 1: ls<año>.<mes> denota un cambio permanente de nivel.

Nota 2: ao<año>.<mes> denota un *outlier* aditivo.