

ESTUDIO LONGITUDINAL SOCIAL DE CHILE (ELSOC 2017)

MANUAL DE USUARIO DE BASE DE DATOS CORTE TRANSVERSAL

true true true true true

2021-07-03



Centro de Estudios de Conflicto y Cohesión Social

**ESTUDIO LONGITUDINAL SOCIAL DE CHILE
(ELSOC 2017)**

MANUAL DE USUARIO DE BASE DE DATOS
CORTE TRANSVERSAL

Editado por:

Roberto González	Matías Bargsted	Daniel Miranda
Benjamín Muñoz	Alejandro Plaza	

Versión 2.0
08 de Enero de 2019

Índice general

1. Presentación	5
2. Diseño del Instrumento	7
2.1. Proceso de Diseño	7
2.2. Diferencias entre Cuestionarios de Ola 1 y 2	9
2.3. Ficha Técnica	12
3. Diseño Muestral del Estudio	15
3.1. Preparación del Marco Muestral	16
3.2. Etapa 1: Selección de Ciudades	16
3.3. Etapa 2: Selección de Bloques (Manzanas)	19
3.4. Etapa 3: Selección de sub - bloques	21
3.5. Etapa 4: Selección de viviendas	21
3.6. Etapa 5: Selección de personas	23
3.7. Etapa 6: Aumento del tamaño muestral	23
3.8. Cambios en el Muestreo en Ola 2017	26
3.9. Diseño de Ponderadores de la Muestra	26
3.10. Supervisión del Trabajo de Campo	29
3.11. Incidencias de Terreno y Códigos de Disposición Final de Casos	30
3.12. Aspectos Longitudinales y Tercera Ola del Estudio	31
4. Registro de Versiones Base de Datos	33
4.1. Casos Falsificados	33
5. Protocolo de Uso de Base de Datos	37
6. Dudas y Consultas	39
7. LIBRO DE CÓDIGOS BASE DE DATOS ELSOC OLA 2017	41
7.1. Variables ELSOC Ola 2	41
7.2. Características Base de Datos ELSOC Ola 2017	44

Capítulo 1

Presentación

El Estudio Longitudinal Social de Chile (ELSOC) es una encuesta desarrollada para analizar intertemporalmente la evolución del conflicto y cohesión en la sociedad chilena, basándose en modelos conceptuales descritos en la literatura nacional e internacional que abordan dichas materias. De este modo, se orienta a examinar los principales antecedentes, factores moderadores y mediadores, así como las principales consecuencias asociadas al desarrollo de distintas formas de conflicto y cohesión social en Chile. Por lo tanto, su objetivo fundamental es constituirse en un insumo empírico para la comprensión de las creencias, actitudes y percepciones de los chilenos hacia las distintas dimensiones de la convivencia y el conflicto, y como éstas cambian a lo largo del tiempo.

Esta encuesta fue diseñada por investigadores pertenecientes al Centro de Estudios de Conflicto y Cohesión Social (COES). El Centro está patrocinado por la Universidad de Chile y la Pontificia Universidad Católica de Chile y cuenta como instituciones asociadas a la Universidad Diego Portales y la Universidad Adolfo Ibáñez. Si desea obtener más información sobre el Centro, visite nuestra página web. COES es una iniciativa que desde 2013 cuenta con el financiamiento¹ del Fondo de Financiamiento de Centros de Investigación en Áreas Prioritarias (FONDAP) de la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT), organismo dependiente del Ministerio de Educación de Chile. El levantamiento de datos de ELSOC (segunda ola de medición 2017) estuvo a cargo del Centro MicroDatos (CMD) de la Universidad de Chile.

El presente documento contiene información relevante para los usuarios de la base de datos de la segunda ola del estudio longitudinal, correspondiente al año 2017. **Este Manual de Usuario enfatiza la dimensión de corte transversal de ELSOC 2017.** Los reportes técnicos relativos a las propiedades longitudinales de ELSOC se encuentran en el Manual de Usuario Longitudinal (ELSOC 2016-2017). De todos modos, algunos aspectos centrales de la evolución inter-olas de ELSOC son descritos. Se recomienda a los usuarios interesados en realizar análisis panel consultar de manera adicional el Manual de Usuario Longitudinal.

El Manual se estructura en torno a secciones temáticas. La siguiente sección describe brevemente el diseño del instrumento, lo cual es complementado con la ficha técnica del estudio. La tercera sección reseña el diseño muestral

¹Proyecto CONICYT/FONDAP/15130009

general del panel como también los detalles específicos de la segunda ola del estudio. En el siguiente apartado se resumen los principales aspectos del trabajo de campo, enfatizando las particularidades del proceso de reentrevista a los participantes. El quinto apartado corresponde al registro de versiones de la base de datos y detalla el protocolo de uso de éstas. Por último, se incluye un apartado con orientaciones básicas para el análisis y el libro de códigos de las variables incluidas en la base de datos de ELSOC 2017.

Capítulo 2

Diseño del Instrumento

A continuación se presenta el desarrollo del cuestionario del ELSOC. El instrumento de recolección de información consiste en un cuestionario estructurado (tipo encuesta) aplicado cara a cara a todos los participantes. Dicha aplicación se hace utilizando el sistema **CAPI** (Encuestas personales asistidas por computadores). El cuestionario fue diseñado para medir una serie de aspectos conceptualmente relevantes que permiten caracterizar los niveles de conflicto y cohesión social en Chile, enfatizando su evolución a lo largo del tiempo. Los principales temas de interés analítico abordados por la encuesta corresponden a los módulos en los cuales se estructura:

1. **Ciudadanía y Democracia.**
2. **Redes sociales e Interacciones inter-grupales.**
3. **Legitimidad y desigualdad social.**
4. **Conflicto social.**
5. **Dimensión barrial y territorial.**
6. **Salud y bienestar.**
7. **Caracterización Socio demográfica.**

2.1. Proceso de Diseño

El proceso de diseño del cuestionario de ELSOC se desarrolló durante el año 2015 y abarcó las olas 2016 y 2017. Por lo tanto, el proceso de trabajo aquí descrito coincide con el presentado en el Manual de Usuario de ELSOC 2016. Las diferencias entre ambos cuestionarios son descritas en el siguiente apartado.

La mayoría de las preguntas, escalas y/o items incluidos en los módulos de ELSOC 2017 provienen de otros estudios de opinión pública, investigaciones -en psicología, sociología, economía, ciencia política- realizadas por académicos nacionales e internacionales y encuestas sociales conducidas en Chile y otros países. En forma complementaria, algunas escalas fueron desarrolladas por los miembros del equipo COES y/o han sido adaptadas de estudios anteriores de éstos. De manera genérica, el cuestionario fue diseñado aprovechando las principales

recomendaciones técnicas y el estado del arte en las distintas áreas de estudio incluidas.

Con el objetivo de satisfacer los criterios y estándares de calidad para cuestionarios y compatibilizar la multiplicidad de agendas de investigación desarrolladas por COES, se optó por elaborar un procedimiento de trabajo para la construcción de éste. Este proceso se desarrolló durante el año 2015 en distintas fases:

1. Se solicitó a los investigadores vinculados a COES proponer proyectos de investigación que contemplen un planteamiento teórico e hipótesis que fundamenten las escalas propuestas para ser incluidas en el cuestionario. Las propuestas podían ser presentadas de manera individual o colectiva y no existían restricciones en el número de ítems a proponer. Sólo se exigía una fundamentación teórica explícita que involucre hipótesis longitudinales y una operacionalización de los constructos a medir en los distintos ítems propuestos.
2. El procedimiento anterior implicó la recepción y organización de un elevado número de agendas de investigación, los que se materializan en más 750 ítems (preguntas únicas o parte de una escala). El equipo ELSOC sistematizó las propuestas recibidas, clasificando los ítems en áreas temáticas. A la vez, se diseñaron mecanismos para reducir el número de ítems y coordinaron reuniones entre los investigadores COES para decidir sobre éstos. Los criterios de selección fueron principalmente teóricos, priorizando las preguntas fundamentales para el análisis longitudinal de los proyectos de investigación como su concordancia con la agenda sustantiva de COES.¹
3. Luego de una selección de las escalas más relevantes para cada tema propuesto, se realizó un estudio piloto del cuestionario desarrollado. Esto implicó pilotear 430 ítems, donde el Centro MicroDatos de la Universidad de Chile se encargó de su ejecución. En base a los resultados de dicho piloto, el equipo panel realizó ajustes a algunos ítems y elaboró una propuesta de reducción de ítems que fue evaluada con los investigadores del Centro.
4. La última etapa de ajuste del cuestionario se centró en identificar los aspectos conceptualmente más relevantes para COES y ponderar los requerimientos metodológicos para su evaluación empírica. De este modo, se clasificaron los ítems según el número de mediciones requeridas, distinguiendo entre ítems permanentes (que serán medidos en todas las olas, ya que constituyen el núcleo analítico del estudio) e intercalados (que serán medidos una única vez o una ola por medio).

La versión final del instrumento de recolección de información consiste en un cuestionario estructurado (tipo encuesta) que se aplica cara a cara a todos los participantes con una periodicidad anual. Sin embargo, debe tenerse presente que existen diferencias entre los cuestionarios de las distintas olas producto de la inclusión de los ítems intercalados.² Luego, en el Cuadro 2.2 se resumen los conceptos que se incluyeron en las distintas secciones del cuestionario de ELSOC 2017. Para más detalles sobre el Cuestionario se debe revisar la sección de Libro de Códigos.

¹ También se incluyeron criterios prácticos, relativos a la eliminación de ítems idénticos o muy semejantes; evidencia previa sobre la calidad de los ítems (encuestas anteriores, especialmente la encuesta de corte transversal desarrollada por COES el año 2014) y el diseño general del estudio (diseño muestral, unidad de análisis, tipo de informante, etc.).

² Algunos ítems aparecen una única vez (características invariantes en el tiempo) y otros son rotados (aparecen en años pares o impares).

2.2. Diferencias entre Cuestionarios de Ola 1 y 2

Para examinar de manera pormenorizada las diferencias entre los instrumentos aplicados en 2016 y 2017 se sugiere a los interesados revisar tanto los cuestionarios como los libros de códigos asociados a cada estudio. A su vez, el **Listado de Variables** incluido es una guía útil para contrastar los cuestionarios, siendo relevante enfocarse en los **items intercalados, áquellos incluidos año por medio o una única vez en ELSOC**. De todos modos, el Cuadro 2.1 señala algunas diferencias entre ambos cuestionarios.

Un tema relacionado es la duración promedio de la entrevista. Durante 2016 la duración promedio por encuestado fue de 55 minutos. CMD solicitó al equipo ELSOC tratar dicho lapso temporal a 45 minutos promedio, motivo por el cual se recomendaba acortar la extensión del estudio. El Equipo ELSOC, en conjunto con investigadores COES definió distintos ajustes que permitieron reducir la extensión del cuestionario. Mientras el cuestionario de la ola 2016 contenía 336 items, en el año 2017 el instrumento contemplaba 317 items.

Cuadro 2.1: Diferencias Temáticas entre Cuestionarios

Temas Agregados	Centro de Estudios de Conflicto y Cohesión Social
Redes Cercanas, Amistad, Elección Residencial, Elecciones Presidenciales, Trayectorias Educativas, Regulación empresarial personalidad, consumo cultural, capital c	

Cuadro 2.2: Temas y Conceptos medidos en ELSCO Ola 2

Sección	Temas
Territorio	Cohesión barrial, elección de hogar/barrio, estigmatización residencial, intención de movilidad residencial, percepción de transformación del barrio, datisfacción con el barrio, satisfacción con la vivienda, seguridad del Barrio, sociabilidad barrial, transporte y violencia barrial
Redes y Actitudes Sociales	Actitudes intergrupales (Ansiedad intergrupar, calidad de Contacto, cantidad de contacto, contacto negativo, normas intergrupales, frecuencia de contacto, percepción de similitud). También se incluye una batería de redes sociales primarias/cercanas que incluye: características sociodemo- gráficas de confidentes, además de creencia religiosa y posición política.

2.3. Ficha Técnica

En este apartado se presenta la Ficha Técnica (Ver Cuadro 2.3, dónde se sintetizan las principales características del Estudio Longitudinal Social de Chile (ELSOC COES) y de la implementación de la primera ola de esta encuesta. Los siguientes apartados entregan mayores detalles sobre el diseño muestral del panel y en específico de su primera ola, como de la ejecución y trabajo de campo.

Cuadro 2.3: Ficha Técnica ELSOC COES, Ola 2 (2017)

Características	ELSOC 2017
Objetivo	Analizar longitudinalmente la evolución del conflicto y cohesión en la sociedad chilena.
Diseño	Estudio cuantitativo por medio de un cuestionario estructurado.
Instrumento	Cuestionario compuesto por preguntas cerradas de carácter simple y múltiple junto a algunas preguntas abiertas. Combina preguntas permanentes (todas las olas) e intercaladas.
Cobertura Temática	Contiene siete módulos temáticos: Territorio, Redes y actitudes sociales, Ciudadanía y democracia, Desigualdad y legitimidad, Conflicto social, Salud y bienestar y Caracterización sociodemográfica.
Unidad de Análisis	Individuos.
Población Objetivo	Hombres y mujeres de 18 a 75 años, residentes habituales de viviendas particulares

Capítulo 3

Diseño Muestral del Estudio

En el siguiente apartado se presenta la descripción del diseño muestral de ELSOC, correspondiente a la primera ola del estudio. **Durante la segunda ola del estudio (2017) el objetivo fundamental es reentrevistar a los encuestados de la primera versión del estudio, por lo que no hay mayores cambios a la estrategia muestral.** Al final de la sección se describen algunos ajustes menores y se discute el nivel de atrición de ELSOC.

El diseño de ELSOC tuvo como objetivo conciliar los múltiples intereses de investigación de los investigadores asociados al Centro. Entre las consideraciones más relevantes destacaron las siguientes:

1. Un diseño muestral que permitiera combinar las variables medidas en el cuestionario con las variables espaciales, registradas a nivel de manzana y comuna, contenidas en las bases de datos desarrolladas por el Centro de Inteligencia Territorial (CIT) de la Universidad Aldofo Ibáñez. Dado que los datos del CIT no están disponibles para todas las manzanas del país, particularmente aquellas ubicadas en localidades rurales, se decidió incorporar en la muestra únicamente zonas urbanas. Esta consideración también coincidió con las preferencias de muchos investigadores del Centro, quienes manifestaron estar principalmente interesados en una muestra de carácter urbano.
2. Algunos investigadores solicitaron un diseño que permitiera estimar modelos multi-nivel (o jerárquicos) agrupados por ciudad y comuna, y por tanto, se estableció que la muestra contuviera un número suficiente de ciudades y comunas, así como un número suficiente de casos dentro de cada ciudad y comuna, que permitiera tal análisis (Snijders & Bosker, Capítulo 10).
3. Otros investigadores estaban interesados en comparar a los habitantes de las tres ciudades más grandes del país, lo que se tradujo en un diseño no proporcional que incrementara el número de encuestados en las zonas del Gran Valparaíso (ciudades de Viña del Mar y Valparaíso) y Gran Concepción (Concepción, Talcahuano y otras).
4. Finalmente, algunos investigadores solicitaron un diseño que permitiera comparar a los encuestados que vivieran en ciudades grandes y pequeñas, lo que favoreció incrementar el tamaño de la muestra de hogares en

ciudades pequeñas (Kish, 1965, Sección 3.5), particularmente aquellas con entre 30 mil y 100 mil habitantes.

Los investigadores de COES trabajaron con la encargada del diseño muestral, Stephanie Eckman, para desarrollar un diseño que pudiera, razonablemente, cumplir con estas necesidades e intereses sustantivos. El diseño muestral final de la ronda 1 de ELSOC COES proporciona una cobertura adecuada de las ciudades más grandes del país (Gran Santiago, Gran Valparaíso y Gran Concepción), así como ciudades más pequeñas, y también asegura la representación de personas en el norte y sur del país. En términos globales, el diseño muestral alcanza una representatividad aproximada del 77 % de la población total del país, y del 93 % de la población urbana. Las siguientes subsecciones detallan los distintos pasos del diseño de la muestra.

3.1. Preparación del Marco Muestral

El proceso de muestreo se realizó en base a los datos del pre-censo del año 2011, los cuales fueron formateados por el CIT. Aunque los recuentos de población del censo de 2012 no son precisos, el trabajo del pre-censo recolectando información sobre los hogares en todos las manzanas (bloques) es de calidad. El conjunto de datos contenía un total de 155.757 bloques, pero se eliminaron cuatro tipos diferentes antes de que comenzara con la selección.

1. Siguiendo los intereses analíticos de los investigadores del Centro, sólo se utilizaron bloques urbanos. Para determinar qué bloques eran urbanos, se empleó la codificación del tipo de localidad (urbana o rural) contenida en la base de datos del pre-censo de 2011. Consecuentemente, 22.188 (14,2 %) bloques fueron excluidos en este paso.
2. Nuevamente, en función de los intereses analíticos de los investigadores del Centro, sólo los bloques que habían sido previamente geo-referenciados por el CIT se conservaron para el muestreo. Esto implica que un total de 1.971 (1,5 % de los bloques urbanos) que no estaban geo-referenciados fueron removidos en este paso.
3. Sólo los bloques que contenían cinco o más hogares (de acuerdo con el pre-censo de 2011) fueron retenidos. 503 bloques (menos del 1 % de los bloques restantes tras los pasos 1 y 2) no alcanzaron este umbral y fueron eliminados.
4. Sólo los bloques en las ciudades con más de 10.000 personas eran elegibles para la selección. 10.238 bloques (7.8 % de los bloques restantes) fueron excluidos del marco muestral.

De esta forma, el marco muestral final contiene 120.857 bloques. La muestra de COES representará solamente estos bloques y no aquellos que fueron excluidos. Las estimaciones derivadas de los datos de la muestra se aplicarán únicamente a esta población objetivo y no deben aplicarse a toda la población chilena. El proceso de selección de entrevistados se desarrolló en cinco etapas, aunque durante el trabajo de campo se añadió una sexta etapa.

3.2. Etapa 1: Selección de Ciudades

El universo de bloques (los 120.857 bloques mencionados) fue agregado al nivel de la ciudad, resultando en 122 ciudades. Las tres ciudades más grandes (Gran Santiago, Viña del Mar - Valparaíso y Concepción - Talcahuano)

fueron seleccionadas con certeza. Las ciudades restantes son estratificadas por la población. La Tabla 3.1 muestra las definiciones de los estratos y los tamaños de población (N) y muestra (n) en cada uno.

Cuadro 3.1: Población por ciudad y tamaños de muestra, por estrato

Estrato	Definición N N habi- tantes	N ciu- dades	n ciu- dades	N Sub-estrato Norte	n Sub-estrato Norte	N Sub-estrato Sur	n Sub-estrato Sur
Gran Santiago		1	1	-	-	-	-
Gran Valparaíso		1	1	-	-	-	-
Gran Concepción		1	1	-	-	-	-
Ciudades Grandes	>100 mil	18	8	8	4	10	4
Ciudades Medianas	>30 mil	28	10	15	6	13	4
Ciudades Pequeñas	>10 mil	73	19	24	6	49	13

Los estratos 4, 5 y 6 fueron estratificados geográficamente por Norte / Sur para asegurar que la muestra contuviera ciudades del norte y sur de Chile. Esto redundó en un total de nueve estratos. La muestra se asignó entre las dos áreas en proporción al tamaño de su población en el universo. Véase la Tabla 3.1 para ver el detalle acerca de los tamaños de población y muestra en cada uno de los estratos norte y sur.

La selección de ciudades dentro de cada uno de estos estratos finales se realizó en forma proporcional al tamaño de la población de cada ciudad. Este método da mayores probabilidades de selección a las grandes ciudades.

La probabilidad de selección de una ciudad i dentro del estrato h fue:

$$\pi_i = \frac{(nc_h)(pop_i)}{\sum_h pop}$$

donde nc_h es el número de ciudades seleccionadas en el estrato h y pop_i es la población de ciudad i .

3.3. Etapa 2: Selección de Bloques (Manzanas)

Las 40 ciudades seleccionadas contenían 87.839 bloques. En la segunda etapa se seleccionaron bloques en cada ciudad con población proporcional al tamaño, donde el tamaño fue determinado a partir del recuento de unidades de vivienda del pre-censo. La selección fue sistemática: la lista de bloques en las ciudades seleccionadas se ordenó según sub-distrito censal y número de bloque para asegurar que los bloques seleccionados se extendieran por toda la ciudad¹.

La Tabla 3.2 muestra el número de bloques seleccionados en cada ciudad. La muestra de bloques se asignó de manera desproporcionada para que las áreas fuera de Santiago estuvieran sobre-representadas en relación con su tamaño en la población objetivo. Varios investigadores COES solicitaron esta asignación para asegurar que la muestra fuera diversa con respecto al tamaño de la ciudad.

La probabilidad de selección de un bloque j en la ciudad i , condicionada a la selección de la ciudad, fue:

$$\pi_{j|i} = \frac{(nb_i)(hu_j)}{\sum_i hu}$$

donde nbi es el número de bloques seleccionadas en la ciudad i y hu_j es la población de la ciudad i .

¹ Los números de bloques y distritos censales fueron entregados por Matías Garretón, investigador de CIT. Los sub-distritos censales son unidades geográficas más pequeñas que la comuna, pero más grandes que los bloques.

Cuadro 3.2: Distribución de los ciudades y bloques por estrato

Estrato	Definición ciu- N habi- tantes	dades	n blo- ques por ciu- dad
Gran Santiago		1	200
Gran Valparaíso		1	100
Gran Concepción		1	100
Ciudades Grandes	>100 mil	8	26
Ciudades Medianas	>30 mil	10	25
Ciudades Pequeñas	>10 mil	19	11
Total		40	1,067

En cuatro ciudades, algunos bloques eran tan grandes que fueron selecciones certeras. Es decir, los recuentos de unidades de vivienda eran mayores que el intervalo de selección y se seleccionarían en cualquier muestra, e incluso podrían seleccionarse dos veces. Para evitar selecciones duplicadas, estos bloques se eligieron primero con certeza y luego se seleccionaron bloques adicionales entre los restantes para aquellas ciudades, de modo de alcanzar el tamaño de muestra total deseado para la ciudad (ver Tabla 3.2. $\pi_{j|i}$ para estas ciudades es 1).

Los 1.067 bloques seleccionados en las 40 ciudades elegidas fueron empadronados en terreno, con la finalidad de realizar la selección de los hogares con la información más actualizada posible. El CIT proporcionó mapas de cada bloque seleccionado. El personal de campo de CMD visitó cada bloque y creó una lista de todas las unidades de vivienda. Los listados fueron revisados cuidadosamente para detectar cualquier error o duplicado.

3.4. Etapa 3: Selección de sub - bloques

Durante el proceso de empadronamiento, el Centro de Microdatos encontró que algunos bloques tenían más de 100 viviendas, lo que dificulta excesivamente el proceso de empadronamiento. Consecuentemente, se dividieron estos bloques en sub-bloques de tamaño aproximadamente igual (40 a 50 viviendas) y seleccionaron uno para ser empadronado. Debido a que los sub-bloques fueron creados para ser casi igual tamaño, estos fueron seleccionados en base a igual probabilidad. En total, 301 bloques fueron submuestreados. Los bloques restantes no se vieron afectados por esta etapa.

3.5. Etapa 4: Selección de viviendas

El número de viviendas seleccionadas en cada bloque varió según el estrato, como se muestra en la Tabla 3.3. Este diseño resultó en 4.001 unidades de vivienda, con lo cual se pretendieron obtener aproximadamente 3.000 entrevistas completas, bajo el supuesto de una tasa de respuesta del 75 % para todos los estratos.

Cuadro 3.3: Distribución de los bloques por estrato

Estrato	Definición HHS	N habi- tantes
Gran Santiago		5
Gran Valparaíso		5
Gran Concepción		5
Ciudades Grandes	>100 mil	3
Ciudades Medianas	>30 mil	3
Ciudades Pequeñas	>10 mil	3
Total		4001

Se realizó una muestra aleatoria simple de viviendas en cada bloque. La combinación de la población proporcional al tamaño de muestreo en las dos primeras etapas y el muestreo aleatorio simple en la tercera y cuarta etapas dio lugar a una muestra de viviendas con aproximadamente igual probabilidad dentro de cada uno de los nueve estratos mostrados en la Tabla 3.1.

La probabilidad de selección de un hogar k en el bloque j en la ciudad i y el estrato h fue:

$$\pi_{k|j,i} = \frac{nh_j}{NH_j}$$

donde nh_j es el número de viviendas seleccionadas en el bloque j , y NH_j corresponde al número de viviendas alistadas en el bloque j .

3.6. Etapa 5: Selección de personas

Los encuestadores visitaron cada hogar seleccionado e intentaron llevar a cabo la entrevista. El primer paso en el proceso de la entrevista fue identificar al entrevistado objetivo. Cuando había más de un adulto en el hogar, uno fue seleccionado usando una muestra aleatoria simple, usando una tabla de Kish.

La probabilidad de selección de una persona en el hogar k fue:

$$\pi_{l|k,j,i} = \frac{1}{NP_j}$$

donde NP_j es el número de adultos (mayores a 18 años y menores de 75 años) que habitan la vivienda j .

3.7. Etapa 6: Aumento del tamaño muestral

Durante el trabajo de campo se observó que el supuesto de una tasa de respuesta del 75% para todos los estratos era incorrecta. En primer lugar, hubo variabilidad significativa por región, y en segundo lugar, la tasa de respuesta fue inferior al 75%. Debido a esta información, se decidió aumentar el número de hogares por bloque para lograr efectivamente las 3.000 entrevistas.

Este aumento en el número de hogares por bloque tiene un efecto limitado sobre la probabilidad de selección de cada vivienda. Sólo afecta a la probabilidad calculada en la Etapa 4, ya que el número de hogares disponibles es menor, pero no hay cambios en las probabilidades calculadas en la Etapa 1 y la Etapa 2. Esto ocurre porque los bloques seleccionados (en la Etapa 2) fueron usados, y no se introdujeron nuevos bloques.

Durante este proceso un total de 1.082 nuevos hogares, ubicados dentro de los bloques seleccionados, se añadieron a la muestra del estudio. La asignación de estos nuevos hogares no fue uniforme en todos los bloques del país. En cambio, se concentraron en cuatro regiones (Coquimbo, O'Higgins, Región Metropolitana, y Biobío) donde los encuestadores tenían muchos problemas para contactar a los encuestados. La tabla 3.4 detalla las comunas en que se aumentó el número de hogares, junto con el número total de hogares incorporados y por bloque.

Cuadro 3.4: Número de hogares agregados a la muestra según región y comuna

Región	Comuna	Total vivien- das agre- gadas	Viviendas agre- gadas por blo- que
Coquimbo	Coquimbo	24	2
	La Serena	28	2
	Salamanca	22	2
O'Higgins	Doñihue	10	1
	Rancagua	42	2
	Santa Cruz	11	1
Biobío	Chiguanyante	24	3
	Concepción	75	3
	Coronel	11	1
	Penco	4	1
	Quillón	6	1
	San Pedro de la Paz	28	2
Metropolitana	Cerrillos	9	3
	Colina	12	3
	Curacaví	14	2
	El Bosque	8	2
	Estación Central	12	3
	Huechuraba	6	2
	Independencia	6	2
	Isla de Maipo	39	3

La Cisterna	9	3
La Florida	24	2
La Granja	6	2
La Pintana	12	2
La Reina	9	3
Las Condes	33	3
Lo Barnechea	9	3
Lo Espejo	6	2
Lo Prado	6	2
Macul	8	2
Maipú	32	2
Ñuñoa	16	2
Padre Hurtado	6	3
Pedro Aguirre Cerda	6	2
Peñaflor	30	2
Peñalolén	14	2
Providencia	7	3
Pudahuel	14	2
Puente Alto	32	2
Quilicura	12	2
San Bernardo	16	2
San Joaquín	6	2
San Miguel	9	3
San Ramón	6	2
Santiago	120	3
Vitacura	9	3

3.8. Cambios en el Muestreo en Ola 2017

Producto de las dificultades observadas en el trabajo de campo de ELSOC 2016 descritas en el apartado previo, se realizaron ajustes en el diseño muestral que contempla heterogeneidad regional en la tasa de no respuesta. No fueron necesarios cambios adicionales para el proceso de trabajo de campo de 2017. Las decisiones metodológicas fueron confirmadas y las viviendas agregadas y los entrevistados seleccionados durante 2016 como resultado de la etapa 6 fueron reentrevistados del mismo modo que los otros encuestados.

3.9. Diseño de Ponderadores de la Muestra

La muestra de ELSOC 2017 no es representativa de la población objetivo del modo que lo era la muestra de 2016, producto de que su objetivo es recontactar a los entrevistados de 2016 y a la presencia de atrición en la muestra (individuos que no logran ser reentrevistados, dejando la muestra). Considerando que a partir de las entrevistas efectivamente realizadas se desea extrapolar a la población objetivo respectiva, es necesario ponderar adecuadamente a cada encuestado según su representación en la población objetivo. De este modo, lo óptimo sería incorporar la atrición en el diseño de los ponderadores muestrales. Los ponderadores provistos con la base de datos de ELSOC ola 2 no ajustan por dichos factores. Se proveen ponderadores similares a los entregados en la primera versión del estudio, cuya finalidad es ajustar por las diferencias en atributos demográficos de la muestra de ELSOC en relación a la población objetivo. **En una etapa posterior se liberarán nuevos ponderadores, ajustados a las particularidades de un diseño longitudinal.**

A continuación se describe el proceso de elaboración de dichos ponderadores² para lograr una correspondencia entre la muestra y la población objetivo. Dicha ponderación corresponde al inverso de su probabilidad de selección o inclusión en la muestra. En este caso la probabilidad de selección del individuo i de la vivienda j que pertenece a la manzana l del estrato k , P_{ijkl} viene dado por:

$$P_{ijkl} = \pi_i | jkl \pi_j | lk \pi_l k$$

Dónde:

- $\pi_i | jkl$ es la probabilidad de que el individuo i sea seleccionado en la muestra dado que la vivienda dónde vive y la manzana dónde se localiza fueron seleccionados.
- $\pi_j | lk$ es la probabilidad de que la vivienda j sea seleccionada en la muestra dado que la manzana l (que contiene a la vivienda j) fue seleccionada.
- $\pi_l k$ es la probabilidad de que la manzana l del estrato k sea seleccionada en la muestra.

Se define el ponderador de diseño o teórico w_{ijkl} al inverso de la probabilidad de selección:

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

²Dichos ponderadores pueden entenderse como “ponderadores de corte transversal,” ya que se ignoran los problemas derivados de su naturaleza longitudinal.

El valor de las antes mencionadas probabilidades es:

$$\pi_{lk} = n_k \frac{M_{kl}}{M_k}$$

$$\pi_{j|lk} = \frac{m_{lk}}{M'_k}$$

$$\pi_{i|jlk} = \frac{1}{N_{jlk}}$$

Dónde se tiene que n_k es el número de manzanas a seleccionar del estrato k , $_{lk}$ es el número de viviendas de la manzana l del estrato k , M_k es el número total de viviendas del estrato k , m_{lk} es el número de viviendas a encuestar dentro de la manzana l , M'_k es el número actualizado de viviendas de la manzana l post-empadronamiento, N_{jlk} es el número de personas de la población objetivo que vive en la vivienda j de la manzana l del estrato k .

En base a lo anterior se diseñaron ponderadores que ajustan en base a las probabilidades base, con ajuste de no respuesta, con ajuste al número de casos³ y con ajuste al número de casos y la variable sexo.⁴

#Implementación de la Encuesta {#implem_encuest}

COES realizó una licitación pública para la ejecución del ELSOC. El Centro Micro Datos (CMD) de la Facultad de Economía y Negocios de la Universidad de Chile se adjudicó dicho concurso y se encargó del desarrollo de la encuesta, incluyendo el trabajo de campo durante las dos primeras olas. El acuerdo con Micro Datos contempla los más altos estándares de calidad y asegura un adecuado resguardo de los datos personales de contacto de los participantes del estudio panel.

La descripción de la implementación de ELSOC 2016 incluye apartados relativos al empadronamiento e inicio del trabajo de campo. Se recomienda revisar dicho documento con el objetivo de formarse una visión detallada de esta fase del estudio. A continuación se describen aspectos relevantes de la ejecución de ELSOC 2017.

##Estrategia de Seguimiento

COES y Centro MicroDatos definieron una estrategia de fidelización y seguimiento de los encuestados de ELSOC 2016 con el objetivo de minimizar la potencial atrición en el estudio. El insumo básico de este procedimiento fue el almacenamiento de información de contacto de los encuestados (además de su nombre de pila, contar con teléfono de contacto y dirección de correo electrónico). Originalmente CMD sólo contaba con correos electrónicos de un 19,3% de los entrevistados, para luego recopilar mayor información de éstos. Se envió un saludo navideño a los participantes de ELSOC por correo electrónico. Posteriormente, en Marzo de 2017 se envió otro saludo. A su vez, se envió información proveniente del equipo de ELSOC para explicar cuál es la utilidad de ELSOC.

³Corresponde a considerar la población total objetivo entre 18 y 75 años en base a las proyecciones de población regional.

⁴Los cuatro ponderadores diseñados son acumulativos. A modo de ejemplo, el segundo ponderador ajusta en base las probabilidades de selección e incorpora **también** el ajuste por no respuesta.

CMD realizó una ronda de recontacto con los participantes de ELSOC previo al trabajo de campo de 2017. El objetivo de dicho trabajo fue actualizar su información de contacto. Esto se realizó por medio de una encuesta web y por vía telefónica. Para esto se contó con un Call Center encargado de contactar a los participantes. Por otra parte, CMD almacenó, con altos estándares de confidencialidad, la información relativa a la dirección de los entrevistados de ELSOC y otra información pertinente derivada del empadronamiento. Dicha información permite contar con un listado de direcciones de residencia de los encuestados ELSOC y la información cartográfica pertinente para facilitar el trabajo de campo del encuestador. Estos insumos fueron cruciales para la planificación del trabajo de campo de 2017⁵.

##Pilotaaje del Cuestionario Y Capacitación

Con el cuestionario correspondiente a la segunda ola de ELSOC previamente elaborado, COES y CMD colaboraron en la programación de éste para que fuera aplicada de manera presencial en tablets (sistema CAPI). En una primera fase, el equipo ELSOC chequeó el flujo del cuestionario programado en tablets y propuso ajustes menores. Luego, se coordinó una versión piloto del estudio con el fin de testear el instrumento y evaluar el funcionamiento operativo del instrumento.

El pilotaaje se realizó durante Junio y Julio del 2017, seguido por una ronda de focus group a los encuestadores con el objetivo de obtener recomendaciones relevantes para el levantamiento final de la encuesta. Esto fue utilizado para introducir cambios menores al cuestionario programado en tablets y para definir ajustes al programa de capacitación de los encuestadores.

El CMD se encargó de la selección y capacitación del personal requerido (coordinadores y encuestadores). La capacitación de coordinadores de campo y encuestadores fue ejecutada por CMD con apoyo técnico de COES para aspectos específicos. CMD realizó dicha tarea de manera centralizada, de modo de que los coordinadores fueron capacitados en dependencias de la Facultad de Economía y Negocios (FEN) de la Universidad de Chile en la Región Metropolitana. Sin embargo, los coordinadores que trabajaron previamente en ELSOC 2016 no tenían que asistir obligatoriamente a dicha sesión de capacitación. En cambio, se elaboró como alternativa una capacitación por sistema web en el cual se repasan los aspectos centrales del estudio y se presentan particularidades del nuevo cuestionario.

Se dispuso de 17 sedes de trabajo administradas por coordinadores de zona. Las sedes permiten recibir a los encuestadores y almacenar los insumos del estudio, contando con conexión a internet para facilitar la carga de las encuestas recolectadas. La capacitación de los encuestadores en el levantamiento se realizó en cada sede y hubo encuestadores apoderados, quienes posteriormente se encargaron de capacitar a los encuestadores en las regiones en que no hubo sedes (aquellas con menos de 400 encuestas a realizar).

#Planificación y Ejecución del Campo

Al igual que en 2016, CMD elaboró un Manual del Encuestador para la implementación de ELSOC y en conjunto con COES preparó documentos orientadores y protocolos para la aplicación del cuestionario. Esto incluye:

⁵Más detalles se encuentran disponibles en el Informe Anual de Servicios para el Levantamiento del Estudio Longitudinal sobre Conflicto y Cohesión Social (COES) elaborado por MicroDatos.

1. **Protocolo de Visita:** primer contacto con la vivienda, primer contacto con el hogar y concertación de entrevistas en casos en que no se puede realizar entrevista al momento de visita.
2. **Protocolo ante Casos Difíciles:** escenario de negaciones, escenario de más de 4 intentos presenciales sin contacto, barreras de acceso a la vivienda, no contacto con la vivienda. También contempla **estrategias de recuperación**.
3. **Protocolo de Aplicación del Cuestionario:** descripción del estudio, consentimiento informado, aplicación de preguntas, manejo de interrupciones, entrega de gift cards.
4. **Protocolo de Campo:** respaldo de información, control de avances, registro de contactos.

Los documentos fueron adaptados de modo tal de ajustarse a los requerimientos del seguimiento de una muestra previamente entrevistada. El levantamiento de información fue realizado en un período de aproximadamente 12 semanas -contadas desde la entrega de material a los encuestadores tras la capacitación-, entre los meses de Julio y Octubre de 2017. Para la ejecución del terreno se contó con 120 encuestadores distribuidos en las 17 sedes de trabajo antes referidas. Todos fueron debidamente capacitados, para posteriormente ser citados para retirar el material de trabajo (hojas de ruta, citas concertadas, croquis de ubicación y tablet).

CMD cuenta con información relevante para garantizar un adecuado recontacto (nombre del entrevistado, información de contacto y aspectos básicos para su identificación) y utiliza dicha información para concertar entrevistas y planificar el campo. Cada encuestador recibió del coordinador de sede o encuestador apoderado los datos sobre las citas concertadas con participantes ELSOC y utiliza un protocolo de entrevista específico para reentrevistas. Las respuestas de los entrevistados fueron almacenadas en tablets provistas por CMD. La información recolectada era subida por los encuestadores a una plataforma web utilizando la tablet de CMD. Al finalizar la entrevista se entrega un incentivo monetario (tarjeta de regalo de 6.000 pesos), un folleto informativo sobre COES y ELSOC y se amplía la base de datos de información de contacto de participantes ELSOC.

3.10. Supervisión del Trabajo de Campo

El Centro MicroDatos desarrolló un proceso de supervisión del trabajo en terreno involucrando cuatro etapas:

1. Supervisión del 100% de las encuestas recopiladas, realizado de manera simultánea por el Coordinador de Grupo. Para esto se utilizó una aplicación web diseñada para estos fines. Esto implicó la verificación del sujeto entrevistado, revisión de variables claves, revisión del resultado de visita y revisión de duración de entrevista.
2. Retroalimentación de encuestadores de manera periódica por parte del jefe de zona como revisión de los datos recolectados por el equipo central de MicroDatos. Esto dio lugar a reportes periódicos de avance del levantamiento de datos. - **Verificación del sujeto entrevistado:** CMD revisa si el instrumento fue aplicado a la persona elegida previamente en ELSOC 2016. - **Revisión de subconjunto de preguntas** (Variables clave), incluyendo preguntas de cadena (por ejemplo “Otro. Especifique cuál”) y condiciones del flujo del cuestionario.
3. Supervisión de encuestas por uso de CAPI. El software desarrollado para el cuestionario permite manejar de

mejor modo la no respuesta, validar internamente el flujo del cuestionario (evitar problemas con filtros) y recolectar información sobre el tiempo de aplicación. La validación interna y el chequeo de rango y consistencia de la información minimiza notablemente el error de medición y es una de las ventajas del diseño de ELSOC.

4. Revisión de casos especiales y validación ex post. Finalmente CMD revisó el funcionamiento del flujo de respuestas y descartó 13 casos con información falsificada, llegándose a un total de 2520 casos.

En paralelo a la revisión y/o eliminación de casos, CMD desarrolló un proceso de validación de los valores de entrevistados a distintas preguntas del cuestionario. Posteriormente el equipo de ELSOC analizó los datos provenientes de la base para evaluar su consistencia interna. Los principales aspectos de dicha revisión se encuentran documentados en el Libro de Códigos. Sin embargo, se detectaron algunos casos de entrevistados con información inconsistente/anómala en comparación a la Ola 2016. Producto de lo anterior se incluye una variable **sospechoso** que señala con un 1 a los 26 casos con información problemática. El usuario decide si utilizar dichos casos. Los usuarios pueden contactarse con el equipo ELSOC con el fin de acceder a otros documentos relativos a la validación de datos del estudio.

3.11. Incidencias de Terreno y Códigos de Disposición Final de Casos

Uno de los principales objetivos del equipo ELSOC es que el campo del estudio no se viera contaminado por el desarrollo de las campañas presidenciales y por otros estudios de opinión levantados por dichos motivos. Por lo tanto, se planificó un período de campo lo suficientemente extenso para minimizar dicho riesgo. De todos modos, el campo superó las nueve semanas planificadas, extendiéndose a doce. Las principales dificultades de recontacto se concentraron en la Región Metropolitana y de Valparaíso. La estructura de pagos a encuestadores sólo fue modificada en la región de Valparaíso, subiendo a 15.000 pesos por encuesta durante la última semana de terreno. En las regiones de Araucanía y Los Ríos se implementó el uso de autos dentro del levantamiento de datos producto de las distancias entre los hogares de los participantes ELSOC y las sedes operativas.

Siguiendo los estándares de la Asociación Americana de Investigación en Opinión Pública (AAPOR por su sigla en inglés), se asoció a cada uno de los casos de la muestra completa un código de disposición final. En base a dicha información, la segunda ola de ELSOC tiene las siguientes tasas de resultados:

- Tasa de Respuesta (RR2) de 82,4 %.
- Tasa de Rechazo (REF2) de 6,1 %.
- Tasa de Cooperación (COOP2) de 93,0 %.
- Tasa de Contacto (CON2) de 88,5 %.

Si desea conocer más detalles sobre la determinación de los códigos de disposición final de casos y el cálculo de las tasas de resultados, dirigirse a http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

Dadas las diferencias observadas a nivel de región en ELSOC 2016, se desagregan los resultados previos (Ver Cuadro ??).

3.12. Aspectos Longitudinales y Tercera Ola del Estudio

Un aspecto esencial de la planificación de ELSOC desde sus orígenes fue la atrición presupuestada. Desde la etapa de diseño se contemplaron distintos escenarios potenciales de atrición y se definió el diseño muestral acorde a las conclusiones de dicho examen (Ver Figura ??). A su vez, los escenarios de atrición fueron parte de los objetivos que guiaron el diseño de la licitación del estudio y el trabajo con CMD.

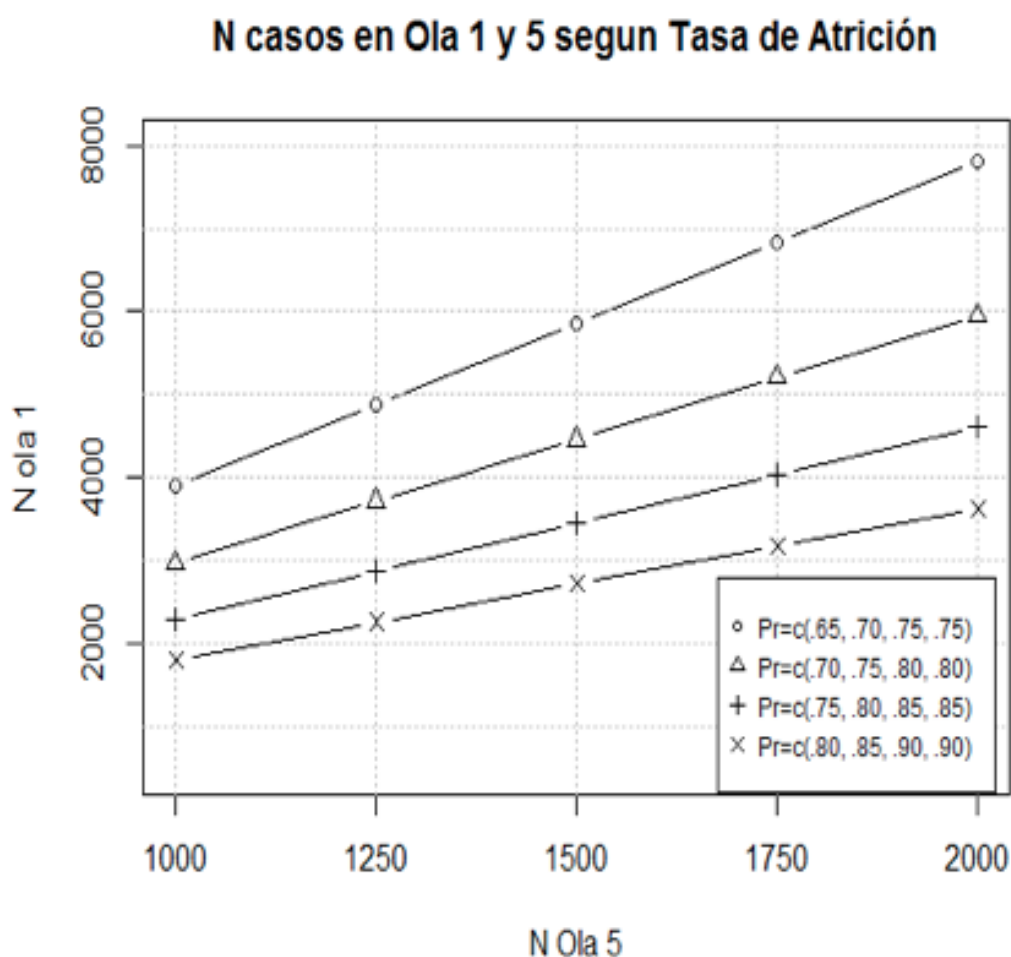


Figura 3.1: \label{fig:atricion}: Diseño Muestral de ELSOC según Ola

El Cuadro ?? describe la recuperación de casos en ELSOC 2017, desagregando en subgrupos claves. Tal como puede observarse, la muestra de 2017 equivale a un 84.49% de la original, resultado mejor a las estimaciones de un 20% de atrición en la primera ola. Luego, se observan diferencias sustantivas en términos de sexo y región de locación del entrevistado.

Según el plan de trabajo contemplado por COES, durante el 2018 se levantará la tercera ola de ELSOC. En esta

tercera ola se volverá a entrevistar a las personas encuestadas en la primera ola (2016), donde además se integrará una muestra de refresco de 1400 casos. El objetivo de la muestra de refresco es corregir la subcobertura de ciertos grupos poblacionales y ajustar por la pérdida de casos dado la atrición. Los resultados de la tercera ola estarán disponibles durante el primer semestre del 2019.

Diseño Estudio Panel Conflicto y Cohesión Social en Chile

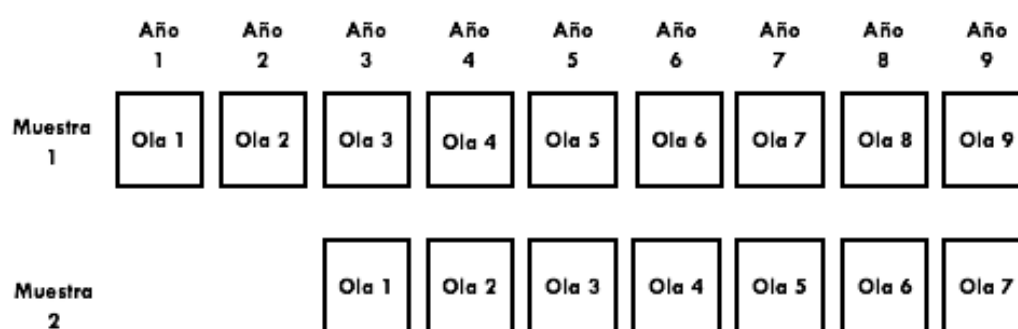


Figura 3.2: \label{fig:atricion2}: Diseño Muestral de ELSOC según Ola

Capítulo 4

Registro de Versiones Base de Datos

Las versiones adjuntadas de la base de datos, como el presente Manual de Usuario fueron elaborados por el Equipo de Encuestas COES. Sin embargo, **no pueden ser consideradas como definitivas**, si no que como un primer producto que se pone a disposición de la Comunidad COES. Se desarrollarán nuevas versiones de la base de datos como del Manual de Usuario, de modo de cumplir con la planificación propuesta. **Ni las bases de datos ni el Manual pueden ser entregados a terceros que no sean parte de COES.** Los datos y documentos deben ser obtenidos desde el sitio oficial del estudio.

En esta sección se resumen las versiones existentes, a la fecha, de la base de datos correspondiente a la primera segunda del estudio longitudinal (Ver Cuadro 4.1). A su vez, se reseñan las diferencias entre dichas versiones y los cambios introducidos.

4.1. Casos Falsificados

Durante la realización del trabajo de campo de ELSOC 2018, el equipo de Centro Micro Datos, junto al equipo de Encuestas COES detectó que 47 entrevistas correspondientes a la ola 2016 fueron falsificadas. Producto de lo anterior, dichas respuestas fueron eliminadas de la base de datos en la versión 3.0 del 28 de Diciembre de 2018. El problema representa menos del 1 % del tamaño muestral original y, en los análisis realizados por el equipo de Encuestas COES, no se detectaron cambios sustantivos en la distribución de las variables tras su eliminación.

```
library(pacman)
pacman::p_load(tidyverse, knitr, kableExtra, dplyr, tinytex)

versiones <- tibble(Versión = c("0.0"),
                    Fecha = c("27/11/2017"),
                    "Nombre Archivo" = c("ELSOC Ola 2"),
                    Características = c("Base original entregada por Centro Micro Datos a COES. Contiene :"))
```

```
add_row(Versión = c("0.1"),
        Fecha = c("12/02/2018"),
        "Nombre Archivo" = c("ELSOC Ola 2 Corregido"),
        Características = c("Base de datos ajustada a solicitudes de COES, incluyendo chequeos"),
add_row(Versión = c("0.2"),
        Fecha = c("16/04/2018"),
        "Nombre Archivo" = c("ELSOC_W02_v1.00"),
        Características = c("Base de datos elaborada por COES uniendo información adicional C

versiones %>%
  kable(booktabs = TRUE, caption="\\label{tab:versiones}Versiones de Base de Datos ELSOC Ola 2 (2017)")
  column_spec(c(1), width = "1cm") %>%
  column_spec(c(2,3), width = "2cm") %>%
  column_spec(c(4), width = "4cm") %>%
  kable_classic("striped", full_width = F)%>%
  landscape()
```

Cuadro 4.1: Versiones de Base de Datos ELSOC Ola 2 (2017)

Versión	Fecha	Nombre Archivo	Características
0.0	27/11/2017	ELSOC Ola 2	Base original entregada por Centro Micro Datos a COES. Contiene respuestas de 2526 entrevistados (N final) a preguntas del cuestionario.
0.1	12/02/2018	ELSOC Ola 2 Corregido	Base de datos ajustada a solicitudes de COES, incluyendo chequeos de datos y entrega de bases adicionales
0.2	16/04/2018	ELSOC_W02_v1.0	Base de datos elaborada por COES uniendo información adicional CMD. Con controles de calidad y ajustes a nombres de variables (homologamiento).

Capítulo 5

Protocolo de Uso de Base de Datos

Desde la liberación de la base de datos de ELSOC COES Ola 2, éstos se encuentran bajo un período de embargo, por lo que el acceso de la comunidad académica general se encuentra restringido. Los datos se irán liberando de manera modular durante el año 2018, en la medida que se realizan las presentaciones de éstos. Concluido lo anterior, se liberará la base de datos completa. Los datos de ELSOC **sólo pueden ser utilizados con fines de docencia y/o investigación**. En caso de que se deseen utilizar para otros fines, es necesario solicitar autorización expresa al Equipo de Encuestas de COES. A su vez, cualquier publicación que incluya para su realización el uso de la base de datos (en cualquiera de sus versiones) de los datos del ELSOC debe citar la fuente de la siguiente forma:

Centro de Estudios de Conflicto y Cohesión Social (2018). Estudio Longitudinal Social de Chile, Primera Ola (ELSOC_W02_v2.00) . [Archivo de datos]. Santiago, Chile: Centro de Estudios de Conflicto y Cohesión Social (COES). www.coes.cl

Por último, en caso de que se desee citar el presente Manual de Usuario

Centro de Estudios de Conflicto y Cohesión Social (2018). Manual de Usuario de Estudio Longitudinal Social de Chile, Segunda Ola (2017). Corte Transversal. Versión 2.0. Santiago, Chile: Centro de Estudios de Conflicto y Cohesión Social (COES).

Capítulo 6

Dudas y Consultas

En caso de que detecte problemas con la base de datos, desee plantear solicitudes y/o tenga dudas sobre aspectos no cubiertos por el presente Manual de Usuario, las cuales no puedan ser resueltas por otros medios, puede comunicarse con el Equipo de Encuestas COES al correo electrónico encuestacoes@gmail.com. El equipo procesará su solicitud y tratará de contestar en el más breve plazo.

Capítulo 7

LIBRO DE CÓDIGOS BASE DE DATOS ELSOC OLA 2017

Para un uso adecuado de la base de datos de la primera ola de ELSOC COES se recomienda a los investigadores trabajar con el libro de códigos, el cual se presenta a continuación. Este apartado detalla el fraseo de cada uno de los ítemes incluidos, las distintas categorías de respuestas asociadas a éste, y el nombre de las variables como las etiquetas incorporadas en la base de datos. Ahora se incorporan recomendaciones generales para el uso de la base de datos y los libros de códigos.

7.1. Variables ELSOC Ola 2

La base de datos de ELSOC Ola 2 (2017) contiene una fila por cada entrevistado (son 2520 casos) y una columna por cada variable. Las variables corresponden a los ítems incluidos en el cuestionario del estudio. Ahora, es necesario que el usuario comprenda plenamente qué es un ítem.

La Figura ?? presenta, a modo de ejemplo, la pregunta M31 del cuestionario de ELSOC, dónde se consulta a los entrevistados si posee alguno de los siguientes bienes y servicios. En esta pregunta, se incluyen 4 ítems distintos, correspondientes a cada uno de los cuatro bienes o servicios. En el ejemplo, el primer ítem corresponde a: "Conexión a TV cable o satelital". Cada ítem corresponde a una variable incluida en la base de datos. El nombre de las variables de la base de datos se construye combinando:

- El código de la variable en el cuestionario, en letras minúsculas (en el ejemplo corresponde a m31).
- Valores numéricos correlativos al orden de los ítems, tal como aparecen en el cuestionario (01, 02, 03 y 04), separado por un guión bajo _ (en el primer caso, sería m31_01) .

El libro de códigos fue diseñado de modo de que sintetice toda la información relevante sobre las variables de la base de datos en un formato común para facilitar su uso. De manera genérica, las variables incluidas en la base de

M31 ¿Podría decirme si su casa tiene alguno de los siguientes bienes y servicios actualmente?

Lea alternativas

<u>Marque con una X</u>	Sí	No	No sabe (no leer)	No responde (no leer)
Conexión a TV cable o satelital				
Computador (ya sea notebook o de escritorio)				
Conexión a internet hogar pagada (no mediante celulares)				
Secadora de ropa				

Figura 7.1: Ejemplo de Pregunta con Múltiples Ítemes en ELSOC

datos tienen el siguiente formato:

variableenbase Fraseo completo de la pregunta y/o el ítem. Incluyendo instrucciones específicas a éste. *[Etiqueta de variable]*

1 Etiqueta categoría de respuesta 1 Frec. Absoluta de 1 % correspondiente a 1

2 Etiqueta categoría de respuesta 2 Frec. Absoluta de 2 % correspondiente a 2

A continuación, se presenta un ejemplo para clarificar dicha estructura. En este caso, **t2_01** es el nombre de dicho ítem del cuestionario en la base de datos (como variable). Por medio de estos códigos se pueden identificar los distintos ítemes incluidos. **Se recomienda rastrear las variables en el Libro de Códigos y en la base de datos según dicho nombre.** Junto a este código aparece la etiqueta de variable incluida en la base de datos ELSOC. En este caso es *Este es el barrio ideal para mí*. Las etiquetas fueron diseñadas por el equipo ELSOC con el objetivo de describir de modo sucinto el fenómeno o dimensión a medir¹, eliminando tildes y otros símbolos no incluidos en todos los softwares estadísticos (por ejemplo, la ñ).

Los distintos valores listados corresponden a las categorías de respuesta asociadas al ítem. En la construcción de la base de datos, dichas categorías fueron ingresadas como valores numéricos y se incluyeron etiquetas de valores. En el ejemplo anterior, si una persona se manifiesta totalmente en desacuerdo con el enunciado, en la encuesta a su respuesta se le asignará un valor numérico de 1, por lo que se podrán realizar operaciones aritméticas con

¹Se intenta preservar el ítem original. Los cambios introducidos se hicieron por motivos de espacio.

esta respuesta. De todos modos, se incluye una etiqueta para indicar que el valor 1 corresponde con la respuesta “Totalmente en desacuerdo”.

t2_01 *Este es el barrio ideal para mí* A continuación voy a leer algunas afirmaciones acerca de su barrio. ¿Qué tan de acuerdo o en desacuerdo está usted con las siguientes afirmaciones? (MOSTRAR TARJETA 2) Este es el barrio ideal para mí.

1&Totalmente en desacuerdo	& 118 & 3.95
2&En desacuerdo	& 419 & 14.04
3&Ni de acuerdo ni en desacuerdo	& 399 & 13.37
4&De acuerdo	& 1621& 54.32
5&Totalmente de acuerdo	& 426 & 14.28
8& No sabe (no leer)	& 1 & 0.03
9& No responde (no leer)	& 0 & 0.00

Por último, es importante también tener en cuenta la información no incluida en este libro de códigos. “A continuación voy a leer algunas afirmaciones acerca de su barrio. ¿Qué tan de acuerdo o en desacuerdo está usted con las siguientes afirmaciones?” corresponde la fraseo exacto de la pregunta, (MOSTRAR TARJETA 2) aparece entre paréntesis y corresponde a una instrucción para el encuestador. Estos elementos son incluidos de manera completa en el Cuestionario y el Listado de Variables correspondiente.

Las variables que poseen categorías residuales de no respuesta (No Sabe y No Responde), generalmente se codifican como 8/9 o 88/99, dependiendo de la anchura de la variable. **La base de datos de ELSOC 2017 homologa todos las respuestas No Sabe a -888 y No Responde a -999** y chequea la naturaleza de los valores perdidos restantes (los cuales corresponden en su totalidad a la aplicación de filtros). Las respuestas No Sabe y No Responde no fueron eliminadas ni transformadas en valores perdidos en la base de datos, de modo que los usuarios deben decidir qué hacer con dichos valores de respuesta. La etiqueta de valores de éstos incluye la instrucción de no ser leídos por el encuestador. En el caso de que la pregunta contenga un filtro, los valores no relevantes (filtrados) aparecen en la última línea y no tienen asignado valores ni etiquetas (son valores perdidos por el sistema).

Las variables de cadena (texto) no presentan códigos, ya que presentan los verbatim literales de las menciones por parte de los encuestados. El único cambio introducido -y fue realizado por motivos de compatibilidad entre softwares- fue la eliminación de símbolos tales como tildes, diéresis y otros caracteres especiales. A continuación se presenta el ejemplo de una de dichas preguntas:

c12_09esp A continuación le voy a leer una lista de organizaciones voluntarias. Para cada una de ellas, ¿podría decirme si es miembro activo, miembro inactivo, o no es miembro de estas organizaciones? **Otra, especifique:** [Especifique]

Variable de Cadena (caracteres).

En el caso de los ítems dónde se pide una respuesta numérica tampoco existen categorías de respuesta, registrándose el valor indicado por el entrevistado. De todos modos se incluyen etiquetas de valores para los valores perdidos. A continuación, un ejemplo de un ítem de este tipo:

m37_01 ¿Cuántos hijos hombres e hijas mujeres ha tenido usted en su vida? (Anote el número respectivo:) 1)
Número de hijos hombres. *[Numero de hijos hombres:]*

Respuesta Numérica 88 No sabe 99 No responde

7.2. Características Base de Datos ELSOC Ola 2017

La versión actual de la base de datos (1.0) contiene información para los 2520 entrevistados (N definitivo tras supervisión) en relación a los 317 ítems del cuestionario (cada una corresponde a una variable en la base de datos) y un conjunto adicional de variables:

1. **Identificadoras de Casos:** id de la encuesta, comuna de residencia y región de residencia.
2. **De Diseño Muestral Complejo:** ponderadores (y factores de expansión), estrato y segmento para incorporar el diseño complejo de la encuesta.

Dentro de las variables de diseño muestral complejo, el estrato identifica a los seis previamente definidos²: Gran Santiago (1), Gran Valparaíso (2), Gran Concepción (3), Ciudades Grandes(4), Ciudades Medianas (5) y Ciudades Pequeñas (6). La variable segmento1 representa las manzanas/bloques, pero los valores son artificiales (generados por MicroDatos y no corresponden a los valores reales) de modo de garantizar la privacidad de los entrevistados.

Se recomienda la utilización de los ponderadores muestrales que ajustan en base a la probabilidad de selección, no respuesta y población objetivo estimada a nivel regional. El ponderador 02 también ajusta en base a la población estimada según sexo.

Aún no se encuentran disponibles en la base de datos las variables georreferenciadas que proveerá el Centro de Inteligencia Territorial (CIT), siendo esto motivo de una nueva versión de la base de datos. Esto implicará un nuevo lanzamiento de la base de datos y todas las facilidades para que los investigadores hagan un uso óptimo de las variables territoriales.

Las bases de datos se encuentran disponibles en formatos .dta (compatibles con versiones Stata 13 y 14), .sav (compatibles con SPSS) y .RData (compatibles con R, los objetos contienen como atributos etiquetas de variables y valores). El Equipo de Encuestas COES también cuenta con bases de datos adicionales sobre la encuesta (en formato .dta y/o .xlsx):

1. Casos supervisados.
2. Capacitaciones a encuestadores.
3. Códigos de disposición final muestra completa.
4. Diseño y selección muestra.
5. Tabla Kish.
6. Construcción de ponderadores.

Sin embargo, los archivos anteriormente listados deben ser solicitados al Equipo de Encuestas, justificando el uso que se darán a los datos requeridos. El Equipo de Encuestas COES decidirá la pertinencia de dichas solicitudes y se

²Los estratos 4, 5 y 6 tienen sub-estratos Norte y Sur.

reserva el derecho para entregar dichos datos.