

**DISEÑO METODOLÓGICO DE LA  
PRIMERA ENCUESTA DE PROTECCIÓN SOCIAL  
(HISTORIA LABORAL Y SEGURIDAD SOCIAL)  
David Bravo<sup>1</sup>**

---

### *Introducción*

Con el fin de satisfacer la necesidad de generar información certera, confiable y desagregada a nivel individual relativa al sistema de pensiones chileno, la Subsecretaría de Previsión Social y el Departamento de Economía acordaron aplicar una encuesta de Protección Social a una muestra representativa de la población afiliada al sistema de pensiones.

La muestra representativa a la que se aplicó la encuesta se obtuvo a partir del universo de afiliados al sistema de pensiones tanto de capitalización individual como del sistema de reparto actualmente en extinción, información proporcionada por la Superintendencia de Administradoras de Fondos de Pensiones y por el Instituto de Normalización Previsional. La metodología aplicada para el diseño del cuestionario de la encuesta y la selección de los afiliados a encuestar se explica en detalle en la próxima sección.

El diseño metodológico del proyecto incluyó dos partes fundamentales: el diseño del cuestionario de la encuesta y el diseño muestral. Ambos procesos se desarrollaron en forma paralela, y en ellos el Departamento de Economía contó con la activa participación de los equipos de la Subsecretaría de Previsión Social y la Superintendencia de AFP. En particular, el diseño del cuestionario estuvo a cargo de un equipo formado por la Subsecretaría y el Departamento de Economía y contó con el valioso aporte de miembros de la Superintendencia de AFP y el Ministerio de Hacienda (Dirección de Presupuestos).

- En el equipo del Departamento de Economía de la Universidad de Chile participaron académicos internacionales de prestigio en sus distintas etapas.

Este informe también reporta información sobre el trabajo de levantamiento de las encuestas y los resultados finales del proceso.

---

<sup>1</sup> Profesor del Departamento de Economía de la Universidad de Chile y Director del Centro de Microdatos perteneciente a la misma institución.

## **1. Diseño del cuestionario**

Ciertamente el cuestionario es parte fundamental de cualquier encuesta que se desee aplicar. Uno de los factores determinantes en el éxito o fracaso de una encuesta, en términos de su utilidad para el posterior análisis de los datos, es el diseño del instrumento mediante el cual se recoge la información.

Dada la importancia de aplicar un cuestionario debidamente diseñado, el equipo del Departamento de Economía desarrolló el instrumento a aplicar en un largo periodo de tiempo, generando múltiples espacios de discusión y sometiendo el cuestionario a diversas revisiones y pruebas preliminares.

Antes de entrar en la etapa de creación de las preguntas y diseño del cuestionario, se definieron los grandes temas sobre los cuales se quería obtener información. Se ha definido como el objetivo principal de este proyecto modelar la participación de los individuos en el sistema de pensiones, la regularidad de las densidades de contribución y estimar la probabilidad de que sean elegibles para obtener subsidios estatales en la vida pasiva. Entonces, el cuestionario debía ser capaz de medir en forma apropiada los principales determinantes del comportamiento previsional de los individuos.

En este contexto, se definieron los siguientes ámbitos como factores explicativos en la participación en el sistema de pensiones: composición del hogar del individuo y principales características de cada uno de los miembros del hogar actual; nivel y características de la educación del individuo; características del hogar donde el afiliado se ha desarrollado y de los familiares (descendientes y ascendientes); comportamiento del individuos con respecto al riesgo y a la planificación temporal; efecto del diseño e incentivos del sistema de pensiones; nivel de conocimiento y percepción del sistema de pensiones; y características del empleo actual y de los empleos anteriores. En base a estos grandes tópicos se diseñó cada uno de los módulos del cuestionario.

Como referencia para el diseño del cuestionario se utilizaron diversas experiencias internacionales vigentes en la actualidad o aplicadas en el pasado reciente. Los principales cuestionarios considerados en este proceso fueron: la encuesta longitudinal Health and Retirement Study, conducida bianualmente a nivel nacional en Estados Unidos por la Universidad de Michigan a la población de 50 años o más; la encuesta de Previsión de Riesgos Sociales diseñada por Packard, T. perteneciente al Banco Mundial y tomada en la Región Metropolitana en noviembre de 1999 por el Departamento de Economía de la Universidad de Chile; y la Encuesta de Caracterización Socio Económica diseñada por el Ministerio de Planificación y Cooperación y aplicada bianualmente, también ejecutada por la División de Encuestas del Departamento de Economía de la Universidad de Chile.

A partir del objetivo de este proyecto, de los principales tópicos que se definieron como determinantes en la participación de los individuos en el sistema de pensiones y de los cuestionarios de encuestas relacionadas, se realizó un intenso trabajo de diseño de las preguntas más adecuadas para obtener la información relevante. Como metodología

de trabajo el equipo a cargo del diseño del cuestionario se reunió regularmente para discutir las propuestas de cada uno de los miembros y los valiosos aportes, revisiones y críticas de un selecto grupo de personas vinculadas a la Seguridad Social desde distintos ámbitos<sup>2</sup>. Este equipo de trabajo tuvo acceso a los borradores del cuestionario y sus comentarios e indicaciones fueron debida y oportunamente discutidos.

El cuestionario final tiene la siguiente estructura:

- Una página de portada, en la que se realizan las observaciones de terreno y está la identificación del individuo a encuestar. Un punto importante dice relación con el texto en el pie de la página, que explicita y solicita el consentimiento del individuo para ser aplicada la encuesta. La instrucción al encuestador indicaba que dicho texto debía ser leído al entrevistado. Parte de este texto es la explicación al encuestado que la identificación del entrevistado se mantiene en estricta reserva a fin de garantizar el secreto estadístico.
- Módulo I: Composición del hogar. Incluye un conjunto de 9 preguntas que el entrevistado debía responder para cada uno de los actuales integrantes de su hogar. Este módulo tiene por objetivo caracterizar el hogar en distintas variables básicas (sexo, condición de jefe de hogar, edad, estado civil y educación, entre otras).
- Módulo II: Información laboral y de ingresos. Incluye un total de 31 preguntas que el entrevistado debía responder para cada uno de los actuales integrantes de su hogar de 15 años o más. Tiene como objetivo caracterizar la situación ocupacional de los miembros del hogar en edad de trabajar. Para ello se recaba información sobre el estatus de actividad laboral, las características del empleo, la realización de cotizaciones provisionales, el monto de remuneración líquida en la ocupación principal y el total de ingresos del hogar.
- Módulo III: Sistema de Pensiones. Este módulo sólo incluye preguntas dirigidas al entrevistado seleccionado. En primer lugar, incluye 9 preguntas generales referidas a temas como expectativas de vida, expectativas de trabajo y retiro, afiliación en los sistemas de pensiones y realización de ahorros entre otras preguntas. En segundo lugar, se agregan 19 preguntas dirigidas fundamentalmente a los entrevistados que tienen una pensión. Aquí las preguntas cubren la edad de pensión, razones para el retiro anticipado y la continuación laboral, además del grado de información sobre la modalidad de pensión vigente y la tasa de reemplazo. En tercer lugar, se incluye un submódulo de pensiones de sobrevivencia destinado a aquellos entrevistados que cuentan con este tipo de pensión. Finalmente se incluyen 32 preguntas destinadas a todos los entrevistados, destinadas a recoger distintos aspectos de conocimiento y percepción del sistema de pensiones.
- Módulo IV: Educación del Entrevistado. Este módulo tiene 16 preguntas destinadas a caracterizar la formación básica, secundaria y superior de los entrevistados.

---

<sup>2</sup> Participaron en este proceso profesionales de la Superintendencia de AFP, de la Dirección de Presupuestos del Ministerio de Hacienda y de la Subsecretaría de Previsión Social, así como también expertos que en la actualidad se desempeñan en distintos ámbitos vinculados con la Seguridad Social.

- Módulo V: Historia Familiar. Este módulo tiene 29 preguntas que apuntan a caracterizar a los padres o tutores de los entrevistados, su vinculación con el mercado laboral y la situación económica del hogar donde crecieron los entrevistados.
- Módulo VI: Historia Individual. Este módulo tiene una variedad de temas referidos al entrevistado en 30 preguntas, que cubren desde los matrimonios y convivencias, condiciones de invalidez o discapacidad, hijos y otras características de prevención de salud y aversión al riesgo.
- Módulo VII: Historias de Condición de Actividad. Este módulo requirió que los entrevistados dieran cuenta cronológicamente de todas las actividades realizadas desde Enero de 1980 a la fecha, a partir de los 15 años. Se trata, por lo tanto, de un autoreporte de la dinámica en el mercado laboral, para lo cual y para cada actividad el entrevistado debía caracterizar sus empleos y períodos de inactividad o desocupación. Una vez revisada su historia laboral, los encuestados debían responder a preguntas de síntesis sobre su situación de cotizaciones y las razones detrás de los períodos de no cotización.
- Módulo VIII: Capacitación. En esta sección los entrevistados debían indicar las tres actividades de capacitación más importantes realizadas desde 1980, incluyendo educación de adultos.

## ***2. Diseño de la muestra representativa***

Durante todo el proceso de implementación del proyecto se ha procurado con especial ahínco el correcto tratamiento de la información en cuanto a la Protección de la Vida Privada establecida en la Ley 19.628 de 1999 y a lo establecido en el artículo N°10 del D.L. 3.500 de 1980.

Para ello, la Superintendencia de AFP y el INP proporcionaron al Departamento de Economía una base de datos *innominada* de todos los afiliados al sistema previsional de capitalización individual, es decir, se omitió todo campo que permitiera la identificación personal de los individuos.

En base a este marco muestral innominado, el Departamento de Economía seleccionó la muestra representativa según los criterios que se explican en la próxima sección.

Una vez generada esta selección innominada de individuos la Superintendencia y el Instituto de Normalización Previsional, construyeron la historia previsional para cada afiliado seleccionado. En forma paralela, el Departamento de Economía aplicó la encuesta a las personas seleccionadas, las cuales en modo alguno están obligadas a proporcionar los datos requeridos en el instrumento y tienen el compromiso de esta institución de no revelar su identificación en virtud del secreto estadístico.

### ***2.1. Diseño Metodológico***

El diseño metodológico de la selección de la muestra fue realizado por el Departamento de Economía de la Facultad de Ciencias Económicas y Administrativas de la Universidad de Chile, institución que tiene una extensa experiencia en el diseño y aplicación de encuestas representativas en Chile. En esta oportunidad, cabe destacar la participación en el equipo del Departamento de Economía, en calidad de consultor, del Sr. Steven Heeringa, experto en diseños muestrales complejos, PhD. y Director de la Unidad de Diseño y Análisis de Encuestas del Institute for Social Research de la Universidad de Michigan. Asimismo, en el desarrollo de esta tarea, destacó la estrecha coordinación de la contraparte técnica conformada por la División de Planificación y Estudios de la Subsecretaría de Previsión Social.

Dado que el objetivo del trabajo a realizar es el diseño de una muestra representativa a nivel nacional que permita la sistematización y análisis de la información sobre densidad de cotizaciones en el sistema de capitalización individual, se consideró como población de referencia a todos los afiliados desde 1981 hasta agosto de 2001 registrados en los archivos administrativos de las 7 AFP que están actualmente vigentes en el sistema así como también en el INP.

Se procedió, entonces, a seleccionar una “muestra histórica” de todos los individuos afiliados al sistema en 1981 o después, que a agosto de 2001 estuvieran en alguna de las siguientes condiciones:

- (1) trabajando,
- (2) cesantes o fuera de la fuerza de trabajo,
- (3) pensionados o
- (4) fallecidos.

Es necesario enfatizar que, conceptualmente, los pensionados y un subgrupo de los afiliados fallecidos debían ser incluidos en la población de referencia, por cuanto éstos formaron parte del universo de afiliados al sistema en algún momento del tiempo después de 1981. Su actual pensión de vejez o de sobrevivencia que causaron es el resultado del desempeño previsional durante la vida activa. Luego, el no incluirlos en la muestra representativa censuraría los datos relativos al patrón de comportamiento de las cohortes de edad al que esos individuos pertenecen, comprometiendo la representatividad de la muestra seleccionada.

### ***2.2. Marco muestral***

La Superintendencia de Administradoras de Fondos de Pensiones y el Instituto de Normalización Previsional proporcionaron al Departamento de Economía información acerca de la totalidad de los afiliados al sistema.

Esta información permitió dimensionar el marco muestral -referencia para la selección de la muestra- en 8.074.673 afiliados, los que se desagregan de acuerdo a lo mostrado en el cuadro N°1:

**Cuadro N°1**  
**Marco muestral Encuesta Protección Social**

Estatus del Afiliado	AFP	INP	Total
Total	6.593.815	1.480.858	8.074.673

### ***2.3. Diseño muestral por clusters, estratificación en dos etapas***

Para obtener una muestra representativa de los afiliados al sistema privado de pensiones, se consideró como mejor opción la utilización de un diseño en dos etapas estratificado por *clusters* (Kish, 1965, *Survey Sampling*). El diseño muestral en dos etapas asegura que cada afiliado elegible en el universo de individuos afiliados *tendrá similar probabilidad de ser seleccionado* y, al mismo tiempo, provee una forma más eficiente de agrupamiento de la muestra.

En consideración con la decisión anterior, y teniendo en cuenta la información disponible para cada afiliado del marco muestral, se consideró la variable *comuna* como la unidad más pequeña y fehaciente para formar los clusters (agrupamientos) de los individuos en los registros del sistema de pensiones.

El muestreo bietápico estuvo basado en una primera etapa de selección de comunas. La segunda etapa estratifica a los afiliados en las comunas seleccionadas en la 1° etapa, según diversas características relevantes, tal como se indica más adelante.

### ***2.4. Primera Etapa: Selección de comunas***

Las comunas pertenecientes al Gran Santiago más San Bernardo y Puente Alto fueron incluidas con probabilidad igual a uno en la primera etapa, así como también las comunas con más de 60.000 habitantes, es decir las que representan las mayores ciudades de Chile. Coyhaique, como capital de la XI región, también fue incluida con certeza, a pesar de tener una población de 45.485 habitantes en el año 2000<sup>3</sup>.

Por lo tanto, 81 comunas fueron seleccionadas con probabilidad uno, las que en conjunto representan aproximadamente un 72% del total de la población chilena<sup>4</sup>.

---

<sup>3</sup> La Selección de comunas auto-representativas fue hecha en base a la población de cada comuna reportada en la encuesta CASEN 2000. Coyhaique se incluyó ya que ninguna ciudad de la región de Aysén cuenta con 60.000 o más.

<sup>4</sup> Según información proveniente de la Encuesta CASEN 1998.

Del resto de las 285 comunas del país incluidas en la encuesta Casen 2000, 32 fueron elegidas en la primera etapa de selección. Para ello, se estratificaron las comunas según región, ingreso per cápita y población. Esto es, todas las comunas pertenecientes a una misma región se ordenaron según el ingreso per cápita de los habitantes en forma descendente.

Con este orden se formaron grupos de comunas que en conjunto tuvieron 100.000 habitantes aproximadamente<sup>5</sup>. Luego, en cada región se formaron entre 1 y 6 estratos de comunas, con una desviación estándar intra-regional del número de habitantes en cada estrato entre 4.134 y 7.893 y una desviación estándar nacional de 15.216. El que los estratos tengan tamaños poblacionales parecidos, hace que la primera etapa sea estadísticamente más eficiente.

De cada estrato se seleccionó 1 comuna que representara a toda la población de las restantes comunas en el estrato. Entonces, el objetivo primero de la definición de los estratos es agrupar comunas similares entre sí. Dadas las diferencias geográficas y de actividad económica entre las regiones, se intentó respetar los límites regionales formando estratos para comunas dentro de una misma región. Asimismo, como variable relacionada con la condición socio-económica de la comuna se utilizó el ingreso per cápita y se agruparon las comunas con ingreso per cápita similares.

Una vez formados los estratos (32), se seleccionó una comuna por estrato con la siguiente metodología: se definió la *probabilidad proporcional al tamaño (PPT)* para cada comuna en un estrato, como la razón entre el tamaño (población) comunal y el tamaño (población) del estrato ( $f_i$ ). Luego, se computó la probabilidad acumulada para cada una de las comunas ordenadas de mayor a menor según el ingreso per cápita promedio y, a partir de una distribución uniforme entre 0 y 1, se eligió un número aleatorio. Finalmente, dentro de cada estrato, se seleccionó la comuna que contuviera la probabilidad elegida aleatoriamente.

Es importante mencionar que las dos variables utilizadas en la primera etapa de selección (población comunal e ingreso per cápita) fueron obtenidas a partir de la encuesta CASEN 2000 y no de los datos de los afiliados al sistema previsional. Se adoptó esta estrategia debido a dos razones: (1) al momento de seleccionar las comunas no se contaba con los registros de los afiliados para conocer el real tamaño de cada comuna; y más importante; (2) los datos institucionales de los afiliados que se obtuvieron para la selección de individuos (segunda etapa de la selección de la muestra), no incluían la condición socio-económica, como por ejemplo, el ingreso, los años de educación, etc.

---

<sup>5</sup> Además del ingreso per cápita promedio de la comuna, se formaron estratos según años de escolaridad promedio de la población y con la tasa de pobreza. Asimismo, como una alternativa a la agrupación en torno a 100.000 habitantes, se intentó agrupar comunas que sumaran 70.000 individuos. Finalmente la combinación ingreso per-cápita y 100.000 habitantes resultó ser la combinación con mejores propiedades estadísticas.

## 2.5. Segunda Etapa: Selección de los afiliados

La segunda etapa de selección consistió en elegir a los afiliados dentro de cada comuna seleccionada, ya sea auto-representativa o no auto-representativa. A diferencia de la primera etapa, en este proceso se utilizaron los datos proporcionados por las Administradoras de Fondos de Pensiones.

Un tamaño muestral del orden de 20.000 casos parecía razonable considerando la indudable y clara importancia para el análisis y diseño de políticas relativas al sistema de pensiones. Esta fue, de hecho, la recomendación de tamaño muestral realizada por el Dr. Steve Heeringa. Con este orden de magnitud es posible llegar al nivel de desagregación y detalle necesario para obtener la representatividad adecuada de ciertos grupos de interés a la hora de fundamentar y evaluar propuestas específicas. El tamaño de la muestra de referencia fue de 30.000 casos para considerar las no respuestas o problemas de direcciones.

Así, considerando el número deseado de afiliados a seleccionar y el total de afiliados en las comunas seleccionadas, se calculó la tasa de muestreo total

( $f = \frac{n \text{ muestral}}{N \text{ poblacional}} = f_1 * f_2 = 0.00367$ ). Ya calculado  $f_1$ , se obtiene  $f_2$  que representa

la tasa a la que se seleccionará a los individuos en cada comuna.

Finalmente, dentro de cada comuna seleccionada se ordena a los afiliados según las siguientes variables:

- (1) condición del afiliado (activo, pensionado, fallecido);
- (2) sexo y
- (3) edad o fecha de afiliación para los afiliados al sistema antes o después de 1981, respectivamente.

Una vez que los afiliados estuvieron ordenados según estos criterios, se seleccionaron aplicando el intervalo de selección para cada comuna, definido como el recíproco de  $f_2$ .

Es importante señalar, que la metodología de selección de la muestra en dos etapas podría ser menos eficiente que una selección aleatoria estratificada. Sin embargo, para este proyecto en particular, el efecto de seleccionar a los afiliados por *clusters* dentro de cada comuna seleccionada en la primera etapa es atenuado por dos características del diseño. Primero, más del 70% de los afiliados residen en las comunas auto representativas y, dado que no se realiza más agrupamiento entre comunas, esta gran mayoría de afiliados no está sujeta a agrupamientos adicionales. Segundo, para el restante 30% de la población, que se encuentra en las comunas no auto-representativas, no se hizo un agrupamiento geográfico localizado dentro de cada comuna, esto es, la comuna es la unidad geográfica más pequeña que se utiliza. En los diseños muestrales multi-etápicas, además de la primera etapa de selección de comunas, se utiliza una segunda etapa de muestreo según áreas locales de residencia dentro de cada comuna (por ejemplo los segmentos en las encuestas Casen). Esta



etapa adicional aumenta la varianza muestral e introduce la correlación entre estas últimas unidades de agrupación.

En resumen, se espera que la pérdida en precisión del diseño en dos etapas sea muy pequeña en relación a las significativas ganancias en términos de costo-eficiencia de la muestra finalmente seleccionada.

### **3. Trabajo de Campo**

#### **(a) Cronograma de Trabajo**

El trabajo de campo para la toma y recolección de las encuestas se realizó entre el 24 de Mayo de 2002 y se extendió hasta el 15 de Enero de 2003, de acuerdo al siguiente calendario:

- Capacitación de supervisores: un total de 40 jefes de grupo y supervisores de todo el país fueron capacitados entre el 10 y el 12 de Mayo de 2002.
- Capacitación de encuestadores: realizada del 14 al 22 de Mayo de 2002 en la Región Metropolitana y del 23 al 30 de Mayo en regiones.
- Inicio de levantamiento de encuestas: 24 de Mayo de 2002.
- Término del levantamiento de encuestas: 15 de Enero de 2003.
- Término de validación y digitación de encuestas: 15 de Marzo de 2003.
- Realización de frecuencias simples para validar la base de datos: término el 31 de Marzo de 2003.
- Cálculo de factores primarios de expansión: Abril de 2003.

#### **(b) Descripción del trabajo de campo de encuestas**

El trabajo de campo se diseñó operativamente pensando en un núcleo central de dirección radicado en Santiago y nueve grupos operativos que administraran una región o agrupación de regiones, cada uno de ellos independiente, compuesto por encuestadores, controladores, jefes de grupo y dirigido por un jefe regional.

En el cuadro 2 se describe la distribución geográfica y la carga de trabajo de cada grupo operativo.

**Cuadro N°2**  
**Distribución Geográfica de la Muestra y Estructura**  
**del equipo de trabajo**

Región	Enc. Nominales	Jefes Regionales	Jefes de Control	Jefes de Grupo
1 y 2	1.901	1	0	3
3 y 4	1.561	1	0	3
5	3.158	1	0	5
6 y 7	3.202	1	0	5
8	3.331	1	0	5
9 y 10a	1.805	1	0	4
10b	1.417	1	0	4
11 y 12	556	1	0	1
13	13.756	1	2	14
<b>Total</b>	<b>30.687</b>	<b>9</b>	<b>2</b>	<b>44</b>

De las 30.687 personas se logró encuestar finalmente a 17.246. (56.2% de la muestra nominal). Las mayores dificultades se presentaron en la ubicación de los individuos dada la gran dispersión geográfica de la muestra y lo antiguo de la base de datos que incidió en que alguna información estuviera obsoleta, así, el 10.8% de los individuos se había cambiado a una dirección desconocida por los actuales moradores, un 5.9% de la muestra correspondía a direcciones que no eran las del afiliado, etc. Para solucionar este déficit se cruzó la información proveniente de las instituciones de previsión con bases de datos del Registro Civil y de Fonasa. Con todo, se llegó a las tasas de respuesta declaradas anteriormente y que se detallan en el cuadro N°3.

Del cuadro N°3 puede apreciarse que un conjunto importante de registros no tenía direcciones apropiadas. En las direcciones registradas no vivía nadie, o no correspondía a las personas buscadas, o bien se trataba de direcciones inubicables. En otros casos, el problema surgió por la antigüedad de la dirección referida, puesto que el entrevistado se había cambiado a una dirección desconocida.

Por lo tanto, para calcular la tasa de respuesta obtenida en el trabajo de campo, se deben excluir de la muestra nominal los casos anteriores. Con ello, **la tasa de respuesta ascendió a 76%**, una proporción elevada para encuestas que buscan seguir a personas específicas.

Otra dificultad importante fue la oportunidad de la encuesta: tomar una encuesta de tal envergadura en invierno, en regiones del sur especialmente, plantea dificultades adicionales por las inclemencias del clima y el poco tiempo de luz solar diaria, lo que acorta el período útil de encuestaje.

**Cuadro N°3  
Estado Final de la Muestra**

Estado de la muestra	Total	%
1. Hecha	17.246	56,2%
2. No hay nadie	1.760	5,7%
3. No vive nadie	314	1,0%
4. No existe la dirección	643	2,1%
5. Se niega a contestar	1.209	3,9%
6. Dirección inubicable	1.459	4,8%
7. Se cambió a dirección desconocida	3.360	10,9%
8. Se cambió a dirección conocida	394	1,3%
9. Dirección no corresponde al afiliado	1.804	5,9%
10. Otro	2.498	8,1%
<b>Total general</b>	<b>30.687</b>	<b>100%</b>

El formulario de la encuesta fue fácil de aplicar, los encuestados se sentían parte del tema, lo que hacía agradable la entrevista, la duración aproximada de 40 minutos en promedio fue algo extensa. En términos de contenido, lo más lento fue reconstruir la historia laboral.

En total se utilizó un número aproximado de 500 encuestadores en todo el proceso de terreno.

### **(c) Capacitación**

#### **Capacitación de supervisores**

En Santiago durante dos días se capacitó a 45 aspirantes a jefes de gupo provenientes de todo el país.

Esta capacitación se realizó en dependencias de la Facultad de Economía de la Universidad de Chile, y fue dirigida por el jefe del proyecto en la Universidad de Chile y contó con la participación de profesionales de la Subsecretaría de Previsión Social expertos en legislación laboral y previsional.

A esta capacitación también asistieron todos los jefes regionales y personal directivo del equipo de trabajo de la Unidad de Encuestas

## **Capacitación de Encuestadores**

Con los conceptos recogidos en la capacitación de jefes de grupo se diseñó la capacitación de encuestadores que consistió en dos jornadas de capacitación de aproximadamente 4 horas cada una distribuidas en dos días no consecutivos dirigidas por los jefes regionales.

En la inscripción de encuestadores se les proporcionó un glosario de los temas que serían tratados en la capacitación, al final de cada jornada se les aplicó un test para medir la comprensión de los elementos entregados, este test también se usó como criterio de selección de encuestadores.

Durante las capacitaciones de los encuestadores también se contó con apoyo de profesionales expertos en el áreas.

En regiones la capacitación de jefes de grupo se difirió en una semana con respecto a la capacitación de la Región Metropolitana

## **Control de las encuestas**

El control del trabajo de terreno se programó de tal forma de tomar contacto telefónico con al menos un 10% de los entrevistados.

A cada encuestador se le controló al menos un 10% de las encuestas que realizó, el control se realizó durante todo el periodo de la encuesta, demás se hizo una muestra aleatoria de aquellas encuestas que no tenían número telefónico para controlarlas mediante visitas a terreno. El porcentaje total de encuestas controladas fue de 11,8%.

El control consistió en repetir preguntas claves del formulario para verificar el correcto uso del instrumento y detectar posibles fraudes.

Las principales deficiencias se detectaron al aplicar la historia laboral, ya que algunos encuestadores no fueron todo lo rigurosos que se requería en el seguimiento de las actividades y episodios laborales. En todo caso fueron faltas leves que pudieron solucionarse mediante visitas posteriores.

## **(d) Digitación y verificación**

El proceso de digitación fue realizado por la Unidad de Computación del Departamento de Economía y se realizó mediante una aplicación de captura de datos especialmente desarrollada para esta encuesta, que impide el ingreso de valores no válidos en los campos respectivos. Entraron los datos ocho digitadores calificados y experimentados, los que fueron supervisados por el programador y un operador jefe .

Para la validación se desarrolló otra aplicación en donde se programaron los chequeos de coherencia y de validez de rango de las variables. La aplicación desplegaba en pantalla una imagen de la encuesta con los valores de las variables y una descripción de los errores permitiendo que un operador experto en el tema de la encuesta (Jefe de grupo) fuera corrigiendo dichos errores y, en un proceso de iteraciones, dejara la encuesta libre de errores.

Las características de los temas de la encuesta: identificación, descripción del hogar, información laboral y de ingresos, pertenencia y opinión del sistema de pensiones, educación, historia familiar, historia individual, historia de condición de actividad y capacitación; hacen que haya múltiples relaciones entre las variables recogidas, por lo tanto el programa de validación de coherencia recorre prácticamente todas las variables por lo que se puede asegurar que la calidad de la base de datos es del mas alto nivel.

Adicionalmente para verificar que el programa de validación estuviera efectivamente funcionando, ayudantes de investigación asignados al proyecto tabularon frecuencias y cruce de variables para mejorar la implementación del mismo, dando lugar finalmente a una base de datos confiable.

El proceso de digitación y validación se realizó durante los meses de enero de 2002 a marzo de 2003