X ENCLA
Encuesta Laboral 2023

Décima Encuesta Laboral
Encla 2023

Informe Metodológico

Encla 2023

Instituto Nacional de Estadísticas • Chile

Dirección del Trabajo
Gobierno de Chile

SUBDIRECCIÓN TÉCNICA
DEPARTAMENTO DE ESTADÍSTICAS DEL TRABAJO
SUBDEPARTAMENTO DE ESTADÍSTICAS ESTRUCTURALES DEL TRABAJO

SUBDIRECCIÓN DE OPERACIONES
SUBDEPARTAMENTO DE IMPLEMENTACIÓN DE DISEÑOS

Décima Encuesta Laboral
Encla 2023

Informe Metodológico
Versión N°1

Instituto Nacional de Estadísticas, Chile.
Fecha de elaboración: agosto de 2024

Contenido

II. Introducción 6

III. Contexto metodológico 7

III.1. Antecedentes 7

III.2. Objetivos 7

IV. Marco de referencia 8

IV.1. Marco conceptual 8

IV.2. Referencias nacionales e internacionales 8

IV.3. Marco normativo 9

V. Diseño estadístico 10

V.1. Población objetivo 10

V.2. Cobertura geográfica 10

V.3. Cobertura temática 11

V.4. Unidades estadísticas 13

V.4.1. Unidad de muestreo 13

V.4.2. Unidad de análisis y de observación 14

V.5. Definición de variables 14

V.6. Clasificaciones estadísticas 15

V.7. Marco estadístico 16

V.7.1. Conformación del marco estadístico 16

V.7.2. Depuración del marco estadístico 16

V.7.3. Estratificación del marco estadístico 18

V.8. Diseño muestral 21

V.8.1. Tipo de muestreo 21

V.8.2. Dominio de estudio 21

V.8.3. Cálculo y distribución del tamaño muestral 21

V.8.4. Selección de las unidades muestrales 26

V.9. Resultado del tamaño muestral 27

V.10. Factores de expansión 32

V.10.3. Factor de expansión para las empresas del tramo de inclusión aleatoria 32

V.10.4. Factor de expansión para las empresas del tramo de inclusión forzosa 43

V.11. Estimación de parámetros 43

VI. Metodología de recolección de datos 46

VI.1. Características generales de la recolección de los datos 46

VI.1.1. Equipo de recolección de datos 48

VI.1.2. Procesos de verificación de la muestra 51

V.1.3. Estrategias de sensibilización de las personas informantes 52

VI.2. Diseño del Instrumento 54

VI.2.1. Programación de cuestionarios 55

VI.2.2. Consulta a personas expertas 55

VI.2.3. Pruebas cognitivas 56

VI.2.4. Caracterización de las versiones finales de los cuestionarios 58
VI.4. Prueba y evaluación de los Instrumentos 61
VI.4.1. Prueba de usabilidad 1 62
VI.4.2. Prueba de usabilidad 2 63
VI.4.3. Prueba de usabilidad 3 64
VI.5. Método de capacitación 65
VI.5.1. Calendarización de la capacitación 65
VI.5.2. Distribución de la capacitación 66
VI.5.3. Módulos de capacitación 66
VI.5.4. Programa de capacitación 67
VI.5.5. Otros elementos de la capacitación 68
VI.6. Método de supervisión 68
VI.6.1. Herramientas para el control de gestión 69
VI.6.2. Tipos de Supervisión 70
VII. Metodología de procesamiento 79
VII.1. Fases del procesamiento 79
VII.2. Codificación de glosas 81
VII.2.1. Codificación de glosas “Otros” 81
VII.2.2. Codificación de Actividad Económica 83
VII.3. Correcciones de Secuencia 88
VII.4. Validaciones Inter-cuestionarios 89
VII.5. Cuadratura de matrices 91
VII.6. Cuadratura Intra Matrices 94
VII.7. Cuadratura Inter-Matrices 98
VII.7.1. Ediciones para relación de equivalencia 99
VII.7.2. Ediciones para relación de cotas 100
VII.8. Imputación de datos 102
VII.8.1. No Respuesta/Inconsistencia 103
VII.8.2. Valores atípicos 105
VII.8.3. Imputación vía RRAA 106
VII.8.4. Imputación vía medianas condicionadas 110
VII.8.5. Ediciones a los cuestionarios a empleadores, sindicatos y trabajadores a partir de Registros Administrativos (RRAA) 112
VII.9. Creación de variables por Registros Administrativos (RRAA) 113
VIII. Metodología de Análisis de Resultados 114
VIII.1. Validación de resultados 114
VIII.2. Aplicación del control de divulgación 116
IX. Indicadores para la medición y aseguramiento de la calidad estadística 117
X. Difusión 123
XI. Glosario 125
XII. Bibliografía 127
XIII. Anexos 128

II. Introducción

La Encuesta Laboral (Encla) es una operación estadística mandatada por la Dirección del Trabajo (DT) cuyo principal objetivo es identificar las condiciones de trabajo y relaciones laborales al interior de las empresas regidas por el Código del Trabajo. Con el fin de generar un producto estadístico que logre cumplir con altos estándares de calidad, desde esta décima versión la encuesta quedó a cargo del Instituto Nacional de Estadísticas (INE). La Encla constituye un instrumento fundamental a la hora de estudiar el mundo del trabajo en la ocupación por la gran cantidad de temas que cubre.

El presente documento da cuenta del proceso de producción estadística realizado en la Encla 2023, para así entregar a las y los usuarios un mayor conocimiento de los datos. Todo proceso estadístico se desarrolla en base a una serie de etapas, decisiones y tareas que deben realizarse a fin de generar datos que cumplan con criterios de calidad adecuados, por lo que documentarlos se hace una parte elemental de la producción.

El manual se compone de los siguientes contenidos:

- Los capítulos III y IV dan cuenta del contexto metodológico y el marco de referencia, dando a conocer los antecedentes de la encuesta (versiones anteriores), sus objetivos, así como los conceptos sobre que trata, además de experiencias similares tanto nacionales como internacionales.

- El capítulo V da cuenta del diseño estadístico de la encuesta, dando a conocer la población objetivo, cobertura geográfica y temática, las unidades estadísticas, variables, clasificaciones estadísticas, el marco estadístico, el diseño muestral y la estimación de parámetros.

- El capítulo VI expone la metodología de recolección de datos. Describe el modo y los períodos de recolección, los flujos de trabajo, el diseño de los instrumentos y las pruebas a las que fueron sometidos, el método de capacitación y de supervisión.

- En el capítulo VII se aborda la metodología de procesamiento de datos, es decir cómo estos se trabajan cuando han sido recepcionados. Esto incluye la codificación de los datos, los procedimientos para detectar inconsistencias (malla de validadores y grilla técnica), los diversos mecanismos para corregirlos (ediciones e imputaciones) y los cálculos y ajustes de ponderadores.

- El capítulo VIII describe los métodos mediante los cuales se analizaron los datos y método de control a la divulgación aplicado.

- En el capítulo IX se exponen los indicadores para la medición y el aseguramiento de la calidad estadística implementados en la operación estadística.

- En el capítulo X se hace un listado de los productos y metadatos que se publicarán, el formato en el que se disponen a las personas usuarias y cómo serán difundidos por los organismos a cargo de la encuesta.

- Los capítulos XI, XII y XIII corresponden al glosario de conceptos relevantes en el contexto de la Encuesta, la bibliografía citada en el texto, y los anexos asociados a este, respectivamente.

III. Contexto metodológico

III.1. Antecedentes

Desde 1998 la Dirección del Trabajo (DT) ha realizado la Encla, que corresponde a una operación estadística por muestreo probabilístico, con periodicidad bienal o trienal, por medio de distintas instituciones públicas y privadas. Mientras que, en sus inicios, la Encla se pensó como un instrumento de fiscalización de las relaciones y condiciones laborales en las empresas del país, hoy constituye una herramienta para diagnosticar el estado de tales aspectos en las unidades económicas de Chile. Si bien en años anteriores aquella entidad ha suscrito convenios con el INE para que este elabore el diseño muestral de la encuesta, en 2019 ambas organizaciones firmaron un convenio mediante el cual se estableció que el organismo estadístico sería también el responsable de las fases de diseño y planificación, construcción, recolección y procesamiento de los datos[1]. La versión 2023 de la Encla es fruto de la continuidad en la colaboración inter-institucional entre el INE y la DT.

A lo largo de sus versiones la encuesta ha tenido diversas modificaciones, entre las que destacan, la ampliación de sus alcances; desde 2008 se aplica en todas las regiones; y el incremento de su tamaño muestral, con el objetivo de lograr una mejor representatividad y expandir sus resultados al universo de las empresas del país[2]. A diferencia de la versión de 2019, la Encla 2023 es representativa a nivel de actividad económica, región, tamaño, y agrupación regional por tamaño; a la vez que su factor de expansión ha sido calibrado al número de trabajadores contratados en las empresas que forman parte de la población objetivo. Por otra parte, como se detallará posteriormente, se ha extendido el uso de registros administrativos para editar e imputar información sobre las estructuras de empleo, las organizaciones sindicales y las negociaciones colectivas que se han realizado en cada unidad. Asimismo, se incluyeron nuevos módulos para medir fenómenos emergentes en el mundo del trabajo, como la digitalización y automatización de procesos, el teletrabajo, los efectos de la pandemia en las empresas, entre otros. Por último, entre otras modificaciones, los cuestionarios a empleadores, trabajadores y sindicatos pasaron de aplicarse en papel a la utilización de dispositivos móviles de captura.

De esta forma, la Encla 2023 constituye una instancia única para obtener información acerca de las transformaciones que ocurren en las relaciones laborales, las condiciones de empleo y trabajo, así como también es un punto de partida para desarrollar nuevos estudios debido a la variedad de temas que cubre en sus distintas unidades de análisis.

III.2. Objetivos

El objetivo general del estudio es diagnosticar el estado y evolución de las condiciones de empleo y trabajo, de las relaciones laborales y la igualdad de género en las empresas regidas por el Código del Trabajo en Chile.

Los objetivos específicos, se detallan a continuación:

- Caracterizar las empresas en relación con su entorno económico

[1] El INE y la DT desarrollaron las fases de análisis y difusión.

[2] Las versiones anteriores fueron realizadas en 1998, 1999, 2002, 2004, 2006, 2008, 2011, 2014 y 2019. Ver informe de resultados de las distintas versiones de la Encla, en el siguiente enlace https://www.dt.gob.cl/portal/1629/w3-propertyvalue-22780.html

- Caracterizar las condiciones de empleo en las empresas
- Caracterizar las condiciones trabajo en las empresas
- Analizar las relaciones laborales que se dan al interior de las empresas
- Analizar la relación entre condiciones de empleo y trabajo, y las relaciones laborales
- Identificar la situación de las empresas en materia de igualdad de género
- Detectar y describir fenómenos laborales emergentes

IV. Marco de referencia

IV.1. Marco conceptual

La Encla tiene como base las definiciones y directrices promovidas por la Organización Internacional del Trabajo (OIT) y las principales normas contenidas en el Código del Trabajo de Chile; base regulatoria de las relaciones laborales entre las personas trabajadoras y empleadoras.

Respecto a sus principales conceptos, la Encla está organizada sobre la base de tres temas: las condiciones de empleo, las condiciones de trabajo y las relaciones laborales (Dirección del Trabajo, 2024).

Las condiciones de empleo dan cuenta de las modalidades de inserción de los trabajadores en el mercado laboral. Este tipo de participación de los individuos es clave en las sociedades modernas pues de ella depende la participación de trabajadores(as) en la distribución del ingreso y el bienestar en general.

Las condiciones de trabajo corresponden al conjunto de aspectos o variables relativas al modo en que se organiza el proceso de trabajo, especialmente la organización de la fuerza de trabajo, los medios y el objeto de trabajo. Estas formas de organización del trabajo permiten conocer los efectos tanto organizativos como subjetivos que producen sobre las personas trabajadoras.

Las relaciones laborales corresponden al conjunto de relaciones que se establecen entre empleadores y trabajadores, en el marco de las normas legales establecidas para regularlas. De este modo, la conflictividad habitual en las relaciones entre los actores que componen el mercado laboral puede ser abordada mediante la regulación de los procesos de búsqueda de acuerdos en torno a las condiciones de empleo y trabajo, y conducida por cauces institucionales mediante procedimientos normados.

IV.2. Referencias nacionales e internacionales

A nivel nacional existe la experiencia de la Encuesta Anual para el Diagnóstico de las Relaciones Laborales en la Administración Central del Estado (ENCLACE), cuyo objetivo es analizar el estado y la evolución periódica de las relaciones laborales al nivel de los servicios de la Administración Central del Estado, excluyendo subsecretarías, intendencias, gobernación y siete servicios particulares debido a que se rigen por el Código del Trabajo (Universidad Alberto Hurtado,

Servicio Civil, 2017). En sus objetivos es una encuesta muy similar a la Encla, pero aplicada al sector público. Esta encuesta se utilizó para contrastar cómo se desenvuelven de las relaciones laborales en los sectores privado y público, destacando las diferencias en lo que respecta a los procesos de negociación colectiva en contextos con marcos normativos diferentes.

A nivel internacional se encuentra la Encuesta sobre Relaciones Laborales en el Lugar de Trabajo (Workplace Employment Relations Study, WERS por sus siglas en inglés) en el Reino Unido que se plantea como principal objetivo mapear las relaciones laborales en Gran Bretaña a lo largo del tiempo. De forma similar a la Encla, recopila datos en una muestra representativa de lugares de trabajo de distintas fuentes: empleadores, representantes de empleados y empleados (Department of Business, Innovation & Skills, 2013). La encuesta ha sido aplicada seis veces desde 1980. Junto con medir variables asociadas a relaciones laborales, la última versión de la WERS incluye preguntas sobre las percepciones y actitudes de las personas informantes respecto de los efectos de la recesión de 2008 en las empresas encuestadas, de forma similar a cómo la presente versión de la Encla consulta respecto de los efectos y medidas tomadas en el contexto de la crisis sanitaria provocada por la pandemia del COVID 19.

Adicionalmente, se tiene como referencia, la Encuesta de Relaciones Profesionales y Negociaciones de Empresa (Relations professionnelles et négociations d'entreprise, REPONSE) busca medir la intensidad y efectividad del diálogo social en la empresa y se realiza cada seis años desde 1993. Se compone de una entrevista con un representante de la gerencia, un representante del personal del establecimiento y otra entrevista a algún trabajador (DARES, 2017). Tal como en la Encla, esta encuesta entiende a las organizaciones sindicales en un ecosistema donde conviven sindicatos de establecimientos, inter-empresa, federaciones, confederaciones, entre otros, y busca ahondar en los vínculos que existen entre estas organizaciones.

Por último, este estudio se ancla a lo señalado en la Declaración de Principios y Derechos Fundamentales en el Trabajo de la OIT (2024), según la cual los miembros de esta organización se obligan a respetar, promover y llevar a cabo la libertad de asociación y el derecho de negociación colectiva entre los trabajadores, con miras a facilitar el cumplimiento de sus derechos laborales, el mejoramiento de sus condiciones laborales, salariales y contractuales, entre otros. La Encla es la principal encuesta del país para medir variables asociadas al derecho de asociación y negociación colectiva de los trabajadores(as).

IV.3. Marco normativo

Los productos estadísticos realizados con el INE se encuentran afectos a la normativa sobre Secreto Estadístico regulada en los artículos 29 y 30 de la Ley N° 17.374, respecto de hechos que se refieran a personas o entidades determinadas de que hayan tomado conocimiento en el desempeño de sus actividades. Dichos artículos indica que

"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 por cualquier persona sujeta a esta obligación hará incurrir en el delito previsto y penado por el artículo 247, del Código Penal, debiendo en todo caso aplicarse pena corporal.

Artículo 30°- 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."

Por tanto, toda información que recopile, genere o entregue, tendrá la característica de ser innominada, indeterminada e indeterminable.

Con base en la normativa vigente, los datos serán puestos a disposición de las y los usuarios tras la aplicación de los procedimientos definidos para la anonimización e indeterminación de las bases de datos, ciñéndose a lo establecido en:

- Compromiso de confidencialidad y reserva de datos protegidos por el secreto estadístico e información sensible y/o sujeta a la ley 19.628 sobre la protección de datos personales.
- Ley 20.285 Sobre acceso a la información pública, en su artículo 1 regula el principio de transparencia de la función pública, el derecho de acceso a la información de los órganos de la Administración del Estado, los procedimientos para el ejercicio del derecho y para su amparo, y las excepciones a la publicidad de la información.

V. Diseño estadístico

V.1. Población objetivo

Empresas formales y vigentes que se encuentran al 2021 con iniciación de actividades en el Servicio de Impuestos Internos (SII) y que al mismo año hayan realizado su actividad principal en algunos de los sectores en estudio según el Clasificador Chileno de Actividades Económicas (CIIU4.CL 2012), desarrollada dentro de los límites del país, excluyendo algunas comunas definidas como difícil acceso[3]. Además, que en noviembre del año comercial 2020-2021 declararon en los registros administrativos del Formulario 1887[4] del SII y de la Superintendencia de Seguridad Social (SUSESO), cuentan con cinco o más trabajadores contratados directamente (excluyendo honorarios) y cuyas relaciones laborales se rigen por el Código del Trabajo.

V.2. Cobertura geográfica

La cobertura geográfica de la Encla 2023 comprende las 16 regiones del país, excluyendo de este estudio a 22 comunas que son definidas como de difícil acceso o que presentan un alto costo de recolección de información en esas zonas. Las comunas excluidas del estudio de la Encla 2023, se describen en el Tabla 1.

[3] Para mayor detalle de las comunas excluidas ver Tabla 1.

[4] Contiene declaraciones sobre rentas del Art. 42 N°1 (sueldos) otros componentes de la remuneración y retención del impuesto único de segunda categoría de la ley de Rentas.

Tabla 1. Comunas excluidas del ámbito de estudio de la Encla 2023

Región	Provincia	Comunas
Arica y Parinacota	Parinacota	General Lagos
Tarapacá	Tamarugal	Colchane
Antofagasta	El Loa	Ollagüe
Valparaíso	Isla de Pascua	Isla de Pascua
	Valparaíso	Juan Fernández
Los Lagos	Llanquihue	Cochamó
	Palena	Chaitén - Futaleufú - Hualaihué - Palena
Aysén del General Carlos Ibáñez del Campo	Aisén	Guaitecas
	Coihayque	Lago Verde
	Capitán Prat	Villa O'Higgins - Tortel
Magallanes y de la Antártica Chilena	Antártica Chilena	Antártica - Cabo de Hornos
	Magallanes	Laguna Blanca - Río Verde - San Gregorio
	Tierra del Fuego	Timaukel - Primavera
	Última Esperanza	Torres del Paine

V.3. Cobertura temática

La Encla 2023 cubre 18 sectores económicos desde el ámbito silvoagropecuario hasta servicios que se encuentran en el Clasificador Chileno de Actividades Económicas (CIIU4.CL 2012) que es una adaptación del Clasificador Industrial Internacional Uniforme de todas las Actividades Económicas. Dicho clasificador permite la comparabilidad de la información a nivel de actividad económica, de acuerdo con estándares internacionales.

Tras la definición de los sectores económicos, para algunos casos en particular, se realiza una agrupación, conformando 13 agrupaciones de actividades económicas, las cuales se detallan en la siguiente tabla.

Tabla 2. Sectores de Actividad Económica según CIIU4.CL 2012 en estudio para la Encla 2023

Sección	Descripción
A	Agricultura, ganadería, silvicultura y pesca
B	Explotación de minas y canteras
C	Industrias manufactureras
D-E	Suministro de electricidad, gas, vapor y aire acondicionado y 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-J	Transporte y almacenamiento e Información y comunicaciones
I	Actividades de alojamiento y de servicio de comidas
K-L	Actividades financieras y de seguros y Actividades inmobiliarias
M-N	Actividades profesionales, científicas y técnicas y Actividades de servicios administrativos y de apoyo
P	Enseñanza
Q	Actividades de atención de la salud humana y de asistencia social
R-S	Actividades artísticas, de entretenimiento y recreativas y Otras actividades de servicios

Como casos particulares, se excluyen del estudio las cooperativas de agua, así como también, las instituciones públicas que, clasificadas en otra actividad económica, pertenecen a municipalidades y corporaciones municipales u otro organismo estatal que se consideran fuera del alcance del estudio, según un trabajo interno de revisión y análisis de identificación de instituciones públicas realizado por el INE para la depuración del marco estadístico.

Para las actividades de enseñanza y de atención de la salud humana y asistencia social, P y Q, respectivamente, se aplican otras exclusiones complementarias al catastro que dispone el INE, la primera de ellas, para las subclases 85311, 86101, 86201 se excluyen directamente, mientras que, para las subclases 85001 y 85420 la exclusión consiste en identificar, a nivel de subclase[5], unidades que declaren pertenecer a sociedades públicas asociadas a las municipales, identificando las corporaciones municipales. Por último, también se excluyen las secciones que O, T y U, las cuales se explicitan a continuación.

[5] Corresponde a la desagregación de actividad económica según CIIU4.CL 2012, a nivel de 5 dígitos.

Tabla 3. Secciones CIIU4.CL 2012 excluidas del ámbito de estudio de Encla 2023

Sección	Descripción Sección	Subclase	Descripción Subclase
P	Enseñanza	85001	Aquí se debe mostrar qué recomendación o buena práctica tomaron de las encuestas mencionadas en este apartado.
		85311	Enseñanza superior en universidades públicas
		85420	Enseñanza cultural
Q	Actividades de atención de la salud humana y de asistencia social	86101	Aquí se debe mostrar qué recomendación o buena práctica tomaron de las encuestas mencionadas en este apartado.
		86201	Actividades de centros de salud municipalizados (servicios de salud pública)
O	Administración pública y defensa; planes de seguridad social de afiliación obligatoria		
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		

V.4. Unidades estadísticas

V.4.1. Unidad de muestreo

La unidad de muestreo es la empresa formal, que desarrolla su actividad económica en el ámbito de estudio y dentro de los límites territoriales del país. Además, que, durante noviembre de 2021, haya declarado tener cinco o más personas trabajadoras asalariadas.

V.4.2. Unidad de análisis y de observación

La unidad de observación y de análisis es la empresa formal que cumple los requisitos para formar parte de la población objetivo. Es importante indicar que los análisis que se desprenden de los datos de la Encla no se limitan necesariamente a un nivel en particular; ya que, dependiendo del fenómeno que se quiera describir la unidad de observación pueden ser las empresas o alguna de sus características (principalmente el número de trabajadores/as); bajo diversas variables de agrupación. No obstante, la unidad económica es la unidad de análisis fundamental, sobre cuya base se desarrollan otros análisis más complejos.

V.5. Definición de variables

Las principales variables de Encla 2023, fueron definidas en torno a dos aristas:

- Variables imprescindibles para estimar el factor de expansión.
- Variables clave para el cumplimiento de los objetivos del estudio, encargadas por la Dirección del Trabajo, previa revisión por parte del INE.

La siguiente tabla describe las variables principales y secundarias que conformaron la grilla de suficiencia; es decir, que se establecieron como variables imprescindibles para establecer el logro de la unidad. Las “variables principales” requirieron un logro del 100% para considerar al cuestionario logrado. Adicional a lo anterior, en el caso de las “variables secundarias” se exigió un logro de al menos el 80% de dichas variables para considerar al cuestionario logrado. El incumplimiento de cualquiera de las dos condiciones anteriores significó considerar la unidad económica como no lograda durante la etapa de recolección.

Tabla 4. Variables principales y secundarias por cuestionario. Encla 2023

Cuestionario	Variables principales	Variables secundarias[6]
Empleadores	M2: Desde el año 2019 a la fecha ¿ha habido alguna negociación colectiva en esta empresa?	D1, F1, J3, J6, J7, J10, J11, J17, R4
Sindicatos	O1: En esta empresa, ¿alguna vez se ha teletrabajado? K2: ¿Cuántos hombres y mujeres se encuentran afiliados a su sindicato al día de hoy?	C2, D1, D2, F3, F5, G2, I2, J4, J5, J6, J12, J13, K1, K4, K6, K8, K9, K10, K15, K16, K19, M6, M16, M21, R6, R7, R8
Trabajadores	01: ¿En esta empresa, ¿alguna vez se ha teletrabajado?	C2, C3, F3, F5, G2, I1, J4, J5, J6, J10, K1, K3, K5, M7,
Autoaplicado	A4, Indique el total de personas contratadas directamente por esta empresa al 30 de noviembre de 2022, de acuerdo a las siguientes categorías de autoridad y sexo: C1_1, Durante el año 2022, ¿cuáles de las siguientes actividades han sido subcontratadas a otras empresas o a terceros? (desarrolladas dentro o fuera de las dependencias de la empresa). Desarrolladas en las dependencias de su empresa Sí - No C14, Durante el año 2022, ¿esta empresa contó con personas suministradas por terceros (empresas de servicios transitorios)? D1, Indique el total de personas contratadas directamente por la empresa con jornada completa (de 40 horas o más a la semana) al 30 de noviembre de 2022 según tramo de remuneraciones y sexo I2, Durante el año 2022, ¿la empresa realizó capacitaciones laborales para su personal?	A7, A8, C2, C3, C10, D2, D4, D5, D6, E2, E3, F1, F8, F11, H1, M2, M4, P1, P5, O18, N1

V.6. Clasificaciones estadísticas

Para facilitar la comparación de la información y de acuerdo con las normas internacionales, se utilizó el Clasificador Chileno de Actividades Económicas CIIU4.CL 2012. Este clasificador se rige por los mismos conceptos y criterios del Clasificador Internacional CIIU Rev.4, elaborado por la División de Estadísticas de las Naciones Unidas y se compone de 18 sectores económicos. Sin embargo, para la Encla estos fueron reclasificados en 13 agrupaciones económicas, para contar con dominios de estudio [7]vinculados a sus objetivos.

[6]La definición de estas variables se encuentra en el anexo 1. "Variables primarias y secundarias GDS Encla 2023".

[7] Dichos sectores y agrupaciones se presentan en la tabla 2.

V.7. Marco estadístico

V.7.1. Conformación del marco estadístico

A diferencia de las versiones anteriores, el marco estadístico se conforma a partir de tres fuentes de información, el Marco Maestro de Empresas INE (MME) año contable 2021; directorio de trabajadores asalariados obtenido desde Formulario N° 1887 (F1887) año comercial 2020, y los registros administrativos de la SUSESO 2021. El detalle de las fuentes de información utilizadas y la depuración que se realiza para la construcción del marco estadístico se presenta a continuación:

V.7.1.1. Marco Maestro de Empresas (MME) 2021

Se utiliza el insumo más actualizado del MME que corresponde al año contable 2021. A partir de este, se seleccionan aquellas empresas que ejercen su actividad económica principal en alguna de las mencionadas en la cobertura temática.

V.7.1.2. Formulario 1887 referencia noviembre 2020

Desde el directorio proveniente del F1887, se obtiene el total de trabajadores asalariados en noviembre de 2020, siendo este el año más actualizado que posee el INE al momento de la construcción del marco estadístico para el estudio. De este insumo, se incluyen las empresas que reportan más de cinco trabajadores en noviembre de 2020.

V.7.1.3. Registros administrativos de la SUSESO referencia noviembre 2021

El registro administrativo GRIS Mutuales SUSESO es proporcionado por esta institución y recibido por el INE periódicamente. Este proporciona mensualmente datos nominados de base de datos, entre ellos se puede obtener el número de trabajadores que reporta cada empresa. Al igual que para el F1887, se incluyen las empresas que reportan más de cinco trabajadores en noviembre de 2021.

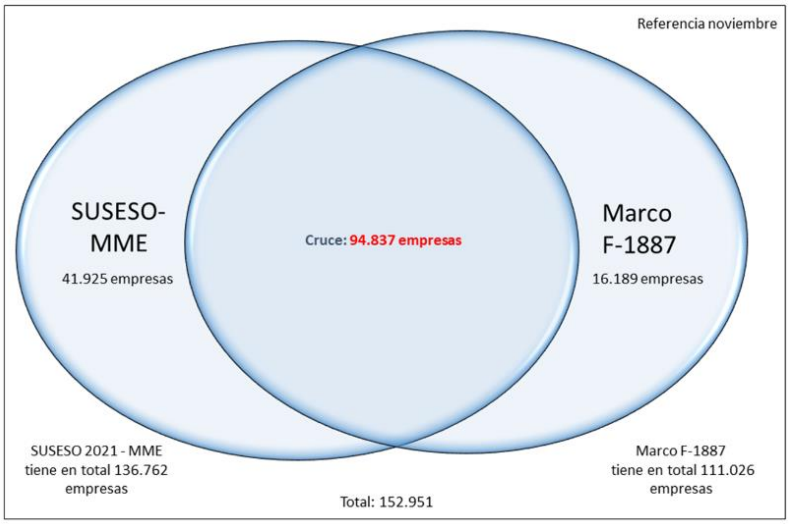
V.7.2. Depuración del marco estadístico

Paso 1: Se cruza el MME con F1887, obteniendo la actividad económica y región desde el MME y el número de trabajadores/as del F1887. A partir de la intersección de ambos directorios y al excluir las empresas con menos de cinco trabajadores/as y las actividades que se encuentran fuera de estudio quedan 111.026 empresas.

Por otra parte, se cruza el MME con SUSESO y se realizan las mismas exclusiones mencionadas anteriormente, es decir, se excluyen las actividades fuera del estudio y las empresas con menos de cinco trabajadores. Como resultado de este cruce se obtienen en total 136.762 empresas.

Paso 2: Se realiza el cruce con las dos bases de datos obtenidas en el paso 1. Para continuar con la construcción del marco estadístico, se analiza tanto el resultado de la intersección como del complemento del cruce. De manera gráfica se muestra el resultado obtenido, en la siguiente figura.

Figura 1. Diagrama del cruce entre las fuentes de información F1887 y SUSESO



Tal como se observa, existen 94.837 empresas que están en ambas fuentes de información (F1887 y SUSESO), 41.925 empresas que solo están en SUSESO y 16.189 que solo están en el F1887. La suma de todas las partes equivale a 152.951 empresas.

Además, con el fin de profundizar en el análisis se verifica la distribución del tamaño de las empresas que están solo en F1887 o solo en SUSESO. Tal como se muestra en la tabla 5, la mayoría de las empresas que están en una u otra fuente son de tamaño micro, seguida por las de tamaño pequeña.

Tabla 5. Distribución de las empresas que no coinciden en ambas fuentes de información, según tamaño de la empresa

Tamaño	Marco SUSESO	Marco F-1887
Micro	31.342	12.165
Pequeña	9.860	3.639
Mediana	657	333
Grande	66	52
Total	41.925	16.189

Considerando que el F1887 es de referencia 2020, es decir, con un año de desfase con respecto de las demás fuentes utilizadas, y además un año atípico debido a la pandemia de COVID-19, es muy probable que estas empresas ya no existan en el país, por lo tanto, se decide junto a la DT excluir del marco estadístico a las empresas pequeñas y micro que solo se encuentran en este insumo.

Paso 3: Finalmente se realiza una depuración de empresas, a través de la información más reciente de recolección de encuestas realizadas por el INE. Para esta depuración se excluyen del marco estadístico todas las unidades que presentan estados de recolección *sin movimiento* y *término de giro* en dichas encuestas. Las encuestas utilizadas corresponden a:

Figura 2. Fuentes de información utilizadas para depurar el marco muestral



Después de la depuración, el marco estadístico queda conformado por 134.035 empresas.

Paso 4: Finalmente, una vez obtenido el listado final de unidades que pertenecen al marco estadístico, se asigna el tamaño de empresa a cada unidad, según los siguientes criterios:

Tabla 6. Criterios de priorización para el tamaño de empresas en el marco estadístico de la Encla 2023

Fuente de información en la que está la empresa	Información que se mantiene en el marco
Solo F1887 2020	F1887
Solo SUSESO 2021	SUSESO 2021
En ambas (F1887 2020 – SUSESO 2021)	SUSESO 2021

Este criterio se establece con la finalidad de priorizar la información más actualizada disponible de la variable número de trabajadores.

V.7.3. Estratificación del marco estadístico

Una vez construido el marco estadístico, se realiza la estratificación. Esta permite reunir elementos de una población en grupos homogéneos previo a la selección, con el propósito de mejorar las medidas de precisión de las estimaciones agrupando las unidades del marco estadístico en clases homogéneas en su interior y que difieran de las características del resto. Los estratos deben ser mutuamente excluyentes y exhaustivos colectivamente, es decir, cada elemento en la población debe ser asignado a un solo estrato y ninguno puede quedar excluido.

Los estratos que se conforman en el marco estadístico corresponden al cruce entre región, agrupación de actividades económicas de acuerdo con la CIIU4.CL 2012, tamaño de empresa según número de trabajadores y tramos de inclusión forzosa (IF) y aleatoria (IA).

Tabla 7. Agrupación de actividad económica según CIIU4.CL 2012 que conforman el ámbito de estudio de Encla 2023

Agrupación actividad económica	Actividad económica	Descripción de la actividad económica
A	A	Agricultura, ganadería, silvicultura y pesca
B	B	Explotación de minas y canteras
C	C	Industrias manufactureras
D-E	D	Suministro de electricidad, gas, vapor y aire acondicionado
	E	Suministro de agua; evacuación de aguas residuales, gestión de desechos y descontaminación
F	F	Construcción
G	G	Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas
H-J	H	Transporte y almacenamiento
	J	Información y comunicaciones
I	I	Actividades de alojamiento y de servicio de comidas
K-L	K	Actividades financieras y de seguros
	L	Actividades inmobiliarias
M-N	M	Actividades profesionales, científicas y técnicas
	N	Actividades de servicios administrativos y de apoyo
P	P	Enseñanza
Q	Q	Actividades de atención de la salud humana y de asistencia social
R-S	R	Actividades artísticas, de entretenimiento y recreativas
	S	Otras actividades de servicios

A continuación, se describen las variables consideradas para conformar los estratos en el marco estadístico de la Encla 2023:

V.7.3.1. Región y agrupación regional

El marco estadístico contiene empresas con casa matriz en las 16 regiones del país, lo que permite conformar el estrato. A su vez estas regiones, se agrupan de acuerdo con sus características geográficas, conformando 9 agrupaciones regionales solicitadas por la DT. Si bien, se construye en el marco estadístico la agrupación regional, esta solo responde a la necesidad de obtener estimaciones a este nivel, siendo no considerada como un nivel de estratificación. En la siguiente tabla se muestra la nueva agrupación a partir de las 16 regiones.

Tabla 8. Agrupación regional a partir de las regiones, solicitadas por la DT.

Agrupación Regional	Nombre de región
Norte Grande	Arica y Parinacota
	Tarapacá
Antofagasta	Antofagasta
Norte Chico	Atacama
	Coquimbo
Valparaíso	Valparaíso
Metropolitana	Metropolitana de Santiago
Centro Sur	Libertador General Bernardo O´Higgins
	Maule
	Ñuble
Bio-Bío	Bio-Bío
Sur	La Araucanía
	Los Ríos
	Los Lagos
Zona Austral	Aysén del General Carlos Ibáñez del Campo
	Magallanes y Antártica Chilena

V.7.3.2. Agrupación de actividades económicas

Las actividades económicas especificadas en la cobertura temática (tabla 2) se agrupan conformando las 13 agrupaciones de actividad económica en estudios, las que forman una variable de estratificación en el marco estadístico.

V.7.3.3. Tamaño de empresas

La clasificación de las empresas según tamaños para la Encla 2023, establece cuatro categorías; grande, mediana, pequeña y micro, definidas según el total del número de personas trabajadoras asalariadas contratadas por la empresa en el período noviembre de 2021. En la Tabla 9 se presentan los límites que definen los tramos de número de personas trabajadoras.

Tabla 9. Tamaño de empresas según número de trabajadores/as asalariados/as

	Número de trabajadores/as asalariados/as	
	Límite inferior	Límite Superior
Grande	200	Más
Mediana	50	199
Pequeña	10	49
Micro	5	9

V.7.3.4. Tramo de inclusión

El tramo de inclusión permite diferenciar las unidades que se incluyen forzosamente (IF) de aquellas en las que su inclusión es aleatoria (IA). Para la Encla 2023, a diferencia de las versiones anteriores de la encuesta, las unidades IF son escogidas por reportar en el período noviembre 2021, un total de personas trabajadoras asalariadas mayor o igual a 2.000, consideradas como unidades autorepresentadas, ya que presentan un alto aporte en el sector económico al cual pertenecen. Esto motiva la conveniencia de crear estratos de inclusión forzosa (IF) dentro de estas actividades, con la intención de mejorar la precisión de las estimaciones de interés.

Respecto al tramo de inclusión aleatoria (IA), este se conforma con el resto de las unidades económicas en la población, con las cuales se estima un tamaño muestral a través de un diseño probabilístico.

V.8. Diseño muestral

V.8.1. Tipo de muestreo

El diseño aplicado corresponde a un muestreo probabilístico y estratificado, donde los estratos se encuentran definidos según el cruce de las variables región, agrupación de actividades económicas de acuerdo con la CIIU4.CL 2012, tamaño de empresa según número de trabajadores y tramos de inclusión; forzosa (IF) y aleatoria (IA).

V.8.2. Dominio de estudio

Los dominios de estudio que permiten asegurar una muestra representativa para la estimación del "Total de trabajadores" con un error muestral esperado que no supere el 30% para la Encla 2023, son los siguientes:

- Nacional
- Región (16 regiones)
- Agrupación Actividad Económica (13 agrupaciones)
- Tamaño de empresa (grande, mediana, pequeña, micro)
- Agrupación regional y tamaño de empresas (9 agrupaciones regionales)

V.8.3. Cálculo y distribución del tamaño muestral

De acuerdo con la metodología propuesta para el diseño muestral de la Encla 2023, se considera un tramo de Inclusión Forzosa (IF), el cual corresponde a un conjunto de unidades que declaran tener contratados mayor o igual a 2.000 personas trabajadoras asalariadas, siendo estas unidades identificas a priori en el marco estadístico y excluidas del cálculo del tamaño de la muestra. En el siguiente capítulo se describe la metodología empleada para definir el tamaño de muestra requerida para el tramo de Inclusión Aleatoria (IA).

Para el cálculo del tamaño muestral se considera como variable de diseño el número de trabajadores/as asalariados/as, cuyo parámetro de interés asociado es el total del número de trabajadores (Tabla 10). Por otro lado, para cubrir posibles cambios, términos de actividad y rechazo en trabajo de campo de las empresas seleccionadas, se utiliza la estrategia de sobremuestreo, que permite gestionar la muestra objetivo y las unidades adicionales simultáneamente en terreno. La sobremuestra se define de acuerdo con la tasa de logro de la Encla 2019 complementada con la tasa de logro obtenida de la Encuesta de Innovación en Empresas (ENI) 2019-2020, esto considerando ajustes del tipo presupuestario y técnico.

Tabla 10. Elementos utilizados en el cálculo del tamaño muestral de Encla 2023

Parámetro	Descripción
Variable de diseño	Número de trabajadores: y
Estimador asociado	Estimador del total del número de trabajadores: $\hat{Y}$
Parámetro asociado	Total, del número de trabajadores:
Niveles de estimación para los cuales se establece el error de muestreo	- Nacional: 5% - Región: 6% - Tamaño de empresas: 6% - Agrupación de actividad económica: 11% - Cruce de agrupación regional y tamaño de empresas: 15%
Nivel de confianza	95%

El diseño muestral de la Encla 2023 es probabilístico y estratificado. Los estratos están conformados, entre otras variables, por el tramo de inclusión forzosa e inclusión aleatoria, donde la primera inclusión corresponde al grupo unidades autorepresentadas, es decir, que solo se representan a sí mismas, y su método de selección se detalla en el apartado sobre la estratificación del marco estadístico. En cambio, para las unidades que pertenecen al tramo inclusión aleatoria, se calcula un tamaño muestral por medio del método que se describe a continuación.

Paso 1: Desde la Encla 2019 se obtuvieron los tamaños de la muestra iniciales calculados por estrato. Este diseño estratificado considera como variable de interés el número de trabajadores, fijando errores relativos del 15% a nivel del estrato, con un coeficiente de confianza del 95%.

$$n_{rhv} = \frac{\left(\sum_{i=1}^{I} N_{rhv_i} \cdot \sigma_{rhv_i}\right)^2 \cdot \left(Z_{1-\frac{\alpha}{2}}\right)^2}{e_r \cdot \sum_{i=1}^{I} Y_i + \left(\sum_{r=1}^{R} N_{rhv_i} \cdot \sigma_{rhv_i}^2\right) \cdot \left(Z_{1-\frac{\alpha}{2}}\right)^2} \quad (1)$$

Donde:

n_{rhv} : Tamaño de la muestra de empresas a nivel de cada estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

N_{rhv_i} : Número total de empresas i en el marco estadístico para el estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

σ_{rhv_i} : Desviación estándar del número de trabajadores de las empresas i en el estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

e_r : Error relativo máximo admisible de 15%.

Y_i : Total del número de trabajadores/as en el marco estadístico para las empresas i en el estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

$Z_{1-\frac{\alpha}{2}}$: Estadístico normal estándar correspondiente a un nivel de confianza del 95% ($Z_{0,975}$=1,96≈2).

Paso 2: El tamaño de muestra objetivo obtenido en cada estrato de región o agrupación regional, agrupación de actividad y tamaño de empresas, se ajusta siguiendo las siguientes reglas:

- Si el tamaño de muestra objetivo (n_{rhv}) calculado en el paso 1 es mayor al número de empresas disponibles en el marco estadístico (N_{rhv}) para el estrato, se incluyen todas las empresas del marco estadístico en la muestra.

- Si el tamaño de muestra objetivo (n_{hvm}) calculado en el paso 1 es menor al número de empresas disponible en el marco estadístico (N_{rhv}) para el estrato, se mantiene el tamaño de la muestral calculada.

Paso 3: Luego para obtener el tamaño de la muestra objetivo a nivel de región o agrupación regional r, agrupación de actividad h y tamaño de empresas v, respectivamente, se suman los tamaños muestrales obtenidos para los estratos a nivel más desagregado.

- Tamaño muestral a nivel de tamaño de empresas:

$$n_v = \sum_{r=1}^{R} \sum_{h=1}^{H} n_{rhv} \quad (2)$$

Donde:

n_v : Tamaño de la muestra de empresas a nivel de tamaño de empresas v.

n_{rhv} : Tamaño de la muestra a nivel de cada estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

- Tamaño muestral a nivel de agrupación de actividad económica:

$$n_h = \sum_{r=1}^{R} \sum_{v=1}^{V} n_{rhv} \quad (3)$$

Donde:

n_h : Tamaño de la muestra de empresas a nivel de agrupación de actividad h.
n_{rhv} : Tamaño de la muestra a nivel de cada estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

- Tamaño muestral a nivel de región o agrupación regional:

$$n_r = \sum_{h=1}^{H} \sum_{v=1}^{V} n_{rhv} \quad (4)$$

Donde:

n_r : Tamaño de la muestra de empresas a nivel de región o agrupación regional r.
n_{rhv} : Tamaño de la muestra a nivel de cada estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

Paso 4. Con la información de la variable número de trabajadores al mes de referencia noviembre 2021, contenida en el marco estadístico, se calculan las varianzas, las sumas totales, coeficientes de variación y errores relativos esperados.

Paso 5. Luego de calculados los errores esperados para los niveles de estimación requeridos, se observan dichos errores y en caso de superar el umbral de error establecido, se debe aumentar el tamaño de la muestra en los estratos correspondientes, por lo tanto, el tamaño de la muestra ajustado en el estrato se denomina n'_{rhv}.

Paso 6. Finalmente, se calcula la sobremuestra a partir de la tasa de logro observada en la Encla 2019 y la Encuesta Nacional de Innovación año de referencia 2019-2020 (ENI 2019-2020). Se decide utilizar ambas tasas debido a los cambios observados en el operativo de campo, luego de la contingencia de salud nacional. Considerando que la recolección de la Encla 2019 se realizó durante el 2019, bajo un escenario de trabajo de campo estable, las tasas de logro obtenidas fueron muy cercanas al 100%. Sin embargo, producto de la emergencia sanitaria ocurrida en el 2020, se evidenció una disminución considerable en las tasas de logro en diferentes encuestas, debido a la dificultad de encontrar a las unidades en el trabajo de campo. Bajo el contexto descrito anteriormente, se incluye la tasa de la ENI 2019-2020 como fuente primaria, asumiendo un escenario más realista de la situación actual en el país puesto que su recolección se realizó en el 2021 y se complementa con la tasa de logro de la Encla 2019, la cual corresponde a la fuente secundaria.

Los criterios aplicados para el uso de las tasas de logro en los tamaños muestrales se detallan a continuación:

Paso 1: para los estratos existentes en ambas encuestas (Encla-ENI), se utiliza la tasa de logro de la ENI 2019-2020 y la Encla 2019, de acuerdo con los siguientes criterios:

- Se incorpora como tasa primaria la obtenida en la ENI 2019-2020.

- En caso de que la tasa de la ENI sea igual a cero o mayor o igual a 1, se utiliza la tasa de la Encla 2019.

Paso 2: para los estratos en que no se cuente con información de las tasas, se utiliza el 30% correspondiente al valor utilizado en reemplazos para versiones anteriores.

Paso 3: en cuanto a la actividad económica Enseñanza (P) se utiliza la tasa de logro de Encla 2019 y en caso de no contar con dicha información se utiliza el 30%.

Así entonces, el tamaño muestral objetivo con la sobremuestra se calcula de la siguiente forma:

$$n^s_{rhv} = \frac{n'_{rhv}}{tl_{rhv}} \quad (5)$$

Donde:

n^s_{rhv} : Tamaño de muestra con sobremuestreo ajustado según la tasa de logro de la encuesta primaria o secundaria por región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

n'_{rhv} : Tamaño de muestra ajustado para el estrato según región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

tl_{rhv} : Tasa de logro según ajuste aplicado, en la región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

Finalmente, los tamaños muestrales objetivos con la sobremuestra a nivel más agregado quedan conformados de la siguiente forma:

- Tamaño muestral a nivel de tamaño de empresas:

$$n^s_v = \sum_{r=1}^{R} \sum_{h=1}^{H} n^s_{rhv} \quad (6)$$

Donde:

n^s_v : Tamaño de la muestra de empresas con sobremuestreo a nivel de tamaño de empresas v.

n^s_{rhv} : Tamaño de la muestra con sobremuestra a nivel de cada estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

- Tamaño muestral a nivel de agrupación de actividad económica:

$$n^s_h = \sum_{r=1}^{R} \sum_{v=1}^{V} n^s_{rhv} \quad (7)$$

Donde:

n^s_h : Tamaño de la muestra de empresas con sobremuestreo a nivel de agrupación de actividad h.

n^s_{rhv} : Tamaño de la muestra con sobremuestra a nivel de cada estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

-

- Tamaño muestral a nivel de región o agrupación regional:

$$n_r^s = \sum_{h=1}^{H} \sum_{v=1}^{V} n_{rhv}^s \qquad (8)$$

Donde:

n_r^s : Tamaño de la muestra de empresas con sobremuestreo a nivel de región o agrupación regional r.

n_{rhv}^s : Tamaño de la muestra con sobremuestra a nivel de cada estrato, región o agrupación regional r, agrupación de actividad h y tamaño de empresas v.

Por último, el incremento de la sobremuestra en forma considerable abrió paso a nuevos ejercicios de sobremuestra, los cuales consistieron en rotar las tasas utilizadas previamente: así entonces, se utilizó como fuente primaria la Encla 2019, y como secundaria la XII ENI. Estos ajustes se realizan en las agrupaciones de actividad económica donde la sobremuestra dobla los tamaños muestrales objetivos, con la meta de reducir el total de unidades a gestionar en el trabajo de campo, para ajustarse de mejor forma al presupuesto disponible. Lo anterior, considerando que una parte del proceso de recolección de la Encuesta Laboral se realiza en forma presencial, aumentando la posibilidad de contactar a la empresa, por lo que el cambio de tasas no debiese generar una disminución del logro.

V.8.4. Selección de las unidades muestrales

Posteriormente la selección de unidades se realiza para cada tramo de inclusión en forma independiente: por un lado, las unidades IF quedaron previamente seleccionadas por el umbral de un número de trabajadores/as mayor o igual a 2.000; por otro, la selección de las unidades de IA se realiza a través del método aleatorio definido con igual probabilidad a nivel del estrato. Al ser una estrategia de muestreo sistemático, la probabilidad de selección de las unidades de cada estrato está definida por la primera unidad en ser seleccionada.

Al igual que el muestreo aleatorio, el muestreo sistemático es un muestreo probabilístico donde cada unidad del marco tiene la misma probabilidad de ser elegida. Al contar con un marco estadístico estratificado a los niveles de región, actividad económica y tamaño, el muestreo sistemático permite garantizar la cobertura de unidades en todos los estratos al realizar un recorrido uniforme completo al interior de cada uno.

Las empresas del tramo IA son seleccionadas aleatoriamente al interior de cada estrato dado por el cruce de región, tamaño de la empresa según número de trabajadores y agrupación de actividad económica.

V.9. Resultado del tamaño muestral

En las tablas siguientes se presenta el número de empresas en el marco estadístico, el tamaño de muestra objetivo con su respectiva sobremuestra, y la muestra efectivamente lograda en la recolección. Se determina un tamaño de muestra objetivo de 4.267 empresas, de las cuales 202 son del estrato de inclusión forzosa (IF), y 4.065 del estrato de selección aleatoria (IA). Luego, tras el trabajo de campo, se logra una muestra de 173 empresas IF y 4.323 unidades IA, sumando una muestra efectiva de 4.496 unidades.

Tabla 11. Tamaño muestral objetivo y efectivo según tamaños de empresas

Tamaños	Marco estadístico	Muestra objetivo			Sobremuestreo	Total, a gestionar	Muestra efectiva		
		Total	IF	IA			Total	IF	IA
Total	**134.035**	**4.267**	**202**	**4.065**	**4.085**	**8.352**	**4.496**	**173**	**4.323**
Grande	3.234	1.028	202	826	289	1.317	981	171	810
Mediana	10.742	972	-	972	811	1.783	1.087	2	1.085
Pequeña	56.483	1.616	-	1.616	2.095	3.711	1.809	-	1.809
Micro	63.576	651	-	651	890	1.541	619	-	619

.

Tabla 12. Tamaño muestral objetivo y efectivo según región

Región	Marco estadístico	Muestra objetivo			Sobremuestreo	Total, a gestionar	Muestra efectiva		
		Total	IF	IA			Total	IF	IA
Total	**134.035**	**4.267**	**202**	**4.065**	**4.085**	**8.352**	**4.496**	**173**	**4.323**
Arica y Parinacota	1.319	185	-	185	133	318	189	-	189
Tarapacá	2.504	217	-	217	194	411	257	-	257
Antofagasta	3.882	221	1	220	192	413	212	1	211
Atacama	1.611	212	1	211	212	424	193	-	193
Coquimbo	4.477	244	-	244	285	529	277	-	277
Valparaíso	12.481	307	7	300	270	577	325	7	318
Metropolitana	65.853	647	174	473	356	1.003	670	148	522
O'Higgins	6.403	273	7	266	358	631	322	6	316
Maule	7.044	270	2	268	295	565	309	2	307
Ñuble	2.691	267	1	266	337	604	258	1	257
Biobío	9.385	284	5	279	298	582	331	4	327
La Araucanía	5.418	273	2	271	305	578	274	2	272
Los Ríos	2.324	228	-	228	236	464	249	-	249
Los Lagos	6.407	280	1	279	272	552	249	1	248
Aysén	690	168	-	168	195	363	174	-	174
Magallanes	1.546	191	1	190	147	338	207	1	206

Tabla 13.Tamaño muestral objetivo y efectivo según agrupación de actividad económica[8]

Actividad económica	Marco estadístico	Muestra objetivo			Sobremuestreo	Total, a gestionar	Muestra efectiva		
		Total	IF	IA			Total	IF	IA
Total	**134.035**	**4.267**	**202**	**4.065**	**4.085**	**8.352**	**4.496**	**173**	**4.323**
A	11.595	392	2	390	600	992	385	4	381
B	870	210	5	205	156	366	159	7	152
C	16.055	427	23	404	309	736	456	16	440
D-E	999	218	2	216	190	408	226	2	224
F	16.636	450	20	430	519	969	447	18	429
G	29.855	457	29	428	386	843	498	27	471
H-J	13.383	361	17	344	314	675	381	19	362
I	8.077	285	15	270	262	547	309	13	296
K-L	5.675	233	9	224	221	454	202	9	193
M-N	14.946	439	54	385	363	802	466	38	428
P	5.439	317	10	307	279	596	482	7	475
Q	3.309	239	13	226	214	453	316	11	305
R-S	7.196	239	3	236	272	511	169	2	167

Tabla 14. Tamaño muestral objetivo y efectivo según región y tamaño

Región	Tamaño	Marco estadístico	Muestra objetivo			Sobremuestreo	Total, a gestionar	Muestra efectiva		
			Total	IF	IA			Total	IF	IA
Total		**134.035**	**4.267**	**202**	**4.065**	**4.085**	**8.352**	**4.496**	**173**	**4.323**
Arica y Parinacota		**1.319**	**185**	**-**	**185**	**133**	**318**	**189**	**-**	**189**
	Grande	12	12	-	12	-	12	8	-	8
	Mediana	69	41	-	41	14	55	39	-	39
	Pequeña	520	88	-	88	81	169	109	-	109
	Micro	718	44	-	44	38	82	33	-	33
Tarapacá		**2.504**	**217**	**-**	**217**	**194**	**411**	**257**	**-**	**257**
	Grande	26	26	-	26	-	26	21	-	21
	Mediana	162	56	-	56	41	97	65	-	65
	Pequeña	1.028	92	-	92	97	189	122	-	122
	Micro	1.288	43	-	43	56	99	49	-	49
Antofagasta		**3.882**	**221**	**1**	**220**	**192**	**413**	**212**	**1**	**211**
	Grande	56	39	1	38	11	50	40	1	39
	Mediana	325	63	-	63	36	99	53	-	53
	Pequeña	1.634	78	-	78	89	167	80	-	80
	Micro	1.867	41	-	41	56	97	39	-	39
Atacama		**1.611**	**212**	**1**	**211**	**212**	**424**	**193**	**-**	**193**

[8] Para las actividades económicas Agricultura, ganadería, silvicultura y pesca (A), Explotación de minas y canteras (B) y Transporte y almacenamiento e Información y comunicaciones (H-J) se observa que el número de empresas clasificadas como inclusión forzosa en la muestra efectiva son mayores al número de empresas de la muestra objetivo, lo cual se debe a un cambio de actividad económica de estas unidades.

Región	Tamaño	Marco estadístico	Muestra objetivo			Sobremuestreo	Total, a gestionar	Muestra efectiva		
			Total	IF	IA			Total	IF	IA
	Grande	37	37	1	36	-	37	24	-	24
	Mediana	118	50	-	50	26	76	50	-	50
	Pequeña	676	82	-	82	120	202	84	-	84
	Micro	780	43	-	43	66	109	35	-	35
Coquimbo		**4.477**	**244**	**-**	**244**	**285**	**529**	**277**	**-**	**277**
	Grande	44	38	-	38	4	42	29	-	29
	Mediana	303	64	-	64	60	124	77	-	77
	Pequeña	1.955	103	-	103	163	266	125	-	125
	Micro	2.175	39	-	39	58	97	46	-	46
Valparaíso		**12.481**	**307**	**7**	**300**	**270**	**577**	**325**	**7**	**318**
	Grande	165	94	7	87	37	131	93	7	86
	Mediana	781	64	-	64	65	129	85	-	85
	Pequeña	5.380	110	-	110	121	231	116	-	116
	Micro	6.155	39	-	39	47	86	31	-	31
Metropolitana		**65.853**	**647**	**174**	**473**	**356**	**1.003**	**670**	**148**	**522**
	Grande	2.259	394	174	220	107	501	390	146	244
	Mediana	5.900	92	-	92	85	177	116	2	114
	Pequeña	27.956	122	-	122	123	245	132	-	132
	Micro	29.738	39	-	39	41	80	32	-	32
O'Higgins		**6.403**	**273**	**7**	**266**	**358**	**631**	**322**	**6**	**316**
	Grande	116	70	7	63	34	104	66	6	60
	Mediana	535	67	-	67	91	158	89	-	89
	Pequeña	2.772	97	-	97	165	262	124	-	124
	Micro	2.980	39	-	39	68	107	43	-	43
Maule		**7.044**	**270**	**2**	**268**	**295**	**565**	**309**	**2**	**307**
	Grande	101	54	2	52	25	79	59	2	57
	Mediana	603	71	-	71	107	178	92	-	92
	Pequeña	2.915	106	-	106	113	219	111	-	111
	Micro	3.425	39	-	39	50	89	47	-	47
Ñuble		**2.691**	**267**	**1**	**266**	**337**	**604**	**258**	**1**	**257**
	Grande	46	37	1	36	3	40	29	1	28
	Mediana	184	62	-	62	55	117	58	-	58
	Pequeña	1.065	129	-	129	211	340	125	-	125
	Micro	1.396	39	-	39	68	107	46	-	46
Biobío		**9.385**	**284**	**5**	**279**	**298**	**582**	**331**	**4**	**327**
	Grande	173	88	5	83	33	121	93	4	89
	Mediana	706	63	-	63	62	125	79	-	79
	Pequeña	3.907	94	-	94	140	234	116	-	116
	Micro	4.599	39	-	39	63	102	43	-	43
La Araucanía		**5.418**	**273**	**2**	**271**	**305**	**578**	**274**	**2**	**272**
	Grande	62	47	2	45	13	60	48	2	46
	Mediana	357	73	-	73	59	132	73	-	73
	Pequeña	2.223	114	-	114	175	289	117	-	117
	Micro	2.776	39	-	39	58	97	36	-	36
Los Ríos		**2.324**	**228**	**-**	**228**	**236**	**464**	**249**	**-**	**249**
	Grande	16	16	-	16	-	16	11	-	11
	Mediana	128	55	-	55	35	90	53	-	53
	Pequeña	903	115	-	115	139	254	138	-	138
	Micro	1.277	42	-	42	62	104	47	-	47
Los Lagos		**6.407**	**280**	**1**	**279**	**272**	**552**	**249**	**1**	**248**
	Grande	103	58	1	57	22	80	51	1	50
	Mediana	438	74	-	74	47	121	76	-	76
	Pequeña	2.615	109	-	109	151	260	101	-	101
	Micro	3.251	39	-	39	52	91	21	-	21
Aysén		**690**	**168**	**-**	**168**	**195**	**363**	**174**	**-**	**174**
	Grande	3	3	-	3	-	3	4	-	4

Región	Tamaño	Marco estadístico	Muestra objetivo			Sobremuestreo	Total, a gestionar	Muestra efectiva		
			Total	IF	IA			Total	IF	IA
	Mediana	33	33	-	33	-	33	29	-	29
	Pequeña	291	91	-	91	132	223	105	-	105
	Micro	363	41	-	41	63	104	36	-	36
Magallanes		**1.546**	**191**	**1**	**190**	**147**	**338**	**207**	**1**	**206**
	Grande	15	15	1	14	-	15	15	1	14
	Mediana	100	44	-	44	28	72	53	-	53
	Pequeña	643	86	-	86	75	161	104	-	104
	Micro	788	46	-	46	44	90	35	-	35

Tabla 15. Tamaño muestral objetivo y efectivo según agrupación de región y tamaño

Agrupación regional	Tamaño	Marco estadístico	Muestra objetivo			Sobremuestreo	Total, a gestionar	Muestra efectiva		
			Total	IF	IA			Total	IF	IA
Total		**134.035**	**4.267**	**202**	**4.065**	**4.085**	**8.352**	**4.496**	**173**	**4.323**
Norte grande		**3.823**	**402**	**-**	**402**	**327**	**729**	**446**	**-**	**446**
	Grande	38	38	-	38	-	38	29	-	29
	Mediana	231	97	-	97	55	152	104	-	104
	Pequeña	1.548	180	-	180	178	358	231	-	231
	Micro	2.006	87	-	87	94	181	82	-	82
Antofagasta		**3.882**	**221**	**1**	**220**	**192**	**413**	**212**	**1**	**211**
	Grande	56	39	1	38	11	50	40	1	39
	Mediana	325	63	-	63	36	99	53	-	53
	Pequeña	1.634	78	-	78	89	167	80	-	80
	Micro	1.867	41	-	41	56	97	39	-	39
Norte chico		**6.088**	**456**	**1**	**455**	**497**	**953**	**470**	**-**	**470**
	Grande	81	75	1	74	4	79	53	-	53
	Mediana	421	114	-	114	86	200	127	-	127
	Pequeña	2.631	185	-	185	283	468	209	-	209
	Micro	2.955	82	-	82	124	206	81	-	81
Valparaíso		**12.481**	**307**	**7**	**300**	**270**	**577**	**325**	**7**	**318**
	Grande	165	94	7	87	37	131	93	7	86
	Mediana	781	64	-	64	65	129	85	-	85
	Pequeña	5.380	110	-	110	121	231	116	-	116
	Micro	6.155	39	-	39	47	86	31	-	31
Metropolitana		**65.853**	**647**	**174**	**473**	**356**	**1.003**	**670**	**148**	**522**
	Grande	2.259	394	174	220	107	501	390	146	244
	Mediana	5.900	92	-	92	85	177	116	2	114
	Pequeña	27.956	122	-	122	123	245	132	-	132
	Micro	29.738	39	-	39	41	80	32	-	32
Centro Sur		**16.138**	**810**	**10**	**800**	**990**	**1.800**	**889**	**9**	**880**
	Grande	263	161	10	151	62	223	154	9	145
	Mediana	1.322	200	-	200	253	453	239	-	239
	Pequeña	6.752	332	-	332	489	821	360	-	360
	Micro	7.801	117	-	117	186	303	136	-	136
Biobío		**9.385**	**284**	**5**	**279**	**298**	**582**	**331**	**4**	**327**
	Grande	173	88	5	83	33	121	93	4	89
	Mediana	706	63	-	63	62	125	79	-	79
	Pequeña	3.907	94	-	94	140	234	116	-	116
	Micro	4.599	39	-	39	63	102	43	-	43
Sur		**14.149**	**781**	**3**	**778**	**813**	**1.594**	**772**	**3**	**769**
	Grande	181	121	3	118	35	156	110	3	107
	Mediana	923	202	-	202	141	343	202	-	202
	Pequeña	5.741	338	-	338	465	803	356	-	356
	Micro	7.304	120	-	120	172	292	104	-	104
Austral		**2.236**	**359**	**1**	**358**	**342**	**701**	**381**	**1**	**380**
	Grande	18	18	1	17	-	18	19	1	18
	Mediana	133	77	-	77	28	105	82	-	82
	Pequeña	934	177	-	177	207	384	209	-	209
	Micro	1.151	87	-	87	107	194	71	-	71

V.10. **Factores de expansión**

Los factores de expansión, en una muestra probabilística, permiten extrapolar el comportamiento observado en la muestra al total de las unidades que forman parte de la población objetivo. Este cálculo es, técnicamente, el inverso de la probabilidad de selección de cada unidad, es decir, se obtienen factores de expansión para las unidades incluidas forzosamente y las aleatorias.

Dado que las unidades de inclusión forzosa se auto-representan, se deben realizar ajustes en el factor de expansión de las 8.150 unidades que corresponden a empresas de inclusión aleatoria, que reflejen las incidencias sucedidas en el trabajo de campo que impidieron la recolección total de la muestra seleccionada. Además, para esta versión se incluye la calibración de los factores de expansión para las unidades de inclusión aleatoria, que permite reconstruir los totales poblacionales a partir de un marco auxiliar elaborado de acuerdo con la información proveniente de SUSESO, considerando el total de trabajadores contratados directamente por las empresas al 30 de noviembre de 2022 y con las exclusiones consideradas en el punto IV. Diseño estadístico.

V.10.3. Factor de expansión para las empresas del tramo de inclusión aleatoria

V.10.3.1. Factor de expansión teórico

El factor de expansión teórico corresponde al inverso de la probabilidad de selección de la empresa i perteneciente al estrato región r, tamaño de la empresa v y agrupación de actividad económica h, que se calcula mediante la siguiente fórmula:

$$P_{rvh_i} = \frac{n_{rvh}^{sm}}{N_{rvh}} \tag{9}$$

Donde:

P_{rvh_i} : Probabilidad de selección de la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

n_{rvh}^{sm} : Tamaño de muestra con sobremuestreo ajustado de la Encla 2023 en el estrato región r, tamaño de la empresa v y agrupación de actividad económica h.

N_{rvh} : Número de empresas en el marco estadístico de la Encla 2023, en el estrato región r, tamaño de la empresa v y agrupación de actividad económica h.

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

$$w_{rvh_i} = \frac{1}{P_{rvh_i}} \quad (10)$$

Donde:

w_{rvh_i} : Factor de expansión teórico de la empresa i en la muestra, perteneciente al estrato región r, tamaño de la empresa v y agrupación de actividad económica h.

P_{rvh_i} : Probabilidad de selección de la empresa i en la muestra, perteneciente al estrato región r, tamaño de la empresa v y agrupación de actividad económica h.

V.10.3.2. Ajuste por elegibilidad desconocida

Durante el trabajo de campo, las empresas se clasifican según su estado de elegibilidad, diferenciando entre unidades "elegibles", "no elegibles" o "elegibilidad desconocida". Tal como se observa en la tabla 16, un 72,1% (5.875) del total de empresas se encuentran clasificadas como elegible, mientras que 15,2% (1.240) es no elegible y 12,7% (1.035) corresponde a elegibilidad desconocida.

Del total de unidades elegibles, la región Metropolitana de Santiago presenta el mayor porcentaje de dichas empresas (8,2%), mientras que la región de Aysén del General Carlos Ibáñez del Campo presenta la menor proporción de empresas elegibles (2,7%). Por otra parte, del total observado de unidades no elegibles la región de Ñuble presenta el menor número de empresas (0,6%), mientras que la región Metropolitana de Santiago, presentan mayor proporción de empresas no elegibles (1,6%). En cuanto a las empresas con elegibilidad desconocida, la región donde se observa la mayor proporción de estas corresponde a Ñuble (2,2%).

Tabla 16. Distribución estado de elegibilidad de las empresas seleccionadas aleatoriamente en la Encla 2023 por región[9]

Región	Elegibles	No elegibles	Elegibilidad desconocida	Total
Arica y Parinacota	242 (2,9%)	58 (0,7%)	18 (0,2%)	318 (3,9%)
Tarapacá	298 (3,6%)	78 (0,9%)	35 (0,4%)	411 (5,0%)
Antofagasta	201 (2,4%)	58 (0,7%)	53 (0,6%)	312 (3,8%)
Atacama	204 (2,5%)	91 (1,1%)	28 (0,3%)	323 (3,9%)
Coquimbo	296 (3,6%)	90 (1,1%)	43 (0,5%)	429 (5,2%)
Valparaíso	461 (5,6%)	68 (0,8%)	41 (0,5%)	570 (6,9%)
Metropolitana de Santiago	671 (8,2%)	130 (1,6%)	28 (0,3%)	829 (10,1%)
Lib. Gral. Bernardo O'Higgins	413 (5,0%)	73 (0,9%)	138 (1,69%)	624 (7,6%)
Maule	370 (4,5%)	83 (1,0%)	110 (1,35%)	563 (6,9%)
Ñuble	364 (4,4%)	53 (0,6%)	186 (2,2%)	603 (7,4%)
Biobío	440 (5,4%)	68 (0,8%)	69 (0,8%)	577 (7,0%)
La Araucanía	387 (4,7%)	117 (1,4%)	71 (0,8%)	575 (7,0%)
Los Ríos	387 (4,7%)	66 (0,8%)	11 (0,1%)	464 (5,6%)
Los Lagos	358 (4,3%)	66 (0,8%)	127 (1,5%)	551 (6,7%)
Aysén del Gral. Carlos Ibáñez del Campo	227 (2,7%)	81 (0,9%)	55 (0,6%)	363 (4,4%)
Magallanes y la Antártica Chilena	256 (3,1%)	59 (0,7%)	22 (0,2%)	337 (4,1%)
Total	5875 (72,09%)	1240 (15,21%)	1035 (12,70%)	8150 (100%)

En la tabla 17 se observa la distribución del estado de elegibilidad según tamaño. El tamaño pequeño presenta el mayor porcentaje de empresas elegibles (31,4%), mientras que el menor porcentaje de empresas elegibles se presenta en las microempresas (11,8%). Asimismo, la mayor proporción de empresas con elegibilidad desconocida y de empresas no elegibles se encuentra en el tamaño pequeño (7,0% y 6,4%, respectivamente).

Tabla 17. Distribución estado de elegibilidad de las empresas seleccionadas aleatoriamente en la Encla 2023, según tamaño de empresas por número de trabajadores/as[10]

Tamaño	Elegibles	No elegibles	Elegibilidad desconocida	Total
Micro	962 (11,8%)	468 (5,7%)	209 (2,5%)	1.639 (20,1%)
Pequeña	2.559 (31,4%)	522 (6,4%)	576 (7,0%)	3.657 (44,8%)
Mediana	1.390 (17,0%)	166 (2,0%)	205 (2,5%)	1.761 (21,6%)
Grande	964 (11,8%)	84 (1,0%)	45 (0,5%)	1.093 (13,4%)
Total	5.875 (72,0%)	1.240 (15,2%)	1.035 (12,7%)	8.150 (100%)

En la tabla 18 se muestra el estado de elegibilidad de las empresas según la agrupación de actividad económica a la que pertenecen. Se observa que, en la actividad de Construcción existe un mayor porcentaje de empresas elegibles (7,9%)

[9] Porcentajes estimados sobre el total de unidades seleccionadas aleatoriamente (8150).
[10] Ídem.

con relación con el resto de las actividades económicas. Así mismo, dicha actividad presenta la mayor cantidad de las empresas no elegibles (2,7%), y Agricultura, ganadería, silvicultura y pesca evidencia la proporción más alta de empresas con elegibilidad desconocida (2,3 %).

Tabla 18. Distribución estado de elegibilidad de las empresas seleccionadas aleatoriamente en la Encla 2023, según agrupación de actividad económica[11]

Actividad económica	Elegibles	No elegibles	Elegibilidad desconocida	Total
A. Agricultura, ganadería, silvicultura y pesca	621 (7,6%)	165 (2,0%)	191 (2,3%)	977 (11,9%)
B. Explotación de minas y canteras	220 (2,7%)	82 (1,0%)	42 (0,5%)	344 (4,2%)
C. Industrias manufactureras	563 (6,9%)	88 (1,0%)	81 (0,9%)	732 (8,9%)
D-E. Suministro de electricidad, gas, vapor, y aire acondicionado; suministro de agua; evacuación de aguas residuales, gestión de desechos y descontaminación	293 (3,6%)	57 (0,7%)	39 (0,4%)	389 (4,7%)
F. Construcción	645 (7,9%)	226 (2,7%)	118 (1,4%)	989 (12,1%)
G. Comercio al por mayor y al por menor; reparación de vehículos automotores y motocicletas	591 (7,2%)	84 (1,0%)	115 (1,4%)	790 (9,6%)
H-J. Transporte y almacenamiento; información y comunicaciones	521 (6,3%)	85 (1,0%)	75 (0,9%)	681 (8,3%)
I. Actividades de alojamiento y servicio de comidas	418 (5,1%)	80 (0,9%)	51 (0,6%)	549 (6,74%)
K-L. Actividades financieras y de seguros; actividades inmobiliarias	272 (3,3%)	80 (0,9%)	42 (0,5%)	394 (4,8%)
M-N. Actividades profesionales, científicas y técnicas; actividades de servicios administrativos y de apoyo	563 (6,9%)	135 (1,6%)	93 (1,1%)	791 (9,7%)
P. Enseñanza	546 (6,7%)	38 (0,4%)	60 (0,7%)	644 (7,9%)
Q. Actividades de atención de la salud humana y de asistencia social	362 (4,44%)	66 (0,8%)	49 (0,6%)	477 (5,8%)
R-S. Actividades artísticas, de entretenimiento y recreativas; otras actividades de servicios	260 (3,1%)	54 (0,6%)	79 (0,9%)	393 (4,8%)
Total	5.875 (72,0%)	1.240 (15,2%)	1.035 (12,7%)	8.150 (100%)

Para realizar el ajuste por elegibilidad desconocida, dentro de cada estrato se distribuye el peso de las empresas con dicho estado entre las empresas con elegibilidad conocida (elegibles y no elegibles), para lo cual es necesario calcular una razón de ajuste mediante la siguiente fórmula:

[11] Ídem

$$\hat{R}_{rvh}^{ed} = \frac{\sum_{i \in \Omega_{rvh}} w_{rvh_i}}{\sum_{i \in \Omega_{rvh}^{ec}} w_{rvh_i}} \quad (11)$$

Donde:

$\hat{R}_{rvh}^{ed}$: Razón de ajuste por elegibilidad desconocida, para las empresas pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

w_{rvh_i} : Factor de expansión teórico de la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

Ω_{rvh} : Conjunto de empresas en la muestra pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

Ω_{rvh}^{ec} : Conjunto de empresas en la muestra pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h (elegibles y no elegibles).

A partir de esta etapa, las empresas clasificadas en condición de elegibilidad desconocida no son consideradas en el factor de expansión y se les asigna una ponderación nula, ya que, su peso se distribuye en aquellas unidades cuya elegibilidad es conocida (elegibles y no elegibles). Para estas últimas, el factor de expansión ajustado por elegibilidad desconocida corresponde al producto entre la razón de ajuste y el factor de expansión teórico. Por lo tanto, el factor de expansión corregido por elegibilidad desconocida se obtiene mediante la siguiente fórmula:

$$w_{rvh_i}^{ed} = \begin{cases} \hat{R}_{rvh}^{ed} \cdot w_{rvh_i} & , si \quad i \in \Omega_{rvh}^{ec} \\ . & , si \quad i \in \Omega_{rvh}^{ed} \end{cases} \quad (12)$$

Donde:

$w_{rvh_i}^{ed}$: Factor de expansión ajustado por elegibilidad desconocida de la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$\hat{R}_{rvh}^{ed}$: Razón de ajuste por elegibilidad desconocida, para las empresas pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

w_{rvh_i} : Factor de expansión teórico de la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

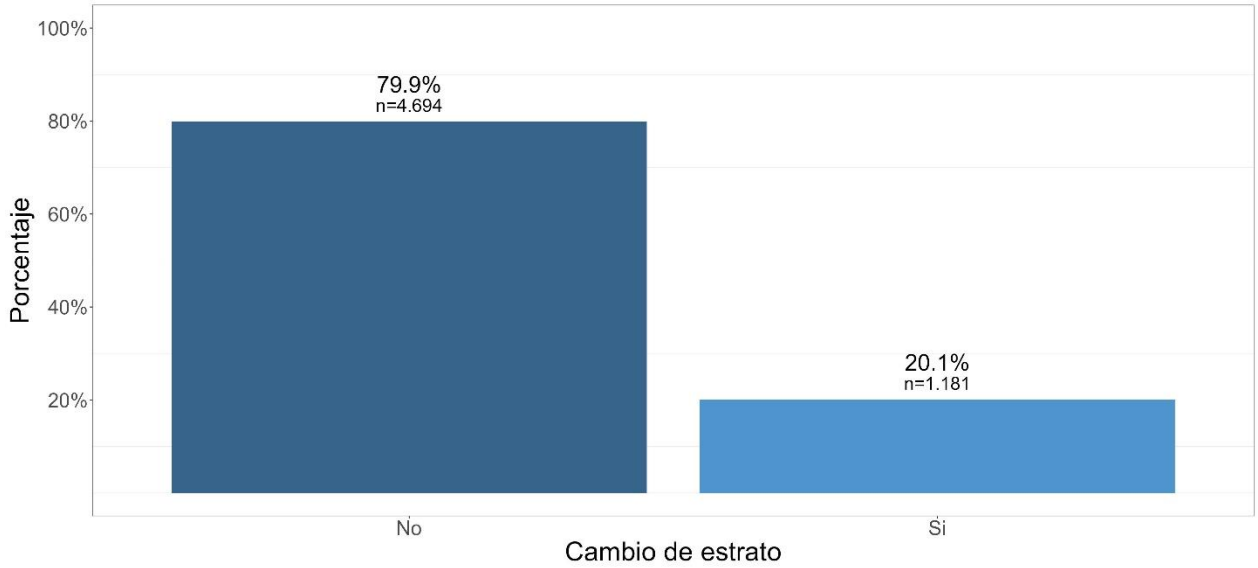
Ω_{rvh}^{ed} : Conjunto de empresas en la muestra pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h (elegibilidad desconocida).

Ω_{rvh}^{ec} : Conjunto de empresas en la muestra pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h (elegibles y no elegibles).

V.10.3.3. Ajuste por cambio de estrato y no elegibilidad

Frecuentemente, en las encuestas a empresas se observa que la información que contienen las variables de estratificación informada en el marco estadístico difiere del dato recolectado en el trabajo de campo, de modo que la empresa no pertenece al estrato en el cual fue seleccionada inicialmente. Para la Encla 2023, se observan cambios de estrato, los cuales pueden darse porque la empresa reportó pertenecer a otra región, ejercer otra actividad económica o tamaño. Siendo así, en la Figura 3, se observa que, a nivel nacional 20,1% de las empresas elegibles cambiaron su estrato inicial.

Figura 3. Proporción de empresas elegibles con cambio de estrato, a nivel nacional



Al analizar los datos de la encuesta se identifican unidades categorizadas como "No elegibles", lo cual se debe, principalmente, al desfase existente entre la información contenida en el marco estadístico y el trabajo de campo, lo cual se corrige a través del cálculo de los factores de expansión realizando un ajuste por no elegibilidad y cambio de estrato[12]. Por lo tanto, para realizar este ajuste se identifican tres casos:

1. La empresa i es elegible, fue seleccionada en el estrato l, y según la encuesta pertenece al estrato l'.

2. La empresa i es elegible y sigue perteneciendo al estrato en donde fue seleccionada.

3. La empresa i es no elegible.

Este ajuste se expresa en la siguiente fórmula:

$$w^{*}_{rvh_i} = \begin{cases} w^{ed}_{rvh_i} & , si \quad i \in \Omega^{el}_{rvh} \wedge i \in \Omega^{l \to l'}_{rvh} \\ \dfrac{N_{rvh}\left(1 - \dfrac{n^{ne}_{rvh}}{n^{ec}_{rvh}}\right) - \sum_{l \neq l'} \sum_{i=1}^{n^{el,l \to l'}_{rvh}} w^{ed}_{rvh_i}}{n^{el,l \to l}_{rvh}} & , si \quad i \in \Omega^{el}_{rvh} \wedge i \in \Omega^{l \to l}_{rvh} \\ . & , si \quad i \in \Omega^{ne}_{rvh} \end{cases} \tag{13}$$

[12] Para la Encla 2023, se adopta el ajuste realizado por el Instituto Nacional de Estadísticas de España en el diseño muestral de las encuestas dirigidas a empresas. Para más información visitar: https://www.ine.es/daco/daco43/metoite2016.pdf

Donde:

$w^{*}_{rvh_i}$: Factor de expansión ajustado por cambio de estrato y no elegibilidad para la empresa i en la muestra, perteneciente a la región r, tamaño de empresa v y agrupación de la actividad económica h.

$w^{ed}_{rvh_i}$: Factor de expansión ajustado por elegibilidad desconocida de la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

N_{rvh} : Número de empresas en el marco estadístico, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

n^{ne}_{rvh} : Número de empresas en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h , categorizadas como no elegibles.

n^{ec}_{rvh} : Número de empresas en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h catalogadas con elegibilidad conocida (elegibles y no elegibles).

$n^{el,l\to l}_{rvh}$: Número de empresas en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h , catalogadas como elegibles y que han sido seleccionadas en el estrato l y después del trabajo de campo se mantienen en el mismo estrato.

Ω^{el}_{rvh} : Conjunto de empresas en la muestra perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h , catalogadas como elegibles.

$\Omega^{l\to l}_{rvh}$: Conjunto de empresas en la muestra que fueron seleccionadas en el estrato l y según la encuesta siguen perteneciendo al estrato l.

$n^{el,l\to l'}_{rvh}$: Número de empresas en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h , catalogadas como elegibles y que han sido seleccionadas en el estrato l y después del trabajo de campo pertenecen a otro estrato l'.

$\Omega^{l\to l'}_{rvh}$: Conjunto de empresas en la muestra que fueron seleccionadas en el estrato l y según la encuesta pertenecen al estrato l' (cambió de estrato).

Ω^{ne}_{rvh} : Conjunto de empresas en la muestra perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h, catalogadas como no elegibles.

. : Indica valor perdido o *missing*.

Es importante resaltar que, al realizar el ajuste por no elegibilidad, se redujo el stock poblacional al que se expande antes del ajuste, debido a que se descuenta el peso de las 1.240 empresas no elegibles. Por lo tanto, después de aplicar este procedimiento, la población de empresas es de 98.054, permitiendo corregir los efectos que pudieran causar estas empresas sobre la estimación.

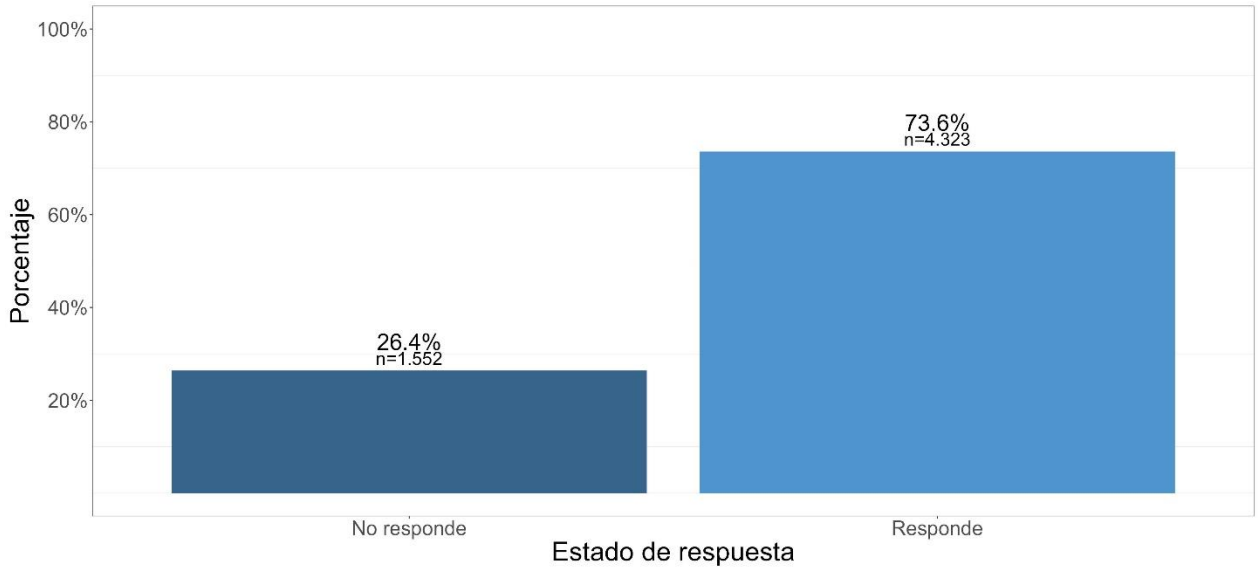
Por otra parte, en cuanto a los cambios de estrato, existe una compensación de empresas dentro de cada uno de ellos, dado que, mientras unas empresas se van del estrato, otras llegan al mismo.

V.10.3.4. Ajuste por no respuesta

Existen algunas unidades muestrales 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.

Por lo tanto, en la Figura 4, se observa que a nivel nacional 73,6% de las empresas elegibles respondieron la encuesta y 26,4% no lo realizó.

Figura 4. Proporción de empresas elegibles que no respondió la Encla 2023, a nivel nacional



Por lo tanto, una vez aplicadas las estrategias operativas necesarias para conseguir la información sin mayor éxito, se aplican 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 sobre la suma de los factores de expansión de las unidades que respondieron la encuesta, tal como se observa en la siguiente fórmula:

$$\hat{R}^{nr}_{rvh} = \frac{\sum_{i \in \Omega^{el}_{rvh}} w^{*}_{rvh_i}}{\sum_{i \in \Omega^{el,R}_{rvh}} w^{*}_{rvh_i}} \quad (14)$$

Donde:

$\hat{R}^{nr}_{rvh}$: Razón de ajuste por no respuesta para las empresas pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$w^{*}_{rvh_i}$: Factor de expansión ajustado por cambio de agrupación de actividad económica o región y no elegibilidad para la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

Ω^{el}_{rvh} : Conjunto de empresas en la muestra pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h, catalogadas como elegibles.

$\Omega_{rvh}^{el,R}$: Conjunto de empresas en la muestra pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h, catalogadas 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*.

$$w_{rvh_i}^{nr} = \begin{cases} \hat{R}_{rvh}^{nr} \cdot w_{rvh_i}^{*} & , si \quad i \in \Omega_{rvh}^{el,R} \\ . & , si \quad i \in \Omega_{rvh}^{el,NR} \end{cases} \tag{15}$$

Donde:

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

$\hat{R}_{rvh}^{nr}$: Razón de ajuste por no respuesta para las empresas pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$w_{rvh_i}^{*}$: Factor de expansión ajustado por cambio de agrupación de actividad económica o región y no elegibilidad para la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$\Omega_{rvh}^{el,R}$: Conjunto de empresas en la muestra pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h, catalogadas como elegibles y que respondieron la encuesta.

$\Omega_{rvh}^{el,NR}$: Conjunto de empresas en la muestra pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h, catalogadas como elegibles y que no respondieron la encuesta.

Al realizar el ajuste por no respuesta, se redujo el marco estadístico a 96.935 empresas, debido a que se encontraron estratos sin ninguna empresa lograda, lo que trae como consecuencia no poder distribuir los pesos muestrales.

V.10.3.5. Suavizamiento del factor de expansión

Después de realizar los ajustes descritos anteriormente, se observa la presencia de valores atípicos en los factores de expansión, de modo que surge la necesidad de realizar un suavizamiento[13].

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 $w_{rvh_i}^{nr}$

[13] El método para el suavizamiento del factor de expansión corresponde al mismo aplicado en ENCLA 2019.

se considera atípico si está a una distancia superior a 2 desviaciones estándares respecto al promedio de su estrato. Si denotamos al factor atípico como A_{rvh_i}, entonces A_{rvh_i} no pertenece al intervalo I_{rvh}, tal como se expresa en la siguiente fórmula:

$$A_{rvh_i} \notin \mathrm{I}_{rvh} = \left[\overline{w}_{rvh}^{nr} - 2 \cdot \sigma_{w_{rvh}}^{nr} \ , \overline{w}_{rvh}^{nr} + 2 \cdot \sigma_{w_{rvh}}^{nr}\right] \tag{16}$$

Donde:

A_{rvh_i} : Factor de expansión de la empresa i en la muestra, considerado atípico, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$\overline{w}_{rvh}^{nr}$: Promedio de los factores de expansión ajustados por no respuesta del conjunto de empresas pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$\sigma_{w_{rvh}}^{nr}$: Desviación estándar de los factores de expansión del conjunto de empresas, pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

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[14]:

$$w_{rvh_i}^{s} = \begin{cases} \dfrac{\left(w_{rvh_i}^{nr} + \overline{w}_{rvh}^{nr}\right)}{2} & , si \quad i \in \left\{\Omega_{rvh}^{el,R} \supset A_{rvh_i}\right\} \\ w_{rvh_i}^{nr} & , si \quad i \in \left\{\Omega_{rvh}^{el,R} \not\supset A_{rvh_i}\right\} \end{cases} \tag{17}$$

Donde:

$w_{rvh_i}^{s}$: Factor de expansión suavizado de la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$w_{rvh_i}^{nr}$: Factor de expansión ajustado por no respuesta de la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$\overline{w}_{rvh}^{nr}$: Promedio de los factores de expansión ajustados por no respuesta del conjunto de empresas de la región r, tamaño de la empresa v y agrupación de actividad económica h.

A_{rvh_i} : Factor de expansión de la empresa i en la muestra, considerado atípico, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$\Omega_{rvh}^{el,R}$: Conjunto de empresas en la muestra que son categorizadas como elegibles y respondieron la encuesta, pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

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 $w_{rvh_{i=1}}^{nr} > w_{rvh_{i=2}}^{nr}$, una vez realizado el suavizamiento $w_{rvh_{i=1}}^{s} > w_{rvh_{i=2}}^{s}$.

[14] 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.

Para esta versión de la encuesta, 4,5% de las unidades contaban con un factor de expansión ajustado por no respuesta fuera del umbral I_{rvh}. Por lo tanto, se realizó el suavizamiento a los estratos donde se encontraban estos valores.

V.10.3.6. 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 seleccionadas respondieron 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 mediante un método de calibración por *raking*[15].

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:

$$G(w^c, w^s) = w^c \log(w^c/w^s) - w^c + w^s \tag{18}$$

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.

[15] Para más información, revisar: Deville & Särndal. (1992). *Calibration Estimators in Survey Sampling*. *Journal of the American Statistical Association.* 87(418), 376-382.

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

$$w^c = w^s \cdot F(w^c, w^s) \tag{19}$$

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 para cada variable: región, actividad económica y tamaño de empresa, según estimación del marco estadístico reconstruido con periodo de referencia noviembre 2022 a partir del Registro Administrativo de SUSESO.
- Total, de número de trabajadores para cada variable: región, actividad económica, tamaño de empresa), según estimación del marco estadístico reconstruido con periodo de referencia noviembre 2022 a partir del Registro Administrativo de SUSESO.

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

V.10.4. Factor de expansión para las empresas del tramo de inclusión forzosa

Las unidades que pertenecen al tramo de inclusión forzosa poseen factores de expansión iguales a 1, ya que, por su relevancia para el sector en estudio, se considera que su comportamiento es único, por lo que no pueden representar ni ser representadas por otras unidades.

V.11. Estimación de parámetros

La estimación estadística corresponde al procedimiento mediante el cual se calculan parámetros de la población a partir de datos muestrales obtenidos. En el contexto de la Encla 2023 se realiza la estimación del total del número de trabajadores, variable que se extrae del formulario autoaplicado, específicamente de la pregunta A3: Indique el total de personas contratadas directamente por esta empresa al 30 de noviembre de 2022, según región del país y sexo. Esta estimación se calcula para cada nivel de dominio del estudio. Para exponer la fórmula de cálculo, esta se expresará en el nivel más desagregado de variables.

La estimación se obtiene mediante la siguiente fórmula:

$$\hat{t}_{rvh} = \sum_{i=1}^{n_{rvh}} y_{rvh_i}^{A1} \cdot w_{rvh_i}^{c} \quad (20)$$

Donde:

$\hat{t}_{rvh_i}$: Estimación del total de trabajadores contratados por las empresas directamente al 30 de noviembre del 2022, pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$y_{rvh_i}^{A1}$ Número de trabajadores contratados directamente por la empresa i al 30 de noviembre del 2022, pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$w_{rvh_i}^{c}$: Factor de expansión calibrado de la empresa i en la muestra, perteneciente a la región r, tamaño de la empresa v y agrupación de actividad económica h.

Por otro lado, para obtener las medidas de precisión, se debe estimar la varianza a nivel de estrato, tal como se especifica en la siguiente fórmula:

$$\hat{v}(\hat{t}_{rvh}) = \sum_{i=1}^{n_{rvh}} \hat{N}_{rvh} \cdot (\hat{N}_{rvh} - n_{rvh}) \cdot \frac{S^2(y)_{rvh}}{n_{rvh}} \quad (21)$$

Donde:

$\hat{v}(\hat{t}_{rvh})$: Varianza de la estimación del total de trabajadores contratados por las empresas directamente al 30 de noviembre del 2022, pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$\hat{N}_{rvh}$: Estimación del número de empresas en la población de la región r, tamaño de empresa v y agrupación de actividad económica h.

n_{rvh} : Número de empresas en la muestra lograda, pertenecientes a la región r, tamaño de empresa v y agrupación de actividad económica h.

$S^2(y)_{rvh}$: Cuasi-varianza del número de trabajadores de la región r, tamaño de empresa v y agrupación de actividad económica h en la muestra.

Por otro lado, el error relativo de la estimación del total de trabajadores contratados directamente por las empresas al 30 de noviembre de 2022 en la Encla 2023, se obtiene a partir de la siguiente fórmula:

$$e\left(\hat{t}_{rvh}^{A1}\right) = \frac{\sqrt{v\left(\hat{t}_{rvh}^{A1}\right)}}{\hat{t}_{rvh}^{A1}} \cdot t_{\mathrm{gl},1-\frac{\alpha}{2}} \quad (22)$$

Donde:

$e\left(\hat{t}_{rvh}^{A1}\right)$: Error relativo de la estimación del total de trabajadores contratados por las empresas directamente al 30 de noviembre del 2022, pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$\hat{t}_{rvh}^{A1}$: Estimación del total de trabajadores contratados directamente por las empresas al 30 de noviembre del 2022, pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$v\left(\hat{t}_{rvh}^{A1}\right)$[16] : Varianza de la estimación del total de trabajadores contratados directamente por las empresas al 30 de noviembre del 2022, pertenecientes a la región r, tamaño de la empresa v y agrupación de actividad económica h.

$t_{gl,1-\frac{\alpha}{2}}$: Cuantil de la distribución T-Student correspondiente a gl grados de libertad y un nivel de confianza $(1 - \alpha)$.

Una vez construidos los factores de expansión calibrados, se han estimado los parámetros y medidas de error según lo detallado en este apartado. Los resultados se presentan en la sección III. Acciones referidas a precisión y confiabilidad del Informe de Calidad, para la estimación del total de trabajadores(as) contratados(as) directamente por las empresas al 30 de noviembre del 2022, según agrupación de actividad económica, tamaño, región, y agrupación regional y tamaño. Los resultados indican un total de 4.923.727 trabajadores(as), con un error relativo del 3,1%.

[16] Los cálculos realizados para la varianza de la estimación del total de trabajadores contratados directamente por las empresas al 30 de noviembre del 2022 fueron realizados en el software R, con el paquete survey en su versión 4.4-2.

VI. Metodología de recolección de datos

VI.1. Características generales de la recolección de los datos

Dado del diseño de la Encla 2023, que contempla la aplicación de tres cuestionarios (autoaplicado, empleadores y sindicatos o trabajadores según cada caso) por unidad, se emplearon los siguientes modos de recolección:

- Entrevista autoadministrada asistida por *web* (CAWI[17] por sus siglas en inglés) para el cuestionario autoaplicado.
- Entrevista telefónica asistida por dispositivo móvil de captura (DMC) como *tablet* o *smartphone* (CATI[18] por sus siglas en inglés) para el cuestionario a empleadores/as.
- Entrevista presencial asistida por dispositivo móvil de captura (DMC) como *tablet* o *smartphone* (CAPI[19] por sus siglas en inglés) para el cuestionario a trabajadores/as.

Se definió en términos operativos que la gestión de las unidades debe iniciar a través del cuestionario a empleadores/as, lo cual permitió recabar información de contacto relevante para la gestión del cuestionario a sindicatos o trabajadores/as. Además, ello permite asignar códigos de no elegibilidad, a modo de evitar aplicar los cuestionarios a trabajadores/as o dirigentes sindicales en unidades no elegibles.

Sin embargo, en caso de que las condiciones lo ameritasen, es posible iniciar la gestión de la unidad aplicando alguno de estos dos cuestionarios. Por otra parte, en esta versión de la Encla la recolección se realizará principalmente a través de DMC o *tablets*, por medio de la aplicación de *Survey Solutions* (SuSo). Por ello, la recolección mediante una entrevista personal aplicada en papel (PAPI[20], por sus siglas en inglés) sólo estuvo considerada para situaciones donde el DMC no funcionara correctamente, por razones técnicas u otras. Los excepcionales casos de recolección de datos en papel fueron digitados en las *tablets* apenas fue posible, para luego sincronizarlos y permitir una descarga oportuna por parte del equipo técnico.

Por último, se definió que la aplicación del cuestionario a empleadores/as se realizaría por vía telefónica, mientras que los cuestionarios a sindicatos y trabajadores/as se recolectarán presencialmente. No obstante, se permitió la aplicación presencial del cuestionario a empleadores/as, así como la aplicación por medio de videollamada de los cuestionarios a sindicatos y trabajadores/as, en caso de que las condiciones así lo requieran. Por su lado, el cuestionario autoaplicado fue completado autónomamente por las personas informantes a través de un formulario web desarrollado por el INE. Estos contaron con un plazo de 15 días para enviar el formulario completado.

El periodo de recolección de la Encuesta Laboral 2023 se planificó para que tuviese lugar entre el 10 de julio de 2023 y el 31 de diciembre de 2023 (cinco meses y medio, 117 días laborales). Respecto al periodo de referencia de los datos, esta es variable según la pregunta. En la mayoría de los casos, el periodo de referencia es al 30 de noviembre de 2022, sin embargo, hay preguntas relativas a situaciones ocurridas antes de la pandemia del COVID-19, por lo que dicho periodo

[17] *Computer assisted Web Interviewing.*
[18] *Computer Assisted Telefonic Interviewing.*
[19] *Computer Assisted Personal Interviewing.*
[20] *Paper Assisted Personal Interviewing.*

de referencia llega a 2019[21]. Por último, en algunas preguntas se hace referencia a lo que la persona informante percibe "actualmente", por lo que este periodo puede variar según la fecha de logro del cuestionario. Si bien esta operación estadística se ha diseñado para realizarse de forma bienal o trienal desde el año 1998, la crisis sanitaria provocada por la pandemia del COVID-19 significó el aplazamiento de la décima versión de la encuesta hasta 2023, lo cual significa una diferencia de cuatro años entre la ejecución de las versiones 2019 y 2023 de la Encuesta.

En lo que concierne a la asignación de la carga de trabajo a las personas recolectoras, los equipos de coordinación regional y central definieron estrategias que permitiesen distribuir la muestra, vehículos y viáticos en consideración de los recursos disponibles para cada región. De este modo, la asignación de la carga se realizó dinámicamente, considerando los resultados obtenidos por cada persona recolectora y las necesidades del proyecto respecto del logro de los estratos incompletos. Con el fin de no sobrecargar los DMC, se definió asignar un máximo de 30 empresas por dispositivo, lo cual significaría un total de 60 cuestionarios disponibles al mismo tiempo, considerando que en cada unidad se deben recolectar dos cuestionarios.

Con el objetivo de controlar el avance y generar cargas de trabajo adecuadas para cada equipo recolector, la Encla 2023 dividió la muestra total en cuatro submuestras operativas, dividiendo la muestra total por cantidades y periodos específicos, como se indica en la siguiente tabla:

Tabla 19. Submuestras operativas a nivel regional

Región	Muestral total	Submuestra 1	Submuestra 2	Submuestra 3	Submuestra 4
Arica y Parinacota	318	80	80	80	80
Tarapacá	411	103	103	103	103
Antofagasta	413	103	103	103	103
Atacama	424	106	106	106	106
Coquimbo	529	132	132	132	132
Valparaíso	577	144	144	144	144
Metropolitana	1003	251	251	251	251
Lib. Gral. Bernardo O'Higgins	631	158	158	158	158
Maule	565	141	141	141	141
Ñuble	604	151	151	151	151
Biobío	582	146	146	146	146
La Araucanía	578	145	145	145	145
Los Ríos	464	116	116	116	116
Los Lagos	552	138	138	138	138
Aysén del Gral. Carlos Ibáñez del Campo	363	91	91	91	91
Magallanes y la Antártica Chilena	338	85	85	85	85

Los periodos en los cuales se dio inicio y término a cada submuestra operativa se detallan en la siguiente tabla:

[21] Por ejemplo, la pregunta O17 del cuestionario auto-aplicado: *En los siguientes periodos, ¿con qué frecuencia se implementó el teletrabajo para la mayoría de quienes teletrabajaron para la empresa?: Antes de la pandemia (2019 hasta febrero de 2020)*

Tabla 20. Calendario de organización de las submuestras operativas

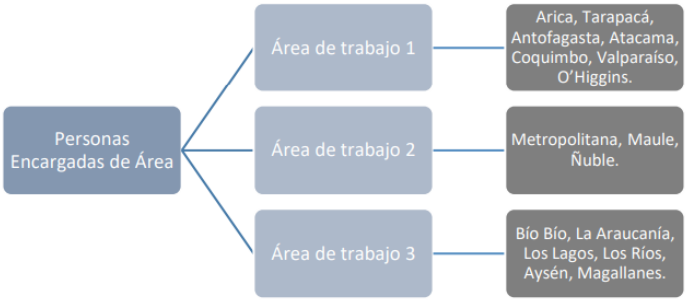
Submuestra operativa	Fecha de inicio	Fecha de término
1	10-07-2023	17-08-2023
2	18-08-2023	14-09-2023
3	15-09-2023	17-10-2023
4	18-10-2023	17-11-2023

VI.1.1. Equipo de recolección de datos

La organización del equipo de recolección de datos de la Décima Encla se generó buscando gestionar la totalidad de unidades en los plazos estimados, y dar cumplimiento a los objetivos muestrales propuestos con los más altos estándares de calidad. Al ser una encuesta estructural, el equipo de recolección se conforma para cada recolección.

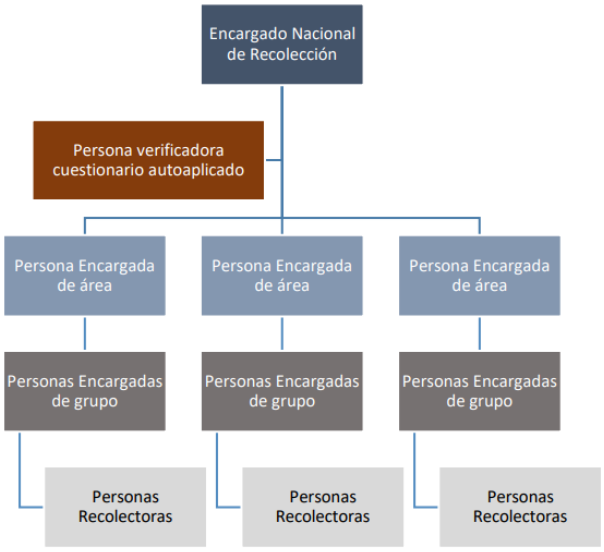
El equipo se encontraba liderado por un Encargado Nacional de Recolección, quien contaba con tres Personas Encargadas de Área de Recolección responsables de gestionar el operativo y asesorar metodológicamente a los equipos de cada región. La figura 5 ilustra la distribución de las áreas de trabajo según macrozonas regionales. Estas cuatro personas formaron el equipo de coordinación de recolección de datos del estudio. A la vez, se incorporó a una verificadora especializada en el cuestionario autoaplicado, para apoyar en el llenado y recuperación de información asociada a este instrumento, así como de traspasar al sistema de gestión operativa de encuestas económicas (SGOE) aquellos cuestionarios enviados en formato Excel o PDF por las empresas.

Figura 5. Distribución de Áreas por Macrozonas Regionales



Por su parte, los equipos de recolección de datos de cada región estuvieron conformados por las personas encargadas de grupo y su equipo de personas recolectoras. Cada equipo regional de recolección fue liderado por las personas encargadas de área de recolección en colaboración con la jefatura de operaciones y/o de proyectos especiales de cada región, exceptuando a la región Metropolitana, donde la coordinación regional fue ejercida por el encargado de área y el encargado nacional de recolección. La figura 6 refleja el organigrama del equipo de recolección en esta versión del estudio.

Figura 6. Organigrama del Equipo de Recolección Encla 2023



La tabla 21 presenta la dotación final del equipo de recolección

Tabla 21. Dotación final del equipo de recolección

Región	Recolector(a) de datos	Encargado(a) de grupo	Verificador(a) Autoaplicado	Encargado(a) Nacional y de Área	Equipo de Recolección
Arica y Parinacota	3	0			3
Tarapacá	3	1			4
Antofagasta	3	1			4
Atacama	3	1			4
Coquimbo	4	1			5
Valparaíso	4	1			5
Metropolitana	8	2	3	4	17
Lib. Gral. Bernardo O'Higgins	3	1			4
Maule	3	1			4
Ñuble	2	1			3
Biobío	4	1			5
La Araucanía	4	1			5
Los Ríos	3	1			4
Los Lagos	4	1			5
Aysén del Gral. Carlos Ibáñez del Campo	2	1			3
Magallanes y la Antártica Chilena	2	1			3
Total	55	16	3	4	78

El área de supervisión de la encuesta se conformó por los perfiles de cargos que se muestran a continuación, cada uno con su rol asignado.

- Persona coordinadora nacional de supervisión:
  - Monitorear los riesgos de incumplimiento para generar ajustes en la planificación y organización de las actividades realizadas en la etapa de supervisión.
  - Sistematizar información acerca de los avances de la supervisión a nivel nacional, informando periódicamente al jefe/a del Subdepartamento de Supervisión y Verificación de Estadísticas Sociales.
  - Generar informes de resultado de la supervisión que permitan reportar a las personas responsables de los procesos, los hallazgos o errores que afecten la calidad de los datos.

- Persona encargada de área de supervisión:
  - Planificar las actividades del equipo de supervisión, informando a la persona supervisora operativa la programación de actividades y tareas a ejecutar durante la recolección de los datos y a la persona coordinadora nacional de supervisión las desviaciones y obstáculos respecto del cumplimiento de lo programado.
  - Monitorear los indicadores de rendimiento y calidad de los datos del proyecto, fomentando las acciones necesarias que garanticen el cumplimiento de las metas establecidas.
  - Colaborar con el proceso de construcción de componentes asociados a las revisiones de flujos operativos, pruebas y puesta en marcha del sistema informático de acuerdo con la metodología contemplada para el proyecto.

- Persona supervisora operativa:
  - Monitorear el correcto cumplimiento de los procesos de recolección, aplicando formularios de supervisión y revisando la calidad de las encuestas realizadas por las personas recolectoras, siguiendo los protocolos y procedimientos definidos en el marco del proyecto.
  - Revisar completitud, coherencia y correspondencia de las encuestas autoaplicadas recepcionadas por el equipo de recolección.
  - Detectar e informar a la coordinación de supervisión del proyecto las debilidades del proceso de recolección, asegurando que todas las actividades se realicen en conformidad con los procedimientos establecidos por el proyecto en forma oportuna.
  - Apoyar al equipo de recolección para incrementar el número de Cuestionarios Autoaplicados mediante contactos telefónicos y/o correos electrónicos a las empresas.

VI.1.2. Procesos de verificación de la muestra

Se realizaron dos procesos de verificación de las empresas incluidas en la muestra de la Encla 2023. La primera de ellas fue realizada por la unidad de verificación telefónica del Subdepartamento de Supervisión y Verificación de Estadísticas Sociales del INE. Este proceso fue coordinado por el Subdepartamento de Investigación Estadística del Departamento de Metodologías e Innovación Estadística, y por el Subdepartamento de Recolección de Estadísticas Sociales de la Subdirección de Operaciones (SDO). El segundo proceso de verificación fue coordinado entre el Subdepartamento de Recolección de Estadísticas Sociales (en SDO) y el Subdepartamento de Estadísticas Estructurales del Trabajo (en SDT). En ambas instancias se realizó la verificación de la muestra total de 8.352 empresas de la Encla 2023.

Los resultados generales de la muestra total a recolectar durante la primera ronda de verificación ejecutada por la unidad de Verificación indican que se logró hacer el contacto efectivo con 1.709 empresas, es decir, en estas empresas existe una persona informante que accede a proporcionar los datos de las variables solicitadas para el producto en cuestión.

Los casos que resultaron diferentes de "Activo" y "Activo con nuevo RUT" (6.643), son candidatos a que sean analizados para un posible tratamiento de verificación en terreno, ya que dichas empresas tienen una ubicación geográfica proveniente desde otros insumos internos del INE. Se debe prestar especial atención para las empresas que se clasifican en los estados "Sin movimiento" y "Término de giro total" (21), ya que corresponden a casos donde un o una informante validó la condición tributaria del RUT, pero se recomienda solicitar mayores antecedentes para determinar su situación final.

Tabla 22. Resultados del primer proceso de verificación muestral Encla 2023

Tipo de Contacto	Cantidad de empresas	Estado de Gestión	Cantidad de empresas
POSITIVO	1.730	Activo	1.705
		Activo con Nuevo RUT	4
		Sin Movimiento	18
		Término de Giro Total	3
		Insolvencia	0
		Paralizada	0
POSITIVO INCOMPLETO	747	Informante No se encuentra	418
		Solicita Gestión Correo	280
		Se Niega	49
NEGATIVO	5.875	Fonos Sin Respuesta	5.051
		Encontrado sin Fonos	37
		Inubicable	787
		Total general	**8.352**

El segundo proceso de verificación, realizado entre los meses de marzo y junio de 2023 contempló la especificidad operativa de la Encla, en el sentido que el cuestionario de verificación generado en esta instancia buscó no sólo obtener datos de contacto para la empresa, sino que también datos de contacto de una/un informante idóneo para contestar el cuestionario a empleadores/as, otro para el cuestionario a trabajadores/as y otro para el cuestionario a dirigentes sindicales. Asimismo, este proceso consolidó la información realizada también por el proceso de verificación 1 para generar el directorio final de unidades con información de contacto actualizada.

Poco menos de la mitad de las empresas consideradas en la muestra lograron ser verificadas de manera efectiva. Concretamente, alcanzan un 42% del total de empresas. Por otro lado, poco más de la mitad considerando "No contacto" (30,6%) y "No ubicable" (23,26%).

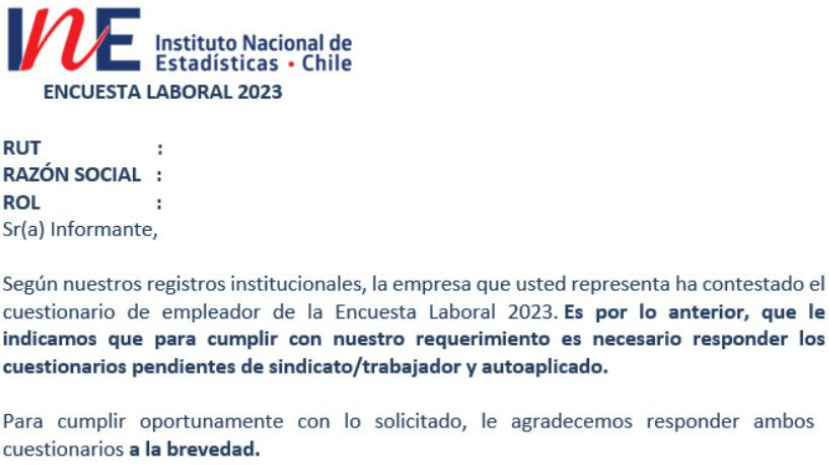
Tabla 23. Resultados del segundo proceso de verificación muestral Encla 2023

Resultado final de la verificación	Cantidad absoluta	Porcentaje (%)
Empresa Verificada (con datos actualizados)	3.494	41,83
Empresa Verificada (sin datos actualizados)	21	0,25
No contacto	2.556	30,6
Rechazo de información	31	0,37
No exigible	247	2,96
No ubicable	1.943	23,26
Sin resultado	60	0,72
Total	**8.352**	**100**

V.1.3. Estrategias de sensibilización de las personas informantes

Como parte de la estrategia de gestión del cuestionario autoaplicado, se realizó el envío constante de correos masivos a las empresas por parte de las personas encargadas de área, con temáticas que van desde las instrucciones asociadas al primer ingreso al cuestionario, a recordatorios de su envío, así como recordatorios para contestar los cuestionarios a empleadores, sindicatos o trabajadores. La siguiente figura muestra un ejemplo de correo masivo enviado a los informantes con los cuestionarios autoaplicados y a sindicatos o trabajadores pendientes.

Figura 7. Ejemplo de correo masivo enviado a los informantes



Instituto Nacional de Estadísticas · Chile

ENCUESTA LABORAL 2023

RUT :
RAZÓN SOCIAL :
ROL :

Sr(a) Informante,

Según nuestros registros institucionales, la empresa que usted representa ha contestado el cuestionario de empleador de la Encuesta Laboral 2023. **Es por lo anterior, que le indicamos que para cumplir con nuestro requerimiento es necesario responder los cuestionarios pendientes de sindicato/trabajador y autoaplicado.**

Para cumplir oportunamente con lo solicitado, le agradecemos responder ambos cuestionarios **a la brevedad.**

Cada envío fue informado a las personas encargadas de grupo de cada región, para que gestionaran el contacto con las personas informantes, con miras a que estos regularicen el estado de su(s) cuestionario(s) pendiente(s). La siguiente tabla muestra la cronología de los correos masivos enviados junto con su contenido.

Tabla 24. Cronología de envío de correos masivos

Fecha	Temática
13-07-2023	Solicitud de información SM1
22-08-2023	Solicitud de información SM2
04-09-2023	Solicitud de información SM3
07-09-2023	Solicitud de información: Notificación plazo vencido SM1 y 2
12-09-2023	Solicitud de información SM4
28-09-2023	Solicitud de información: Notificación plazo vencido SM3 Y 4
05-10-2023	Recordatorio autoaplicado pendiente SM1 y 2
12-10-2023	Recordatorio a todas las submuestras: autoaplicado en estado “Completo” e “Incompleto”.
16-10-2023	Solicitud de información empresas que no habían recibido correo masivo antes, de todas las submuestras (autoaplicado en estado “No Ingresado”, “Completo” e “Incompleto”)
19-10-2023	Recordatorio autoaplicado pendiente de todas las submuestras en estado “No Ingresados, “Completo” e “Incompleto”, a 1 día de vencer plazo.
25-10-2023	Recordatorio varias temáticas de todas las submuestras: 1. Cuestionarios autoaplicado en estados “Completo” e “Incompleto” que no habían recibido correo masivo antes. 2. Cuestionarios autoaplicado en estado “No Ingresados” que no habían recibido correo masivo antes. 3. Cuestionarios autoaplicado en estados “Completo” e “Incompleto” a 1 día de vencer plazo. 4. Cuestionarios autoaplicado en diferentes estados con plazo vencido.
31-10-2023	Recordatorio a empresas con autoaplicado pendiente y cuestionarios a sindicato o trabajador pendiente, todas las submuestras.
08-11-2023	Dos temáticas: 1. Recordatorio a empresas con autoaplicado pendiente y cuestionarios a sindicato o trabajador pendiente, todas las submuestras. 2. Recordatorio a autoaplicado en estado “No ingresado, “Completo” e “Incompleto”, todas las submuestras.
15-11-2023	Recordatorio autoaplicado pendiente de todas las submuestras en estado “No Ingresados, “Completo” e “Incompleto”, vence plazo ese mismo día.
28-11-2023	Recordatorio varias temáticas de todas las submuestras: 1. Empresas solo con autoaplicado pendiente. 2. Cuestionario autoaplicado y a sindicato o trabajador pendientes. 3. Cuestionario autoaplicado y a empleador pendientes. Empresas con los tres cuestionarios pendientes.
14-12-2023	Envío de proceso de aviso de multa mediante sistema institucional Bizagi.
26-12-2023	Recordatorio varias temáticas de todas las submuestras: 1. Recordatorio a empresas con autoaplicado pendiente incluidas en proceso de Bizagi anterior. 2. Recordatorio a empresas con sindicato o trabajador pendiente. 3. Recordatorio a empresas con autoaplicado pendiente. 4. Recordatorio a empresas con autoaplicado pendiente y cuestionarios a sindicato o trabajador pendiente, todas las submuestras. 5. Recordatorio a empresas con los tres cuestionarios pendientes.

Complementando la estrategia de envío de correos masivos, se realizó un proceso de aviso de multas a través del sistema institucional Bizagi, que consistió en notificar a todas aquellas empresas con cuestionarios a empleadores y a sindicatos

o trabajadores logrados, pero que aún mantuviesen el cuestionario autoaplicado pendiente, y a las empresas de inclusión forzosa que contaran con al menos uno de los tres cuestionarios pendientes. Este proceso se realizó el día 14 de diciembre de 2023, notificando a 741 unidades. De este modo, se envió a las empresas un correo electrónico informando el aviso de multa por no suministrar información sobre la Encla, que no sería cursada en caso de remitir la información solicitada. Este proceso generó una carta y un acta de visita por empresa, que fueron impresas por los equipos regionales de recolección y que, en caso de ser posible, se entregaron presencialmente por las personas recolectoras para notificar a la empresa.

VI.2. Diseño del Instrumento

Con el fin de realizar análisis sobre cambios en las características de las relaciones laborales en Chile, los cuatro cuestionarios de la Encla 2023 replican en gran parte las preguntas realizadas en cuestionarios de la Encla de versiones anteriores. Sin embargo, también se incluyeron preguntas adicionales y/o modificaciones a preguntas existentes, esto se revisó durante el segundo semestre del 2022 en reuniones de trabajo conjuntas entre la Unidad de Estudios de la Dirección del Trabajo y el equipo Técnico Encla del INE. Todos estos cambios fueron evaluados en el marco de un plan de diseño de cuestionarios que incluyó diversas etapas que se abordarán en el próximo apartado.

El proceso comenzó la segunda semana de diciembre de 2022, a partir de la elaboración del protocolo de las pruebas cognitivas y la verificación de los directorios de sindicatos y empleadores. En enero de 2023, ya definidos los cuestionarios preliminares, se realizó la selección del conjunto de preguntas a testear, con base a los requerimientos de la Dirección del Trabajo y las recomendaciones emanadas por el INE, expresados a través de un listado de los principales indicadores que se elaborarán con los datos recogidos por la Encla e indicaciones realizadas sobre la inclusión de preguntas específicas por el mismo servicio durante las reuniones de revisión de cuestionarios. La siguiente tabla resume las instancias de revisiones y pruebas que han tenido los cuestionarios:

Tabla 25. Instancias de revisiones y pruebas de cuestionarios Encla 2023

Cuestionario	Pruebas cognitivas	Consulta a expertos	Prueba de usabilidad 1	Evaluación y ajuste de cuestionarios	Prueba de usabilidad 2	Evaluación y ajuste final de cuestionarios
Autoaplicado	✓	✓	--	--	--	✓
Empleadores	✓	✓	✓	✓	✓	✓
Trabajadores	✓	✓	✓	✓	✓	✓
Sindicatos	✓	✓	✓	✓	✓	✓

VI.2.1. Programación de cuestionarios

Los cuestionarios Encla fueron programados en dos plataformas distintas. Por un lado, el cuestionario autoaplicado se programó en un sistema propio (a través de desarrolladores pertenecientes a la Subdirección de Tecnologías de la Información del INE), ya que de esta forma se pudo tener un mejor control del proceso de llenado del formulario y recabar información sobre las interacciones que tiene el usuario/a con la plataforma. Los cuestionarios de Empleadores, Sindicatos y Trabajadores fueron programados por el equipo técnico en la plataforma SuSo, que es un *software* gratuito desarrollado por el grupo de datos del Banco Mundial. Esta plataforma permite contar con un *software* diseñador de cuestionarios y un servidor para gestionar la asignación de encuestas y monitorear el avance de la recolección.

En general, la principal ventaja de la programación de encuestas en plataformas computacionales, en comparación con una recolección de datos PAPI, es que permite realizar entrevistas CATI y CAPI. La programación de reglas de validación directamente en los cuestionarios en estas plataformas permite superar muchos de los problemas que se generan al momento de digitar encuestas que fueron llenadas en papel, tales como inconsistencias, valores no admisibles o saltos entre preguntas mal ejecutados. Este elemento es una de las principales ventajas con las que cuenta la recolección Encla 2023 respecto a versiones anteriores en términos del proceso de recolección de información.

VI.2.2. Consulta a personas expertas

Los cuestionarios fueron sometidos a revisión en mesas de trabajo conjuntas entre el INE y la DT. Este proceso de revisión se extendió por alrededor de tres meses entre diciembre de 2022 y febrero de 2023 e incluyó un examen exhaustivo de cada una de las preguntas involucradas en el cuestionario, considerando tanto sus objetivos para el estudio, como el uso que se les otorgará en análisis posteriores. Como resultado de este proceso de revisión, se obtuvieron cuestionarios diferentes respecto a su versión original, que incluyen un conjunto considerable de modificaciones que reflejan el proceso de discusión y reflexión de las mesas de trabajo. Estas modificaciones fueron de diferente naturaleza, incluyendo aspectos de mejoramiento de los enunciados, incorporación de instrucciones específicas a módulos, grupos de preguntas o preguntas específicas, perfeccionamiento de las alternativas de respuesta, cambios en el orden y en el flujo de las preguntas y módulos, e incluso la eliminación e inclusión de preguntas.

La primera instancia de evaluación para mejorar los cuestionarios de la Encla en su versión 2023 fue la consulta a personas expertas. El equipo de la Unidad de Estudios de la DT realizó una consulta a expertos/as que alimentó los análisis para evaluar y proponer modificaciones a los cuatro cuestionarios que componen la encuesta. Esta consulta se efectuó durante la segunda quincena del mes de enero de 2023 e involucró a 10 profesionales expertos/as representantes de siete instituciones públicas relacionadas con el mundo del trabajo, la producción de estadísticas oficiales y normativas atingentes al ámbito laboral. Estos profesionales recibieron un listado de preguntas de los cuestionarios relacionadas a su ámbito de experticia y se les solicitó realizar una revisión para sugerencias y comentarios.

En concreto, los expertos/as generaron una serie de sugerencias referidas a distintos niveles de los cuestionarios, como enunciados, alternativas de respuesta o instrucciones. La siguiente tabla muestra el número de modificaciones sugeridas por los expertos/as consultados que fueron aceptadas por el equipo de la DT por considerar que implicaban una mejora a los respectivos instrumentos de recolección:

Tabla 26. Número de modificaciones sugeridas por consulta a Expertos/as admitidas por equipo Dirección del Trabajo

Cuestionario	N° de cambios sugeridos por expertos/as
Autoaplicado	47
Empleadores	25
Sindicatos	24
Trabajadores	20

VI.2.3. Pruebas cognitivas

De forma paralela a la consulta de expertos, el equipo técnico de la Encla llevó a cabo pruebas cognitivas. Luego de una selección acordada de preguntas a evaluar por el equipo técnico INE en conjunto con contraparte DT, dichas preguntas fueron sometidas al proceso de pruebas cognitivas; lo que implicó evaluarlas en términos de dimensiones tales como comprensión, memoria, estimación y respuesta[22]. La tabla muestra el listado de preguntas evaluadas en esta etapa:

Tabla 27. Preguntas sometidas a evaluación de pruebas cognitivas

Cuestionario	Módulo	N° preguntas evaluadas
Autoaplicado	E: Trabajadores extranjeros	2
	G: Salud y seguridad	2
	P: Digitalización y automatización	4
	O: Teletrabajo	7
Empleadores	D: Remuneraciones	2
	E: Trabajadores extranjeros	1
	F: Organización del tiempo de trabajo	1
	G: Seguridad y salud	1
	L: Incidencia de conflictos	1
	M: Negociación colectiva	1
	N: Crisis sanitaria	1
Sindicatos	N: Crisis sanitaria	3
	E: Trabajadores extranjeros	1
	F: Organización del tiempo de trabajo	1
	O: Teletrabajo	5
	G: Salud y seguridad	2
	J: Relaciones laborales	1
	L: Incidencia de conflictos	1
	M: Negociación colectiva	2
	P: Digitalización y Automatización	5
Trabajadores	N: Crisis Sanitaria	1
	O: Teletrabajo	5
	G: Salud y Seguridad	2
	L: Incidencia de conflictos	1
	M: Negociación colectiva	1
	P: Digitalización y automatización	4
TOTAL		58

La aplicación de las pruebas cognitivas se llevó a cabo entre la última semana de diciembre de 2022 y la segunda semana de enero de 2023. La elaboración de la muestra contó con criterios diferentes según el tipo de informante idóneo de

[22] El detalle de los cambios realizados a las preguntas de los cuatro cuestionarios, según los resultados obtenidos en las pruebas cognitivas, puede ser revisado en el siguiente enlace: Informe final entrevistas cognitivas Encla 2023.docx

cada cuestionario. En el cuestionario a trabajadores se utilizó el nivel educacional de los informantes como variable clave para la definición de cuotas, según se presenta en la tabla 28.

Tabla 28. Distribución de la muestra de trabajadores para la aplicación de pruebas cognitivas

Nivel educacional	Número de trabajadores	Número de entrevistas logradas
Bajo: educación básica completa	6	1
Medio: educación media completa	7	7
Alto: educación terciaria completa	7	7
Total	20	15

En el caso de los dirigentes sindicales, la variable de cuoteo fue el sector económico en que se desempeña su empresa, en el entendido de que la realidad de las organizaciones de trabajadores difiera debido a la actividad económica principal que realiza la unidad. La tabla 29 muestra la distribución de la muestra para estos informantes.

Tabla 29. Distribución de la muestra de dirigentes sindicales para la aplicación de pruebas cognitivas

Sector de la economía	Número de trabajadores	Número de entrevistas logradas
Sector primario	6	7
Sector secundario	7	7
Sector terciario	7	7
Total	20	21

Por último, la variable de cuoteo para los informantes de los cuestionarios autoaplicados y a empleadores es el tamaño de la empresa según número de trabajadores. La siguiente tabla presenta la distribución de la muestra para los empleadores que participaron de las pruebas cognitivas.

Tabla 30. Distribución de la muestra de empleadores para la aplicación de pruebas cognitivas

Tamaño de la empresa	Número de trabajadores	Número de entrevistas logradas
Microempresa (5-9 trabajadores)	5	2
Pequeña (10-49 trabajadores)	5	4
Mediana (50-199 trabajadores)	5	4
Grande (200 o más trabajadores)	5	8
Total	20	18

VI.2.4. Caracterización de las versiones finales de los cuestionarios

Los cuatro cuestionarios de la Encla (autoaplicado, empleadores, dirigentes sindicales y trabajadores) fueron sometidos a revisión en mesas de trabajo conjuntas entre el INE y la DT. Este proceso de revisión se extendió por alrededor de tres meses e incluyó un examen exhaustivo de cada una de las preguntas involucradas en el cuestionario, considerando tanto sus objetivos para el estudio, como el uso que se les otorgará en análisis posteriores.

VI.2.4.1. Cuestionario autoaplicado

El cuestionario autoaplicado se compone de 14 módulos que, en total contienen 112 preguntas, de las cuales 15 fueron elegidas para formar parte de las entrevistas cognitivas. La entrevista se aplicó a través de un formulario *web.*

Tabla 31. Distribución preguntas del cuestionario autoaplicado, según módulo

Módulo		Número de preguntas del módulo	Prueba cognitiva	
			Preguntas	Número
A: Aspectos generales	Contiene la identificación de la empresa	9	-	0
B: Contratos	Información respecto de contrataciones, desvinculaciones y tipos de contratos	9	-	0
C: Subcontratación	Información sobre actividades económicas sub-contratadas, características de dichos subcontratos y personal subcontratado	15	-	0
D: Remuneraciones	Información sobre distribución de remuneraciones en la empresa	10	-	0
E: Trabajadores extranjeros	Información sobre trabajadores/as extranjeros/as, países de origen, remuneraciones	5	E1, E4	2
F: Organización del tiempo de trabajo	Información sobre jornadas, turnos, horas extra, entre otros	11	-	0
G: Salud y seguridad	Información sobre incidencia de accidentes y/o enfermedades profesionales; conformación de Comité Paritario de Higiene y Seguridad	5	G2, G3	2
H: Protección a la maternidad y paternidad	Información sobre trabajadores/as que acceden a beneficios de protección a la paternidad o maternidad en la empresa	10	-	0
I: Capacitación	Número de capacitaciones realizadas, fuentes de financiamiento, personal capacitado	4	-	0
K: Acción sindical	Número de sindicatos y número de afiliados	3	-	0
M: Negociación colectiva	Número de instrumentos colectivos y/o acuerdos de grupo negociador vigentes	4	-	0
P: Digitalización y automatización	Tecnologías de digitalización y/o automatizaciones implementadas en la empresa y su impacto sobre los empleos	7	P1, P2, P3, P7	4
O: Teletrabajo	Personas que realizan teletrabajo en la empresa. Cultura de teletrabajo. Incidencia del teletrabajo antes, durante y después de la pandemia.	20	O1, O2, O4, O8, O10, O19, O20	7
N: Crisis sanitaria	Número de trabajadores de la empresa que se acogieron a Ley N° 21.227 de protección al empleo.	2	-	0
Total de preguntas		**112**		**15**

VI.2.4.2. Cuestionario dirigido a empleadores/as

El cuestionario dirigido a empleadores/as está conformado por once módulos y un total de 67 preguntas, de las cuales ocho fueron elegidas para ser probadas a través de entrevistas cognitivas. La entrevista se aplicó de forma telefónica o presencial.

Tabla 32. Distribución preguntas prueba cognitiva del cuestionario a empleadores, según módulo

Módulo	Número de preguntas del módulo	Prueba cognitiva	
		Preguntas	Número
A: Aspectos generales	2	-	0
D: Remuneraciones	4	D2, D3	2
E: Trabajadores extranjeros	5	E5	1
F: Organización del tiempo de trabajo	3	D14	1
G: Seguridad y salud	8	G8	1
I: Capacitación	7	-	0
J: Relaciones laborales	15	-	0
L: Incidencia de conflictos	4	L1	1
M: Negociación colectiva	13	M4, M13	1
N: Crisis sanitaria	2	N1	1
R: Antecedentes personales	4	-	0
Total preguntas	**67**	-	**8**

VI.2.4.3. Cuestionario dirigido a sindicatos

El cuestionario dirigido a sindicatos contiene 14 módulos (cuatro más que en su versión anterior) y 142 preguntas en total, de las que trece fueron escogidas para ser probadas por medio de la prueba cognitiva. La entrevista se aplicó de forma presencial.

Tabla 33. Distribución preguntas prueba cognitiva del cuestionario a dirigentes sindicales, según módulo

Módulo	Número de preguntas del módulo	Prueba cognitiva	
		Preguntas	Número
N: Crisis sanitaria	6	N1, N4, N6	3
C: Externalización y temporalidad	5	-	-
D: Remuneraciones	4	-	-
E: Trabajadores extranjeros	4	E4	1
F: Organización del tiempo de trabajo	9	F1	1
O: Teletrabajo	11	O1, O6, O7, O8, O11	5
G: Salud y seguridad	12	G1, G10	2
I: Capacitación	6	-	-
J: Relaciones laborales	10	J6	1
K: Acción sindical	28	-	-
L: Incidencia de conflictos	4	L1	1
M: Negociación colectiva	27	M5, M23	2
P: Digitalización y automatización	7	P1, P2, P3, P5, P6	5
R: Antecedentes personales	9	-	-
Total preguntas	**142**		**21**

VI.2.4.4. Cuestionario dirigido a trabajadores

El cuestionario dirigido a trabajadores/as está organizado en 14 módulos, considerando el de "antecedentes personales", y se compone de 96 preguntas, de las cuales 14 fueron seleccionadas para participar de la prueba cognitiva. La entrevista se aplicó de forma presencial. Si no existe ningún sindicato en la empresa, se realiza el cuestionario a trabajadores/as.

Tabla 34. Distribución preguntas del cuestionario trabajadores, según módulo

Módulo	Número de preguntas del módulo	Prueba cognitiva	
		Preguntas	Número
Sin módulo	1	-	-
N: Crisis Sanitaria	3	N1	1
C: Externalización y temporalidad	5	-	-
D: Remuneraciones e incentivos	2	-	-
E: Trabajadores Extranjeros	4	-	-
F: Organización del tiempo de trabajo	9	-	-
O: Teletrabajo	11	O1, O6, O7, O8, O11	5
G: Salud y Seguridad	12	G1, G10	2
I: Capacitación	6	-	-
J: Relaciones Laborales	9	-	-
K: Acción Sindical	5	-	-
L: Incidencia de conflictos	4	L1	1
M: Negociación colectiva	12	M11	1
P: Digitalización y automatización	7	P1, P2, P3, P5	4
R: Antecedentes personales	6	-	-
Total preguntas	**96**	**-**	**14**

VI.4. Prueba y evaluación de los Instrumentos

Como parte de las tareas asociadas a la construcción de los cuestionarios Encla 2023, se realizaron pruebas de usabilidad que permitieron evaluar el uso de los cuestionarios en DMC; así como el funcionamiento de las validaciones automáticas y el flujo en general de la aplicación de las entrevistas. En concreto, las pruebas de usabilidad consistieron en la aplicación real (en el caso de la primera prueba) y simulada (en las pruebas 2 y 3) de los cuestionarios autoaplicados (solo en la prueba 1), a empleadores, trabajadores y sindicatos. Mientras la primera prueba de usabilidad se orientó a testear la aplicación de los cuestionarios en un escenario real, evaluando tanto la duración como la comprensión y el proceso de respuesta de los informantes; la segunda se abocó a medir la duración de los cuestionarios en sus flujos más extensos; mientras que la tercera estuvo enfocada en el testeo de la hoja de ruta desarrollada para cada cuestionario.

VI.4.1. Prueba de usabilidad 1

La primera prueba de usabilidad fue llevada a cabo una vez que se implementaron los cambios provenientes de las pruebas cognitivas y de las consultas a expertos. Es importante indicar que, en el caso del cuestionario autoaplicado, los cambios introducidos en esta etapa fueron considerados finales, ya que fue necesario entregar este formulario al Departamento de Implementación de Operaciones Estadísticas del INE para su programación en el sistema de gestión operativa para encuestas económicas. Sin embargo, para el caso de los cuestionarios a empleadores, trabajadores y sindicatos, se realizó esta primera prueba de usabilidad, cuyo objetivo principal fue el de evaluar los formularios en cuanto a su correcta programación (evaluación de flujos y saltos), así como su duración. Esta prueba fue aplicada a 26 personas informantes con perfiles representativos de diversas empresas.

Tabla 35. Personas informantes encuestadas en pruebas de usabilidad 1

Cuestionario	Actividad económica	Tamaño	Cargo
Trabajadores	Otras actividades de servicios	Grande	Auxiliar de aseo
	Actividades financieras y de seguros	Grande	Account manager
	Comercio	Pequeña	Cajera
	Actividades de atención de la salud humana y de asistencia social	Grande	Trabajadora de la salud
	Comercio	Grande	Jefe F&I
	Alojamiento y de servicio de comidas	Grande	Supervisor de operaciones
	Actividades profesionales, científicas y técnicas	Pequeña	Analista de datos
	Actividades profesionales, científicas y técnicas	Grande	Administrador
	Alojamiento y de servicio de comidas	Grande	Chef ejecutiva
	Transporte y almacenamiento	Grande	Analista de plan. estratégica
Sindicatos	Industrias manufactureras	Grande	Presidente del sindicato
	Agricultura, ganadería, silvicultura y pesca	Grande	Presidente del sindicato
	Suministro de electricidad, gas, vapor y aire acondicionado	Grande	Secretario del sindicato
	Industria manufacturera	Grande	Presidente del sindicato
	Industria manufacturera	Grande	Presidente del sindicato
	Comercio	Grande	Presidente del sindicato
	Comercio	Grande	Secretario del sindicato
	Suministro de electricidad, gas, vapor y aire acondicionado	Grande	Presidente del sindicato
	Industria manufacturera	Grande	Presidente del sindicato
Empleadores	Industria manufacturera	Pequeña	Gerente de Operaciones
	Industria manufacturera	Mediana	Jefe/a de personal
	Industria manufacturera	Grande	Sub-gerente de RRHH
	Transporte y almacenamiento	Pequeña	Gerente de Operaciones
	Actividades profesionales, científicas y técnicas	Pequeña	Socio/a
	Comercio	Grande	Sub-gerente de RRHH

Los resultados de la primera prueba de usabilidad revelaron dos problemas principales:

1. Elevado tiempo promedio de aplicación de cuestionarios (sobre todo en el caso de trabajadores y sindicatos). En el caso de sindicatos esta duración alcanzó un promedio de 96 minutos (cerca de una hora y media), con módulos tales como el G (Seguridad y salud) y el K (Acción sindical) con una duración promedio de 20 minutos. Este hallazgo

fue considerado por el equipo técnico del INE y por el equipo de la DT como un problema necesario de resolver, ya que duraciones elevadas afectan la viabilidad de los cuestionarios en tanto implican un agotamiento tanto de la persona informante como del recolector/a, generando el riesgo de obtener entrevistas incompletas y/o sesgadas en sus respuestas.

2. **Mejoras necesarias en formulación de preguntas.** También se detectaron problemas relacionados a la formulación y comprensión de preguntas por las personas informantes. En ese sentido, esta prueba de usabilidad fue útil para evaluar la totalidad de preguntas de los cuestionarios en lugar de una selección de preguntas como se hizo en las pruebas cognitivas. Se detectaron mejoras necesarias en usabilidad (facilitar acceso a preguntas organizadas en matrices), formulación de preguntas y mejoramiento de instrucciones para el/la encuestador/a.

El resultado principal de este análisis fue un listado de mejoras y de eliminación de preguntas para mejorar el tiempo de aplicación de los cuestionarios.

VI.4.2. Prueba de usabilidad 2

Una vez aplicadas las mejoras detectadas en la prueba de usabilidad 1, se testeó el rendimiento del cuestionario en una segunda prueba de usabilidad, orientada fundamentalmente a verificar que los tiempos de aplicación efectivamente hubiesen sido reducidos a un nivel aceptable en términos metodológicos. El equipo técnico definió en una hora la duración máxima de un cuestionario para considerarse operativamente factible. Esta segunda prueba de usabilidad fue realizada a personas informantes colaboradoras dentro del Departamento de Estadísticas del Trabajo del INE y fuera de la institución, ya que su objetivo principal no estuvo directamente orientado a evaluar comprensión, sino la duración. Por ello, se testearon las aplicaciones más largas posibles en cada cuestionario, abriendo todos los flujos posibles. Se realizaron en total 16 aplicaciones de cuestionarios que buscaron evaluar el tiempo promedio de aplicación preferentemente en flujos que implicaban contestar una mayor cantidad de preguntas. El detalle se encuentra en la tabla 36 a continuación:

Tabla 36. Número de cuestionarios realizados en prueba de usabilidad 2

Cuestionario	N° de pruebas realizadas
Trabajadores	5
Dirigentes Sindicales	5
Empleadores	6
Total	16

El resultado de los tiempos promedio de duración mostró que los cuestionarios redujeron su tiempo de duración de manera relevante. Al comparar con los tiempos promedio observados en la primera prueba, es posible apreciar que si bien el cuestionario a empleadores aumentó en 3 minutos promedio; el cuestionario a trabajadores se redujo en 27 minutos y el cuestionario a dirigentes sindicales se redujo a un promedio de 41 minutos. Sin perjuicio de que en esta

instancia se observaron principalmente mejoras para reducir el tiempo de aplicación de las entrevistas, de todas formas, se aprovechó la instancia para introducir nuevos ajustes.

VI.4.3. Prueba de usabilidad 3

La tercera prueba de usabilidad de los cuestionarios se realiza una vez definida la Hoja de Ruta (ver detalle de este instrumento en el punto de la sección siguiente de este documento). La complejidad de las situaciones de contacto en la Encla, al tratarse de múltiples modos de recolección y con distintos escenarios de disponibilidad de datos de contacto implican probar los distintos flujos posibles para determinar que la programación entregue el código de disposición final correcto asociado a la encuesta.

Es importante recordar que el proceso de verificación del directorio puede terminar con tres escenarios posibles de disponibilidad del contacto a la persona informante: contacto de la o el informante idóneo; contacto a otro/a informante dentro de la empresa o no existencia de datos de contacto. Para cada uno de estos escenarios puede efectuarse un flujo de contacto definido por dos modos de contacto: presencial o telefónico. Se definieron un total de 82 posibles flujos de contacto presencial o telefónico que pueden seguir las entrevistas, considerando tanto la cooperación como el rechazo de las personas informantes, así como la (no) elegibilidad de las unidades. En el caso del cuestionario a trabajadores/sindicatos el modo prioritario de contacto es presencial y en el caso del cuestionario a empleadores es telefónico; sin embargo, para cada caso existe la opción de utilizar el modo alternativo en casos excepcionales (por solicitud de la persona informante o dificultad de acceso/lejanía de la faena).

La tabla 37 muestra los distintos escenarios que fueron testeados en la tercera prueba de usabilidad para probar la correcta asignación de los códigos de disposición:

Tabla 37. Flujos de Hoja de Ruta probados en tercera prueba de usabilidad

Cuestionario	Tipo de contacto	Número de situaciones de contacto testeadas[3]
Empleadores	Telefónico	26
	Presencial	25
Trabajadores/Sindicatos	Telefónico	17
	Presencial	14

En síntesis, los procedimientos anteriormente mencionados permitieron mejorar los cuestionarios en su conjunto, adecuando las preguntas al proceso de respuesta de los informantes idóneos de cada cuestionario, así como ajustando los tiempos de aplicación para viabilizar procesos de entrevista que reduzcan el error de medición, especialmente en el cuestionario a sindicatos.

VI.5. Método de capacitación

"La capacitación es el proceso educativo, aplicado de manera sistemática y organizada, por medio del cual las personas adquieren conocimientos, desarrollan habilidades y competencias en función de objetivos definidos" (Chiavenato, I., 2017). En este sentido, la capacitación permite que los equipos estén dotados de conocimientos relevantes y de herramientas que les permitan realizar su trabajo de forma eficiente y bien orientada; y que conozcan el manejo de las plataformas que utilizarán para la recolección de datos.

En este marco, para establecer un proceso de recolección con altos estándares de calidad, que cuente con una medición consistente y con menos fallas, se diseñó y se entrenó a todos los cargos del equipo operativo respecto al plan de capacitación. De esta forma, se promovió una estandarización de protocolos operativos y homogenización en la entrega de conocimientos técnicos, operativos, y administrativos.

Entendiendo las características propias de este estudio, se realizó una capacitación por etapas, en la que se consideró, primero, de forma separada, tres grupos participantes que tenían relación con el perfil del cargo y sus responsabilidades asociadas:

- Persona recolectora
- Persona encargada de grupo
- Persona supervisora

VI.5.1. Calendarización de la capacitación

En cuanto a la estrategia para la capacitación, esta se realizó de forma sistemática y escalonada, iniciando con todo el personal perteneciente al equipo de operaciones (personas recolectoras, personas encargadas de grupo y personas supervisoras) por un periodo de cinco días hábiles. Al finalizar esta etapa de capacitación general, se realizó una capacitación específica a los equipos de personas encargadas de grupo y supervisoras, en salas separadas; esta capacitación específica contempló un día completo de trabajo.

Asimismo, en función de la gran cantidad de participantes que tuvo la Encla 2023 se definió realizar la capacitación de forma telemática (mediante la aplicación *Microsoft Teams*), a excepción del equipo operativo que desarrolló funciones en la Región Metropolitana, quienes tuvieron una modalidad mixta; la capacitación se realizó de manera remota, pero las personas capacitadas se reunieron de manera presencial en una sala ubicada en las dependencias del INE, en las que contaron con el apoyo del equipo de personas encargadas de área. Vale destacar que la capacitación por la marcha blanca sólo se ejecutó en las regiones de Valparaíso y Metropolitana, mientras que la capacitación por recolección oficial se impartió al resto de regiones del país. La calendarización de cada instancia se presenta en la tabla 38.

Tabla 38. Calendarización de capacitaciones marcha blanca y recolección oficial, por etapa

Etapa	Perfil	Inicio	Término	Tipo de capacitación	Cantidad de asistentes	Capacitadores
Marcha blanca	Personas Encargadas de grupo y recolectoras	13-06-2023	19-06-2023	*Microsoft Teams*	29	Personas Encargadas de Área/Equipo Técnico
	Personas Encargadas de grupo	20-06-2023	20-06-2023	*Microsoft Teams*	17	Personas Encargadas de Área /Equipo Técnico
Recolección oficial	Personas Supervisoras y recolectoras	03-07-2023	07-07-2023	*Microsoft Teams*	64	Personas Encargadas de Área/Equipo Técnico
	Personas Supervisoras	10-07-2023	10-07-2023	*Microsoft Teams*	7	Personas Encargadas de Área /Equipo Técnico

VI.5.2. Distribución de la capacitación

Otra de las estrategias que se estableció para este proceso, en particular para la capacitación de las personas recolectoras y supervisoras, tiene relación con la distribución del equipo capacitador.

La presente tabla muestra la organización que tuvieron las capacitaciones en cuanto a la separación por grupos, fechas y personas capacitadoras:

Tabla 39. Distribución de capacitación marcha blanca y recolección oficial, por etapa

Etapa	Sala	Región	Cantidad Personas Capacitadas	Capacitadores	Modalidad
Marcha blanca	1	Metropolitana - Valparaíso	29	Personas Encargadas de Área Recolección / Analistas SDT / SDA / DT	Hibrida/ Telemática
Recolección oficial	1	Regiones varias	32	Personas Encargadas de Área SDO / Analistas SDT / Analistas SDA / DT	Telemática
	2		32		

VI.5.3. Módulos de capacitación

En cuanto a los contenidos de la capacitación, se distribuyeron en 14 módulos temáticos desarrollados en un trabajo colaborativo entre las unidades de coordinación, Subdirección de Operaciones (SDO) y Subdirección Técnica (SDT), y su mandante, la DT. Además, la capacitación se realizó de forma dinámica y con un modelo de reforzamiento sistemático mediante ejercicios prácticos que permitieron un aprendizaje tangible de los contenidos.

Tabla 40. Módulos de capacitación

Ítem	Módulo	Objetivo	Equipo Responsable
1	Contexto y Antecedentes	Presentar los objetivos y las principales características del proyecto. Además, se revisa la forma de organización del equipo de recolección.	SDT – DT
2	Aspectos Administrativos	Entregar un panorama general de la institución, señalando el marco legal en que se desarrolla el trabajo de recolección, revisando las medidas de protección personal y el protocolo en caso de accidentes, licencias, entre otros.	SDA
3	Conceptos centrales de la recolección	Presentar los principales conceptos asociados a la recolección de datos de empresas: unidad estadística, empresa, actividad económica, razón social, entre otros.	SDO – SDT
4	Control de calidad y supervisión	Presentar las funciones del equipo de supervisión en relación con el control de calidad de la recolección de datos.	SDO
5	Personas Recolectoras de datos	Presentar las funciones del personal recolector y revisar el rol que cumplen en el proceso de producción de datos, señalando sus deberes y enfatizando el trabajo de apertura de empresas.	SDO
6	Personas Encargados de grupo	Presentar las funciones de las personas encargadas de grupo y enfatizar en el rol que cumplen en el proceso de producción de datos, monitoreando y organizando el trabajo del equipo de recolectores.	SDO
7	Personas Supervisoras	Presentar las funciones de las personas supervisoras y enfatizar en el rol que cumplen en la calidad del proceso de producción de datos.	SDO
8	Sistema de Gestión Operativa (SGO)	Presentar las características y funcionalidades del sistema SGO. Asignación, revisión y envío de encuestas.	SDO
9	*Survey Solutions* (SuSo)	Presentar las características y funcionalidades del sistema SuSo. Sincronización, llenado de encuestas, completar y enviar.	SDO
10	Hoja de ruta	Presentar la hoja de ruta con sus objetivos, estructura del instrumento, correcto uso, explicar los códigos de disposición final.	SDT
11	Cuestionario a Empleador	Presentar cuestionario a empleador, con sus objetivos, conceptos fundamentales, flujo y forma de llenado.	SDT/DT
12	Cuestionario a Trabajador y Sindicato	Presentar cuestionario a trabajador y sindicato, con sus objetivos, conceptos fundamentales, flujo y forma de llenado.	SDT/DT
13	Cuestionario Autoaplicado	Presentar cuestionario autoaplicado con sus objetivos, conceptos fundamentales, flujo y forma de llenado.	SDT/DT
14	Prueba de usabilidad	Realizar prueba de usabilidad, para detectar posibles errores y verificar saltos de la encuesta.	SDO

VI.5.4. Programa de capacitación

El plan de capacitación para personas recolectoras de datos y personas encargadas de grupo fue una experiencia enriquecedora que se extendió a lo largo de cinco días intensivos, seguidos de un día adicional exclusivo para los y las encargadas de grupo. Cada jornada estuvo meticulosamente estructurada en bloques temáticos distintos, con pausas estratégicas de 15 minutos para permitir la reflexión y la recarga de energías.

Durante esa semana, se exploró una amplia gama de temas fundamentales para el éxito en la recolección de la encuesta. Desde los aspectos metodológicos del proyecto hasta la ejecución de una entrevista a informantes ficticios en un ambiente de prueba, cada día proporcionó a las personas participantes conocimientos especializados y habilidades prácticas que pueden aplicar en sus roles dentro del proyecto.

Además, el día adicional dedicado a las personas encargadas de grupo les brindó la oportunidad de profundizar en temas de liderazgo, uso de sistemas, gestión de equipos y coordinación. Esta sesión exclusiva les equipó con las herramientas y estrategias necesarias para liderar con éxito a sus equipos y garantizar una óptima ejecución de las encuestas.

VI.5.5. Otros elementos de la capacitación

Previo al inicio de las capacitaciones telemáticas, se entregaron los manuales correspondientes a cada perfil de cargo enfocados en las tareas específicas y generales de cada uno. En esa línea, los insumos generados fueron el Manual de recolección de datos, el Manual de supervisión y el Manual de persona encargada de grupo. El objetivo principal de estos insumos fue minimizar los errores y estandarizar las instrucciones al equipo operativo.

Las presentaciones de los módulos presentados en cada sesión se enviaron por correo electrónico para su revisión por el equipo capacitado. Además, las capacitaciones fueron grabadas[23], lo que permitió la posterior revisión en caso de ser necesario.

Asimismo, se incluyeron evaluaciones que permitieron revisar los niveles de avance en el aprendizaje de quienes participaron, y de esta forma, se generaron retroalimentaciones y jornadas de resolución de dudas al equipo durante la misma capacitación. Estas evaluaciones se realizaron mediante formularios de *Microsoft Forms*, en los que se realizaron preguntas relacionadas a las temáticas de las jornadas de capacitación, esto con la finalidad de reforzar los conocimientos adquiridos, lo que, a su vez, permitió observar quiénes tuvieron un aprendizaje significativo e intensificar la enseñanza de ciertos contenidos a nivel individual y de grupo.

Así, el proceso de capacitación fue esencial para el buen desarrollo de los participantes del equipo de recolección de datos, donde se fomentó la participación de todo el equipo durante las jornadas, sin importar el papel que iba a cumplir. Esto con el objetivo de integrar los conocimientos transmitidos de una manera transversal y entender la recolección de datos como un proceso interdependiente en el que se agrupan los esfuerzos de distintos equipos, y así alcanzar las metas del mandante, la DT, dentro de los estándares de calidad institucionales.

VI.6. Método de supervisión

La supervisión es el proceso que garantiza la corrección del trabajo de recolección de datos mediante la identificación de errores que lo pueden invalidar total o parcialmente, con el objetivo de asegurar que la persona recolectora está realizando en forma correcta la aplicación de los cuestionarios y, así, confirmar que los datos registrados son fidedignos

[23] Previo consentimiento verbal de todas las personas participantes de la capacitación.

El modelo de supervisión implementado consideró dos aspectos: el primero, enfocado en el momento que se llevó a cabo la recolección de los datos a través del Acompañamiento Operativo; y el segundo, orientado en el trabajo posterior de la recolección de los datos mediante la Supervisión Indirecta, la Supervisión del Cuestionario Autoaplicado y la Supervisión de No Elegibles. Además, a medida que se avanzó en la recolección de los datos, se sumaron Supervisiones Específicas que no estaban consideradas en el diseño inicial de la metodología, puesto que surgieron como solicitudes de la Subdirección Técnica y la Subdirección de Operaciones.

VI.6.1. Herramientas para el control de gestión

El proceso de supervisión de la Encla necesitó la implementación de diferentes estrategias de monitoreo, razón por la que contó con tres sistemas de gestión complementarios que estuvieron destinados a funciones específicas de la recolección y supervisión de la encuesta:

- *Survey Solutions (SuSo)*

En la supervisión Encla se utilizó para gestionar los diferentes Formularios *Survey Solutions*, en modalidades de recolección CATI, desde computadores instalados en la oficina. Esto permitió tener información en tiempo real, sin la dificultad de la sincronización de dispositivos de captura. El sistema contó con diferentes perfiles de trabajo:

- Perfil Sede: es el controlador general del proceso de recolección de datos. Establece el directorio de información y posee el acceso a las bases de datos. Este perfil lo utilizó la persona coordinadora nacional de supervisión.

- Perfil Supervisor: es el que coordina las cargas de trabajo, asignando y revisando los formularios. En el se revisa la completitud y coherencia de la información. Este perfil lo utilizó la persona encargada de área de supervisión.

- Perfil Entrevistador: es el que utiliza de forma directa los formularios con información. Este perfil lo utilizó la persona supervisora operativa.

El equipo de supervisión al no contar con una persona analista de datos, un panel de control *(Dashboard)* ni bases de datos estandarizadas, tuvo que construir sus indicadores de forma directa desde la información recopilada en SuSo.

- Sistema de Gestión Operativa (SGO).

El equipo de supervisión utilizó el Sistema de Gestión Operativa para facilitar el acceso a los formularios de empresas a supervisar y disponer la información de gestión operativa de la encuesta. Los perfiles utilizados por el equipo de supervisión fueron:

- Encargado Nacional Supervisión: este perfil es el que crea y administra usuarios, distribuye las áreas de trabajo y tiene acceso al directorio de empresas. Este perfil lo utilizó la persona coordinadora nacional de supervisión.

- Encargado Macrozona Supervisión: este perfil es el que asigna la carga de trabajo utilizando filtros que permiten identificar las unidades, al igual revisa la completitud y coherencia de los formularios. Este perfil lo utilizó la persona encargada de área de supervisión.

- Supervisor (ES): este perfil solo puede visualizar la carga de trabajo asignada por el perfil "Encargado Macrozona Supervisión". Este perfil lo utilizó la persona supervisora operativa.

El Sistema de Gestión Operativa no permitió descargar una base de datos, en consecuencia, no se pudo utilizar para construir indicadores.

- **Sistema de Gestión Operativa de Empresas (SGOE).**

El equipo de supervisión utilizó el Sistema de Gestión Operativa de Empresas para facilitar el acceso a los Cuestionarios autoaplicados y disponer de la información de gestión operativa de la encuesta. Los perfiles utilizados por el equipo de supervisión fueron:

- Coordinador SDO: este perfil es el que crea y administra usuarios, reasigna cuestionario a analistas SDO, revisa el *Tracking* encuesta y directorio de empresas, puede visualizar el avance total de cuestionarios. Este perfil lo utilizó la persona coordinadora nacional de supervisión y la persona encargada de área de supervisión.

- Supervisor: este perfil puede revisar el *Tracking* encuesta, directorio de empresas y avance total de cuestionarios, al igual que permite dejar comentarios de supervisión a cada cuestionario autoaplicado.

El Sistema de Gestión Operativa de Empresas permitió descargar dos bases de datos del cuestionario autoaplicado, donde la primera contenía la información registrada por las personas informantes y la segunda la información modificada por la persona analista SDO. Aquello permitió realizar los procesos de supervisión de autoaplicado y de supervisión de la recuperación de esto.

VI.6.2. Tipos de Supervisión

VI.6.2.1. Acompañamiento Operativo.

El acompañamiento operativo tuvo por objetivo observar los protocolos de recolección de datos a partir del acompañamiento en terreno a las personas recolectoras. Fue una actividad que además buscó constatar los protocolos establecidos de selección de la empresa y la persona informante, observar la aplicación de los cuestionarios, evaluar cualitativamente la jornada de trabajo y relacionar el cuestionario entrevistado con la actividad económica de la empresa.

Aquí se estableció que todas las personas recolectoras debían ser observadas al menos una vez en terreno por las personas encargadas de grupo, ocupándose para ello, el formulario de acompañamiento operativo implementado en *SuSo.* Dicho formulario, se dividió en tres secciones en las que se buscó recoger diferentes momentos del trabajo en terreno (Tabla 41) haciéndose uso de respuestas numéricas, abiertas y de tipo Likert. Si bien el acompañamiento operativo lo planificaron y ejecutaron las personas encargadas de grupo, fue el equipo de supervisión quien revisó la planificación de las salidas a terreno, asignó los formularios y analizó la información registrada por las personas encargadas de grupo.

Tabla 41. Secciones del formulario de acompañamiento operativo

FORMULARIO DE ACOMPAÑAMIENTO OPERATIVO		
Secciones del formulario	Contenido	Dimensiones que evaluó
Antecedentes generales de la jornada de acompañamiento.	Metodología.	Confirmación de la unidad.
		Presentación de la encuesta.
		Selección de la persona informante.
	Desempeño de la persona recolectora.	Actitud ante la persona informante.
		Uso de la vestimenta institucional.
		Conocimientos de la encuesta.
	Recolección de datos.	Aplicación del cuestionario.
		Dispositivo de captura móvil.
		Cierre de la entrevista.
Contexto general de la jornada de acompañamiento.	Empresa visitada y entrevistada.	Clasificar qué cuestionario se utiliza en cada entrevista.
	Actividad económica de las empresas entrevistadas.	Rol(es) de todas las empresas entrevistadas para correlacionar su tamaño y actividad.
	Tipo de formulario utilizado.	Forma en que se realiza la recolección de datos: DMC o papel.
Cierre del acompañamiento operativo.	Descripción cualitativa de la jornada.	Fortalezas de la persona recolectora.
		Oportunidades de mejora.
		Observaciones de la jornada de acompañamiento.
	Evaluación de la jornada de acompañamiento.	Excelente, bien, necesita otra jornada de acompañamiento, necesita un reforzamiento metodológico y no se logra observar una entrevista.

Respecto a los criterios de asignación de acompañamiento operativo, a la persona encargada de grupo se le asignó un formulario de acompañamiento operativo por cada persona recolectora a través de *SuSo,* ya que, lo que se buscó plasmar fue una observación general de toda la jornada. Esto significó que, si en los meses de recolección de datos, la persona encargada de grupo acompañó tres veces a la misma persona recolectora, esta tuvo tres formularios de acompañamiento operativo asignados en diferentes fechas. Con respecto a los criterios de acompañamiento, se puede señalar que existieron de dos tipos: generales y específicos (Tabla 42). Los primeros aludieron a que toda persona recolectora tuvo

que ser acompañada en una jornada de trabajo durante el primer mes de ejecución del proyecto[24] (lo que no se concretó), mientras que los segundos estuvieron enfocados en los casos donde fue necesario volver a planificar una nueva salida a terreno.

Tabla 42. Criterios de asignación[25] del Acompañamiento Operativo.

Criterios generales	Criterios específicos
Todas las personas recolectoras alguna vez en el proyecto.	Personas recolectoras mal evaluadas en la primera jornada de acompañamiento.
Nuevos ingresos.	En cualquier momento por necesidades del proyecto, por ejemplo: persona informante reticente o requerimiento de observar la recolección de datos por parte de la persona encargada de grupo o de la persona encargada de área de recolección.

VI.6.2.2. Supervisión Indirecta.

La supervisión indirecta tuvo por objetivo verificar la información registrada por las personas recolectoras en los Cuestionarios de Empleadores y Dirigentes Sindicales/Trabajadores, mediante llamadas telefónicas a personas informantes para asegurar la veracidad y calidad de los datos. Asimismo, fue una instancia que buscó verificar la selección de la empresa y la persona informante, re-entrevistar informantes para volver a aplicar las preguntas seleccionadas de los cuestionarios de recolección y verificar la aplicación de los protocolos de recolección de datos.

Para esta supervisión se consideraron todas las empresas que se entrevistaron a nivel nacional, siendo las personas supervisoras operativas las responsables de recontactar a las mismas personas informantes que entrevistó el equipo de recolección en un máximo de 7 intentos (diferentes días y horarios)[26]. Como la Encla 2023 trabajó con diferentes perfiles de personas informantes, se diseñaron dos Formularios de Supervisión Indirecta Empleadores y Dirigentes Sindicales/Trabajadores que fueron implementados en *SuSo.* Estos se estructuraron en torno a cuatro secciones[27] (Tabla 43) donde se incorporaron respuestas abiertas y de selección, incluyéndose también la codificación de hallazgos (Tabla

[24] Eso no impidió que dependiendo de las necesidades de cada región se siguieran efectuando acompañamientos operativos hasta diciembre de 2023, previa planificación con las personas encargadas de área de recolección y supervisión.

[25] Hubo cambios si se compara con lo indicado en el Manual de Supervisión. En específico, este indica que, en los criterios específicos para asignar acompañamiento operativo a nuevos ingresos, incluyen una baja tasa de respuesta de encuestas entrevistadas por la persona recolectora.
26 Adicionalmente, la persona encargada de área realizó supervisiones Indirectas a empleadores y dirigentes sindicales/trabajadores entre el 25 de enero y el 20 de febrero de 2024.

[27] La selección de las preguntas a reentrevistar surgieron de una revisión analítica de los módulos, preguntas y tipos de respuestas de los cuestionarios de recolección, determinándose que debían: 1) Estar vinculadas al objetivo general de la encuesta; 2) Hacer referencia a una acción/actividad/situación que las personas informantes recordaran fácilmente; 3) Tener una respuesta de tipo dicotómica (si, no, no aplica, no sabe o no responde); y 4) Ser coherentes con la grilla de suficiencia que estableció la SDT.

44). Ambos Formularios de Supervisión Indirecta tuvieron precargadas las respuestas de los cuestionarios de recolección en la sección de "Reentrevista", por lo tanto, esta supervisión apuntó a la búsqueda de la coincidencia del dato de entrevista y re-entrevista.

Tabla 43. Secciones del Formulario de Supervisión Indirecta a Empleadores y Dirigentes Sindicales/Trabajadores

FORMULARIO DE SUPERVISIÓN INDIRECTA	
Secciones del formulario	Contenido
1. Registro de contacto.	Fecha y hora.
	Rut persona supervisora operativa.
	Nombre y apellido persona supervisora operativa.
	Teléfono tiene tono.
	Se contesta la llamada.
	Se logra contactar a la persona informante.
	La persona informante está dispuesta a responder inmediatamente.
2. Preguntas a reentrevistar.	Selección de empleador, dirigente sindical y/o trabajador.
	Reentrevista de preguntas por tipo de cuestionario.
	Reentrevista de temáticas con respuesta declarativa por tipo de cuestionario.
3. Aplicación de protocolos de recolección de datos.	Protocolos de recolección de datos.
4. Resultado de la supervisión.	Estado de la supervisión.
	Franja horaria en que se realiza la supervisión.
	Observaciones de la supervisión.

Tabla 44. Codificación de hallazgos en la Supervisión Indirecta a Empleadores y Trabajadores/Dirigentes Sindicales

Tipo de hallazgo	
Mejora de información	Cuando la persona informante entrega un dato de mejor calidad, por lo tanto, rectifica la información entregada a la persona recolectora asumiendo que sí había entendido la pregunta al momento de la entrevista.
Error de registro	Cuando la persona informante responde una información distinta a la registrada en el Cuestionario de Recolección, sosteniendo que en la entrevista había declarado la misma respuesta de la reentrevista.
Otro	Cuando la persona informante no entrega mayores detalles respecto al cambio de su respuesta, impidiéndole a la persona supervisora operativa determinar si esta es una mejora de información o un error de registro.

Las asignaciones de los Formularios de Supervisión Indirecta se realizaron mediante el SGO, estando sujetas a los criterios establecidos en la metodología (Tabla 45), y a los indicadores de avance del equipo. Esto último significó que, si una persona supervisora operativa tenía más logro, por ejemplo, en el Formulario de Supervisión Indirecta a Empleadores se detenía la asignación de ese formulario priorizando por unos días solo el de Dirigentes Sindicales/Trabajadores. Aquí se comprometió supervisar el 20% de las empresas entrevistadas por el equipo de recolección, sin embargo, ese porcentaje se elevó debido a la activación del protocolo de posible falsificación (Supervisiones Dirigidas).

Tabla 45. Criterios de asignación de la Supervisión Indirecta

Selección dirigida	Selección aleatoria orientada
Posible falsificación detectada por la persona encargada de grupo.	Persona recolectora.
Posible falsificación detectada por la persona encargada de área de recolección.	Inconsistencia en aquellas preguntas que impiden continuar con la Supervisión Indirecta.
Posible falsificación detectada por el equipo técnico.	Cuando más del 50% de las preguntas de reentrevista presentan inconsistencias.
Posible falsificación detectada en la reentrevista a la persona informante.	Cuando la persona informante señala que no respondió la encuesta.
Cuestionarios de Dirigentes Sindicales o Trabajadores correspondientes a empresas de Inclusión Forzosa.	Aplicar intencionalmente la encuesta en un RUT distinto al indicado en el directorio, siendo que la persona informante menciona ese cambio de RUT tanto a la persona recolectora como a la persona supervisora operativa.
	Cuando el Registro Sindical Único (RSU) no correspondía al indicado en la portada o el sindicato entrevistado no era el mayoritario[28].

VI.6.2.3. Supervisión del Cuestionario Autoaplicado[29].

La supervisión del Cuestionario Autoaplicado tuvo por objetivo revisar la completitud, coherencia y correspondencia de la información que se declaró en las preguntas A1[30] y A2[31], para asegurar que se cumplieran los estándares de calidad de

[28] A partir de la submuestra operativa 3 estos casos ya no activaron el protocolo de posible falsificación, por lo que solo fueron notificados a la persona encargada de área de recolección.
[29] Para más información, se anexa "Formulario de supervisión de gabinete" en su versión en papel.
[30] A1. Señale la o las actividades realizadas por la empresa en orden de importancia (1°, 2° y 3° lugar de ingresos y/o utilidades). Marque como máximo 3 actividades diferentes.
[31] A2. Describa con sus propias palabras las actividades económicas realizadas por su empresa y los productos o servicios principales entregados a sus clientes, según importancia de ingresos y/o utilidades.

la grilla de suficiencia que exigió el proyecto. Igualmente, fue una instancia que buscó corroborar la coherencia y la correspondencia entre la(s) actividad(es) económica(s) de la empresa registrada por la persona informante y la(s) establecida(s) en el directorio.

Para esta supervisión se consideraron todos los Cuestionarios Autoaplicados recibidos por el equipo de recolección, siendo las personas supervisoras operativas las responsables de revisar dichas preguntas en base a lo establecido en el "Clasificador Chileno de Actividades Económicas, CIIU4.CL 2012". Para registrar las inconsistencias se utilizó el SGOE, pero ante la tardanza en su entrega, se tuvo que diseñar e implementar el Formulario de Supervisión de Gabinete en *SuSo*. Este último estuvo estructurado a partir de cinco secciones (Tabla 46) en las que se incorporaron respuestas abiertas y de selección como también la codificación de hallazgos (Tabla 47).

Tabla 46. Secciones del Formulario de Supervisión Indirecta a Empleadores y Dirigentes Sindicales/Trabajadores

FORMULARIO DE SUPERVISIÓN DE GABINETE	
Secciones del formulario	Contenido
1. Registro de contacto.	Fecha y hora.
	Rut persona supervisora operativa.
	Nombre y apellido persona supervisora operativa.
	Estado del cuestionario autoaplicado.
2. Revisión Autoaplicado	Revisión del cuestionario autoaplicado.
	Corrección del cuestionario autoaplicado.
3. No Elegible	CDT[32]de la empresa.
	Confirmación de CDT.
4. Intento de contacto cuestionario autoaplicado	Número de revisión.
	Intentos de contacto.
	Último intento de contacto.
5. Resultado de la encuesta	Estado de los cuestionarios.
	Estado de la encuesta.

Tabla 47. Clasificación de error en la Supervisión del Cuestionario Autoaplicado.

Tipo de error	
Inconsistencia interna	Cuando la información declarada en una pregunta y/o módulo se contradice con la información declarada en otra pregunta y/o módulo.
Campo en blanco	Cuando no hay información declarada en una pregunta y/o módulo.
Verifique información	Cuando la información declarada en una pregunta y/o módulo genera duda y/o inconsistencia al tomar en cuenta las características de la empresa.

Las asignaciones de los Formularios Supervisión Recuperación de Datos Autoaplicado se realizaron mediante *SuSo*, los que no estuvieron adscritos a ningún criterio de asignación en específico. Esto significó que se consideraron todas las

[32] CDT: código de disposición transitoria.

inconsistencias que enviaron las personas verificadoras de datos a las empresas. Aquí se comprometió supervisar el 20% de las empresas contactadas durante enero y febrero de 2024.

VI.6.2.4. Supervisión de No Elegibles.

La Supervisión de No Elegibles tuvo por objetivo corroborar el estado de no elegibilidad[33] de las empresas seleccionadas, a través de la revisión del Código de Disposición Transitorio (CDT) del Cuestionario de Empleadores y la revisión de la actividad económica en el Servicio de Impuestos Internos (SII). Del mismo modo, fue una instancia que buscó anexar documentación que respaldara el CDT de no elegibilidad, ya sea con un pantallazo de la página web del SII o de las observaciones del Cuestionario de Empleadores.

Para esta supervisión se consideraron todas las empresas a nivel nacional que se categorizaron como no elegibles, siendo las personas supervisoras operativas las responsables de verificar lo indicado por las personas recolectoras. Las inconsistencias fueron registradas en el SGOE, pero ante la tardanza en su entrega, se tuvo que utilizar el mismo Formulario de Supervisión de Gabinete (*SuSo*) descrito anteriormente (Tabla 46).

Al igual que la supervisión anterior, las asignaciones de los Formularios de Supervisión de Gabinete se realizaron mediante *SuSo*, los que dejaron de utilizarse una vez operativo el SGOE. Ya con este sistema en uso, no fue necesario realizar procesos de asignación, puesto que las personas supervisoras operativas trabajaron con las mismas empresas asignadas en los Formularios de Supervisión de Gabinete. Aquí no se establecieron criterios de asignación, comprometiéndose supervisar el 100% de las empresas categorizadas como no elegibles a nivel nacional.

VI.4.2.5. Supervisiones específicas de la Subdirección Técnica:

- **Supervisión a empresas con Registro Sindical Único (RSU) sin sindicato.**

Surge como parte del análisis que la SDT realizó a los datos entregados por la Dirección del Trabajo y el tipo de persona informante que entrevistó el equipo de recolección. Esto significó que, si el Cuestionario de Dirigentes Sindicales/Trabajadores traía en la portada un RSU, se debía entrevistar a una persona representante del sindicato, pero en algunas empresas aquello no ocurrió y se entrevistó el Cuestionario de Trabajadores. Ante tal inconsistencia, el objetivo de esta supervisión fue verificar la existencia de sindicato por medio de Supervisiones Indirectas para ratificar que el Cuestionario de Trabajadores estaba correctamente entrevistado, independiente si la empresa tenía asociado un RSU.

En la mayoría de estas supervisiones no hubo un proceso de asignación con un sistema de gestión, puesto que el número de casos inconsistentes se enviaron en planillas Excel. En este sentido, los casos supervisados siempre se distribuyeron equitativamente para no aumentar las cargas de trabajo de las personas supervisoras operativas. En el único caso donde sí hubo una asignación a través de SuSo fue en la "supervisión a empresas con RSU sin sindicato", ya que se abordaron

[33] En la Encla 2023 la no elegibilidad se refirió a que la empresa no se encontraba con actividad económica vigente, considerándose los dos últimos años anteriores al período de referencia de la encuesta.

como supervisiones indirectas. En este proceso no se comprometió un porcentaje de supervisión, pues se definió la gestión total de la nómina producida por el equipo técnico.

VI.6.2.6. Supervisiones específicas de la Subdirección de Operaciones

- **Supervisión a la recuperación de datos del Cuestionario Autoaplicado.**

Esta supervisión surgió como una solicitud debido a que las personas verificadoras de datos intervinieron un cuestionario que fue respondido exclusivamente por las personas informantes. Por tal motivo, el objetivo de esta supervisión fue revisar el flujo de comunicación y corrección de inconsistencias entre las personas verificadoras de datos y las personas informantes del Cuestionario Autoaplicado, mediante la revisión de los correos electrónicos enviados a las empresas y los cambios realizados en el SGOE.

Para esta supervisión se consideraron las empresas que fueron contactadas por las personas verificadoras de datos durante enero y febrero de 2024, siendo la persona encargada de área la responsable de revisar los correos electrónicos y las bases de datos del SGOE (informante y analista) para comprobar los cambios en los módulos que arrojaron inconsistencia. Para llevar a cabo esta supervisión, se generó una base de datos con las empresas a recuperar y se diseñó el Formulario de Supervisión Recuperación de Datos Autoaplicado, el cual fue implementado en *SuSo*.

VI.6.2.7. Falsificación de encuestas

En la Encla 2023 se entendió por falsificación a la desviación intencional y no informada, por parte de la persona recolectora, de las pautas o instrucciones que se establecieron para el flujo operativo, práctica que afectaría directamente la validez de los datos recolectados (INE, 2023, p. 4). Bajo esa lógica, se dio curso al flujo de posible falsificación desde el instante que la persona supervisora operativa identificó una inconsistencia en aquellas preguntas que impedían continuar con la Supervisión Indirecta[34] o cuando más del 50% de las preguntas de re-entrevista presentaron discrepancias de información. También se dio curso al flujo de posible falsificación cuando el equipo de recolección o técnico solicitó una supervisión dirigida.

Del momento en que se le informó al equipo de recolección sobre la activación de este protocolo (supervisión dirigida) y, por lo tanto, de la asignación del 100% de la carga de la persona recolectora en estado “entrevistada” y “aprobada por supervisor”, siempre se habló de posible falsificación dejando atrás esa figura cuando se terminó de gestionar el 100% de la carga y se envió el “Informe Final de Supervisión Dirigida”. Recién ahí se determinó si hubo o no una falsificación, tomándose en cuenta los criterios que determinó la Encla 2023:

- Aplicar intencionalmente la encuesta en un RUT distinto al indicado en el directorio.

[34] Las preguntas del Formulario de Supervisión Indirecta estuvieron categorizadas, es decir, algunas necesitaban clasificación y observación de hallazgos (mejora de información, error de registro u otro) activando el flujo de inconsistencias (por ejemplo, preguntas de reentrevista); mientras que otras finalizaban la supervisión activando el flujo de posible falsificación (por ejemplo, preguntas sobre la selección de la empresa, sindicato y/o persona informante).

- Aplicar intencionalmente la encuesta sin respetar el orden de prioridad de los cargos establecidos en cada cuestionario.
- Aplicar intencionalmente la encuesta a una persona cuyo cargo no está considerado en el orden de prioridad del cuestionario.
- Aplicar solo una parte del cuestionario a la persona informante (omisión intencional de preguntas) y completar las preguntas restantes en base a deducciones y/o criterio de la persona recolectora o en un momento distinto a la situación de entrevista.

El incurrir en alguna de estas faltas perjudicaría al estudio en los siguientes ámbitos:

- No reflejaría la realidad de las empresas, afectando la representatividad regional y nacional.
- Podrían tomarse medidas inadecuadas respecto a políticas públicas.
- Afectaría a la imagen y credibilidad institucional.

VII. Metodología de procesamiento

Los principales hitos del procesamiento de datos de la Encla 2023, se enfoca en las tareas llevadas adelante sobre la versión de los datos que incluyen solo a empresas logradas. El logro estuvo definido como el cumplimiento de los criterios de suficiencia técnica, acordados entre el equipo técnico del INE y el equipo del área de Estudios de la Dirección del Trabajo. En general, el trabajo de procesamiento estuvo orientado a generar un proceso de aseguramiento de la calidad de los datos obtenidos a través de todas las fases de ordenamiento, codificación, depuración, corrección, imputación y edición de datos.

VII.1. Fases del procesamiento

La estructura de los siguientes capítulos obedece a las etapas sucesivas del procesamiento que se aplicó a los datos de los cuestionarios pertenecientes a empresas logradas de la Encla 2023. Estas fases pueden resumirse en los siguientes elementos:

a. *Transformación de los RRAA en RREE (Fuente DT):* En esta etapa se procesan los registros administrativos descritos más adelante, a modo de hacerlos compatibles con la estructura de los datos de la Encla.

b. *Integración de datos*: los cuestionarios a empleadores y a sindicatos y trabajadores fueron unidos por filas en sus dos versiones de SuSo, filtrando las observaciones cuyo código de estado de la entrevista correspondiese a 100 (completada), 120 (aprobada por encargado de grupo) o 130 (aprobado por encargado de área). Después, los cuestionarios a sindicatos y trabajadores se dividieron, filtrando por la pregunta A01, que mide presencia o ausencia de sindicato y abre los flujos para responder cada cuestionario. Tras el procesamiento, y antes del etiquetado, ambos cuestionarios se unificaron nuevamente. En el caso del cuestionario autoaplicado, se contaba con dos versiones: una versión analista, que incluye los cambios ingresados por sistema por parte de las personas analistas en los procesos de supervisión y recuperación de información; y una versión bruta, que sólo incluye la información enviada por las personas informantes. Dado que se constató que 8 casos contaban con información duplicada en la versión analista por problemas informáticos en la descarga de los datos, esos casos fueron trabajados con la versión bruta de la base de datos. En todo el resto de los casos, se trabajó con la versión analista. En ambas versiones se trabajó sólo con aquellos cuestionarios con estado completa y enviada. Este proceso también consideró la medición del logro de la grilla de suficiencia definida para cada cuestionario, por lo que sirvió de insumo para la recuperación de información.

c. *Codificación de glosas "otros".* En esta etapa se codifican las glosas entregadas a partir de las preguntas abiertas de categorías "otros". Se evaluó si las respuestas ahí proporcionadas ameritan la creación de nuevas categorías, forzar su codificación como una ya existente, o bien eliminar la glosa en caso de que no guarde relación alguna con la pregunta. Esta tarea se realizó para los cuatro cuestionarios.

d. *Corrección de secuencias y no respuesta.* En esta etapa se corrigen los problemas de secuenciaque

ocurrieron en el formulario autoaplicado. Este formulario, como su nombre lo dice, fuerespondido por uno o más informantes de la empresa de manera autónoma (con el apoyo de funcionarios INE del área de recolección en caso de ser solicitado).

e. *Cuadratura de matrices.* Esta etapa se ocupa de asegurar consistencia en las preguntas que recaban información numérica tipo matriz en el Autoaplicado. En ese sentido, se busca corregir valores según reglas de "cuadratura" dentro de las matrices, así como también entrematrices relacionadas.

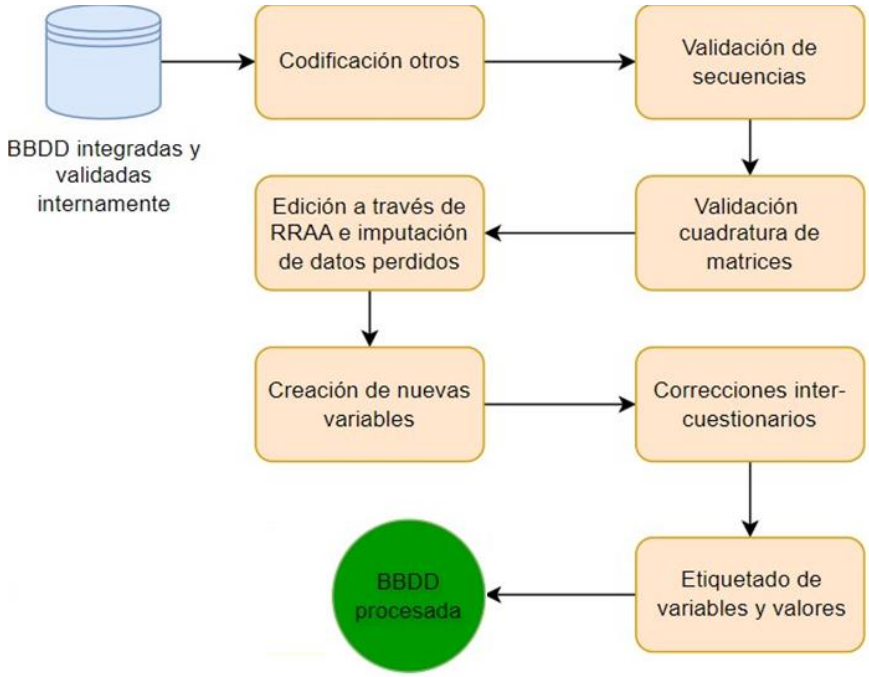
f. *Edición e imputación de datos*. Luego de las correcciones realizadas a partir de las fases b y cdescritas arriba, en el caso de no contar con información del formulario Autoaplicado en ciertas variables numéricas, estas se imputan o se editan según una serie de condiciones queson descritas en la sección correspondiente. Este paso incluye la edición e imputación a partir de registros administrativos.

g. *Creación de variables analíticas*. Se generan las variables provenientes de los registros de huelgas facilitadas por la Dirección del Trabajo.

h. *Validaciones Inter-cuestionario*. En esta etapa se validan y corrigen preguntas relacionadas entre el formulario Autoaplicado y el de Empleadores, con el fin de asegurar consistencia en las respuestas entregadas en una misma empresa.

i. *Etiquetado de datos*. Se realiza el etiquetado de variables y de valores en las distintas versiones de bases de datos generadas.

Figura 8. Flujo de procesamiento de bases integradas de la Encla



VII.2. Codificación de glosas

VII.2.1. Codificación de glosas “Otros”

Esta etapa del procesamiento consistió en la codificación de las respuestas dadas por las personas informantes a aquellas preguntas que incluyen la posibilidad de especificar otros motivos, razones, situaciones, etc., no consideradas en las categorías originales de cada pregunta[35]. Las y los informantes declararon glosas al indicar el valor 77 (“Otro(a)”) en cada una de las preguntas, cuyo resumen puede verse en la siguiente tabla:

Tabla 48. Detalle de preguntas codificadas y editadas por cuestionario

Cuestionario	Número de preguntas codificadas	Detalle de preguntas editadas
Autoaplicado	21	D7, D9, D10, G3, H6, H7, 8, I4, O6, O7, O8, O9, O11, O13, O15, O16, O19, P1, P5
Empleadores	9	F1, F2, F3, I3, J5, L2, L3, M3, N2
Sindicatos	13	J13, K6, K8, K9, K10, K16, K17, L2, L3, M3, M17, O4
Trabajadores	10	J11, K2, K4, K5, L2, L3, M3, M8, O3, O4

La labor de codificación constó de una etapa manual y una etapa automática. La primera comenzó con la asignación de un código numérico a las glosas capturadas en cada uno de los cuatro cuestionarios por parte del equipo técnico del INE. El trabajo puede resultar en cuatro situaciones:

a) Asignar a la glosa un código perteneciente a alguna de las categorías ya existentes de la pregunta.

b) Asignar a la glosa una nueva categoría no considerada en las respuestas del cuestionario, para las cuales se asignan valores numéricos correlativos a la cantidad de categorías originales. Para crear nuevas categorías, al menos cinco glosas deben se asignadas al mismo código.

c) Asignar el valor 77, que denota la categoría “Otro(a)”, a aquellas glosas que, aun siendo una respuesta válida a la pregunta formulada, no son categorizables en los códigos originales ni en los nuevos creados.

d) Eliminar la glosa, cuando se ha determinado que la respuesta del informante no guarda relación con el objetivo de la pregunta. Estas glosas fueron descartadas del proceso de edición.

Mientras que en las preguntas de selección única solo se asignó un código a cada glosa, en las preguntas de selección múltiple y las grillas se admitió la asignación de más de un código a cada glosa, considerando las características de cada tipo de pregunta. Semanalmente, el equipo técnico del INE envió a la DT una planilla con el detalle de los códigos

[35] Por ejemplo, se solicita que se especifiquen elementos como otras actividades subcontratadas, otros criterios para aumentar remuneraciones, otros daños o problemas de salud que se hayan presentado, otras razones por las cuales no hay sindicato en la empresa, otras razones por las cuales no se ha negociado colectivamente, entre otros.

asignados a cada glosa de cada pregunta, así como con propuestas de nuevas categorías. Luego, la DT revisaba ese insumo y volvía a enviar esta planilla revisada, haciendo comentarios y correcciones a la asignación de códigos y a la creación de nuevas categorías. Con base en ese insumo proporcionado por la DT, el equipo técnico del INE asignó los códigos definitivos a cada glosa y consolidó las nuevas categorías finales. Con ello se inició la asignación automática de los códigos asignados a cada unidad. Para ello, se combinaron los microdatos de cada cuestionario con la planilla de codificación manual, utilizando las glosas como llave. Posteriormente, se editaron las respuestas de los informantes para que estas reflejen lo declarado en las preguntas abiertas.

El algoritmo de edición es diferente para, por un lado, las preguntas de selección única (ej. pregunta I3 del cuestionario a empleadores) y, por otra, las preguntas de selección múltiple (ej. pregunta F3 del cuestionario a empleadores) y las grillas (ej. pregunta F1 del cuestionario a empleadores).

En el primer caso existen dos posibilidades. Si la glosa fue codificada como una categoría original distinta a 77 o como una nueva categoría, se transfirió el valor de 77 a la categoría correspondiente. Por otra parte, si la glosa fue codificada como 77, se mantuvo el valor original.

En el segundo caso, si la glosa fue codificada como alguna categoría original, se modificó el valor de la respectiva variable a 1 (presencia del ítem). Si la glosa fue codificada como 77 o alguna categoría nueva, se mantuvo el valor en la variable que refleja la categoría 77, y se crearon nuevas variables que detallan el o los "Otros(as)" señalados en la glosa. Estas variables siguen la nomenclatura XX_otro_cY, donde XX refiere a la pregunta en cuestión, e Y a un valor entre 1 a 3, que especifica el número máximo de "otros" distintos en las que fue codificada una glosa para cada pregunta.

Tómese por ejemplo la siguiente glosa declarada en la pregunta g3 del cuestionario autoaplicado:

Tabla 49. Ejemplo procedimiento de creación de nuevas variables

Glosa declarada en g3_otro	Variables generadas en base de datos			
	g3_77	g3_otro_c1	g3_otro_c2	g3_otro_c3
Agresión de terceros, asaltos, atropello, choque, colisión, golpes, mordedura de perros, torceduras	1	9	10	11
Caídas mismo nivel, golpes con balones o pelotas deporte, corte por manipulación basura, Covid	1	8	10	NA

En el primer caso, el texto capturado se ha codificado en las nuevas categorías *9. Accidentes de tránsito en el trabajo, 10. Lesiones traumáticas* y 11*. Picaduras, mordeduras y/o ataques de animales.* En consecuencia, el valor de g3_77 sigue siendo 1, en tanto se declararon problemas de salud adicionales a los consultados originalmente; y los valores de g3_otro_c1, g3_otro_c2 y g3_otro_c3 permiten especificar cuáles fueron esos otros daños o problemas de salud experimentados por los trabajadores de aquella empresa en el período consultado. Asimismo, dado que la segunda glosa solo se puede codificar en dos nuevas categorías *–8. Covid-19* y *10. Lesiones traumáticas–*, el valor de g3_otro_c3 se mantiene vacío. Lo mismo aplica para g3_otro_c2 en caso de que alguna glosa solo pueda asignarse a una nueva categoría.

Vale destacar que los valores asignados a las variables XX_otro_c1, XX_otro_c2 y XX_otro_c3 se asignan en orden ascendente. Así, en el ejemplo anterior, el valor más bajo es asignado a g3_otro_c1 (9 y 8, respectivamente); y el más alto, a g3_otro_c3 (11) y g3_otro_c2 (10), respectivamente.

Para las variables C1, C5, H7 e I4 del cuestionario autoaplicado, dadas sus características, se ha desarrollado un algoritmo de edición ad-hoc. En concreto:

- C1: esta pregunta distingue la subcontratación de actividades desarrolladas dentro (c1_1_X) y fuera (c1_2_X) de la empresa. Sin embargo, la variable que captura las glosas es la misma para ambas. Por ello, se ha utilizado el

valor de c1_1_77 y c1_2_77 para definir si la glosa refiere a actividades dentro o fuera de la empresa, y editar los valores de c1_1_X, c1_2_X en consecuencia.

- C5: en el caso de esta variable numérica, las glosas solo se codificaron con los valores 3. Contrato escriturado civil o comercial, 77. Otro tipo de contrato, y 96. Sin dato. Cuando la glosa fue codificada como 3, el valor de c5_3 corresponde a la suma del valor original de c5_3 y el valor de c5_77, siendo este último llevado a 0. Cuando el valor corresponde a 77 o 96, se mantiene el valor en c5_77, a la vez que c5_otro_c1 detalla si ese otro tipo de contrato era codficable en algún otro tipo de contrato (77), o si la información proveída no fue suficiente para distinguir el tipo de contrato realizado (96).

- H7: esta pregunta involucra una matriz, y fue trabajada verticalmente. Si el valor de la glosa era codificado en 1. Sala cuna de la empresa, 2. Sala cuna externa o 3. Bono, se sumó al valor originalmente declarado el valor declarado en la fila 77. Otra, el cual fue posteriormente llevado a 0. Por su parte, si la glosa era codificable en los valores 4. Cuidados de un externo o tercero o 77. Otra, se mantuvo el valor de h7_1_77, h7_2_77 o h7_3_77, respectivamente, y se detalló en h7_1_otro_c1, h7_2_otro_c1 o h7_3_otro_c1 el valor 4 o 77, según corresponde.

- I4: En este caso, la asignación de los códigos “Otros” se hizo desde i4_1 a i4_3. Así

    - Si i4_1 era igual a 77. Otra; i4_2 e i4_3 están vacíos o son iguales a 6. Ninguna otra y son distintos de 77. Otra; y la glosa fue codificada en un valor válido, se asigna el código de la glosa a i4_1. De lo contrario, se mantiene su valor original.

    - Si i4_2 era igual a 77; i4_1 tiene información válida distinta de 77. Otra, 6. Ninguna otra y del código asignado a la glosa; e i4_3 está vacía o es igual a 6. Ninguna otra, se asigna el código de la glosa a i4_2. De lo contrario, se mantiene su valor original.

    - Si i4_3 era igual a 77. Otra, y tanto i4_1 como i4_2 son distintas a 77. Otra, 6. Ninguna otra y al código asignado a la glosa, se asigna el código de la glosa a i4_3. Si no, se mantiene su valor original.

Por último, en la variable F2 del cuestionario a empleadores, se infirió el valor específico (que refleja si las medidas informadas se encuentran vigentes o si se implementarán en el corto plazo) a partir de las glosas informadas.

VII.2.2. Codificación de Actividad Económica

La clasificación de actividad económica de las unidades seleccionadas es una de las variables propias del diseño muestral del estudio. El trabajo constó de dos procedimientos: uno de codificación manual, y otro de codificación automática de las glosas declaradas por los informantes en la pregunta A2 del cuestionario autoaplicado, siguiendo la clasificación CIIU4.CL 2012 a nivel de Sección (un dígito).

El trabajo de codificación manual de la actividad económica fue ejecutado por tres integrantes del equipo técnico del INE. Semanalmente, se asignó a cada integrante entre 50 y 150 unidades por codificar, según el avance en la recolección del cuestionario autoaplicado. El trabajo de codificación consistió en asignar la actividad económica de la unidad a nivel de sección según CIIU4.CL 2012, así como una a tres palabras claves que permitieron a los codificadores identificar la sección a la que corresponde la unidad (que sirvieron de insumo para el modelo de codificación automática),

identificando las glosas problemáticas. Luego, una vez por semana, el equipo se reunió para discutir respecto de aquellas glosas cuya información no fue lo suficientemente clara como para asignarle una Sección de forma unívoca. Esta operación se llevó a cabo hasta lograr 1883 unidades codificadas manualmente.

El 29 de septiembre, el equipo técnico envió mil glosas codificadas a la Sección de Nomenclaturas perteneciente al Subdepto. de Calidad y Estándares del INE, que se encargó de auditar las codificaciones realizadas por el primero. Este proceso permitió definir la calidad de las codificaciones manuales por parte de expertos en el área, identificando tres estados:

a. Conforme (72,9% de las glosas enviadas): la codificación se acepta sin observaciones. Para ello, la información declarada en A2 ha de ser clara, precisa y consistente; y, además, debe existir consistencia entre el código de actividad económica asignado por el equipo técnico de la Encla y el analista de codificación de la Sección de Nomenclaturas.

b. Conforme con observaciones (6,7% de las glosas enviadas): dado que la información entregada por el informante en A2 no es suficiente para la codificación, se utiliza información complementaria (como la razón social, información obtenida de la página web, la actividad secundaria declarada, entre otros), que permitió codificar la actividad económica de la unidad en coincidencia con lo definido por el equipo técnico de la Encla.

c. No conforme (20,4% de las glosas enviadas): lo declarado por el informante en A2 no fue suficiente para codificar la actividad económica de la unidad, o la codificación realizada por el equipo técnico de la Encla no coincide con lo definido por el analista de codificación de la Sección de Nomenclaturas.

Una vez recibidos los resultados de la auditoría, el equipo técnico de la Encla revisó una por una las codificaciones de las 271 unidades categorizadas como "Conforme con observaciones" y "No conforme". Para ello, se repitió la labor de codificación manual de cada caso, considerando los comentarios (que, en la mayoría de los casos, indicaba que la glosa permitía clasificar a la unidad en más de una actividad económica, y en cuáles) y la nueva glosa provista por la Sección de Nomenclaturas. Esta información fue triangulada, en la mayoría de los casos, con la información presente en las páginas web de las empresas codificadas de forma insuficiente, a modo de asegurar una codificación final satisfactoria. Luego de este proceso individualizado de revisión y corrección, el equipo técnico definió mantener la totalidad de las codificaciones auditadas originalmente en estado "Conforme con observaciones", y el 43.13% (88/204) de aquellas con estado "No conforme".

Por su parte, la codificación automática de la actividad económica se realizó a partir de un modelo de regresión logística multinomial con regularización de red elástica, utilizando el paquete parsnip de R. A través de técnicas *bagging* de aprendizaje automático supervisado, se ensamblan una serie de modelos que buscan predecir la actividad económica de las unidades con base en diversas covariables, cuyos resultados se promedian para obtener la predicción final. Los datos se dividieron en muestra de entrenamiento y muestra de *test* en la razón 5:1. El modelo descrito ha sido desarrollado por el equipo técnico de la Encuesta Estructural de Remuneraciones y Costos Laborales (EERCL) para los mismos fines, obteniendo resultados similares a la codificación manual respecto de la precisión de las predicciones. Se debe destacar que el modelo solo fue estimado con cuestionarios autoaplicados que se han catalogado como logrados con base en la grilla de suficiencia (n = 4.631[36]).

[36] Dado que la razón de la construcción de las muestras de entrenamiento y test es de 1:5, el tamaño de la muestra de entrenamiento es igual a 927 observaciones.

La variable respuesta del modelo se construyó a partir de la codificación manual (considerando las correcciones posteriores a la auditoría de la Sección de Nomenclaturas) y el Directorio Nacional de Empresas (DNE). Específicamente, se asignó la sección codificada manualmente a los 1.883 casos que fueron sometidos a esta labor, y se imputó la sección del resto de casos a partir de la información proveniente del directorio. Por ello, el modelo siempre predecirá una sección para cada unidad, independiente del nivel de ajuste de tal predicción.

Por otra parte, el modelo incluye un total de 48 predictores. En específico:

1. Tamaño de la unidad según el DNE.
2. Actividad económica de la unidad según el DNE.
3. Actividad económica a un dígito, obtenida de la API del Servicio de Impuestos Internos (SII).
4. Actividad económica predicha a uno y dos dígitos por el modelo Support Vector Machine (SVM) utilizado para estos efectos en la Encuesta Nacional de Empleo (ENE), recodificada a un dígito.
5. Actividad económica predicha por el modelo de la API[37] de Servicios Compartidos del INE, a un dígito.
6. Número de palabras clave con las que coincide la glosa informada en A2, por cada una de las 13 secciones económicas según el DNE.
7. Número de palabras clave con las que coincide la razón social de la unidad, por cada una de las 13 secciones económicas según el DNE.
8. Número de palabras clave con las que coincide el nombre de fantasía de la unidad, por cada una de las 13 secciones económicas según el directorio.
9. Actividad económica con mayor número de coincidencias entre las palabras clave de cada sección y la glosa informada en A2.
10. Actividad económica con mayor número de coincidencias entre las palabras clave de cada sección y la razón social de las unidades.
11. Actividad económica con mayor número de coincidencias entre las palabras clave de cada sección y el nombre de fantasía de las unidades.

De esta manera, se aprovecha la información recolectada de fuentes secundarias, como el SII, así como el trabajo de codificación manual para dar un peso central a la glosa informada en A2 en la codificación automática. Siguiendo las indicaciones provistas por la Sección de Nomenclaturas en el informe de su auditoría, el modelo no incluye la actividad económica autoreportada en A1 como predictor. Esto se debe principalmente al error de medición asociado a esa forma de capturar la actividad económica de las unidades. Concretamente, forzar una respuesta al informante a partir de las secciones de una clasificación especializada que no conocen puede generar confusiones al declarar su actividad

37 Interfaz de programación de aplicaciones (Application Programming Interface en inglés).

económica, como puede ser el codificarse en comercio cuando la unidad es una panadería, que corresponde a la sección de industrias manufactureras.

Una vez cerrado el proceso de recuperación de información de las unidades, el equipo técnico de la Encla llevó a cabo una auditoría interna de las codificaciones definidas por el modelo anteriormente descrito. En concreto, se revisó la sección en que se codificaron aquellos casos donde a) la codificación del modelo difiere de la sección con mayor recurrencia entre la sección de la variable dependiente (que incluye codificaciones manuales y secciones provenientes del directorio), la sección definida por la API de Servicios Compartidos del INE, la sección predicha por SVM a uno y dos dígitos, la sección obtenida desde el SII, la sección definida a partir del mayor número de coincidencias de la glosa capturada en A2 con el diccionario de palabras clave para cada sección, y la sección definida por el modelo; o b) la sección predicha por el modelo difiere de la sección codificada manualmente u obtenida desde el directorio, según corresponda. Así, se auditaron 814 casos, es decir, un 18.1% de la muestra lograda. La labor de auditoría siguió el método de la codificación manual detallado previamente. La siguiente tabla desglosa, para cada sección según la variable predicha, el total de unidades auditadas, así como el número y porcentaje de unidades donde la sección de la variable predicha y la sección del modelo difieren de la sección producto de la auditoría interna.

Tabla 50. Resumen por actividad económica de diferencias entre auditoría interna y codificación original

Sección original	Número de unidades	Diferencia entre Auditoría interna y Sección original (n y %)	Diferencia entre Auditoría interna y Sección modelo (n y %)
A	44	32 (72.7%)	15 (34.1%)
B	29	14 (48.3%)	16 (55.2%)
C	100	45 (45%)	54 (54%)
D-E	27	17 (65.4%)	18 (69.2%)
F	75	39 (52%)	48 (64%)
G	103	68 (66%)	69 (67%)
H-J	64	35 (54.7%)	44 (68.8%)
I	37	12 (32.4%)	23 (62.2%)
K-L	69	54 (78.3%)	46 (66.7%)
M-N	110	62 (56.4%)	67 (60.9%)
P	15	4 (26.7%)	14 (93.3%)
Q	21	7 (33.3%)	13 (61.9%)
R-S	120	85 (70.8%)	78 (65%)
Total	814	474 (58.3%)	505 (62.1%)

Siguiendo los datos es posible indicar que, en estos casos identificados como problemáticos, tanto la sección de la variable predicha (que incluye codificaciones manuales y datos del directorio) como la sección predicha por el modelo tienden a diferir de forma similar con la sección definida manualmente en el proceso de auditoría interna. Exceptuando las secciones A, K-L y R-S, las diferencias entre los resultados de la auditoría interna y la sección predicha por el modelo tienden a ser mayores que las diferencias entre aquella y la sección de la variable predicha. Esta labor permitió corregir los casos más inconsistentes predichos por el modelo.

Luego de la auditoría interna, se llevó a cabo una re-estimación del modelo, imputando los resultados de tal labor en la variable predicha. Así, un total de 2.420 (52.25% del total de unidades logradas) fueron codificadas manualmente, consignando ese valor en la variable predicha.

Para medir el rendimiento del modelo en su capacidad de predecir la actividad económica de las unidades a nivel de sección, se ha estimado su precisión (verdaderos positivos sobre verdaderos y falsos positivos) y *recall* (verdaderos positivos sobre verdaderos positivos y falsos negativos) en comparación a la API de Servicios Compartidos del INE, y el modelo SVM a un dígito. Se asume que el valor verdadero corresponde a la variable respuesta del modelo, que incluye la sección identificada a partir de la codificación manual, y la sección identificada en el directorio para aquellos casos que no fueron codificados manualmente.

Así, una precisión de 1 indica que el 100% de las unidades codificadas en determinada sección corresponden a ella en su valor verdadero, sin falsos positivos en la predicción; y un *recall* de 1 indica que el 100% de las unidades codificadas en determinada sección correspondan a ella en su valor verdadero, sin falsos negativos en la predicción. En conjunto, ambas medidas permiten conocer qué tan bien se clasifica, y en qué proporción el modelo se equivoca en clasificar respecto del valor real. Como se presenta en las figuras 9 y 10, el modelo multinomial estimado presenta medidas de precisión y *recall* iguales o mejores que el clasificador de la API de Servicios Compartidos en todas las secciones de actividad económica.

Figura 9. Comparación de la precisión estimada para el modelo multinomial y el clasificador de la API de Servicios Compartidos (n = 927)

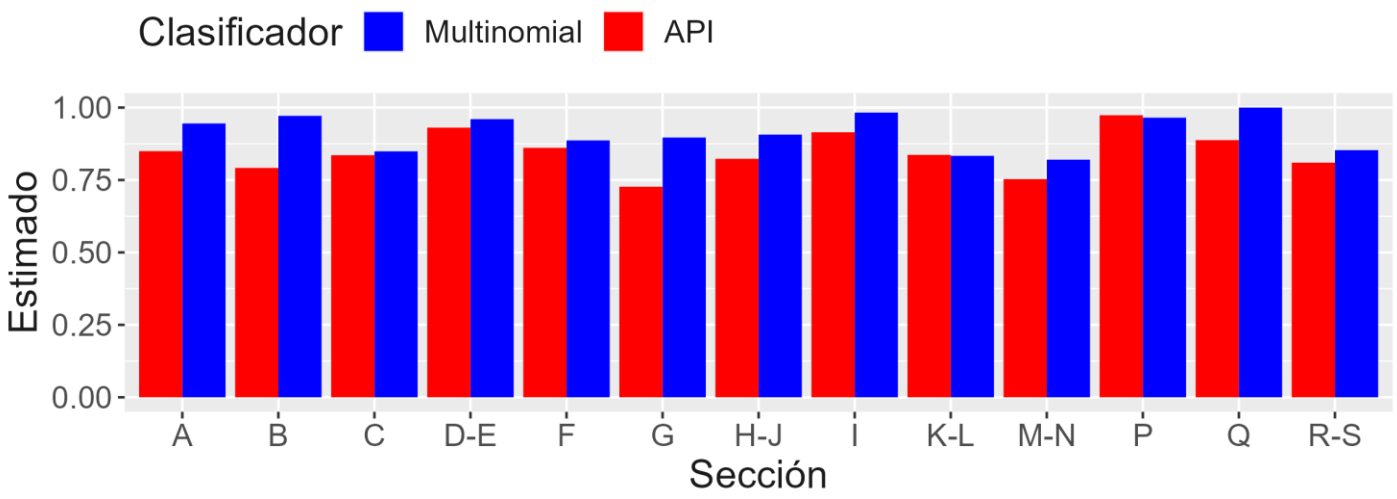
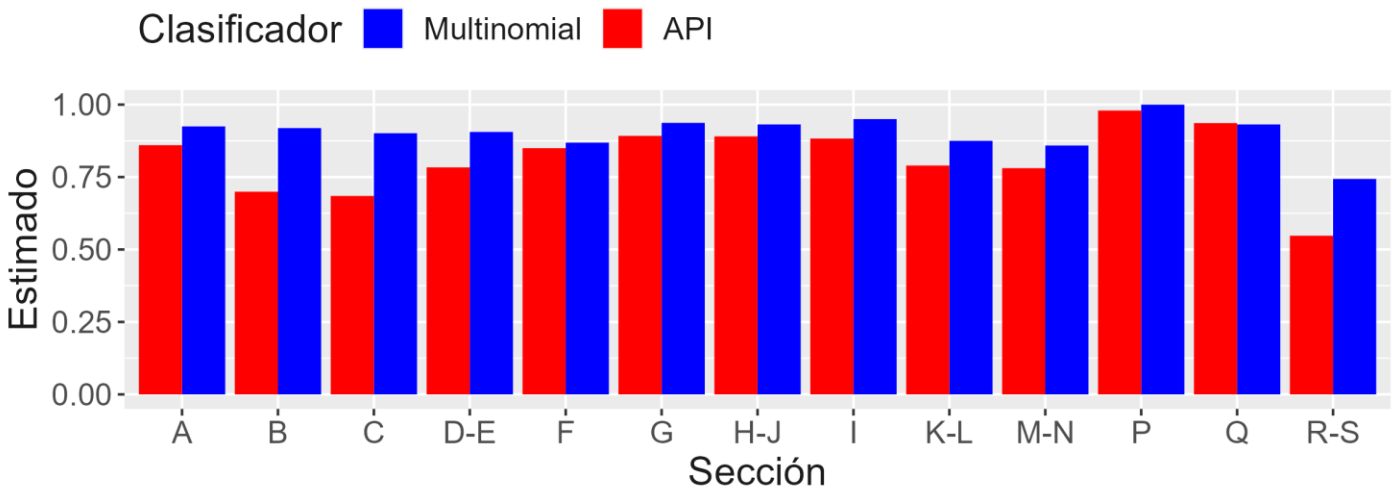


Figura 10. Comparación del *Recall* estimado para el modelo multinomial y el clasificador de la API de Servicios Compartidos (n = 927)



VII.3. Correcciones de Secuencia

Las validaciones y correcciones de secuencia se realizaron para corregir situaciones donde los datos presentaron alguna situación de inconsistencia, ya sea por falta de información o por saltos mal implementados. Cabe destacar, que estos problemas se presentaron mayormente en el formulario Autoaplicado. Esto se explica por los siguientes motivos:

- Los formularios a Empleadores y a Trabajadores/Sindicatos fueron generados en la plataforma *Survey Solutions*; lo que le permitió al equipo técnico la programación de validadores que podían ser probados de manera instantánea durante su escritura.

- Asimismo, estos cuestionarios fueron aplicados, ya sea de manera telefónica o presencial, por recolectores/as de datos capacitados en el marco del proyecto; que se aseguraban de que la entrevista fuera realizada de manera completa.

- El formulario Autoaplicado, por el contrario, fue programado pensando en dar fluidez al proceso de respuesta por parte del o los informantes; de esta forma, hubo dos tipos de validaciones programadas en la plataforma: errores y advertencia. Los primeros estaban asociados principalmente a variables de suficiencia técnica y llenado de información clave y su resolución era condición necesaria para que la persona informante pudiese completar el formulario y cerrar el proceso. Las advertencias, por otro lado, mostraban un mensaje de alerta, generalmente sobre inconsistencia de información, pero su resolución no era obligatoria para que la persona informante pudiese enviar de vuelta el formulario y cerrar el proceso de llenado.

- Junto con lo anterior, el formulario Autoaplicado contiene muchas preguntas numéricas dispuestas en matrices; lo cual presenta el desafío de lograr una cuadratura en la suma de filas, columnas y totales.

- Adicionalmente, durante las primeras semanas de aplicación de la encuesta, el sistema informático no eliminaba información asociada a un cambio de flujo.[38]

- Por último, el formulario Autoaplicado es un cuestionario compuesto de 12 secciones y 1.032 variables; muchas de ellas de tipo numérico y que tienen, en algunos casos, relaciones lógicas y/o conceptuales que no estuvieron necesariamente contempladas al momento de la programación del cuestionario.

El listado de validaciones aplicadas en cada una de las bases de datos mostró que no se detectan errores de No Respuesta, Secuencia o de Valor Admisible en las bases de Empleadores, Trabajadores, ni en la de Sindicatos.

Las validaciones utilizadas para identificar la información de los datos del Autoaplicado que presentan problemas de secuencia, son fundamentalmente tres; y se aplican en casos donde hay un salto de pregunta involucrado:

· Pregunta de origen habilita variables y hay al menos 1 NA en cualquier variable habilitada.

· Pregunta de origen inhabilita variables y hay al menos 1 valor distinto de NA en cualquier variable inhabilitada.

· Pregunta de origen se encuentra vacía.

· Otras reglas ad-hoc a la pregunta.

Es importante notar que las modificaciones realizadas en el marco de las correcciones de secuencia hacen referencia a variables que luego podrían ser modificadas en el proceso de anonimización (ya sea recodificándolas o eliminándolas).

VII.4. Validaciones Inter-cuestionarios

Las validaciones inter-cuestionarios refieren fundamentalmente a preguntas de un cuestionario de la encuesta que puede ser validado por la respuesta en una pregunta de otro cuestionario que son idénticas. Luego de haber analizado los cuatro cuestionarios en función de lo anterior, tenemos el siguiente número de preguntas que se pueden validar a partir de otros cuestionarios, por cuestionario:

Tabla 51. resumen de preguntas a validar por cuestionario

Cuestionario	Autoaplicado	Empleadores	Sindicatos	Trabajadores
Número de preguntas	15	20	25	17

No obstante, solo se realizaron ediciones entre los cuestionarios auto aplicado y empleadores/as en virtud de la admisibilidad de diferencias para los casos de trabajadores/as y sindicatos y por tratarse de personas informantes diferentes. Por esta razón, solo se expondrá el detalle de variables editables de los cuestionarios mencionados que efectivamente se editaron (empleadores y autoaplicado).

[38] Por ejemplo, si al responder la alternativa 1 en la pregunta A se bloqueaba la pregunta B, la persona informante podía responder ambas si habilitaba B y luego se devolvía a A para deshabilitarla. Esta acción no generaba el borrado de información en B.

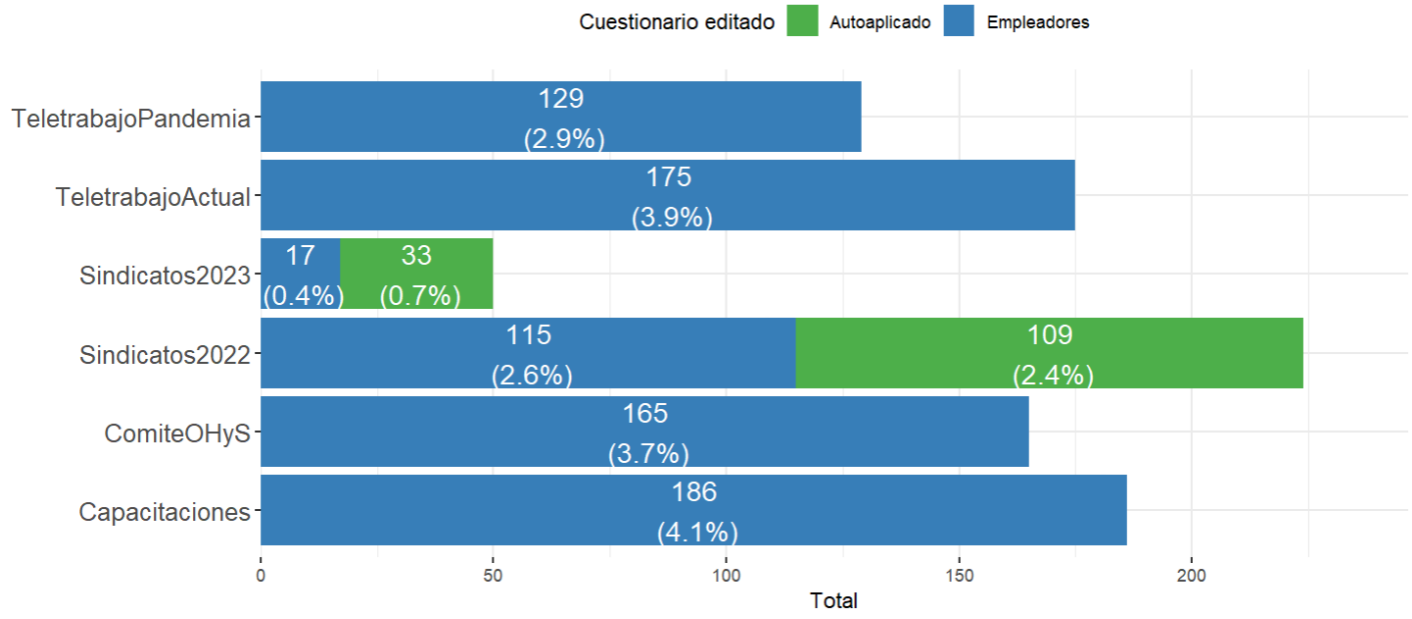
En términos concretos, se editaron 6 pares de variables, realizando ediciones cuando la proporción de cuestionarios editados es igual o menor al 5% del tamaño muestral (225 casos). Por otra parte, las preguntas cuyo flujo se "abrió" a partir de una edición (por ejemplo, la grilla de G4 en empleadores) fueron imputadas con 96 en caso de ser categóricas y –96 en caso de ser numéricas. La tabla a continuación resume los seis pares de variables involucradas.

Tabla 52. Pares de variables de cuestionarios a empleadores y autoaplicado validadas y editadas

Concepto	Empleadores	Autoaplicado
Teletrabajo en pandemia	F1_01	O17_2
Teletrabajo actualmente	F2_01	O1
Presencia Comité Paritario	G3	G4
Capacitaciones en 2022	I2	I2
Sindicatos en 2022	J8	K1
Sindicatos actualmente	J9	K2

El resultado de las ediciones realizadas y el cuestionario en el que ocurrieron se pueden visualizar en el siguiente gráfico:

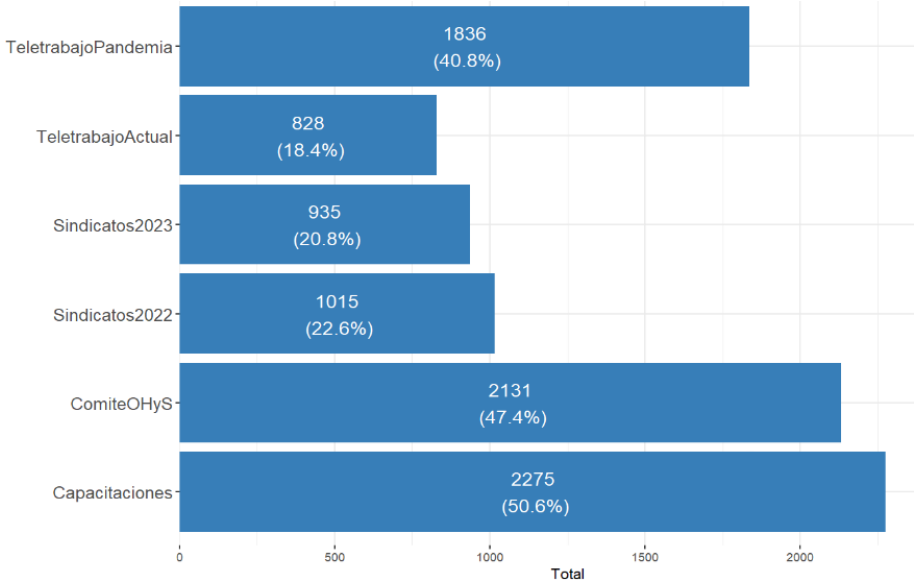
Figura 11. Ediciones por concepto en cuestionarios a empleadores y autoaplicado



Así, las dos variables con mayor cantidad de ediciones inter-cuestionarios corresponden a la existencia de sindicato durante el 2022 y a la realización de capacitaciones.

Como consecuencia de las ediciones, los resultados en las preguntas involucradas sufren algunos cambios, que se ven reflejados en el gráfico desplegado a continuación:

Figura 12. Unidades con cada atributo después de las ediciones inter-cuestionarios



VII.5. Cuadratura de matrices

En el contexto del formulario autoaplicado de la Encla 2023, una matriz es una fuente de información numérica que contiene al menos una suma de elementos. Un caso particular de información inconsistente es la existencia de matrices “no cuadradas” (cuyas sumas totales o parciales no coinciden totalmente dentro de una matriz, o bien no es concordante con otra matriz relacionada del cuestionario).

La razón por la que es necesario realizar correcciones sobre matrices se debe a dos motivos principales:

a) Comportamiento del informante al omitir llenado de información en casos donde no existen datos para reportar: algunos informantes prefirieron omitir llenar con un cero en celdas donde no existen registros por reportar, por lo tanto, la celda quedó vacía. Por ejemplo, cuando en A3 se pide el detalle de trabajadores por región/sexo. Para una empresa que tiene la totalidad de sus trabajadores localizados en la Región Metropolitana, todas las regiones, excepto la RM, deberían venir con 0. Sin embargo, el llenado de todas las celdas de dicha matriz con ceros podría resultar tedioso para un informante. En estos casos, el sistema de ingreso de información del Autoaplicado permitió solo llenar celdas "relevantes" y totales, controlando, a través de advertencias, que la suma de los ítems ingresados fuera consistente con los totales de filas/columnas y total general. Por lo tanto, en este caso, es necesario reemplazar los NA por ceros.

En este sentido, el primer proceso de imputación que se realiza a las matrices del formulario Autoaplicado es el siguiente:

- Si un total de matriz es 0 y todos los componentes de dicho total son NA, entonces se llevan todos los componentes de ese total a 0.

- Si un total de matriz es NA y todos los componentes de dicho total son NA y existe respuesta distinta de NA en otro ítem de la matriz (esto último para no incluir casos donde podría haber omisión completa de información), entonces se llevan a cero el total y los componentes de dicho total.

Estas ediciones tienen el objetivo de imputar ceros en las celdas de las matrices donde este valor tiene sentido lógico.

b) Inconsistencia de información: Para permitir un flujo eficiente en el llenado y envío de la información por parte de los informantes, las validaciones obligatorias (que impedían el envío del formulario completo frente a omisión o inconsistencia de información) se limitaron a las matrices que contenían variables pertenecientes a la Grilla de Suficiencia. En todo el resto de los casos, la omisión y/o inconsistencia en la información reportada se presentaba al informante en forma de "advertencias", cuya resolución no fue obligatoria para el envío de información. En este caso, es necesario corregir y cuadrar matrices interna y externamente inconsistentes.

Antes de detallar el procesamiento que corrige estas inconsistencias, se hace la siguiente definición. Las matrices serán clasificadas en matrices tipo 1 y tipo 2:

- Tipo 1: Matrices con datos individuales, totales parciales y total general:

Figura 13. Disposición de datos numéricos en matrices Tipo 1

Datos individuales (DI)	Totales parciales (TP) de fila
Totales parciales (TP) de columnas	Total general (TG)

Un ejemplo de matriz tipo 1 puede observarse en la matriz B2; compuesta por 7 filas y 6 columnas, como se observa en la figura a continuación:

Figura 14. Ejemplo de matriz Tipo 1

B2 **Indique el total de personas contratadas directamente por esta empresa al 30 de noviembre de 2022, según tipo de contrato y tramo etario.**

Tipo de contrato		Menores de 18 años	De 18 a 24 años	De 25 a 29 años	De 30 a 59 años	De 60 años y más	Total
01	Indefinido						
02	Plazo fijo						
03	Por obra o faena						
04	De aprendizaje						
05	Honorarios						
06	Sin escrituración (contrato "de palabra")						
07	**Total**						

Las siguientes matrices del formulario Autoaplicado son matrices tipo 1: A3, A4, B1, B2, B3, B5, B6, B8, D1, D4, E2, E3, E4, E5, F1, F8, G5, H1, H9 y O3.

- Tipo 2: Matrices con datos individuales y una clase de total (sólo filas o sólo columnas)

Figura 15. Disposición de datos numéricos en matrices Tipo 2

Ejemplo de matrices tipo 2 son las preguntas C2 y C3 del formulario Autoaplicado. Ambas se componen de solo una fila (aunque como se ve en la figura anterior, estas pueden también estar dispuestas en columnas):

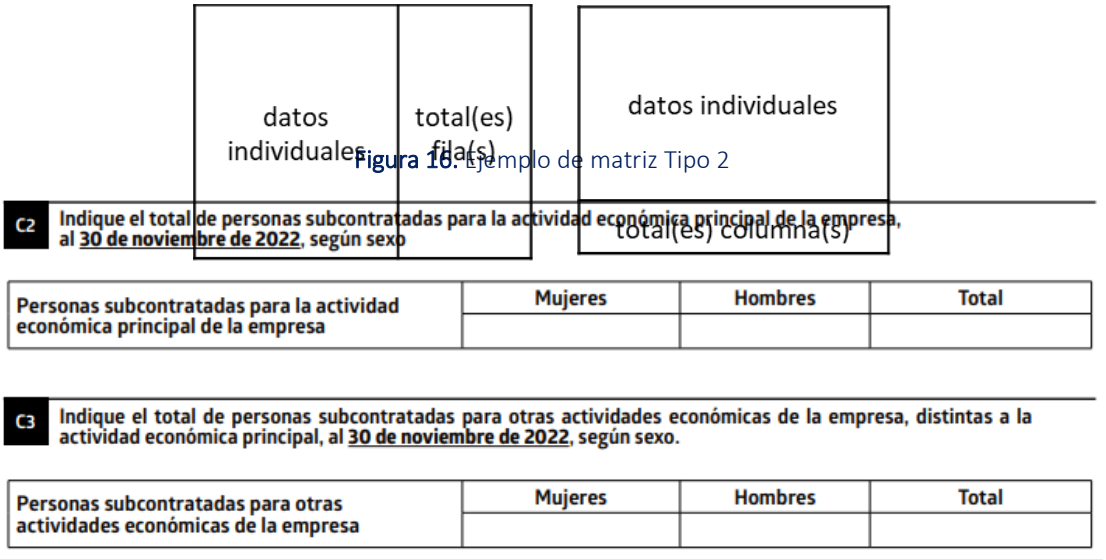


Figura 16. Ejemplo de matriz Tipo 2

C2 **Indique el total de personas subcontratadas para la actividad económica principal de la empresa, al 30 de noviembre de 2022, según sexo**

Personas subcontratadas para la actividad económica principal de la empresa	Mujeres	Hombres	Total

C3 **Indique el total de personas subcontratadas para otras actividades económicas de la empresa, distintas a la actividad económica principal, al 30 de noviembre de 2022, según sexo.**

Personas subcontratadas para otras actividades económicas de la empresa	Mujeres	Hombres	Total

Las siguientes matrices del formulario Autoaplicado son matrices tipo 2: A5, A6, B7, B9, C2, C3, C5, C15, D3, F3, F5, F6, F11, G2, H4, H7, H10, I3, K3, M2, O18, N1 y N2.

VII.6. Cuadratura Intra Matrices

La cuadratura intra-matrices se refiere a las ediciones de datos que se realizan para mantener la consistencia interna de los datos numéricos ingresados en las matrices del Autoaplicado. Concretamente, se trata de la aplicación sucesiva de 8 reglas de edición, las cuales se describen a continuación:

0. Regla 0: Si una celda es editada de forma que, al quitar el último dígito se da origen a una matriz cuadrada, entonces se edita celda, conservando número sin último dígito original. Esta regla aplica para matrices tipo 1 y tipo 2.

Ejemplo:

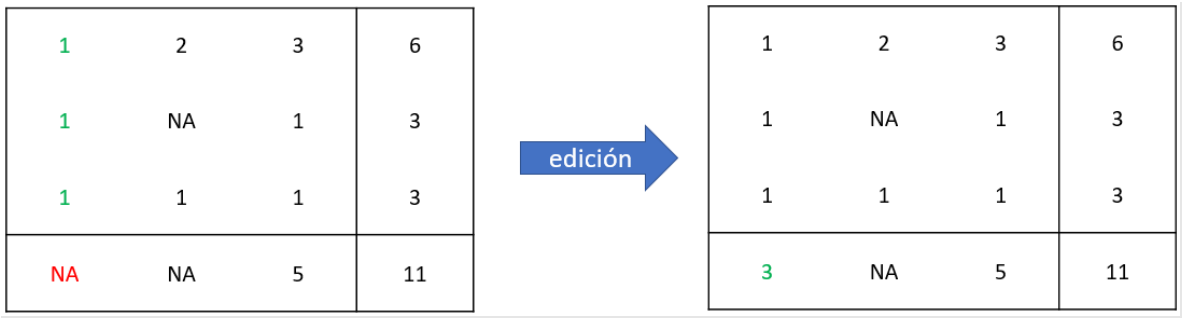
Figura 17. Esquema de corrección aplicada por Regla 0



1. Regla 1: Si (i) el total de fila o columna es NA y (ii) todos los elementos que componen ese total son distintos de NA, entonces se edita total de fila/columna como la suma de los elementos que componen ese total. Esta regla aplica para matrices tipo 1 y tipo 2.

Ejemplo:

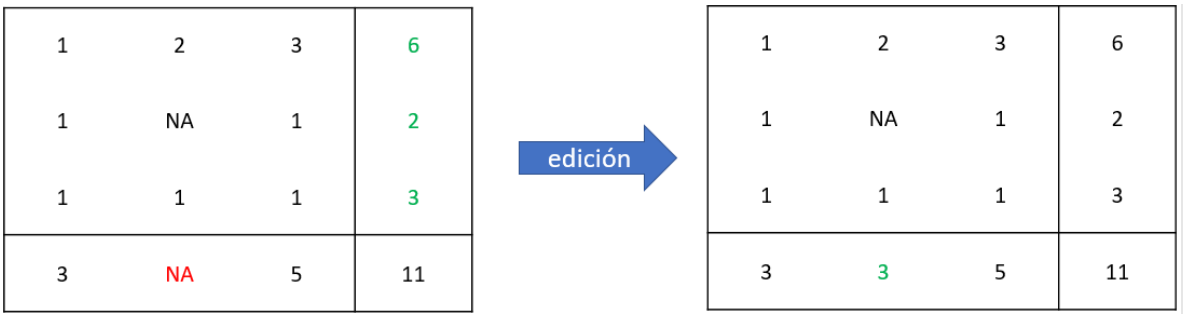
Figura 18. Esquema de corrección aplicada por Regla 1



2. Regla 2: Si (i) el total parcial de columna o el total general de columna es NA, (ii) el total parcial de filas están cuadrados, y (iii) suma de datos individuales es mayor a 0, entonces se edita total parcial de columna como suma de los elementos que componen ese total. Esta regla aplica solo para matrices tipo 1.

Ejemplo:

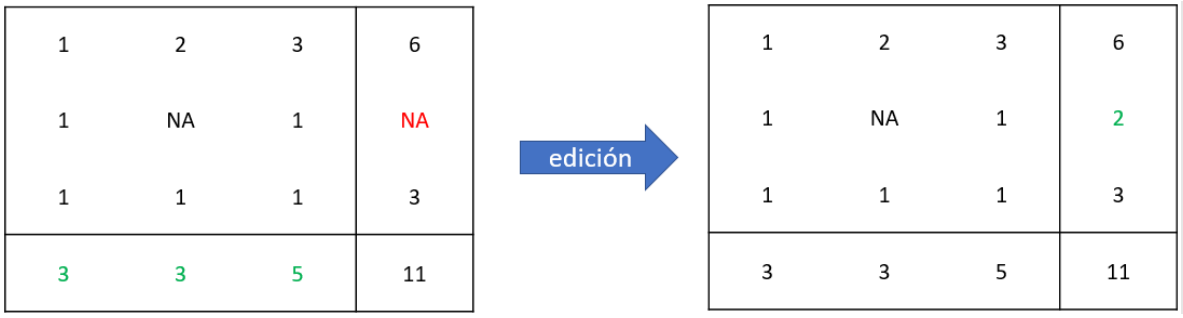
Figura 19. Esquema de corrección aplicada por Regla 2



3. Regla 3: Si (i) el total parcial de filas o el total general de filas es NA, (ii) el total parcial de columnas están cuadrados, y (iii) suma de datos individuales es mayor a 0, entonces se edita total parcial de fila como suma de los elementos que componen ese total. Esta regla aplica solo para matrices tipo 1.

Ejemplo:

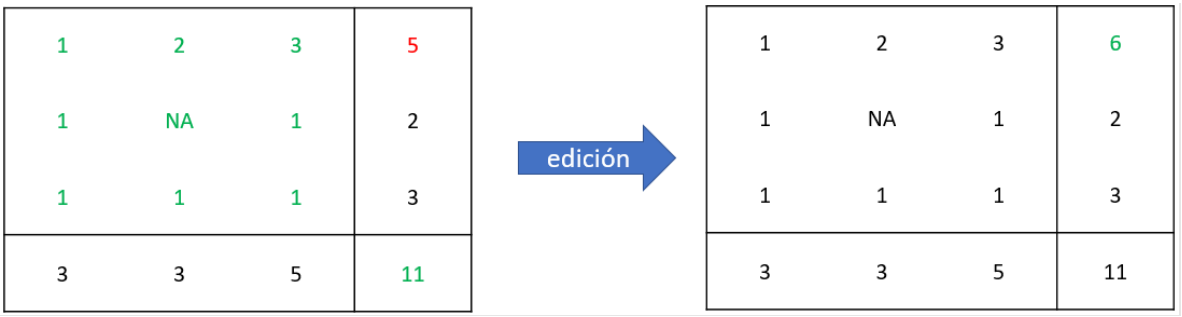
Figura 20. esquema de corrección aplicada por Regla 3



4. Regla 4: Si suma de datos individuales es congruente con total general de la matriz, entonces se editan totales parciales de filas y columnas. Esta regla aplica solo para matrices tipo 1.

Ejemplo:

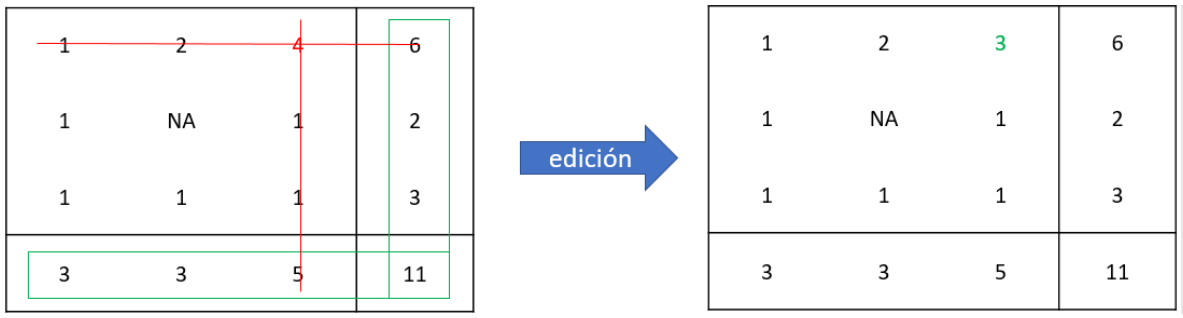
Figura 21. Esquema de corrección aplicada por Regla 4



5. Regla 5: Si (i) totales parciales son congruentes con el total general y (ii) no existe congruencia en solo una fila y solo una columna, entonces se edita celda donde intersecta la fila y la columna problemática. Esta regla aplica solo para matrices tipo 1.

Ejemplo:

Figura 22. esquema de corrección aplicada por Regla 5



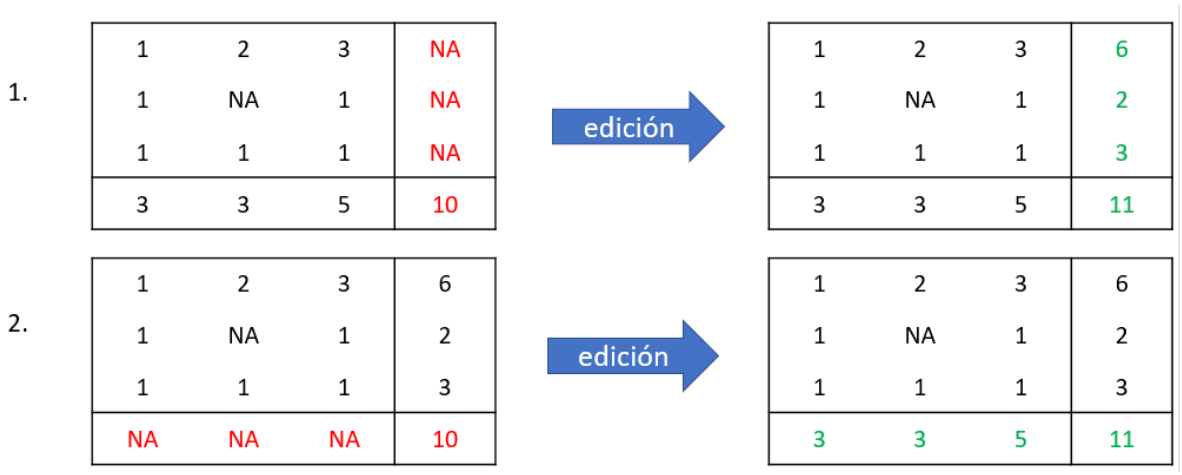
6. Regla 6: Considerando una matriz tipo 1 de J columnas x I filas. Esta regla aplica solo para matrices tipo 1.

a. Si en las primeras J-1 columnas existe congruencia en los totales de columna, entonces se edita cada elemento de la columna de totales como la suma de los elementos de su fila respectiva.

b. Si en las primeras I-1 filas existe congruencia en los totales de fila, se edita cada elemento la fila de totales como la suma de los elementos de su columna respectiva.

Ejemplo:

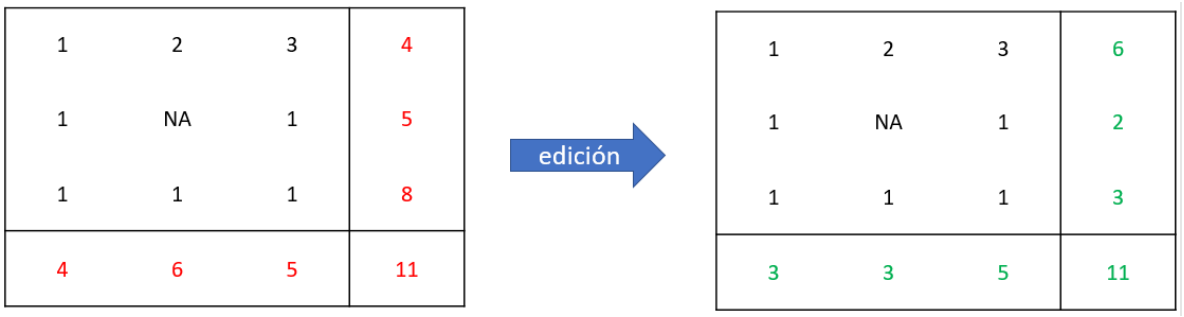
Figura 23. Esquema de corrección aplicada por Regla 6



7. Regla 7: Se editan los totales de la matriz como la suma de los datos individuales involucrados en ese total. Esta regla aplica para matrices tipo 1 y tipo 2.

Ejemplo:

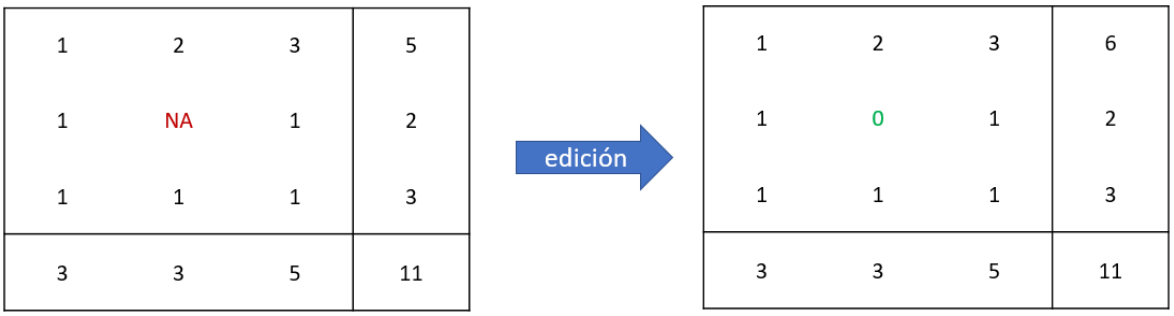
Figura 24. Esquema de corrección aplicada por Regla 7



8. Regla 8: Los valores dentro de las matrices que quedan con NA y que están involucradas en sumas de filas o columnas congruentes con sus respectivos totales, se llevan a cero.

Ejemplo:

Figura 25. Esquema de corrección aplicada por Regla 8



VII.7. Cuadratura Inter-Matrices

La cuadratura Inter-Matrices se refiere a los procesamientos realizados para asegurar la consistencia de datos entre matrices que presentan información relacionada y/o equivalente.

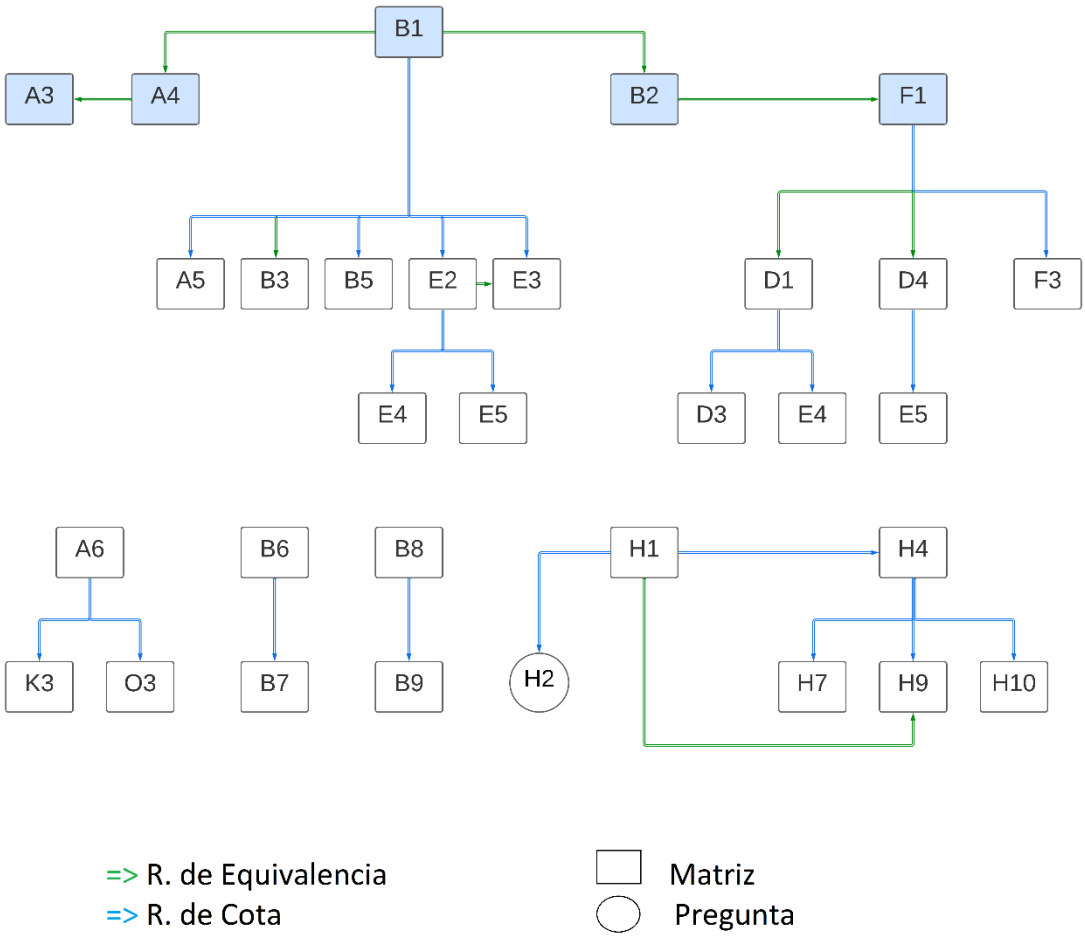
En ese sentido, existen matrices que apuntan al total de trabajadores para una misma fecha (más comúnmente, 30 de noviembre 2022). Estas matrices, para efectos de esta parte del procesamiento de datos, se llamarán matrices primarias (MP). Concretamente, se trata de las matrices A3, A4, B1, B2 y F1. El resto de las matrices, las llamamos matrices secundarias (MS).

Las matrices se vinculan entre sí por (1) relaciones de equivalencia (RE), o (2) restricciones de cotas (RC). Las relaciones de equivalencia apuntan a que el número reportado en una matriz debe coincidir con el número reportado en otra. Por ejemplo, el total de la matriz A3 debería coincidir con el total de la matriz B1, ya que ambas reportan el número de personas contratadas directamente por la empresa al 30 de noviembre de 2022 (la primera según región y sexo; mientras que la segunda, según tipo de contrato).

Las restricciones de cotas se refieren a que la cifra de una matriz debe ser menor o igual (o bien mayor o igual) que la cifra de otra matriz (ejemplo: total A6 [total de trabajadores en la empresa] debe ser mayor o igual (pero no menor) al total de K3 [número de trabajadores afiliados al sindicato], si esta es respondida).

La siguiente figura presenta un mapa de relaciones entre matrices (en verde las RE y en azul las RC).

Figura 26. Relaciones entre matrices del Formulario Autoaplicado



VII.7.1. Ediciones para relación de equivalencia

Como se mencionó previamente, las matrices primarias se vinculan entre sí en tanto coinciden en el total medido, correspondiendo al total de trabajadores en la empresa al 30 de noviembre de 2022. En caso de que este valor no coincida en las distintas matrices primarias, se corrige a través de los siguientes pasos:

1. Obtener el número de trabajadores que se utilizará como referencia para el resto de las matrices primarias. Para ello, se obtiene total de las matrices primarias: A3, A4, B1, B2 y F1. El número de trabajadores corresponde a la moda de los 5 valores obtenidos. Con dos excepciones:
    a. Cuando el total de B1, excluyendo a los trabajadores con contratos de palabra y honorarios, es igual a la moda y además se reporta un número de trabajadores con contratos de palabra u honorarios mayor a cero. En ese caso, el número de trabajadores es igual al total de B1. Esta excepción está pensada para casos donde el informante no incluye a trabajadores con contrato de palabra y honorarios en A3, A4 o F1.
    b. Cuando no existe moda. En ese caso, el número de trabajadores también es igual al total de B1.

Luego se chequea que este dato (número total de trabajadores), sea congruente en todas las matrices primarias (MP). De no serlo, se ajustan las matrices que no se cuadran con esta cifra (expandiendo o reduciendo proporcionalmente; resguardando la consistencia en los totales generales, así como también en los totales por sexo, cuando corresponde).

Para los casos en que el número total (N) de las MP difiere en el número de trabajadores, se realiza lo siguiente:

1. Se construye un factor de escala: $factor = \frac{NT}{N}$

2. Se multiplica cada celda (excepto totales) de la MP por este factor.
3. Se redondea cada número al entero más cercano.
4. Se revisa si la suma de los números aproximados (Napx) coincide con NT.
5. Si existe diferencia, se aumenta/disminuye en 1 tantas celdas como unidades de diferencia existan.

Ejemplos: Si NT=6 y Napx=5, entonces +1 a una celda

Si NT=6 y Napx=8, entonces -1 a dos celdas

6. Las celdas que aumentar/disminuir se definen en base a la parte decimal de las celdas distintas de 0.

Lógica: decimales más próximos por debajo a 0.5 (exclusive) => mayor prioridad para +1 decimales más próximos por arriba a 0.5 (inclusive) => menor prioridad para +1

decimales más próximos por debajo a 0.5 (exclusive) => menor prioridad para -1 decimales más próximos por arriba a 0.5 (inclusive) => mayor prioridad para -1

7. Una vez realizado el ajuste (si es necesario), se editan totales de la MP.

VII.7.2. Ediciones para relación de cotas

Con ediciones por cotas, nos referimos fundamentalmente a una relación numérica entre valores de matrices del autoaplicado donde el valor de una celda/variable de una matriz no puede ser mayor (o menor) que el valor de una celda/variable en otra matriz.

Un resumen de las validaciones de cota en todo el cuestionario puede verse en la figura 26, reflejados en las líneas azules del esquema. En general, la corrección por cotas implica que, si el dato evaluado es mayor que el dato de referencia, entonces el primero se reduce hasta ser igual al dato de referencia.

Tabla 53. relaciones de cotas evaluadas y corregidas

Matrices evaluadas	Descripción
B1-A5	Los valores de a5_1, a5_2 y a5_3 no pueden ser mayores a b1_1_7, b1_2_7 y b1_3_7, respectivamente
B1-B5	Los valores de b5_1_5, b5_2_5 y b5_3_5 no pueden ser mayores a b1_1_7, b1_2_7 y b1_3_7, respectivamente
B1-E2	Los valores de e2_1_9, e2_2_9 y e2_3_9 no pueden ser mayores a b1_1_7, b1_2_7 y b1_3_7, respectivamente
B1-E3	Los valores de e3_1_9, e3_2_9, e3_3_9, e3_4_9, e3_5_9 y e3_6_9 no pueden ser mayores a b1_3_1, b1_3_2, b1_3_3, b1_3_4, b1_3_5, b1_3_6 respectivamente
B1-F3	Los valores de f3_1, f3_2 y f3_3 no pueden ser mayores a f1_1_7, f1_2_7 y f1_3_7, respectivamente
F3-F4	El valor de f4_4 no puede ser mayor al valor de f3_3.
F1-F6	Los valores de f6_1, f6_2 y f6_3 no pueden ser mayores a f1_1_7, f1_2_7 y f1_3_7, respectivamente
F1-F8	Los valores de f8_1_3, f8_2_3 y f8_3_3 no pueden ser mayores a f1_1_7, f1_2_7 y f1_3_7, respectivamente
F1-F11	Los valores de f11_1, f11_2 y f11_3 no pueden ser mayores a f1_1_7, f1_2_7 y f1_3_7, respectivamente
E2-E4	Los valores de e4_1_10, e4_2_10 y e4_3_10 no pueden ser mayores a e2_1_9, e2_2_9 y e2_3_9, respectivamente
E2-E5	Los valores de e5_1_10, e5_2_10 y e5_3_10 no pueden ser mayores a e2_1_9, e2_2_9 y e2_3_9, respectivamente
D1-D3	La suma de d3_1_1 y d3_1_2 no puede ser mayor que d1_1_10. La suma de d3_2_1 y d3_2_2 no puede ser mayor que d1_2_10. La suma de d3_3_1 y d3_3_2 no puede ser mayor que d1_3_10.
D1-E4	Los valores de e4_1_10, e4_2_10 y e4_3_10 no pueden ser mayores a d1_1_10, d1_2_10 y d1_2_10, respectivamente
D4-E5	Los valores de e5_1_10, e5_2_10 y e5_3_10 no pueden ser mayores a d1_1_10, d1_2_10 y d1_2_10, respectivamente
E2-E4+E5	La suma de e4_1_10 y e5_1_10 no puede ser mayor que e2_1_9. La suma de e4_2_10 y e5_2_10 no puede ser mayor que e2_2_9. La suma de e4_3_10 y e5_3_10 no puede ser mayor que e2_3_9.
A6-K3	Los valores de k3_1, k3_2 y k3_3 no pueden ser mayores a a6_1, a6_2 y a6_3, respectivamente
A6-O3	Los valores de o3_1_7, o3_2_7 y o3_3_7 no pueden ser mayores a a6_1, a6_2 y a6_3, respectivamente
H1-H2	El valor de h2 no puede ser menor al valor máximo entre h1_2_2, h1_2_4, h1_2_5, h1_2_6
H1-H4	El valor de h4_1 no puede ser menor al valor máximo entre h1_1_3, h1_1_4, h1_1_5, h1_1_6 Además, el valor de h4_2 no puede ser menor al valor máximo entre h1_2_2, h1_2_4, h1_2_5, h1_2_6
H4-H7	Los valores de

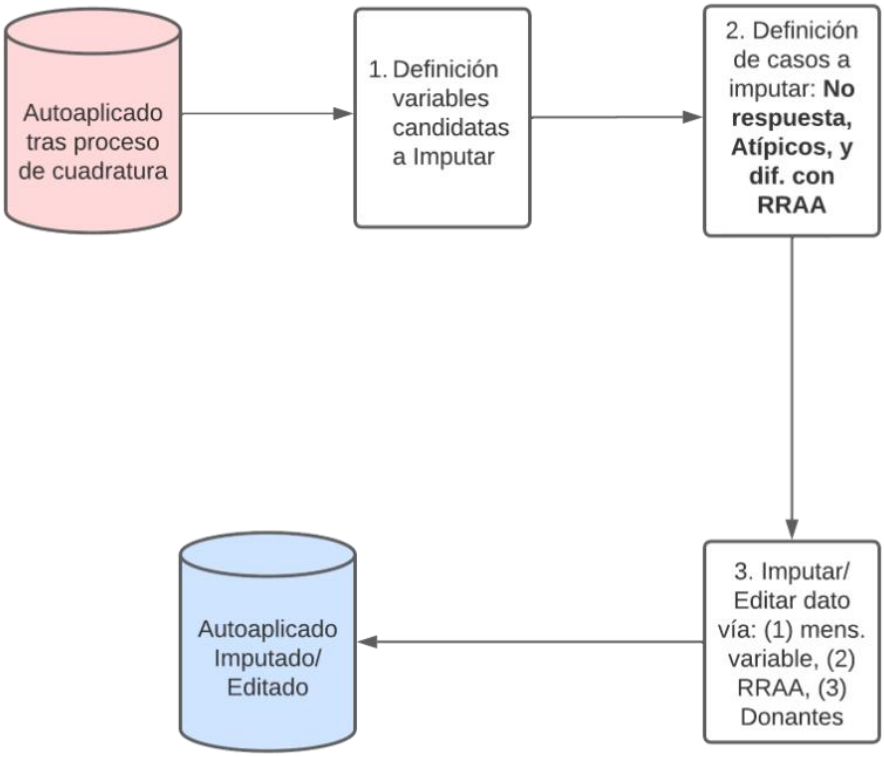
H4-H9	Los valores de h9_1_3, h9_2_3 y h9_3_3 no pueden ser mayores a h4_1, h4_2 y h4_3, respectivamente
H4-H10	Los valores de h10_1, h10_2 y h10_3 no pueden ser mayores a h4_1, h4_2 y h4_3, respectivamente
B6 interno	La cantidad de trabajadores no puede ser mayor que la cantidad de contratos celebrados en cada categoría de tipos de contrato
B7 interno	La cantidad de trabajadores no puede ser mayor que la cantidad de contratos celebrados en cada categoría de tipos de contratos previos y renovaciones
B8 interno	La cantidad de trabajadores no puede ser mayor que la cantidad de contratos a los que se puso término en cada categoría de tipos de contrato
B6-B7	La suma de la cantidad de tipos de contratos previos renovados no puede ser mayor a la categoría de referencia en b6. La suma de la cantidad de trabajadores según tipos de contratos previos no puede ser mayor a la categoría de referencia en b6.
B8-B9	La suma de b9_1_1 y b9_1_2 no puede ser mayor que b8_4_5 La suma de b9_1_2 y b9_2_2 no puede ser mayor que b8_5_5
G1-G2	Los valores de g2_3_1 y g2_3_2 no pueden ser mayores a g1_1 y g1_2, respectivamente

Ahora bien, puede ocurrir que una corrección de cota genere distorsiones en la tabla que afecten su cuadratura intra o inter matriz (por relación de equivalencia). Para solucionar este problema, las correcciones de cota incluyen un ajuste para cuadratura intra matriz e inter-matriz posterior a la reducción por cota.

VII.8. Imputación de datos

El flujo de trabajo para la imputación de datos parte con las bases de datos cuadradas luego del procedimiento de cuadratura que se describió previamente. En un primer momento se define qué variables son candidatas para imputar que cumplen ciertos requisitos como, basado principalmente en el tipo de variables que se trate (solo se imputarán variables numéricas excluyendo las categóricas) y que efectivamente tengamos casos perdidos en dichas variables. Luego de definir las variables, se determina en qué casos se imputará, los cuales incluyen casos de no respuesta, valores atípicos, y cuando hay diferencias con registros administrativos. Luego de este paso, imputa/edita el dato que corresponda a partir de tres formas diferentes: (1) mensualización de la variable (se divide por 12 el valor pensando en que la variable que pregunta por un mes específico se haya respondido por el año completo), (2) vía registros administrativos en los casos que se disponga información y (3) donantes de acuerdo con las variables de estratificación de la muestra de la ENCLA: Región, Tamaño, y Actividad Económica. Un resumen de todo el proceso de imputación puede resumirse en el siguiente diagrama de flujos.

Figura 27. Diagrama de flujo del proceso de imputación



VII.8.1. No Respuesta/Inconsistencia

Las variables que son potencialmente imputables son todas variables del cuestionario autoaplicado. El resumen, asociado a no respuesta (NR)/ inconsistencia, puede verse a continuación:

Tabla 54. Resumen de variables candidatas a imputar

Cuestionario	Variable/Matriz	N° de casos
Autoaplicado	a3	0
Autoaplicado	a4	20
Autoaplicado	a5	4
Autoaplicado	a6	8
Autoaplicado	a8	0
Autoaplicado	b1	0
Autoaplicado	b2	0
Autoaplicado	b3	13
Autoaplicado	b5	0
Autoaplicado	b6 (b6_3_5, b6_6_5)	15
Autoaplicado	b7 (b7_3_x, b7_6_x)	53
Autoaplicado	b8 (b8_3_5, b8_6_5)	30

Autoaplicado	b9 (b9_3_1, b9_3_2)	58
Autoaplicado	c2	22
Autoaplicado	c3	1
Autoaplicado	c4	31
Autoaplicado	c5	2
Autoaplicado	c11	0
Autoaplicado	c12	7
Autoaplicado	c15	0
Autoaplicado	d1	0
Autoaplicado	d2	830
Autoaplicado	d3	180
Autoaplicado	d4	0
Autoaplicado	d5	37
Autoaplicado	e2	0
Autoaplicado	e3	14
Autoaplicado	e4	3
Autoaplicado	e5	15
Autoaplicado	f1	0
Autoaplicado	f3	3
Autoaplicado	f4 (f4_x)	71
Autoaplicado	f5 (f5_x)	13
Autoaplicado	f6 (f6_x)	12
Autoaplicado	f8	1
Autoaplicado	f11	0
Autoaplicado	g1	22
Autoaplicado	g2	27
Autoaplicado	g5	356
Autoaplicado	h1	0
Autoaplicado	h2	303
Autoaplicado	h4	1
Autoaplicado	h7	649
Autoaplicado	h9	15
Autoaplicado	h10	15
Autoaplicado	i3	604
Autoaplicado	k1	9
Autoaplicado	k2	1
Autoaplicado	k3	2
Autoaplicado	m2	0
Autoaplicado	m3	22
Autoaplicado	m4	0

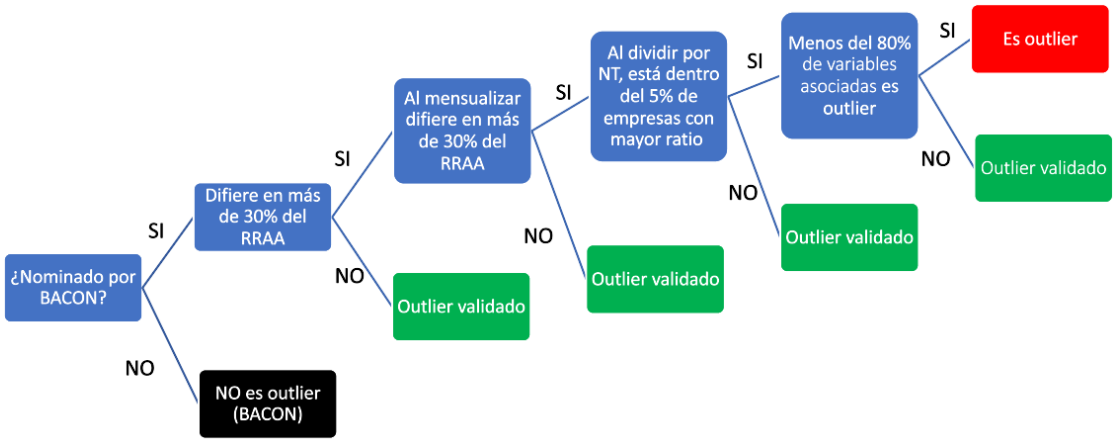
Autoaplicado	o3	30
Autoaplicado	o18 (o18_x_3)	0
Autoaplicado	n1 (n1_1_3_x, n1_2_3_x)	4
Autoaplicado	n2 (n2_1_3_x, n2_2_3_x)	50

VII.8.2. Valores atípicos

En el caso de los valores atípicos tenemos el siguiente procedimiento para determinarlo. Se realiza una nominación de *outliers* a través del método BACON (cuyas siglas significan *blocked adaptative computationally efficient outlier nominators*) teniendo como covariables, el tamaño (según número de trabajadores), la macrozona, y el sector económico de la empresa.

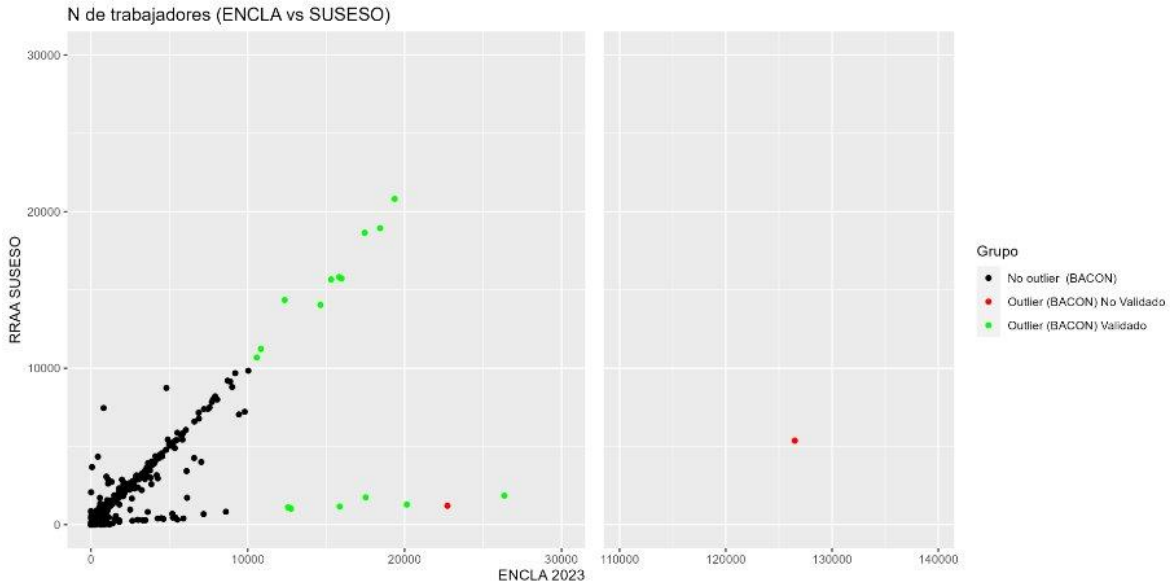
Para lo anterior, se sigue un árbol de decisiones que clarifica cada paso en el proceso de imputación de datos atípicos, resumido en el siguiente esquema:

Figura 28. Árbol de decisión imputación de datos atípicos



A modo de ejemplo, para la matriz B1 del cuestionario autoaplicado podemos ver la siguiente distribución de los valores de acuerdo con el procedimiento anteriormente mencionado.

Figura 29. Gráfico de dispersión entre el total de trabajadores declarados en B1 y el total de trabajadores registrados en los datos de la SUSESO, identificando *outliers* con BACON



El resumen de aplicar el árbol de decisión puede verse en la siguiente tabla resumen:

Tabla 55. Casos atípicos matriz B1 del autoaplicado

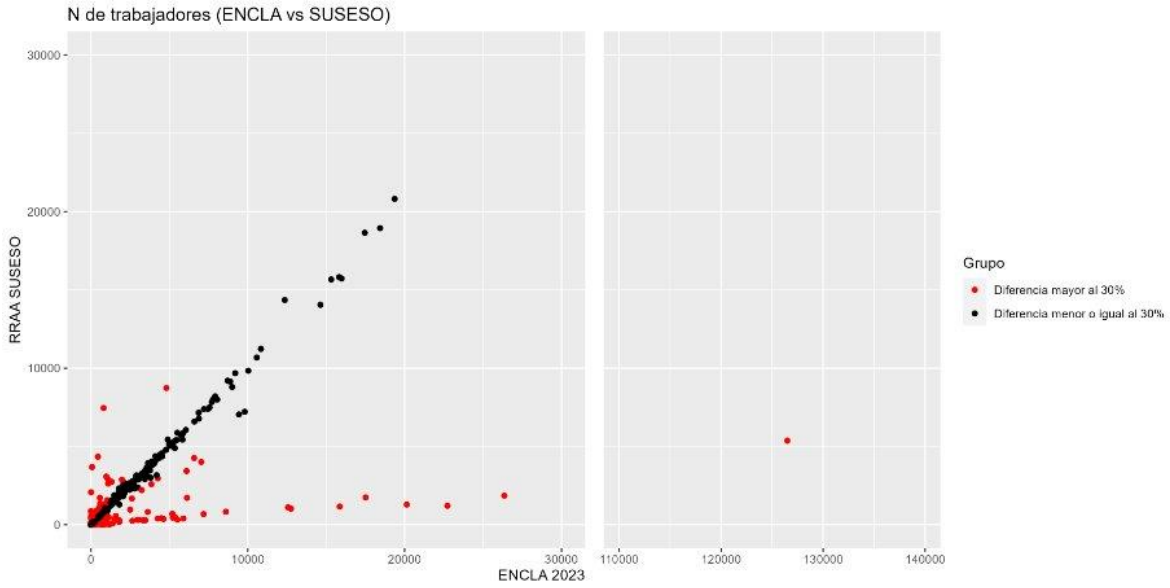
N	4496
Nominados por BACON	18
Más de 30% de diferencia con RRAA	8
Más de 30% de diferencia con RRAA (mensualizado)	2
En el 5% de casos con mayor ratio	2 (No Aplica)
Menos de 80% de outliers en variables asociadas	2

VII.8.3. Imputación vía RRAA

VII.8.3.1. Imputación B1

Siguiendo el ejemplo, vemos las diferencias de los valores observados del cuestionario (en este caso pregunta B1) y lo comparamos con los datos existentes en los registros administrativos. Para el caso de B1, podemos verlo del siguiente modo:

Figura 30. Gráfico de dispersión entre el total de trabajadores declarados en B1 y el total de trabajadores registrados en los datos de la SUSESO, según la magnitud de la diferencia entre ambos



Lo que en términos numéricos se traduce en las siguientes cantidades:

Tabla 56. Diferencias de matriz B1 con registros administrativos

Diferencia mayor al 30%	511	11.4%
Diferencia menor o igual al 30%	3970	88.6%

Siguiendo el flujo de imputación descrito, dentro del subconjunto de *outliers* no validados con diferencia mayor al 30% tenemos dos caminos para la edición con el siguiente orden de prevalencia: (1) Si la variable mensualizada (total de matriz B1*/12) difiere en más del 30% del registro administrativo, lo que se hará será editar el total considerando lo siguiente:

b1_3_7 final (total de personas contratadas directamente por la empresa) = N(RRAA) + b1_3_5 (Total de honorarios que no vienen incluidos en el registro administrativo) + b1_3_6 (total de contratos sin escrituración).

(2) Si la variable mensualizada (total de la matriz B1*/12) difiere en menos del 30% del registro administrativo, entonces:

b1_3_7final = (b1_3_7/12)

Como resultado de lo anterior, tenemos las siguientes proporciones de casos editados para el ejemplo mencionado:

Tabla 57. resultados de aplicar flujo de imputación a matriz b1

Se edita con RRAA	481	10.7%
Se edita mensualizando b1_3_7	30	0.7%
Validado	3970	88.6%

VII.8.3.2. Imputación A6

Análogamente a la imputación por RRAA de B1, la pregunta A6 se imputa con registros administrativos de SUSESO si existe una diferencia mayor al 30% entre ambas fuentes de información por empresa. Debido a que A6 pregunta por el total de trabajadores en la actualidad, la información a contrastar es contingente al momento en que el cuestionario autoaplicado es enviado.

La siguiente tabla resume los resultados de la comparación:

Tabla 58. Diferencias de matriz A6 con registros administrativos

Diferencia mayor al 30%	580	12.9%
Diferencia menor o igual al 30%	3916	87.1%

VII.8.3.3. Imputación G5

Con el fin de corregir inconsistencias y valores atípicos, e imputar información en unidades sin datos al respecto, en las variables de la matriz G5 sobre la composición de la representación de las empresas y los trabajadores en los Comités Paritarios de Orden, Higiene y Seguridad, la DT entregó datos agregados sobre el número de comités y representantes de cada actor en 2.659 empresas. De ellas, 405 (15.2%) corresponden a empresas logradas en la Encla 2023.

Como instituye la legislación laboral, los establecimientos con más de 25 trabajadores deben contar con un Comité Paritario de Orden, Higiene y Seguridad, el que a su vez debe estar constituido por tres representantes de la empresa y tres representantes de los trabajadores. En la Encla 2023 no se recolecta información sobre el número de Comités Paritarios ni de establecimientos con los que se cuenta a nivel nacional, por lo que no es posible estimar con precisión el máximo de representantes que puede haber en cada empresa. Sin embargo, para efectos del procesamiento, se trabajaron dos situaciones que fueron consideradas necesarias de corregir:

a) el número de representantes registrado en la variable g5_3_3 es mayor a seis; y

b) la razón de g5_3_3 entre 6, que sirve de proxy para estimar el número de Comités Paritarios en la empresa, supera la razón de b1_3_7 entre 25, que sirve de proxy para el máximo de Comités Paritarios que podría haber en la empresa (o sea, el número máximo de establecimientos con, por lo menos, 25 trabajadores).

Para los casos donde se presentó alguno de los dos problemas:

- Si se cuenta con registros administrativos para la unidad, se asigna la información del registro.
- En aquellos casos sin registros, se chequea si la razón entre b1_3_7 (el número de trabajadores en la empresa) y el número de Comités Paritarios informados (g5_3_3/6) es mayor a 25.
    - Si la condición se cumple, se valida la información declarada por los informantes.
    - Si la condición no se cumple, se asigna el valor –96.

Además, para los casos editados a partir de las validaciones inter-cuestionarios, se ha editado con el valor del registro administrativo cuando ha sido posible; en caso contrario, se ha asignado el valor –96. Hay que recordar que en este caso se tomó como referencia la variable G3 del cuestionario a empleadores. En caso de existir inconsistencia; se modifica G4 del Autoaplicado para mantener consistencia en caso de que el Empleador haya manifestado la existencia de CPHS en la empresa.

En los casos editados con registros administrativos, se han mantenido las proporciones informadas originalmente para las desagregaciones por sexo. Además, los casos con valores perdidos fueron imputados a través del método de medianas condicionadas que se detallará en el siguiente subapartado. En este caso, hay dos situaciones:

a) Toda la matriz viene con –96: en este caso, se utiliza el método de medianas condicionadas para imputar la totalidad de la matriz G5.
b) La matriz cuenta con información para los totales por entidad representada, pero viene con –96 en las desagregaciones por sexo: se ha imputado mediante medianas condicionadas la proporción de representantes por sexo y tipo de representante, manteniendo los totales obtenidos del registro.

VII.8.3.4. Imputación de módulo K

En sus revisiones, la DT identificó una serie de inconsistencias y casos sin dato por revisar en las preguntas del módulo K del cuestionario autoaplicado, las cuales fueron corregidas recurriendo a la información del SIRELA entregada por la DT. Concretamente:

- Pregunta K1:
    - En aquellos casos donde el número de sindicatos en la empresa durante 2022 (k1) fuese mayor que el número de trabajadores sindicalizados en la empresa (k3_3), se ha editado con la información del registro administrativo.
    - Los casos con valor –96 que cuentan con registro administrativo fueron editados con esa información; en caso de no contar con registro administrativo, se asignó el valor 0.
- Pregunta K2:
    - En los casos donde se declaró por lo menos la existencia de al menos un sindicato actualmente (k2), y donde el número de sindicatos informados es igual al número de trabajadores sindicalizados declarados en k3_3, se ha editado con la información del registro administrativo.
    - Los casos con valor –96 que cuentan con registro administrativo fueron editados con esa información; en caso de no contar con registro administrativo, se asignó el valor 0.
- Pregunta K3:
    - Los casos con valor –96 fueron editados con la información del registro administrativo.
    - En los casos donde el valor de k3_1, k3_2 y k3_3 era igual a 0, y se contaba con registros administrativos, se editó aquella información; en los casos donde la unidad no cuenta con registros administrativos para el módulo K, se editó con la información capturada en la matriz k2 del cuestionario a sindicatos.
    - En los casos donde el valor de k3_3 era igual a 0, pero el valor de k3_1 o k3_2 era mayor a 0, se editó en k3_3 la suma de k3_1 y k3_2.
    - En los casos donde el valor de k3_3 es menor a 8, se editó con la información del registro administrativo en las variables de toda la matriz; en caso de no contar con registro administrativo, se asignó la información capturada en la matriz k2 del cuestionario a sindicatos para editar toda la matriz k3 del cuestionario autoaplicado.

VII.8.3.5. Imputación de módulo M

Tal como en las preguntas del módulo K, las observaciones realizadas por el equipo de la DT a la base de datos significó la edición de las variables que forman parte del módulo M. En particular:

- M2_3, m3_1_4 y m3_2_4: se editó con la información del registro administrativo en los casos donde se contaba con ella; en el resto, se asignó el valor 0.
- M3_1_1, m3_1_2, m3_1_3 y m3_1_4: en aquellos casos donde la suma de las variables indicadas fuese igual a m2_5, se imputó la información del registro administrativo. Para las variables m3_1_1 y m3_1_2, en caso de no contar con información en el registro administrativo, se editó con la información capturada en k3_3, asumiendo que todos quienes están sindicalizados han suscrito al contrato o convenio colectivo. Asimismo, en aquellos casos donde la cantidad de trabajadores que suscribieron instrumentos colectivos y acuerdos de grupo negociador (sumatoria de m3_1_1, m3_1_2, m3_1_3 y m3_1_4) dividida entre el total de entidades negociadoras registradas (sumatoria de m4_1, m4_2, m4_3, m4_4 y m4_5) es menor a cinco, se ha mantenido el valor original cuando m2_1 a m2_4 y su correspondiente m3_1_X eran mayores a 0 y 1, respectivamente; luego, en caso de no cumplir con esa condición, se ha editado con la información del registro; en caso contrario, se ha asignado el valor 0. Para m3_1_1 y m3_1_2, en caso de no contar con información en el registro, se ha imputado k3_3.
- M2_1, m2_2, m2_3 y m2_4: luego de las correcciones implementadas en m3_1_1 a m3_1_4, se han llevado a 0 los valores correspondientes para las preguntas m2_1 a m2_4, en aquellos casos donde el nuevo valor en m3_1_X sea igual a cero.
- M3_2_1, m3_2_2, m3_2_3 y m3_2_4: en los casos donde m3_1_X fuese igual a m3_2_X, y tanto m3_1_X como m3_2_X fuesen distintos de 0 y –96, se asignó el valor –96 en m3_2_X, conservando el valor original en el resto de los casos.
- M4_1, m4_2, m4_3, m4_4 y m4_5: se han editado aquellos casos donde la cantidad de trabajadores que suscribieron instrumentos colectivos y acuerdos de grupo negociador (sumatoria de m3_1_1, m3_1_2, m3_1_3 y m3_1_4) dividida entre el total de entidades negociadoras registradas (sumatoria de m4_1, m4_2, m4_3, m4_4 y m4_5) es menor a cinco. Cuando se contó con información en el registro, se asignó esa información. Para m4_1, m4_2, m4_4 y m4_5, cuando se identificó el problema, no se cuenta con registro administrativo, y hay un número mayor a uno en m3_1_1 o m3_1_2, se ha mantenido el valor original.
- M2_5: se identificaron casos donde la razón entre el número de trabajadores de la empresa (b1_3_7) y la cantidad de instrumentos colectivos o acuerdos de grupo negociador vigentes es menor a 8; sin embargo, las correcciones detalladas anteriormente corrigieron este problema.

VII.8.4. Imputación vía medianas condicionadas

En caso de que un valor susceptible de imputación (variable cuantitativa que presenta no respuesta o valor atípico) no haya podido ser imputado vía registros administrativos, se realizará una imputación vía mediana condicionadas (si es posible).

En concreto, el procedimiento de imputación consiste en los siguientes pasos:

1. **Definición de set de donantes**: para cada variable a imputar el set de donantes corresponde a todas las observaciones que presenten información válida en esa pregunta, es decir: que sea una pregunta que debía ser respondida por flujo y que no sea un valor susceptible a imputación.

2. **Definición de subgrupos de donantes**: se definen como los subgrupos resultantes del cruce de 3 variables relevantes. En un comienzo, las 3 variables son: actividad económica (13 categorías), decil de número de trabajadores (10 categorías), y región (16 categorías).
3. **Obtención de medianas condicionadas y n de donantes (1ra ronda)**: para cada uno de los subgrupos de donantes se obtiene la mediana de la variable a imputar (o set de variables a imputar, como en el caso de matrices), además del número de donantes en cada subgrupo.
4. **Imputación (1ra ronda)**: se identifica el subgrupo al que pertenece la observación que debe ser imputada, asignando la mediana de ese subgrupo siempre que haya un mínimo de 3 donantes. En caso de haber menos de 3 donantes, se pasa a una siguiente ronda de imputación.
5. **Obtención de medianas condicionadas y n de donantes (2da ronda):** en este paso se calcula la o las medianas condicionadas y el número de donantes para los subgrupos, dados por el cruce de las variables de actividad económica y decil de número de trabajadores
6. **Imputación (2da ronda)**: en aquellos valores no imputados en la primera ronda se asigna la mediana del subgrupo definido en el paso inmediatamente anterior, siempre y cuando hay un mínimo de 3 donantes. En caso de haber menos de 3 donantes, el valor no se imputa, asignándose 96 (dato perdido).
7. **Ajuste de efectos colaterales**: tras la aplicación de los pasos previos en cada variable, se realizan chequeos (y ajustes de ser necesarios) de cuadratura y consistencia con otras preguntas del cuestionario.

La siguiente tabla presenta un resumen de las imputaciones realizadas por módulo (para mayor detalle, en anexos se presenta el número de imputaciones por variable):

Tabla 59. resumen de imputaciones por módulo vía medianas condicionadas

Módulo	Casos a Imputar	Imputados en R1	Imputados en R2	No imputados
A	31	22	9	0
B	1.276	686	580	10
C	65	11	51	3
D	1.427	255	901	271
E	33	10	23	0
F	222	42	120	60
G	484	138	252	94
H	1.388	448	882	58
I	1.280	265	935	80
K	0	0	0	0
M	1	0	1	0
O	69	27	36	6
N	416	289	127	0

La no imputación de información se debe a la ausencia de un número suficiente de donantes, según los criterios de definición de subgrupos de donantes descritos anteriormente.

VII.8.5. Ediciones a los cuestionarios a empleadores, sindicatos y trabajadores a partir de Registros Administrativos (RRAA)

En sus observaciones, el equipo de la Dirección del Trabajo también propuso correcciones a los cuestionarios a empleadores, sindicatos y trabajadores, utilizando los datos del Sistema Informático de Relaciones Laborales (SIRELA); así como ediciones generales por inconsistencias entre preguntas.

VII.8.5.1. Cuestionario a empleadores

En el caso del cuestionario a empleadores, junto con corregir problemas de etiquetado, se han implementado las siguientes ediciones:

- R2: los valores menores a 18 años, o iguales a 99, se han recodificado como –96.
- F1_01, f1_02, f1_03 y f1_04: cuando la variable correspondiente en la pregunta F2 registra un valor igual a 85, se asignó el valor 85 en f1_0X.
- F2_01, f2_02, f2_03 y f2_04: cuando la variable correspondiente en la pregunta F1 registra un valor igual a 85, se asignó el valor 85 en F2_0X.
- G4_01 a g4_06: en aquellos casos en que G3 es igual a 1 y no se cuenta con información válida en la matriz G4, se ha imputado el valor 96.
- I3 a i7_99: en los casos donde I2 es igual a 1, y no existe información válida en las variables comprendidas entre las preguntas I3 e I7, se ha asignado el valor 96 en estas últimas.
- J12_veces: los valores mayores a 70 se han recodificado como 96.
- M4: los valores fuera de rango se han recodificado utilizando el número de semanas comprendido en el registro administrativo, entre la presentación de la propuesta al empleador y la suscripción del instrumento colectivo o acuerdo de grupo negociador.

VII.8.5.2. Cuestionario a sindicatos

En el caso del cuestionario a sindicatos, junto con corregir problemas de etiquetado, se han implementado las siguientes ediciones:

- R2: el caso con valor igual a 99 ha sido recodificado como 96.
- G11_02: se ha asignado el valor 85 en aquellas unidades con un valor igual a uno en G10, y sin información válida en G11_02.
- J7: cuando el valor en la variable j7_frec fuese igual a 0, se asignó el valor 8 en j7.
- J7_frec: cuando el valor de la variable era igual a 0, se eliminó la información dada la edición realizada en J7; cuando el valor registrado fuese igual a 99, se asignó el valor –96.
- K11: se han corregido los valores atípicos identificados por la DT, asignando la información del registro administrativo.
- M5_m, m5_h y m5_t: se ha aplicado el procedimiento de identificación y edición de casos atípicos, tal como se ha implementado en la base de datos del cuestionario autoaplicado. Un caso no pudo ser imputado, por lo que su información fue editada con el valor 96.

Adicionalmente, se han realizado las siguientes correcciones a problemas identificados por el equipo técnico del INE:

- K2_m, k2_h y k2_t: se han corregido aquellos casos donde se declararon mujeres u hombres como parte del directorio sindical (K4), pero se declararon cero mujeres u hombres en las preguntas k2_m y k2_h, respectivamente; o donde se declararon cero mujeres u hombres en k2_m y k2_h, pero se indicó que hubo mujeres u hombres participando en la comisión negociadora en la pregunta M5. En aquellos casos, se ha imputado el número de mujeres u hombres afiliados al sindicato con el registro administrativo; en los casos que no cuenten con esta información, se ha comparado el número de mujeres u hombres declarados en K4 y M5, asignando el valor más alto entre ambas preguntas. Luego, se ha imputado k2_t con la suma de los nuevos valores asignados en k2_m y k2_h.
- K3_ch y k3_t: la modificación realizada en k2_m, k2_h y k2_t puede haber significado una alteración en el total, lo cual redunda en una diferencia entre k2_t y k3_t.
- K5: en aquellos casos donde se declararon personas extranjeras en el directorio sindical, pero se informaron cero socios o socias del sindicato de nacionalidad extranjera, se ha modificado el valor de K5 a 2.
- M6: se identificó que el número de fallos arbitrales originados por la última negociación colectiva registrados en el cuestionario a sindicatos era mayor a lo esperable. Por ello, se ha mantenido el valor 3 (fallo arbitral) en aquellos casos que coinciden con el registro administrativo; en el resto de los casos, se ha modificado el valor 3 por los valores 1 (contrato colectivo) o 2 (convenio colectivo), según lo consignado en el registro administrativo.

VII.8.5.3. Cuestionario a trabajadores

En el caso del cuestionario a trabajadores, junto con corregir problemas de etiquetado, se ha implementado la siguiente edición:

- G11_02: se ha asignado el valor 85 en aquellas unidades con un valor igual a uno en G10, y sin información válida en G11_02.

Por último, como regla general para todos los cuestionarios, los valores de las variables de selección múltiple, que originalmente registraban un valor 0 para “No” y un 1 para “Sí”, han sido recodificados para que el “No” corresponda al número 2, tal como en el resto de las preguntas de los cuestionarios.

VII.9. Creación de variables por Registros Administrativos (RRAA)

La DT solicitó la creación de una serie de variables asociadas a la ocurrencia de huelgas, fiscalizaciones y sanciones en los últimos años en las empresas que forman parte de la muestra de la Encla 2023 para ser incorporadas en las bases de datos oficiales de la encuesta. Sin embargo, su incorporación fue evaluada durante el proceso de anonimización de bases de datos, ya que el nivel de detalle que muestran estas variables hacía inviable su inclusión conservando niveles de riesgo de identificación razonables para el producto estadístico.

Las variables que finalmente se incorporaron a las bases de datos solo se presentan en la versión de uso interno y se excluyeron de las bases de uso público. Estas son descritas en la tabla a continuación:

Tabla 60. Variables a crear en las 4 bases de datos Encla 2023

Variables creadas en Encla 2023
h2020. Existencia de huelgas en la empresa en 2020
h2021. Existencia de huelgas en la empresa en 2021
h2022. Existencia de huelgas en la empresa en 2022
h2023. Existencia de huelgas en la empresa en 2023
fd2022. Número de fiscalizaciones por denuncia realizadas a la empresa en 2022
fp2022. Número de fiscalizaciones por programa realizadas a la empresa en 2022
sa2022. Número de sanciones impuestas a la empresa en 2022

Las variables descritas en la tabla anterior son vinculadas a las bases de datos a través del RUT presente en ambos registros (a saber, datos recolectados durante el proceso estadístico y registro administrativo provisto por la Dirección del Trabajo). Es importante destacar que estos datos en particular no fueron parte del diseño muestral, y tampoco fueron sometidos a tratamientos de calibración posteriores a su captura(más allá de las recodificaciones realizadas para resguardar la anonimidad de las unidades económicas). Por lo tanto, pueden existir diferencias entre los stocks poblacionales de la distribución de respuestas en estas variables en la Encla en comparación a los totales capturados en los registros administrativos.

Estos datos fueron traspasados al INE en el marco del Convenio de Producción Estadística "Diseño de Cuestionarios, Verificación y de la Muestra Objetivo, Ejecución de la Marcha Blanca y el Levantamiento Oficial, Encla 2023" celebrado entre ambas instituciones durante enero de 2023.

VIII. Metodología de Análisis de Resultados

VIII.1. Validación de resultados

- *Estándar para la evaluación de la calidad de las estimaciones de encuestas económicas*: el análisis de resultados de la Décima Encuesta Laboral sigue los lineamientos definidos por el Departamento de Metodología e Innovación Estadística con apoyo de la Subdirección Técnica, y que establece criterios para definir la fiabilidad de las estimaciones y los lineamientos para publicación de tabulados.

En resumen, el *Estándar para la evaluación de la calidad de las estimaciones de encuestas económicas* define un conjunto de requisitos estadísticos para clasificar una estimación en tres posibles categorías de fiabilidad: estimaciones fiables, estimaciones poco fiables, y estimaciones no fiables. Los tabulados estadísticos deben presentar más del 50% de estimaciones fiables para ser publicables. En caso contrario, las variables y/o las categorías que las conforman deben reformularse o descartarse del análisis.

Se generó un listado original de 65 indicadores para formar parte de la Síntesis de Resultados Encla 2023. Estos indicadores luego fueron revisados en virtud del Estándar descrito arriba y, según el cumplimiento en términos de fiabilidad estadística, se fueron reformulando o descartando. Finalmente, el listado de indicadores para Síntesis de Resultados Encla 2023 fue de 50 indicadores. Pueden encontrarse más detalles en el Reporte de Calidad, sección *III.2.5. Calidad de las Estimaciones*.

- *Análisis de comportamiento de resultados*: con el objetivo de explorar los datos a nivel agregado, se realizó un análisis descriptivo de las principales variables pertenecientes a los cuatro cuestionarios en estudio. El propósito de este ejercicio fue observar el comportamiento de los resultados en términos a nivel muestral y poblacional. En el caso del cuestionario autoaplicado, se llevó a cabo un análisis de seguimiento de las ediciones ejecutadas en la muestra en los diversos subprocesos del procesamiento de datos, así como un análisis univariado de los datos finales post-procesamiento a nivel poblacional. En los cuestionarios a empleadores, trabajadores y sindicatos, sólo se elaboraron análisis poblacionales de la distribución de las respuestas a cada una de las preguntas de los cuestionarios.

- *Comparación con registros administrativos*: con el objetivo de ejecutar un análisis de consistencia, cuando fue posible, se contrastaron los resultados de las variables principales de cada cuestionario con información proveniente de los registros administrativos del Sistema Informático de Relaciones Laborales (SIRELA) de la Dirección del Trabajo. Este registro incluye información sobre organizaciones sindicales y procesos de negociación colectiva en las empresas del país. La consistencia de los resultados con la información contenida en dichas fuentes fue considerada como un indicador de coherencia. Es preciso señalar que este análisis fue realizado durante la ejecución del procesamiento, de manera tal que su resultado contribuyó con la validación de datos.

Los registros de SIRELA fueron un insumo fundamental para validar la correcta recolección de los cuestionarios a trabajadores y a sindicatos. En primer lugar, esta información se utilizó durante los procesos de verificación y recolección, de modo que las personas recolectoras contasen con antecedentes de existencia previa o actual de sindicatos en cada empresa. Asimismo, se definió supervisar aquellas unidades donde existiesen registros de sindicatos y se hubiere recolectado el cuestionario a trabajadores.

Por otra parte, los registros de SIRELA se utilizaron para corregir inconsistencias halladas en el procesamiento de datos. Se debe señalar que la comparación con los registros no sirvió como fuente de identificación de inconsistencias en estos casos. En específico, se identificaron inconsistencias y valores de rango en información declarada por los informantes[39], que fueron contrastados con la información del registro de SIRELA, manteniendo o editando la información original según cada caso.

[39] Por ejemplo, negociaciones colectivas con una duración mayor a 60 semanas; sindicatos con más socios(as) que personas contratadas directamente por la empresa; más trabajadores suscritos a contratos colectivos que trabajadores(as) asociados a sindicatos, entre otros.

VIII.2. Aplicación del control de divulgación

La Encla 2023 busca ser un aporte tanto a la discusión de políticas públicas relacionadas al mundo del trabajo en Chile, como también, un insumo para las y los investigadores y académicos de las distintas realidades laborales del país. Para esto, tanto el Instituto Nacional de Estadística (INE) como la Dirección del Trabajo (DT) deben conciliar la transparencia de sus datos, para que estos efectivamente puedan ser utilizados por investigadores y ciudadanos en general, con el anonimato e indeterminación de sus unidades de análisis y de información.

Calibrar estos dos aspectos resulta importante debido a que la normativa nacional los regula: por una lado la ley N° 20.285 en su artículo 11 establece que "(…) toda persona goza del derecho a acceder a la información que obre en poder de los órganos de la Administración del Estado (…)", mientras que, por otro lado, se establece el deber de los particulares de responder las encuestas del INE y el derecho de que se mantenga su identidad en el anonimato, mediante las normativas del secreto estadístico (ley N° 17.374, artículo 29) y la protección de la vida privada (ley N° 19.628, artículo 2).

Pero hay también consideraciones éticas y prácticas que llevan a la necesidad de anonimizar efectivamente, como la protección del patrimonio nacional estadístico: "(la anonimización) es importante porque garantiza la calidad de los datos y las tasas de respuesta en futuras encuestas. Si los encuestados sienten que los productores de datos no están protegiendo su privacidad, es posible que no estén dispuestos a participar en futuras encuestas" (Benschop et al. 2019: 12).

A fin de cumplir con los estándares existentes, se siguieron una serie de procedimientos enfocados en garantizar el anonimato de las empresas y de las personas informantes en las bases de datos. Para esto se utilizaron métodos del Control Estadístico del Riesgo (Statistical Disclosure Control, SDC) que permiten conocer la probabilidad de que las identidades de las unidades de las bases de datos sean descubiertas con variables de identificación indirectas, en base a escenarios hipotéticos en los que intrusos cuentan con datos externos que pueden operar como "llaves" (keys variables) para acceder a la identidad de las empresas de la muestra (Benschop et al. 2019).

Por ejemplo: la DT publica diariamente un registro de huelgas legales, en base al cuál puede elaborarse una lista de las 341 empresas que tuvieron huelgas entre 2020 y 2022. Por otro lado, la Encla 2023 contiene una variable en donde se pregunta si en la empresa ocurrieron huelgas entre los años 2020-2022, la cual fue respondida positivamente por 56 informantes (variable l1_08 en el caso del cuestionario a sindicatos). Considerando algunos datos adicionales externos públicos que también son preguntados por la Encla 2023, los métodos SDC calculan la probabilidad de identificar las 56 empresas que respondieron positivamente a la huelga en Encla con las 341 registradas con huelgas entre 2020 y 2022 según el dato externo de la DT. Los métodos SDC permiten ser conscientes de estos riesgos y ofrecen una serie de formas para disminuirlos, a fin de que se liberen bases de datos "seguras", en el sentido de que resguardan la identidad de sus unidades de análisis y las personas informantes.

IX. Indicadores para la medición y aseguramiento de la calidad estadística

IX.1. Indicadores operativos de contacto.

Para evaluar la gestión de las unidades durante el proceso de recolección, se estimaron los siguientes indicadores, cuyos resultados son presentados en el informe de calidad:

IX.1.1. Tasa de respuesta

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

$$TR_{np} = \left[\frac{(E+P)}{(E+P+R+NC+O+ED)}\right] * 100$$

Donde:

TR_{np} : Tasa de respuesta.

P : Número de unidades con entrevista parcial.

R : Número de unidades que rechazaron la entrevista.

NC : Número de unidades no contactadas.

O : Número de unidades no entrevistadas por otros motivos.

ED : Número de unidades con elegibilidad desconocida.

IX.1.2. Tasa de contacto

La tasa de contacto es la proporción de casos de la muestra en los que el encuestador/a logra algún grado de contacto (hablar o captar la atención de la persona entrevistada), independiente del resultado final de la encuesta. Se calcula a través de la siguiente ecuación:

$$TCON = \left[\frac{(E+P)+R+O}{(E+P)+R+NC+O+ED}\right] * 100$$

Donde:

$TCON$: Tasa de contacto.

E : Número de unidades con entrevista completa.

P : Número de unidades con entrevista parcial.

R : Número de unidades que rechazaron la entrevista.

O : Número de unidades no entrevistadas por otros motivos.

NC : Número de unidades no contactadas.
ED : Número de unidades con elegibilidad desconocida.

IX.1.3. Tasa de rechazo

Describe el número de unidades donde se obtiene un rechazo, sobre el total de unidades muéstrales a encuestar que son potencialmente elegibles (entrevistadas, rechazadas, no contactadas y no entrevistadas por otros motivos) y aquellas de elegibilidad desconocida, en un período de tiempo determinado.

$$TREC = \left[\frac{R}{(E+P)+R+NC+O+ED}\right] * 100$$

Donde:

$TREC$: Tasa de rechazo.

R : Número de unidades que rechazaron la entrevista.

E : Número de unidades con entrevista completa.
P : Número de unidades con entrevista parcial.

NC : Número de unidades no contactadas.
O : Número de unidades no entrevistadas por otros motivos.
ED : Número de unidades con elegibilidad desconocida.

IX.1.4. Tasa de cooperación

Este indicador describe el número de unidades donde se realizó una entrevista completa y/o parcial, sobre el total de unidades muestrales elegibles y contactadas de la muestra (incluyendo entrevistadas, rechazos y otras unidas no entrevistadas que pueden haber sido contactadas) en un período de tiempo determinado. Se calcula con la siguiente fórmula:

$$TCOOP = \left[\frac{(E+P)}{(E+P)+R+O}\right] * 100$$

Donde:

$TCOOP$: Tasa de cooperación.

E : Número de unidades con entrevista completa.

P : Número de unidades con entrevista parcial.

R : Número de unidades que rechazaron la entrevista.

O : Número de unidades no entrevistadas por otros motivos.

IX.2. Indicadores de supervisión.

Por otra parte, para evaluar el trabajo realizado en lo que respecta a la supervisión, se planificó estandarizar dos indicadores operativos para los diferentes procesos de supervisión, subdividiéndolos cuando sea necesario:

- Indicador de gestión de supervisión: se definió como la cantidad de unidades asignadas a las personas supervisoras sobre el estimado correspondiente de unidades de ese proceso de supervisión expresado en porcentaje.
- Indicador de logro de supervisión: se definió como la cantidad de unidades logradas por las personas supervisoras sobre el estimado correspondiente de unidades de ese proceso de supervisión expresado en porcentaje.

IX.2.1. Acompañamiento Operativo:

Como se mencionó anteriormente, el equipo de supervisión no contó con panel de control ni base de datos estandarizada. Para visualizar los indicadores de Acompañamiento Operativo (Tablas 61 y 62) se crearon planillas de control en Excel que fueron alimentadas por las bases de datos de *SuSo* y las gestiones de la persona encargada de área.

Tabla 61. Indicador de gestión de acompañamiento operativo

Indicador	Definición
Indicador de gestión de acompañamiento operativo	Cantidad de Formularios de Acompañamiento operativo asignados a las personas encargadas de grupo sobre cantidad de personas recolectora activas en el proyecto.
$x = \frac{Formularios\ de\ acompañamiento\ operativo\ asignados}{\text{Personas recolectora activa}} * 100$	

Tabla 62. Indicador de logro del Acompañamiento Operativo

Indicador	Definición
Indicador de logro de acompañamiento operativo	Cantidad de Formularios de Acompañamiento operativo logrados por las personas encargadas de grupo sobre cantidad de formularios de acompañamiento operativo asignados.
$x = \frac{Formularios\ de\ acompañamiento\ operativo\ logrados}{\text{Formularios de acompañamiento operativo asignados}} * 100$	

IX.2.2. Supervisión de Autoaplicado:

A diferencia de los otros dos procesos de supervisión, esta no contó inicialmente con un formulario en *SuSo*, por lo que se debieron implementar los indicadores directamente desde el SGOE, pero con el retraso de esto, se tuvo que diseñar un formulario en *SuSo* que supliera las funcionalidades del SGOE, por lo que los indicadores establecidos (tablas 63 y 64), al igual que los otros indicadores– se expresaron en una planilla Excel.

Tabla 1. Indicador de gestión de supervisión de autoaplicado

Indicador	Definición
Indicador de gestión de supervisión de autoaplicado	Cantidad de Formularios de supervisión gabinete revisión autoaplicado asignados a las personas supervisoras sobre cuestionario de autoaplicado en estado de enviadas, recepcionadas, en análisis, terminada por procedimiento y terminada en SGOE.
$x = \frac{Formularios\ de\ supervision\ gabinete\ revision\ autoaplicado\ asignados}{\text{Cuestionarios de autoaplicado Enviadas, Recepcionadas, En Analisis, Terminada por Procedimiento y Terminada}} * 100$	

Tabla 64. Indicador de logro de supervisión de autoaplicado

Indicador	Definición
Indicador de logro de supervisión de autoaplicado	Cantidad de Formularios de supervisión gabinete revisión autoaplicado revisado por las personas supervisoras sobre cuestionario de autoaplicado en estado de enviadas, recepcionadas, en análisis, terminada por procedimiento y terminada en SGOE.
$x = \frac{Formularios\ de\ supervision\ gabinete\ revision\ autoaplicado\ logrado}{\text{Cuestionarios de autoaplicado Enviadas, Recepcionadas, En Analisis, Terminada por Procedimiento y Terminada}} * 100$	

IX.2.3. Supervisión de unidades No elegibles

Al igual que la supervisión anterior, esta no contaba con un formulario en *SuSo* y se debía implementar directamente el SGOE, pero con el retraso de esto, se implementó un formulario en *SuSo* que supliera las funcionalidades del SGOE y así lograr implementar los indicadores establecidos (Tabla 65) en una planilla Excel.

Tabla 65. Indicador de gestión de supervisión de no elegible

Indicador	Definición
Indicador de gestión de supervisión de No elegible	Cantidad de Formularios de supervisión gabinete revisión no elegible asignados a las personas supervisoras sobre cuestionario de empleador con código 4110 (ruptura), 4120 (reestructuración del grupo de empresas), 4130 (fusión), 4140 (cese definitivo de actividades), 4210 (paralizada temporalmente), 4221 (cambio de actividad económica principal), 4222 (otros cambios en la unidad para considerarla fuera de ámbito), 4230 (duplicada) y 4300 (otra razón para ser considerada no elegible) aprobados por supervisor.
$x = \frac{Formularios\ de\ supervision\ gabinete\ revision\ \ No\ elegible\ asignados}{\text{Cuestionarios empleador con codigo } 4110, 4120, 4130, 4140, 4210, 4222, 4230 \text{ y } 4300} * 100$	

IX.3. Indicadores de edición e imputación

Finalmente, se estimaron indicadores de imputación y edición para medir el impacto de las reglas definidas sobre los datos recolectados.

IX.3.1. Tasa de ediciones de secuencia

La tasa de ediciones de secuencia describe el número de ediciones de secuencia implementadas en cada módulo, dividido en el producto entre el número de variables por módulo y el número de unidades logradas, multiplicado por 100. La fórmula es la siguiente:

$$TE_s = \left[\frac{Es}{(V * L)}\right] * 100$$

Donde:

TE_s : Tasa de ediciones de secuencia.

E_s : Número de ediciones de secuencia.

V : Número de variables por módulo.

L : Número de unidades logradas.

IX.3.2. Tasa de ediciones de cuadratura

La tasa de ediciones de cuadratura describe el número de ediciones de cuadratura[40] implementadas en cada módulo, dividido en el producto entre el número de variables que forman parte de matrices por módulo y el número de unidades logradas, multiplicado por 100. La fórmula es la siguiente:

$$TE_c = \left[\frac{Ec}{(Vm * L)}\right] * 100$$

Donde:

TE_c : Tasa de ediciones de cuadratura.

E_c : Número de ediciones de cuadratura implementadas en cada módulo.

V_m : Número de variables que forman parte de matrices por módulo.

L : Número de unidades logradas.

IX.3.3. Tasa de ediciones intercuestionarios[41]

La tasa de ediciones intercuestionarios describe el número de ediciones intercuestionarios implementadas en cada módulo, dividido en el producto entre el número de variables consideradas en las ediciones intercuestionarios por módulo y el número de unidades logradas, multiplicado por 100. La fórmula es la siguiente:

$$TE_i = \left[\frac{Ei}{(Vi * L)}\right] * 100$$

Donde:

TE_i : Tasa de ediciones intercuestionarios.

E_c : Número de ediciones de cuadratura implementadas en cada módulo.

V_m : Número de variables que forman parte de matrices por módulo.

L : Número de unidades logradas.

IX.3.4. Tasa de imputación

La tasa de imputación describe el número de imputaciones implementadas en cada módulo, dividido en el producto entre el número de variables numéricas por módulo y el número de unidades logradas, multiplicado por 100. La fórmula es la siguiente:

[40] La edición de una celda equivale a una edición de cuadratura.

[41] Como se indicó previamente, las ediciones intercuestionarios sólo se implementaron entre los cuestionarios autoaplicado y a empleadores.

$$TI = \left[\frac{I}{(Vn * L)}\right] * 100$$

Donde:

TI : Tasa de imputación.

I : Número de imputaciones implementadas en cada módulo.

V_n : Número de variables numéricas por módulo.

L : Número de unidades logradas.

X. Difusión

Tanto los resultados como los metadatos serán puestos a disposición de las personas usuarias a través del sitio web del Instituto Nacional de Estadísticas como del sitio de la Dirección del Trabajo[42].

Bases de datos: a diferencia de versiones anteriores, la Décima Encuesta Laboral Encla 2023 es la primera vez que cuenta con tres bases de datos oficiales, en lugar de cuatro. La primera corresponde a la base de datos del formulario autoaplicado; la segunda a la base de datos del cuestionario a empleadores; y la tercera a la base de datos del cuestionario a Trabajadores y Sindicatos. Estas dos últimas fueron unidas para facilitar el proceso de uso de los datos; sobre todo en vista que entre ambos cuestionarios existe una gran cantidad de preguntas compartidas.

De esta forma, ya no es necesario para las y los usuarios unir ambas bases (Trabajadores y Sindicatos) para realizar análisis que incluya al total de empresas en la muestra sobre preguntas provenientes de los cuestionarios a Trabajadores y/o Sindicatos.

Las bases serán puestas a disposición de los usuarios y usuarias en los siguientes formatos: .rds, .dta y .sav. La conversión de los formatos de las bases de datos se realizó mediante el *software R Studio*, específicamente con el paquete `haven` (Wickham y Miller 2019) y `labelled` para el etiquetado.

[42] Según lo definido en el convenio de colaboración interinstitucional, la publicación se realizará primeramente en el sitio web de la DT y, posteriormente, en la web del INE.

Síntesis de resultados: Se publicará en fecha por confirmar. Existen dos versiones de la síntesis de resultados; una realizada por la Dirección del Trabajo, que se aloja en su micrositio Encla[43] y otra realizada por el INE, que también hará disponible a los usuarios y usuarias en su micrositio respectivo.[44]

Los metadatos que acompañan los resultados de la operación estadística entregando información sobre los datos producidos, describiéndolos, caracterizándolos y permitiendo una mejor comprensión de estos son los siguientes:

- Manuales de usuarios de bases de datos: enfocado en orientar el uso de las bases de datos, contiene una descripción general de las bases de datos y de las variables e indicadores en ellas contenidos.
- Manual metodológico: documentación del proceso estadístico.
- Ficha metodológica: formulario que resume el diseño metodológico del estudio.
- Reporte de calidad: documentación de indicadores sobre el aseguramiento de la calidad del proceso y producto estadístico.
- Cuestionarios (ver anexo II)

[43] https://www.dt.gob.cl/portal/1629/w3-propertyvalue-188830.html
[44] https://www.ine.gob.cl/estadisticas/sociales/mercado-laboral/condiciones-de-empleo-y-relaciones-laborales

XI. Glosario

La fuente de estas definiciones proviene del Centro de Consultas de la Dirección del Trabajo https://www.dt.gob.cl/portal/1628/w3-channel.html

- Comité bipartito de capacitación Tiene por objetivos acordar y evaluar el o los programas de capacitación ocupacional de la empresa, así como asesorar a la dirección de la misma en materias de capacitación. Toda empresa con más de 15 trabajadores debe constituir un comité bipartito de capacitación, integrado por representantes de la empresa y de los trabajadores.

- Comité paritario de higiene y seguridad Es un comité que tiene por objetivos asesorar e instruir a los trabajadores en materias de higiene y seguridad (correcta utilización de los instrumentos de protección, vigilancia del cumplimiento de las medidas de prevención, higiene y seguridad; investigación de las causas de los accidentes del trabajo y enfermedades profesionales que se produzcan en la empresa, entre otros). Todas las empresas con más de 25 trabajadores deben organizar un comité paritario de este tipo, el que estará integrado por representantes de la empresa y de los trabajadores.

- Conflicto colectivo Corresponde a un conflicto laboral que involucra a varios trabajadores, quienes actúan colectivamente para enfrentar al empleador. Conflicto laboral Corresponde a un conflicto entre trabajador y empleador por causas propias del proceso de trabajo. Puede involucrar a uno o más trabajadores.

- Contrato colectivo Es un acuerdo o documento suscrito por los trabajadores y la empresa, en el marco de una negociación, y en el que se estipulan las condiciones de trabajo y de remuneraciones pactadas por la empresa y los trabajadores. Este instrumento está sometido a un procedimiento reglado, está sometido a plazos y la ley establece los mecanismos a los que los trabajadores pueden recurrir en caso de no llegar a acuerdo con el empleador, como es el caso de la huelga.

- Contrato individual de trabajo Es una convención por la cual el empleador y el trabajador se obligan recíprocamente, este a prestar servicios personales bajo dependencia y subordinación del primero, y aquel a pagar por estos servicios una remuneración determinada. Es consensuado entre trabajador y empleador y debe constar por escrito.

- Convenio colectivo Es un documento que tiene los mismos efectos del contrato colectivo, pero se diferencia por el hecho de que a él se llega mediante un procedimiento menos regulado que en el caso del contrato colectivo. Por ejemplo, un convenio puede negociarse en cualquier momento, no da derecho a fuero, no da derecho a huelga, etc.

- Mecanismos de participación Corresponde a instancias o iniciativas mediante las cuales la empresa permite la participación de los trabajadores en diversos temas (comités de bienestar, comités de cultura, comités de recreación, comités de capacitación, entre otros).

- Negociación colectiva Es un procedimiento regulado por la legislación laboral mediante el cual los trabajadores y los empleadores negocian o acuerdan condiciones de trabajo y de remuneraciones por un tiempo determinado. Los trabajadores pueden estar representados por uno o más sindicatos, o por grupos de trabajadores (grupo negociador). El resultado de una negociación colectiva es un instrumento colectivo (contrato o convenio.

- Reglamento interno de higiene y seguridad Es un reglamento al que están obligadas todas las empresas, independiente del número de trabajadores que tengan, por la Ley 16.744 (Art. 67) sobre seguro de accidentes del trabajo y enfermedades profesionales, y cuyo contenido corresponde a normas sobre prevención y seguridad en el trabajo.

- Reglamento interno de orden, higiene y seguridad Es un reglamento al que están obligadas las empresas que tengan 10 trabajadores o más, como resultado del artículo 153 del Código del Trabajo. Debe contener las obligaciones y prohibiciones a que deben estar sujetos los trabajadores, en relación con sus labores, permanencia y vida en las dependencias de la respectiva empresa o establecimiento, así como las normas que deben garantizar un ambiente laboral digno y de mutuo respeto entre los trabajadores.

- Relaciones laborales Se llama así al conjunto de relaciones que establecen en la empresa los trabajadores y los empleadores, a través de los órganos de representación que ellos se den (sindicatos, comités, grupos negociadores, etc.).

XII. Bibliografía

- Chiavenato, I. (2017). Administración de Recursos Humanos (10 ed.). McGraw-Hill.

- DARES. (2017). *Relations professionnelles et négociations d'entreprise*. https://dares.travail-emploi.gouv.fr/enquete-source/relations-professionnelles-et-negociations-dentreprise-2017

- Department for Business, Innovation & Skills. (2013, 23 enero). Workplace Employment Relations Study (WERS). GOV.UK. https://www.gov.uk/government/collections/workplace-employment-relations-study-wers

- Deville & Särndal, (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association.* 87(418), 376-382.

- Dirección del Trabajo (2024). *Estudios y estadísticas. Encuesta Laboral (Encla)*. https://www.dt.gob.cl/portal/1629/w3-propertyvalue-22780.html

- Instituto Nacional de Estadísticas (INE). (2021). *Guía para el control de divulgación estadística de microdatos.* https://www.ine.gob.cl/docs/default-source/buenas-practicas/directrices-metodologicas/guias-y-orientaciones-metodologicas/documentos/gu%C3%ADa-control-divulgaci%C3%B3n-estad%C3%ADstica-microdatos.pdf?sfvrsn=b6bdd28f_2

- Instituto Nacional de Estadísticas (INE). (2023). Instructivo para casos de detección de inconsistencias en operaciones de estadísticas sociales. Santiago de Chile: Instituto Nacional de Estadísticas.

- Organización Internacional del Trabajo. (2024). *Declaración de la OIT relativa a los principios y derechos fundamentales en el trabajo. https://www.ilo.org/es/declaracion-de-la-oit-relativa-los-principios-y-derechos-fundamentales-en*

- REPONSE. (2017). "*Relations professionnelles et négociations d'entreprise*". https://dares.travail-emploi.gouv.fr/dares-etudes-et-statistiques/enquetes/reponse2017

- Universidad Alberto Hurtado; Servicio Civil (2017). "*Encuesta anual para el diagnóstico de las relaciones laborales en la Administración Central del Estado (ENCLACE). Informe Resultados. Primera Aplicación.*" http://biblioteca.digital.gob.cl/bitstream/handle/123456789/806/2017%20Encuesta%20anual%20para%20el%20diagnostico%20de%20las%20relaciones%20laborales%20en%20la%20AC%20del%20Estado%20%28ENCLACE%29.pdf?sequence=1&isAllowed=y

XIII. Anexos

Anexo 1: Variables primarias y secundarias GDS Encla 2023

1. Variables primarias (logro exigido 100%)

Pregunta 2023	Cuestionario 2023	Enunciado	Observaciones
A4	**Auto aplicado**	Indique el total de personas contratadas directamente por esta empresa **al 30 de noviembre de 2022**, de acuerdo a las siguientes categorías de autoridad y sexo	
C1_1	**Auto aplicado**	**Durante el año 2022,** ¿cuáles de las siguientes actividades han sido subcontratadas a otras empresas o a terceros? (desarrolladas dentro o fuera de las dependencias de la empresa)	
C14	**Auto aplicado**	**Durante el año 2022**, ¿esta empresa contó con personas suministradas por terceros (empresas de servicios transitorios)?	
D1	**Auto aplicado**	Indique el total de personas contratadas directamente por la empresa con **jornada completa** (de 40 horas o más a la semana) **al 30 de noviembre de 2022** según tramo de remuneraciones y sexo	
I2	**Auto aplicado**	**Durante el año 2022**, ¿la empresa realizó capacitaciones laborales para su personal?	Agregada por la DT, **reemplaza a I3,** que no queda ni como variable primaria ni como variable secundaria. Lo que principalmente nos interesa es el porcentaje de empresas que capacitan.
O1	**Sindicatos / Trabajadores**	En esta empresa, ¿alguna vez se ha teletrabajado?	Dado que, como se pudo detectar en las entrevistas cognitivas, los trabajadores y dirigentes pueden desconocer las formas de trabajo de ciertas áreas de la empresa, las respuestas **"No sabe"** deben ser consideradas como respuestas válidas. Es decir, la **unidad NO** debe considerarse como **"no lograda"** si es que en esta pregunta se responde "No sabe".
M2	**Empleadores**	**Desde el año 2019 a la fecha** ¿ha habido alguna negociación colectiva en esta empresa?	En la propuesta de grilla de suficiencia enviada por INE, la **pregunta M2** incluida como variable primaria era aquella perteneciente al cuestionario de **Sindicatos**. Para efectos de contar con esta **información en todas las empresas**, con o sin sindicato, la variable que proponemos incluir corresponde al cuestionario de **Empleadores.**

K2	**Sindicatos**	¿Cuántos hombres y mujeres se encuentran afiliados a su sindicato **al día de hoy?**	

Las siguientes preguntas se eliminaron como variables primarias (y tampoco fueron convertidas a variables secundarias):

I3 (Auto aplicado); M2 (Sindicatos); K7 (Sindicatos).

2. Variables secundarias (logro exigido 80%)

Pregunta 2023	Cuestionario 2023	Enunciado	Observaciones
A7	Auto aplicado	**Durante el año 2022,** esta empresa, ¿exportó sus productos o servicios?	
A8	Auto aplicado	**Durante el año 2022,** ¿qué porcentaje de ventas exportó con respecto al total de ventas de ese año?	
C2	Auto aplicado	Indique el total de personas subcontratadas para la actividad económica principal de la empresa, **al 30 de noviembre de 2022,** según sexo.	
C3	Auto aplicado	Indique el total personas subcontratadas para otras actividades económicas de la empresa, distintas a la actividad económica principal, **al 30 de noviembre de 2022,** según sexo.	
C10	Auto aplicado	**Durante el año 2022,** ¿esta empresa fue subcontratada por otras empresas?	
D2	Auto aplicado	Señale a cuánto asciende el monto de la remuneración bruta mínima mensual y de la remuneración bruta máxima mensual pagados en la empresa **al 30 de noviembre de 2022**. Sólo considere al personal con jornada completa (de 40 horas o más a la semana) y contratado directamente.	
D4	Auto aplicado	Indique el total de personas contratadas directamente por la empresa con **jornada parcial** (30 horas o menos a la semana) **al 30 de noviembre de 2022,** según tramo de remuneraciones brutas y sexo.	
D5	Auto aplicado	De la planilla de remuneraciones mensuales correspondientes al mes de noviembre de 2022, indique el porcentaje que representan las remuneraciones fijas y variables brutas, respecto a la suma total de remuneraciones de personas contratadas directamente por la empresa	Esta pregunta la **cambiamos de categoría, desde variable primaria a variable secundaria**. Si bien es relevante, no lo es hasta el punto de que la empresa deba considerarse como "no lograda" si es que no es respondida.

D6	Auto aplicado	**Durante el año 2022,** ¿esta empresa pagó regularmente incentivos, como bonos, premios u otros para aumentar la producción y/o la productividad?	
E2	Auto aplicado	Indique el número de personas extranjeras contratadas directamente por la empresa al **30 de noviembre de 2022,** según nacionalidad y sexo.	
E3	Auto aplicado	Indique el número de personas extranjeras contratadas directamente por la empresa **al 30 de noviembre de 2022**, según nacionalidad y tipo de contrato o relación laboral	**Agregada por DT.** Debido a que consideramos el tipo de contrato de personas extranjeras como un indicador clave en la encuesta.
F1	Auto aplicado	Indique el total de personas contratadas directamente por la empresa, **al 30 de noviembre de 2022,** que se desempeñaban bajo las siguientes jornadas, según sexo	Esta pregunta la **cambiamos de categoría, desde variable primaria a variable secundaria.** Si bien es relevante, no lo es hasta el punto de que la empresa deba considerarse como "no lograda" si es que no es respondida.
F8	Auto aplicado	Respecto a las personas contratadas directamente por la empresa que trabajaron **los días domingo del mes de noviembre de 2022**, indique el número de mujeres y hombres que lo hicieron en las siguientes labores:	**Agregada por DT**. Debido a que consideramos el trabajo en día domingo un indicador clave de la encuesta.
F11	Auto aplicado	Indique el número de personas **contratadas directamente por la empresa** que trabajaron durante la noche **en el mes de noviembre de 2022, según sexo.**	
H1	Auto aplicado	**Durante el año 2022** ¿cuántas personas contratadas directamente por la empresa hicieron uso de los siguientes derechos de protección a la maternidad y paternidad?	
M2	Auto aplicado	¿Cuántos instrumentos colectivos o acuerdos de grupo negociador vigentes existen en la empresa?	
M4	Auto aplicado	Para los instrumentos colectivos o acuerdos de grupo negociador **actualmente** vigentes, indique el número de entidad(es) negociadora(s), según tipo.	
P1	Auto aplicado	¿Cuál de las siguientes formas de digitalización utiliza la empresa, e indique si estas se implementaron como consecuencia de la pandemia del Coronavirus (entre marzo de 2020 y diciembre de 2021)?	**Agregada por DT**. Considerar como variable secundaria sólo la pregunta "¿La empresa la utiliza? y NO la pregunta ¿se implementó como consecuencia de la pandemia?
P5	Auto aplicado	¿Cuál de las siguientes formas de automatización utiliza la empresa, e indique si estas se implementaron como consecuencia de la pandemia del Coronavirus (entre marzo de 2020 y diciembre de 2021)?	Considerar como variable secundaria sólo la pregunta "¿La empresa lo ha implementado? y **NO** la pregunta ¿se implementó a **causa de la pandemia**?

O18	Auto aplicado	Del personal contratado directamente por la empresa, ¿cuántos trabajadores y trabajadoras realizaban sus labores mediante la modalidad de teletrabajo en los siguientes periodos?	
N1	Auto aplicado	**Para los años 2020 y 2021**, indique el número de trabajadoras y trabajadores contratados directamente por la empresa que se acogieron a la Ley N° 21.227 de protección del empleo, de acuerdo con la situación en la que se encontraban:	
D1	Empleadores	**Durante el año 2022**, y sin tomar en cuenta el reajuste legal de los salarios mínimos, ¿la empresa realizó algún reajuste de las remuneraciones para la mayoría (más del 50%) del personal?	**Agregada por DT.** Debido a que consideramos el reajuste legal de los salarios un indicador clave de la encuesta.
F1	Empleadores	**De marzo de 2020 a diciembre de 2021,** indique cuáles de las siguientes medidas que flexibilizan la jornada de trabajo fueron implementadas durante la pandemia del coronavirus:	
J3	Empleadores	Indique si actualmente tiene contratadas personas de los siguientes grupos y, de tenerlas, señale si ha tomado medidas para favorecer su integración en la empresa.	Esta pregunta proponemos cambiarla de categoría, **desde variable primaria a variable secundaria** debido a que en esta versión la formulación es más compleja que la realizada en 2019, lo cual conlleva un riesgo mayor de no respuesta.
J6	Empleadores	¿Cuánto confía usted en…	
J7	Empleadores	Pensando en los sindicatos en general, usted cree que:	
J10	Empleadores	Pensando en el sindicato mayoritario (con mayor número de trabajadores/as afiliados/as), usted cree que:	
J11	Empleadores	Durante el año 2022, ¿algún representante de la empresa sostuvo reuniones con el(los) sindicato(s) para abordar alguna de las siguientes cuestiones?	Esta pregunta proponemos **cambiarla de categoría, desde variable primaria a variable secundaria**. Si bien es relevante, no lo es hasta el punto de que la empresa deba considerarse como "no lograda" si es que no es respondida.
J17	Empleadores	La calificación de servicios mínimos que tiene la empresa fue obtenida:	
R4	Empleadores	¿Cuál es su cargo en esta empresa?:	
C2	Sindicatos/Trab.	Ese personal, ¿fue empleado para realizar alguna labor similar a las del personal contratado directamente por la empresa?	
C3	Sindicatos/Trab.	El personal subcontratado y/o suministrado por terceros que realizó laborales similares a las del personal contratado directamente por la empresa, tuvo:	
D1	Sindicatos	**Durante el año 2022,** ¿el sindicato realizó alguna evaluación sobre el nivel de las remuneraciones del personal contratado directamente por esta empresa en comparación con el personal	

		de otras empresas del mismo rubro y similar tamaño?	
D2	Sindicatos	Respecto a la pregunta anterior, usted diría que esta empresa comparada con otras del mismo rubro y similar tamaño tiene remuneraciones:	
F3	Sindicatos/Trab.	Indique la frecuencia en que se trabaja en los siguientes tipos de labores **los días domingo**.	
F5	Sindicatos/Trab.	Una vez asignados los turnos de trabajo al personal, ¿con qué frecuencia son modificados unilateralmente por la empresa?	
G2	Sindicatos/Trab.	**Durante el año 2022**, ¿cuáles de las siguientes acciones de prevención de riesgos ocupacionales ejecutó esta empresa?	
I2/I1	Sindicatos/Trab.	**Durante el año 2022**, ¿la empresa realizó capacitaciones laborales para su personal?	
J4-J5	Sindicatos/Trab.	**Durante el año 2022,** ¿esta empresa **informó/consultó** formalmente a su sindicato sobre las siguientes materias?	
J6	Sindicatos/Trab.	**Durante el año 2022**, ¿esta empresa consultó formalmente a su(s) sindicato(s) sobre las siguientes materias relacionadas a las condiciones de trabajo en la empresa?	
J12-J10	Sindicatos/Trab.	Una vez adquiridos acuerdos con los trabajadores y las trabajadoras, usted diría que, la mayoría de las veces, esta empresa los:	**Agregada por DT J12** de sindicatos para ser coherentes con la **inclusión de J10 de trabajadores**.
J13	Sindicatos	**Durante el año 2022**, ¿presenció o supo de alguna de las siguientes acciones en la empresa? Recuerde que toda la información entregada es confidencial.	
K1	Sindicatos	**Actualmente** el sindicato al que usted representa es un:	
K4	Sindicatos	Indique el sexo de las personas que ocupan los siguientes cargos en el directorio sindical	INE propone eliminar esta pregunta de la GDS. DT está de acuerdo en que no amerita considerarla como variable primaria, **pero sí es lo suficientemente relevante** como para **considerarla variable secundaria.**
K6	Sindicatos	**Durante el año 2022,** ¿cuáles de las siguientes actividades realizó su sindicato?	
K8	Sindicatos	¿De qué forma colaboró la empresa con su sindicato?:	
K9	Sindicatos	**Durante el año 2022,** de las siguientes acciones de defensa de los derechos de los trabajadores y las trabajadoras, ¿qué acciones realizó su sindicato?	
K10	Sindicatos	**Durante el año 2022,** de las siguientes acciones de participación política y social, ¿qué acciones realizó su sindicato?	

K15	Sindicatos	¿Cuál es la actitud más frecuente de esta empresa respecto de su sindicato?	
K16	Sindicatos	**Durante el año 2022,** ¿la empresa ha realizado alguna de las siguientes prácticas antisindicales o desleales?	
K19	Sindicatos	En su opinión, ¿por qué no hay más trabajadores y trabajadoras en su sindicato?:	
M6	Sindicatos	La última negociación colectiva de su sindicato dio origen a un:	
M16	Sindicatos	**Sin considerar el reajuste legal de los salarios mínimos,** ¿cómo fue el reajuste de las remuneraciones realizado por la empresa como resultado de la última negociación colectiva que llevó a cabo su sindicato?	
M21	Sindicatos	¿Esta empresa tiene calificación de servicios mínimos?	
R6 - DI01_i	Sindicatos	¿Cuánto tiempo lleva trabajando en esta empresa?	
R7 - DIO1_j	Sindicatos	¿Cuánto tiempo lleva ejerciendo como dirigente sindical? En cualquier empresa o institución.	
R8 - DI01_k	Sindicatos	¿En qué año se afilió a este sindicato?	
K1	Trabajadores	¿Hubo sindicato anteriormente en esta empresa?	
K3	Trabajadores	**Desde 2022 a la fecha,** ¿se han hecho intentos para formar un sindicato?	
K5	Trabajadores	En su opinión, ¿cuál es **la razón principal** por la que no se ha formado un sindicato? Piense en la más importante:	
M7	Trabajadores	¿Cómo fue el reajuste de las remuneraciones realizado por la empresa como resultado de la última negociación colectiva?	

Las siguientes preguntas consideradas variables secundarias en 2019 se eliminan de la grilla de suficiencia:

P6 (Auto aplicado); D2 (Empleadores); K12 (Sindicatos).

Anexo II

Cuestionarios:

Por tratarse de cuatro cuestionarios y bastante extensos, se presenta a continuación las caratulas de cada uno de ellos, y el enlace a su ubicación en micrositio Encla:

Autoaplicado

Empleadores

Dirigentes Sindicales

Trabajadores

INSTITUTO NACIONAL DE ESTADÍSTICAS
CONVENIO DIRECCIÓN DEL TRABAJO - INE
Subdirección de Operaciones
Departamento de Estadísticas Sociales
Morandé 801, piso 22, Santiago - Chile
Teléfonos: (56) 2 892 4049
https://www.ine.gob.cl

IDR

ENCLA

ENCUESTA LABORAL 2023 (ENCLA) AÑO 2023

ENCUESTA LABORAL 2023
AUTOAPLICADO (ENCLA 2023)

ENCUESTA LABORAL 2023
ENCLA 2023

PRESENTACIÓN

La Encuesta Laboral (ENCLA) 2023 tiene como propósito principal caracterizar las condiciones de trabajo y las relaciones laborales al interior de las empresas del país. La presente versión de la encuesta es fruto de un convenio de colaboración entre la **Dirección del Trabajo y el Instituto Nacional de Estadísticas (INE)**; este último es el responsable del levantamiento de la información en terreno a una muestra de empresas seleccionadas aleatoriamente.

Toda la información necesaria para este estudio se obtiene mediante la aplicación de tres cuestionarios: uno dirigido a un(a) representante de la empresa con conocimiento detallado sobre la planta de trabajadores(as), otro dirigido a empleadores(as) y un cuestionario dirigido a un(a) representante de los(as) trabajadores(s) de esta. Las respuestas son totalmente anónimas y los datos de la empresa están protegidos por secreto estadístico (Ley N°17.374).

El presente cuestionario es el que está dirigido a un(a) representante de la empresa y tiene por objetivo caracterizar a los trabajadores(as) de esta. El llenado del cuestionario lo debe realizar directamente la persona informante, para lo cual puede contar con la asistencia de un(a) analista del INE, quien se contactará con la persona responsable de gestionar el llenado de este cuestionario y la guiará en lo que sea necesario para asegurar que sea completado en los plazos indicados desde el momento en que el INE tomó contacto con esta empresa.

INSTITUTO NACIONAL DE ESTADÍSTICAS
CONVENIO DIRECCIÓN DEL TRABAJO - INE
Subdirección de Operaciones
Departamento de Estadísticas Sociales
Morandé 801, piso 22, Santiago, Chile
Teléfono (56) 2 892 4049
www.ine.cl

IDR

CONFIDENCIAL

Esta encuesta está autorizada por la ley N° 17.374. Toda información se mantendrá con carácter estrictamente confidencial, bajo secreto estadístico.

ENCLA

ENCUESTA LABORAL 2023 (ENCLA) AÑO 2023

ENCUESTA LABORAL 2023 - SINDICATOS ENCLA 2023

ENCUESTA LABORAL 2023
SINDICATOS (ENCLA 2023)

IDENTIFICACIÓN DE LA EMPRESA

RUT EMPRESA									-	

LUGAR DE LA ENTREVISTA	
CALLE, AVENIDA, PASAJE / NÚMERO	**COMUNA**

IDENTIFICACIÓN DEL INFORMANTE RESPONSABLE	
NOMBRE	
TELÉFONO DE CONTACTO	
CORREO ELECTRÓNICO	

INSTITUTO NACIONAL DE ESTADÍSTICAS
CONVENIO DIRECCIÓN DEL TRABAJO - INE
Subdirección de Operaciones
Departamento de Estadísticas Sociales
Morandé 801, piso 22, Santiago - Chile
Teléfonos: (56) 2 892 4049
https://www.ine.gob.cl

IDR

ENCLA

ENCUESTA LABORAL 2023 (ENCLA) AÑO 2023

CONFIDENCIAL

Esta encuesta está autorizada por la ley Nº17.374. Toda información se mantendrá con carácter estrictamente confidencial, bajo secreto estadístico.

ENCUESTA LABORAL 2023 - EMPLEADORES ENCLA 2023

ENCUESTA LABORAL 2023
EMPLEADORES (ENCLA 2023)

IDENTIFICACIÓN DE LA EMPRESA

RUT EMPRESA									-	

LUGAR DE LA ENTREVISTA	
CALLE, AVENIDA, PASAJE / NÚMERO	**COMUNA**

IDENTIFICACIÓN DEL INFORMANTE RESPONSABLE	
NOMBRE	
TELÉFONO DE CONTACTO	
CORREO ELECTRÓNICO	

INSTITUTO NACIONAL DE ESTADÍSTICAS
CONVENIO DIRECCIÓN DEL TRABAJO - INE
Subdirección de Operaciones
Departamento de Estadísticas Sociales
Morandé 801, piso 22, Santiago, Chile
Teléfono (56) 2 892 4049
www.ine.cl

IDR

CONFIDENCIAL

Esta encuesta está autorizada por la ley Nº17.374. Toda información se mantendrá con carácter estrictamente confidencial, bajo secreto estadístico.

ENCLA
ENCUESTA LABORAL 2023 (ENCLA)
AÑO 2023

ENCUESTA LABORAL 2023 - TRABAJADORES ENCLA 2023

ENCUESTA LABORAL 2023
TRABAJADORES (ENCLA 2023)

IDENTIFICACIÓN DE LA EMPRESA

RUT EMPRESA									-	

LUGAR DE LA ENTREVISTA	
CALLE, AVENIDA, PASAJE / NÚMERO	**COMUNA**

IDENTIFICACIÓN DEL INFORMANTE RESPONSABLE	
NOMBRE	
TELÉFONO DE CONTACTO	
CORREO ELECTRÓNICO	