



**METODOLOGÍA 2007
COMERCIO Y SERVICIOS
ENCUESTA ESTRUCTURAL**

Versión 1.0

ÍNDICE

| | Páginas |
|--|----------------|
| 1. Introducción | 3 |
| 2. Decisión de realizar una encuesta | 4 |
| 3. Diseño de la encuesta | 5 |
| 4. Recogida de datos | 16 |
| 5. Captura y tratamiento de los datos | 17 |
| 6. Análisis de los datos y calidad de los resultados | 18 |
| 7. Documentación y difusión | 20 |
| 8. Ciclo de mejora | 20 |

1. INTRODUCCIÓN

a. Población Objetivo.

La población objetivo de la encuesta está constituida por todas las empresas del país que desarrollan actividades de Comercio al por menor, Comercio al por Mayor, Servicios Inmobiliarios, Empresariales, de Alquiler, de Alojamiento Turístico y de Alimentación. Estas actividades corresponden a aquellas clasificadas en las secciones G, H, K y O del compendio CIIU Rev. 3.

b. Cobertura geográfica y temática.

La información que cubre un año calendario, es recopilada a través de una muestra que tiene representación a nivel nacional, seleccionada a partir del directorio del Servicio de Impuestos Internos, SII, que incluye las empresas o contribuyentes situados dentro de los límites territoriales del país.

Se excluyen del estudio las siguientes actividades: Intermediación Financiera, Investigación y Desarrollo, Administración Pública y Defensa, Planes de Seguridad Social de Afiliación Obligatoria, Enseñanza, Actividades de Servicios Sociales y de Salud, Hogares Privados con Servicio Doméstico, Organizaciones y Órganos Extraterritoriales.

c. Periodo de referencia y periodicidad

El periodo de levantamiento de la encuesta es anual y el periodo de referencia corresponde al año contable anterior. El desfase de la publicación es de 17 meses después de terminado el periodo analizado, esto está acorde con los estándares internacionales ya que la encuesta sólo se puede realizar una vez cerrados y depurados los años contables de las empresas. La publicación se realiza el 30 de Mayo de cada año.

d. Utilidad y oportunidad de la información a recoger.

Los datos obtenidos a través de la encuesta son de utilidad para el desarrollo de políticas públicas y privadas de mediano y largo plazo. Específicamente, la encuesta tiene como objetivo describir la estructura económica y contable de cada actividad. Esto permite ocupar la información obtenida como base para el cálculo de los índices que miden las variaciones coyunturales producidas en cada sector y desarrollados por el mismo INE.

Además, los datos recopilados alimentan como información de referencia al cálculo de las Cuentas Nacionales, complementando el estudio de las matrices de Oferta y Utilización desarrolladas por el Banco Central de Chile.

La actualización del diseño de los formularios y el diseño muestral se analizan en profundidad cada 5 años, en un proceso denominado "Compilación de referencia". El objetivo de este proceso es actualizar los datos estadísticos que representan la estructura de la función de producción de cada sector.

2. DECISIÓN DE REALIZAR UNA ENCUESTA

a. Procedimientos de interrelación con usuarios.

El resumen de la información recogida por la encuesta se encuentra disponible en la página web del INE, con acceso al público en general. Además, existe la posibilidad de que cualquier institución pública o privada, como por ejemplo: universidades, alumnos tesistas, centros de investigación, ministerios, etc., accedan al detalle de los datos innominados e indeterminados si lo solicitan al INE a través de su sección de comercialización.

Sin embargo, el usuario principal de los datos capturados es el Banco Central de Chile. La relación con dicha organización es estrecha y continua. Se constituye de reuniones de trabajo ocasionales para trabajar en los formularios y en el enfoque de las preguntas y por los envíos mensuales de las bases de datos preliminares y finales de las bases de datos y su respectiva retroalimentación.

En todos los casos la información es entregada utilizando un Rut ficticio para cada empresa presente en la muestra, acorde con el concepto de "secreto estadístico" bajo el cual trabaja el INE.

b. Análisis de las necesidades de los usuarios.

El diseño de los formularios, las preguntas que se incluyen en ellos y los conceptos detrás de cada pregunta son producto de un proceso constante de análisis y trabajo en equipo donde el objetivo principal es el aumento de la calidad de la información capturada y su utilidad.

Estas mejoras son causa principalmente del conocimiento empírico capturado por el personal del Departamento de Comercio y Servicios y la interacción con diversas instituciones que utilizan dicha información. Así, el análisis principal de estas mejoras continuas se realiza en conjunto con el BCCh., durante las reuniones de trabajo nombradas anteriormente.

c. Objetivos de la encuesta.

El objetivo del estudio es obtener cifras agregadas anuales a nivel nacional, en relación a la situación de las siguientes actividades:

Comercio

- Comercio al por mayor
- Comercio al por menor

Servicios

- Servicios de alimentación y alojamiento
- Servicios personales y servicios a empresas

La información anterior debe permitir configurar un cuadro esquemático de la situación por la que atraviesan estos sectores de la actividad económica del país, y también conocer, de alguna manera, la composición básica de la oferta y distribución de los bienes y servicios.

Entre las variables que se investigan se encuentran, por ejemplo:

- Ventas e Ingresos
- Remuneraciones
- Empleo
- Compras
- Gastos y Activo Fijo
- Otras

d. Planeación de recursos.

La Encuesta Estructural de Comercio, Servicios y Turismo es un producto permanente del INE por lo que sus recursos económicos son determinados por la institución. Dichos recursos son administrados por el coordinador del departamento.

La administración relacionada con los recursos humanos, los tiempos y procesos que deben desempeñar estos, es de responsabilidad del jefe de producto.

3. DISEÑO DE LA ENCUESTA

a. Conceptos de la encuesta y delimitación del campo en estudio.

El campo en estudio de la encuesta se limita al ámbito contable de la empresa, el cual es profundizado por preguntas enfocadas a contrastar datos, a definir la actividad económica de la empresa o a regionalizar la información.

Los conceptos más relevantes a tratar en la encuesta son:

Comercio, es la reventa (venta sin transformación) habitual de bienes nuevos o usados a usuarios intermedios o finales.

Comercio al por Mayor, es la reventa (venta sin transformación) de productos nuevos o usados a comerciantes minoristas, usuarios industriales o comerciantes, instituciones y profesionales.

Comercio al por Menor, es la reventa (venta sin transformación) de productos nuevos o usados al público en general, para consumo, uso personal o doméstico, actuando algunos comerciantes como agentes y vendiendo en consignación o a comisión.

Servicios, es aquella actividad económica en donde la venta corresponde a una prestación de desarrollo intelectual o bienes intangibles.

Servicio de Alimentación, se define como los establecimientos que expenden alimentos y bebidas preparadas para el consumo inmediato.

Servicio de Alojamiento, se incluye el suministro, a cambio de una retribución, de hospedaje temporal y de lugares y servicio para acampar.

Otros Servicios, se refiere a los servicios que se proporciona a los hogares, empresas o personas y establecimientos que pueden describirse como oficinas, teatros, talleres, salones y otros.

Ventas e Ingresos, este concepto comprende el valor de ventas de todos los bienes cuya propiedad o usufructos se ha cedido a terceros con miras a su compra efectiva, y el valor de todos los servicios prestados durante el período que mide la encuesta, cualquiera sea el plazo o la forma de pago.

Compras, es el costo en que incurre la empresa por la adquisición de mercaderías destinadas a la venta sin transformación no incluyendo el impuesto al valor agregado, pero sí en algunas actividades incluye los impuestos específicos, adicionales y sobretasas.

Gastos Operacionales, son aquellos en que incurre la empresa para la realización de su actividad, y que se pueden dividir en bienes y servicios.

b. Diseño muestral (Sólo para encuesta por muestreo)

Unidad Estadística

La unidad de estadística es la empresa, organización que tiene iniciación de actividades independientes, es decir, RUT y contabilidad propia, y cuyo giro lo puede realizar en uno o más locales comerciales, que desarrollan en forma principal la actividad de Comercio o Servicios.

Marco Muestral

El marco muestral se construyó a partir de los directorios que registra el Servicio de Impuestos Internos, S.I.I. (Declaración Anual de Impuestos a la Renta, formulario 22 y Declaración Mensual del IVA, formulario 29), de los contribuyentes que desarrollaron actividades de Comercio, Servicio, Hoteles y Restaurantes, en el año 2007.

Para generar el marco definitivo, se consideraron las empresas que en conjunto acumulan el 95% de las ventas para cada CIIU Rev. 3 (a tres dígitos). Esto debido al bajo aporte que presentan algunas empresas en el nivel de ventas, y lo dificultoso que resulta la investigación de unidades extremadamente pequeñas que además tienen una gran inestabilidad en el tiempo.

Estratificación

De acuerdo a los objetivos del estudio, se procedió a realizar una estratificación según la CIIU Rev. 3 (a tres dígitos), quedando definidos los siguientes estratos de interés.

Cuadro 1. Descripción de las Clasificaciones Consideradas

| División y Grupo | Descripción |
|-------------------------|--|
| 50 | Venta, Mantenimiento y Reparación de Vehículos Automotores y Motocicletas; Venta al por Menor de Combustible para Automotores |
| 501 | Venta de vehículos automotores |
| 502 | Mantenimiento y reparación de vehículos automotores |
| 503 | Venta de partes, piezas y accesorios de vehículos automotores |
| 504 | Venta, mantenimiento y reparación de motocicletas y de sus partes, piezas y accesorios |
| 505 | Venta al por menor de combustible para automotores |
| 51 | Comercio al por Mayor y en Comisión, excepto el Comercio de Vehículos Automotores y Motocicletas |
| 511 | Venta al por mayor a cambio de una retribución o por contrata |
| 512 | Venta al por mayor de materias primas agropecuarias, animales vivos, alimentos, bebidas y tabaco |
| 513 | Venta al por mayor de enseres domésticos |
| 514 | Venta al por mayor de productos intermedios, desperdicios y desechos no agropecuarios |
| 515 | Venta al por mayor de maquinaria, equipo y materiales |
| 519 | Venta al por mayor de otros productos ^{1/} |
| 52 | Comercio al por Menor, excepto el Comercio de Vehículos Automotores y Motocicletas; Reparación de Efectos Personales y Enseres Domésticos |
| 521 | Comercio al por menor no especializado en almacenes |
| 522 | Venta al por menor de alimentos, bebidas y tabaco en almacenes especializados |
| 523 | Comercio al por menor de otros productos nuevos en almacenes especializados |
| 524 | Venta al por menor en almacenes de artículos usados |
| 525 | Comercio al por menor no realizado en almacenes |
| 526 | Reparación de efectos personales y enseres domésticos |
| 55 | Hoteles y Restaurantes |
| 551 | Hoteles; campamentos y otros tipos de hospedaje temporal |
| 552 | Restaurantes, bares y cantinas |
| 63 | Actividades de Transporte Complementarias y Auxiliares; Actividades de Agencias de Viajes ^{1/} |
| 630 | Actividades de transporte complementarias y auxiliares; actividades de agencias de |
| 70 | Actividades Inmobiliarias |
| 701 | Actividades inmobiliarias realizadas con bienes propios o arrendados |
| 702 | Actividades inmobiliarias realizadas a cambio de una retribución o por contrata |

| | |
|-----------|---|
| 71 | Alquiler de Maquinaria y Equipo Sin Operarios y de Efectos Personales y Enseres Domésticos |
| 711 | Alquiler de equipo de transporte |
| 712 | Alquiler de otros tipos de maquinaria y equipo |
| 713 | Alquiler de efectos personales y enseres domésticos n.c.p. |
| 72 | Informática y Actividades Conexas |
| 721 | Consultores en equipo de informática |
| 722 | Consultores en programas de informática y suministro de programas de informática |
| 723 | Procesamiento de datos |
| 724 | Actividades relacionadas con bases de datos |
| 725 | Mantenimiento y reparación de maquinaria de oficina, contabilidad e informática ^{1/} |
| 729 | Otras actividades de informática ^{1/} |
| 74 | Otras Actividades Empresariales |
| 741 | Actividades jurídicas y de contabilidad, teneduría de libros y auditoría; asesoramiento en materia de impuestos; estudio de mercados y realización de encuestas de opinión pública; asesoramiento empresarial y en materia de gestión |
| 742 | Actividades de arquitectura e ingeniería y otras actividades técnicas |
| 743 | Publicidad |
| 749 | Actividades empresariales n.c.p. |
| 90 | Eliminación de Desperdicios y Aguas Residuales, Saneamiento y Actividades Similares |
| 900 | Eliminación de desperdicios y aguas residuales, saneamiento y actividades similares |
| 91 | Actividades de Asociaciones N.C.P. |
| 911 | Actividades de organizaciones empresariales, profesionales y de empleadores |
| 912 | Actividades de sindicatos |
| 919 | Actividades de otras asociaciones |
| 92 | Actividades de Esparcimiento y Actividades Culturales y Deportivas |
| 921 | Actividades de cinematografía, radio y televisión y otras actividades de entretenimiento |
| 922 | Actividades de agencias de noticias |
| 923 | Actividades de bibliotecas, archivos y museos y otras actividades culturales |
| 924 | Actividades deportivas y otras actividades de esparcimiento |
| 93 | Otras Actividades de Servicios |
| 930 | Otras actividades de servicios |

1/ Clasificación incorporada en el levantamiento 2008, para la obtención de datos año 2007.

Al examinar el comportamiento de la variable de interés, que es la venta por empresa, se puede apreciar su heterogeneidad y gran asimetría, esto es, un gran número de ellas con pequeñas ventas y unas pocas con grandes volúmenes de ventas, lo que determinó la conveniencia de realizar una segunda estratificación por tamaño (grandes, medianas y pequeñas), de acuerdo a las ventas al interior de cada grupo CIIU (Rev. 3), de interés.

De acuerdo al estudio estadístico realizado a la variable ventas en los sectores de Comercio, Servicios y Turismo, se determinaron los siguientes intervalos para los tamaños:

Cuadro 2. Descripción de los cortes por intervalos de ventas según tamaño.

| Tamaño de ventas | Ventas (\$ 2007) | |
|------------------|------------------|-----------------|
| | Limite Inferior | Limite Superior |
| Grande | 3.115.354.005 | Más |
| Mediano | 794.630.133 | 3.115.354.004 |
| Pequeño | Menos | 794.630.132 |

Departamento de Investigación y Desarrollo, INE

El objetivo de realizar las mencionadas estratificaciones, es homogenizar la variable de interés dentro de cada CIIU Rev.3, de forma tal de optimizar la muestra, lo que no implica que se obtengan estimaciones a nivel de tamaño,

dado que la muestra fue elaborada para obtener estimaciones según los estratos definidos en el Cuadro 1.

Como una forma de obtener una mejor representación y disminuir el error de muestreo que proviene de la dispersión observada en las ventas, se consideró, una vez ordenados los estratos por la variable de interés, integrar las unidades más grandes como Tramo a Censar (TC), y el resto como Tramo a Muestrear (TM).

El punto de corte para estos tramos, se determinó según el aporte de las ventas de las empresas a censar y el error de muestreo que resulta de restar al marco muestral las empresas que afectan notoriamente la precisión de las estimaciones. Estas empresas adquieren carácter censal o de Inclusión Forzosa (IF).

Unidades de Muestreo y de Información

La unidad de muestreo es la empresa que cuenta con iniciación de actividades para comercializar bienes y servicios, pertenecientes al comercio por mayor, comercio por menor, servicios, hoteles y restaurantes, que desarrollan su actividad comercial dentro de los límites territoriales del país.

La variable de interés que se considera para el estudio muestral respectivo, son las ventas anuales declaradas en el formulario 29 del SII, del año 2007, de cada contribuyente que realizó algún tipo de actividad comercial o de servicios descrita en el Cuadro 1.

Tamaño de la Muestra

La estimación del tamaño de la muestra a nivel nacional, fue obtenida a partir de un muestreo probabilístico estratificado, según las agrupaciones CIIU Rev. 3 (tres dígitos), con un coeficiente de confianza del 95%, alcanzando a 5.282 empresas que corresponde al 5% del total de las unidades del marco muestral. Para la determinación del tamaño muestral de cada estrato, se distribuyó el total de unidades muestrales en forma proporcional al nivel de ventas de cada uno de ellos. Se aplicaron los siguientes algoritmos:

Tamaño de la Muestra:

$$n = \frac{(Z^2 * N * Q)}{((Z^2 * Q) + (P * N * e_r^2))}$$

Donde:

n : Tamaño de la muestra.

Z : Estadístico normal estándar correspondiente a un nivel de confianza 95%.

re : Error relativo.

P : Valor 0,5, asumiendo una varianza máxima.

Q : Valor 0,5, asumiendo una varianza máxima.

N : Tamaño de la población.

Distribución del tamaño muestral en forma proporcional de acuerdo al nivel de ventas de cada estrato.

$$n_h = n * \frac{\sum_{i=1}^k y_{hi}}{\sum_{h=1}^H \sum_{i=1}^k y_{hi}}$$

Donde:

n_h : Tamaño del estrato h.

n : Tamaño de la muestra.

$\sum_{i=1}^k y_{hi}$: Sumatoria de las ventas de las empresas en el tramo aleatorio del estrato h.

$\sum_{h=1}^H \sum_{i=1}^k y_{hi}$: Sumatoria de las ventas de las empresas del tramo aleatorio de todos los estratos.

Sin embargo, el tamaño definitivo de cada uno de los estratos estuvo sujeto a los costos versus precisión, por lo cual se debió aumentar o disminuir el número de empresas de algunos estratos para fortalecer su representación y precisión necesaria para el estudio.

Para determinar la precisión muestral respectiva, se utilizaron los siguientes algoritmos:

Coefficiente de variación del estrato h

$$CV_h = \frac{\sqrt{V(\hat{Y})_h}}{\sum_{i=1}^k y_{hi}}$$

Donde:

CV_h : Coeficiente de variación del estrato h.

$V(\hat{Y})_h$: Varianza muestral estimada de las ventas del estrato h.

$\sum_{i=1}^k y_{hi}$: Sumatoria de las ventas del estrato h.

Varianza muestral de las ventas de las empresas

$$V(\hat{Y})_h = \sum_h^H N_h * (N_h - n_h) * \frac{S_h^2}{n_h}$$

Donde:

- $V(\hat{Y})_h$: Varianza muestral estimada de las ventas.
 S_h^2 : Cuasi-varianza de las ventas de las empresas del estrato h.
 N_h : Número de empresas en el estrato h.
 n_h : Tamaño muestral del estrato h.

Cuasi-varianza de las ventas de las empresas del estrato h

$$S_h^2 = \frac{N_h}{(N_h - 1)} * \sigma_h^2$$

Donde:

- S_h^2 : Cuasi-varianza que mide la dispersión de las ventas de las empresas de la muestra en el estrato h.
 N_h : Número de empresas en el estrato h.
 σ_h^2 : Varianza de las ventas del estrato h.

Varianza de las ventas del estrato h

$$\sigma_h^2 = \frac{\sum_{i=1}^k (y_{hi} - \bar{y}_h)^2}{N_h}$$

Donde:

- σ_h^2 : Varianza de las ventas del estrato h.
 y_{hi} : Venta de la i-ésima empresa del estrato h.
 \bar{y}_h : Media aritmética de las ventas en el estrato h.
 N_h : Número de empresas en el estrato h.

Cuadro 3. Descripción Tamaño Muestra Efectiva

| Clasificación CIIU Rev.3 (tres Dígitos) | Marco Muestral Nº | Muestra efectiva | | | | Coeficiente de Variación (%) |
|---|-------------------------|------------------------------|-------------------|-------------------------|----------------------|---------------------------------------|
| | | Total muestra efectiva | Inclusión forzosa | | Muestra Aleatoria | |
| | | | Nº | Aporte Ventas (%) | Nº | |
| Total | 106.421 | 5.282 | 635 | 47,24 | 4.647 | 0,71 |
| 501 | 274 | 140 | 58 | 74,66 | 82 | 1,18 |
| 502 | 2.069 | 170 | 10 | 41,49 | 160 | 5,95 |
| 503 | 841 | 163 | 21 | 38,97 | 142 | 2,08 |
| 504 | 61 | 38 | 13 | 73,73 | 25 | 2,41 |
| 505 | 782 | 155 | 14 | 34,41 | 141 | 1,87 |
| 511 | 846 | 46 | 17 | 59,21 | 29 | 4,94 |
| 512 | 1.565 | 291 | 33 | 49,90 | 258 | 2,06 |
| 513 | 1.097 | 226 | 24 | 35,44 | 202 | 1,94 |
| 514 | 806 | 261 | 24 | 60,67 | 237 | 1,97 |
| 515 | 802 | 221 | 35 | 55,32 | 186 | 1,86 |
| 519 | 4.086 | 29 | | 0,00 | 29 | 14,26 |
| 521 | 16.189 | 263 | 15 | 63,92 | 248 | 2,02 |
| 522 | 19.503 | 342 | 11 | 8,97 | 331 | 7,60 |
| 523 | 12.995 | 416 | 16 | 34,21 | 400 | 2,56 |
| 524 | 477 | 44 | 5 | 16,48 | 39 | 18,62 |
| 525 | 575 | 18 | 7 | 44,26 | 11 | 17,49 |
| 526 | 3.075 | 97 | 7 | 9,81 | 90 | 16,51 |
| 551 | 1.797 | 163 | 17 | 29,36 | 146 | 4,94 |
| 552 | 8.510 | 336 | 19 | 32,89 | 317 | 3,89 |
| 630 | 193 | 65 | 6 | 45,65 | 59 | 2,14 |
| 701 | 2.307 | 174 | 17 | 25,90 | 157 | 3,07 |
| 702 | 408 | 77 | 14 | 59,04 | 63 | 2,94 |
| 711 | 919 | 78 | 12 | 53,02 | 66 | 5,66 |
| 712 | 397 | 75 | 19 | 47,80 | 56 | 1,86 |
| 713 | 367 | 71 | 5 | 44,06 | 66 | 5,56 |
| 721 | 83 | 41 | 6 | 28,10 | 35 | 3,45 |
| 722 | 237 | 105 | 28 | 72,41 | 77 | 1,14 |
| 723 | 43 | 36 | 13 | 78,26 | 23 | 0,39 |
| 724 | 56 | 16 | 7 | 54,49 | 9 | 5,25 |
| 725 | 205 | 20 | 4 | 13,27 | 16 | 17,25 |
| 729 | 10 | 5 | 2 | 49,42 | 3 | 16,76 |
| 741 | 4.385 | 186 | 13 | 27,41 | 173 | 3,84 |
| 742 | 1.902 | 163 | 19 | 31,21 | 144 | 2,79 |
| 743 | 1.232 | 148 | 20 | 32,40 | 128 | 3,82 |
| 749 | 6.729 | 238 | 20 | 25,28 | 218 | 3,25 |
| 900 | 69 | 23 | 8 | 51,91 | 15 | 3,65 |
| 911 | 124 | 31 | 6 | 27,71 | 25 | 10,91 |
| 912 | 27 | 6 | 3 | 49,18 | 3 | 9,18 |
| 919 | 229 | 38 | 8 | 42,25 | 30 | 9,60 |
| 921 | 528 | 91 | 18 | 60,71 | 73 | 2,45 |
| 922 | 65 | 24 | 13 | 77,84 | 11 | 4,20 |
| 923 | 18 | 12 | 5 | 76,77 | 7 | 2,99 |
| 924 | 396 | 70 | 13 | 61,99 | 57 | 3,39 |
| 930 | 9.142 | 70 | 10 | 7,50 | 60 | 17,28 |

Departamento de Investigación y Desarrollo, INE.

El aumento en el Coeficiente de Variación y la disminución en el aporte de empresas IF en algunas clasificaciones, específicamente 524; 525; 526; 911; 912; 919; 930, se debe al aumento de sesgos no atribuibles a la muestra y que dependen de la dinámica y la complejidad del sector.

Los sesgos no atribuibles a la muestra, y que dependen de la dinámica y la complejidad del sector se pueden resumir de la siguiente manera:

- Calidad del directorio del universo
- Los cambios de actividades
- Formas de comercialización
- Localización geográfica de las empresas

Como una forma de cubrir los posibles cambios o cese de actividad, que pudieran presentar algunas empresas seleccionadas, entre la fecha de referencia de los marcos muestrales y la fecha del levantamiento, se considero un porcentaje de empresas para eventuales reemplazos. En este caso alcanzó a un 15% adicional de unidades muestrales.

Método de Selección

Con el objetivo de mantener un panel de empresas levantado el año anterior y a solicitud del Banco Central de Chile, se realizó el correspondiente pareo de empresas para ser incluidas en la selección respectiva, teniendo en consideración los nuevos tamaños obtenidos a partir del marco muestral actualizado al año 2007.

Considerando lo descrito anteriormente, se procedió a efectuar la primera selección a las empresas que pertenecen al tramo censado, las cuales son incluidas forzosamente en la muestra.

La segunda selección corresponde al tramo a muestrear, la cual se realizó en forma independiente para cada sector y tamaño, aplicando una selección sistemática, con el objetivo de asegurar un recorrido completo dentro de cada uno de los estratos.

Factores de Expansión

La información recolectada, es representada por medio de una muestra de empresas, por lo cual, para obtener una visión completa de las actividades que componen el sector, se debe realizar una expansión de la información levantada.

Para realizar la mencionada expansión, se debe considerar en primer lugar las empresas de inclusión forzosa que son unidades muestrales autorepresentadas, ya que tienen carácter censal. En el caso que una empresa no conteste y continúe desarrollando su actividad durante el periodo de referencia, el algoritmo de cálculo es el siguiente:

Factor de expansión de las empresas de inclusión forzosa (IF).

$$FE(IF)_h = \frac{\sum_{i=1}^K y(IF)_{h,i}}{\sum_{i=1}^K y(IFe)_{h,i}}$$

Donde:

- $FE(IF)_h$: Factor de expansión de las empresas en el estrato h .
- $\sum_{i=1}^K y(IF)_{h,i}$: Sumatoria de las ventas del marco muestral correspondiente a las empresas IF del estrato h .
- $\sum_{i=1}^K y(IFe)_{h,i}$: Sumatoria de las ventas del marco muestral que corresponden a las empresas IF efectivamente logradas, del estrato h .

Recuérdese que para homogenizar la variable en estudio se realizó una segunda estratificación en el tramo aleatorio, de acuerdo al tamaño según el nivel de ventas que declararon las empresas. Esto es relevante al momento de realizar la expansión, debido a que se debe efectuar con la misma desagregación, así se obtendrá una mejor estimación de las ventas totales, ya que las empresas de distintos tamaños tendrán su propio tratamiento.

La expresión algebraica del factor de expansión de los establecimientos pertenecientes al Tramo Aleatorio (TA), tiene la siguiente forma:

$$FE(TA)_{h,m} = \frac{\sum_{i=1}^K y(TA)_{h,m,i}}{\sum_{i=1}^K y(TAe)_{h,m,i}}$$

Donde:

- $FE(TA)_{h,m}$: Factor de expansión del tramo aleatorio del estrato h , de tamaño n .
- $\sum_{i=1}^K y(TA)_{h,m,i}$: Sumatoria de las ventas del marco muestral correspondiente a las empresas del TA del estrato h y tamaño m .
- $\sum_{i=1}^K y(TAe)_{h,m,i}$: Sumatorias de las ventas del marco muestral que corresponden a las empresas efectivamente logradas en el TA del estrato h y tamaño n .

Nivel de Estimación

Los niveles de estimación para los sectores comercio y servicios, están de acuerdo a las divisiones CIIU Rev.3 (tres dígitos), con representación nacional.

Estimadores

La estimación estadística corresponde al procedimiento mediante el cual se calculan parámetros de la población a partir de datos muestrales. De este modo, se realizan estimaciones para totales de la población a partir de la información obtenida en la muestra.

La estimación de totales a nivel de CIIU Rev. 3, corresponde al total del estrato expandido, para lo cual se consideran con factor de expansión todas aquellas empresas que fueron incluidas forzosamente en la muestra, más el factor de expansión de los estratos del tramo aleatorio. Así se obtiene el siguiente algoritmo:

$$\hat{Y}_h = FE(IF)_h * \sum_{i=1}^K Y(YF)_{h,i} + FE(TA)_h * \sum_{i=1}^L Y(TA)_{h,i}$$

Donde:

- \hat{Y}_h : Estimación del total del estrato h .
- $FE(IF)_h$: Factor de expansión de inclusión forzosa del estrato h .
- $\sum_{i=1}^K Y(YF)_{h,i}$: Sumatoria de las ventas de las empresas del estrato h , del tramo inclusión forzosa.
- $FE(TA)_h$: Factor de expansión de inclusión aleatoria del estrato h .
- $\sum_{i=1}^L Y(TA)_{h,i}$: Sumatoria de las ventas de las empresas del estrato h , del tramo aleatorio.

La estimación de totales a nivel de grandes divisiones se obtiene sumando las estimaciones de total descrita anteriormente, correspondiente a cada uno de los estratos pertenecientes a una división, Comercio al por mayor, Comercio al por menor, Servicios, etc., obteniendo el siguiente algoritmo:

$$\hat{Y}_{gd} = \sum_{h=1}^W [FE(IF)_h * \sum_{i=1}^K Y(YF)_{h,i} + FE(TA)_h * \sum_{i=1}^L Y(TA)_{h,i}]$$

Donde:

- \hat{Y}_{gd} : Estimación del total para grandes divisiones.
- $FE(IF)_h$: Factor de expansión de inclusión forzosa del estrato h .
- $\sum_{i=1}^K Y(YF)_{h,i}$: Sumatoria de las ventas de las empresas del estrato h , del tramo inclusión forzosa.
- $FE(TA)_h$: Factor de expansión de inclusión aleatoria del estrato h .
- $\sum_{i=1}^L Y(TA)_{h,i}$: Sumatoria de las ventas de las empresas del estrato h , del tramo aleatorio.

c. Diseño Cuestionario de entrevista.

Los cuestionarios que forman parte de la Encuesta Estructural de Comercio y Servicios, viven un proceso de correcciones y mejoras cada año, de acuerdo con el conocimiento empírico adquirido por los profesionales del departamento con el objetivo de mejorar la calidad de las respuestas. Además, se consideran las necesidades de los principales usuarios para aumentar la utilidad de la información.

Los cambios de forma en el diseño de los formularios son trabajados de manera conjunta con el Departamento de Ediciones del INE.

d. Desarrollo y puesta a prueba de los instrumentos de medida (prueba piloto).

Cuando los cambios realizados en el formulario se encuentran en el contexto de un Cambio de Año Base, los formularios pasan por una prueba piloto donde se le pide a las empresas más importantes de la muestra de IF que respondan el nuevo formulario y otorguen sus apreciaciones. Además se realizan comparaciones entre las respuestas obtenidas con el nuevo y el antiguo formulario.

4. RECOGIDA DE DATOS

a. Fuentes de datos.

Para formular el marco muestral y obtener información de contacto con las empresas se utiliza un Directorio entregado por el Servicio de Impuestos Internos.

b. Distribución del trabajo y capacitación de los encuestadores.

El jefe de producto asigna un número de encuestas de responsabilidad de cada analista, considerando su experiencia y habilidades, sin olvidar que la responsabilidad de la entrega final es de todo el departamento. Los analistas son capacitados constantemente por aquellos que poseen más experiencia, por los supervisores y jefes de productos. Además, si es necesario se les capacita de forma formal frente a cambios profundos que afecten a la encuesta.

c. Reducción de la falta de respuesta.

Para aumentar la tasa de respuesta que obtiene el INE frente a todas sus encuestas, existe una legislación que obliga a los informantes seleccionados a responder los formularios.

Para prevenir faltas de respuestas debido a errores en el directorio o a cambios o cese de actividades, se considera un 15% de unidades muestrales adicionales.

Si aún así existen empresas de las cuales no se puede obtener información, afectando la representatividad de la muestra, éstas son reemplazadas por otras de similares características.

d. Trabajo de campo y requerimientos cartográficos.

El envío de los formularios se hace a través de Correos de Chile y el contacto con cada informante se realiza vía telefónica, para verificar la recepción de la encuesta y entregar los lineamientos generales. Si la encuesta no fue recepcionada por el informante se realiza un nuevo envío de formularios con las direcciones ya verificadas.

e. Controles de supervisión y evaluación del levantamiento.

Para controlar el destino de las encuestas que se enviaron vía Correos de Chile, dicha organización envía un informe donde reporta qué encuestas fueron exitosamente entregadas en la dirección de destino y cuáles no. Con esta información, se puede comenzar a determinar que empresas no podrán ser ubicadas y por ende deberán ser reemplazadas en la muestra.

Además, cada semana se publica en el diario mural del departamento un informe con el número de encuestas que se encuentran en cada etapa del levantamiento, calculando su variación. Así, los analistas son guiados para enfocarse en una acción en particular, por ejemplo, en capturar encuestas, en validarlas, en terminarlas, etc.

Conjunto a esto, se publican informes semanales del rendimiento de cada analista, considerando el número de encuestas terminadas por cada uno de ellos en comparación a las terminadas por el equipo en su conjunto.

5. CAPTURA Y TRATAMIENTO DE LOS DATOS

a. Captura de datos.

Una vez que los formularios son enviados por Correos de Chile y recepcionados por el informante, éste puede entregar los datos requeridos a través de varias opciones: por correo, en forma personal a las dependencias del INE, por correo electrónico o por web, donde existe un sitio especializado para completar la información.

b. Recepción y acopio de cuestionarios.

Los datos entregados por las empresas deben guardarse en el sistema URCE, que es el software de recopilación y análisis de datos perteneciente al INE. Cuando la encuesta es respondida por la web el traspaso de datos es automático, mientras que cuando la información es enviada por otro medio es el analista quién debe incluirla en el sistema.

c. Reglas de validación e imputación.

La encuesta posee varias etapas de validación y depuración. Las primeras validaciones se le hacen al informante cuando éste digita los datos por la web, o cuando el analista digita los datos directamente en la URCE dependiendo de cómo haya sido enviada la información. En esta etapa se validan aspectos de lógica, como sumatorias o campos que deben ser contestados por obligación.

La segunda validación se realiza una vez que los datos ya han sido guardados en el sistema, en esta etapa se verifican aspectos más complejos relacionados con la coherencia de la información y sus incidencias contables. En esta misma etapa se realizan validaciones intertemporales que tienen por objetivo identificar datos que varían excesivamente respecto de la información entregada el año anterior.

Después de realizadas estas validaciones, las encuestas son terminadas y sus datos pasan a formar parte de la base de datos que es depurada por el jefe de producto.

d. Procedimientos de depuración.

La depuración se realiza en forma conjunta a las validaciones hechas por cada analista o por el jefe de producto, el que a su vez puede requerir que alguna encuesta sea nuevamente analizada.

e. Elaboración de manuales.

Actualmente existen dos tipos de manuales que intentan mejorar la calidad de la información capturada. El primero es el "Manual de Procedimientos" que está enfocado a ayudar al analista, en él se explican las etapas de trabajo que se realizan durante el año relacionadas con la encuesta y los acuerdos. Además, existen manuales para cada uno de los formularios enfocados a responder las posibles inquietudes que tenga el informante al momento de completar la encuesta.

6. ANÁLISIS DE LOS DATOS Y CALIDAD DE LOS RESULTADOS

a. Calidad.

La calidad de los datos obtenidos depende de la calidad de los procesos y de la calidad del personal encargado. Como se ha nombrado anteriormente, el personal encargado mantiene un alto nivel de profesionalización, de experiencia y capacitación, los procesos por su parte son sujetos constantes de análisis y mejoras (mejoras relacionadas con los recursos y tiempos de los procesos, el diseño de la encuesta, las tasas de respuesta, el tiempo en la entrega de información, etc.)

b. Acuracidad.

Como se detalla en el punto 3, la muestra se obtiene considerando un nivel de confianza del 95%, un porcentaje de error del 5% y una varianza máxima. Los coeficientes de variación para cada estrato se pueden observar en el cuadro 3.

Además, el departamento está conciente que el procedimiento de reemplazo de encuestas podría ser una fuente de errores no muestrales. Cuando existen empresas dentro de la muestra inicial que no se podrán capturar se buscan empresas de similares características para que las reemplacen, la idea es mantener el tamaño de la muestra y su representatividad. La similitud es importante ya que se intentan mantener los mismos coeficientes de variación y se ocuparán los mismos factores de expansión. Sin embargo, el hecho de que las empresas no sean exactamente iguales implica que exista un error no muestral involucrado.

c. Oportunidad y puntualidad.

El resultado de toda la información obtenida se publica sólo una vez al año, cada 30 de Mayo. El año 2009 se publicaron las estimaciones correspondientes al año contable 2007. No existen publicaciones anticipadas ni con datos preliminares. No existieron retrasos en la publicación.

d. Accesibilidad.

El resumen de la información recopilada se publica anualmente en un documento denominado "Comercio y Servicios. Informe Anual". Este documento incluye un análisis general y comparativo, un enfoque de género, un análisis financiero, uno intertemporal, un resumen metodológico y los cuadros estadísticos que incluyen los datos agregados de cada variable para cada estrato. El documento se encuentra a disposición del público general como publicación impresa y en la página web del INE (<http://www.ine.cl>), en la sección *Estadísticas Económicas*, en el link *Comercio y Servicios*, bajo el nombre de Anuario de Comercio y Servicios 2007.

e. Transparencia.

La publicación "Comercio y Servicios. Informe Anual" incluye un resumen metodológico del diseño muestral de la encuesta, además de los formularios que reciben las empresas y las fórmulas de cálculo y formato de las tabulaciones.

f. Comparabilidad.

Las estimaciones obtenidas de los datos capturados son continuamente comparadas con las variables de la misma empresa para años anteriores, con el promedio de la variable para el resto de la muestra y finalmente con los datos internacionales, todo en pos de conocer los órdenes de magnitud correspondientes y evitar datos que se alejen excesivamente de ellas. Algunas de estas comparaciones son publicadas en el anuario, otras sirven para imputar los datos faltantes.

g. Coherencia.

Las estimaciones publicadas son producto de los datos obtenidos a través de las encuestas hechas a las empresas, sin incluir otras fuentes adicionales. Los conceptos utilizados en los formularios y en las tabulaciones intentan ser coherentes con otras fuentes utilizadas en el cálculo de las cuentas nacionales, para evitar inconsistencias en los datos agregados, sin embargo, esta labor corresponde al BCCh. El INE por su parte se responsabiliza de la coherencia entre las encuestas estructurales y coyunturales. Por la misma razón, se publican las fórmulas de cálculo de todas las estimaciones y todas las preguntas, para identificar las posibles diferencias.

h. Exhaustividad.

Las estimaciones de la encuesta tienen representatividad a nivel nacional, por cada estrato antes normado según su nivel de ventas. Algunos usuarios tienen la necesidad de obtener dichas estimaciones con representación regional, lo que es muy difícil en el sector encuestado, para ello se necesitaría que la unidad muestral fuera el establecimiento en vez de la empresa. Dado que el sector de Comercio, Servicios y Turismo es muy volátil, es muy difícil conocer el domicilio

de los establecimientos que utiliza una empresa, los cuáles se encuentran en constante cambio.

A nivel internacional, se requiere que las estimaciones sean representativas por tramo del nivel de ventas así como por tramo del número de trabajadores. Actualmente se contempla analizar los cambios que implicarían trabajar con esos niveles de representatividad.

i. Análisis de los datos.

Los datos son analizados como se describe en los puntos anteriores, el objetivo es obtener una base de datos de gran calidad. Sin embargo, el análisis que se realiza de la información es solamente descriptiva, las estimaciones publicadas se limitan a entregar la agregación de las variables encontradas en cada sector, visualizadas a través de cálculos estadísticos básicos, nada inferencial.

j. Control de la desagregación.

La representatividad de la muestra está dada por la estratificación hecha por actividad al nivel de 3 dígitos de la CIIU rev. 3. Los coeficientes de variación presentes en el cuadro 3 dan cuenta de dicha representatividad. Durante el proceso de recopilación de datos, el jefe de producto observa cómo dichos coeficientes van cambiando con el fin de responder a los niveles de desagregación deseados.

7. DOCUMENTACIÓN Y DIFUSIÓN

a. Documentación de los metadatos.

El metadato de la encuesta se encuentra publicado en el Sitio Web Plan de Recopilación Estadística (<http://lingue.ine.cl/metadatos/>)

b. Estrategias de difusión.

El departamento de difusión del INE es el encargado de difundir nuestros productos. El departamento a su vez envía anualmente un tríptico a las empresas que participaron de la encuesta. Dicho tríptico contiene información general de los datos agregados obtenidos durante el periodo capturado.

c. Gestión o almacenamiento de los datos.

Los datos de cada periodo se encuentran almacenados en los servidores respectivos, que se encuentran en las dependencias del INE. Adicionalmente, las encuestas recibidas en formato físico también son almacenadas en dichas dependencias.

8. CICLO DE MEJORA

a. Adaptabilidad / Flexibilidad.

La muestra de empresas contiene un porcentaje de unidades IF (inclusión forzosa) y otro porcentaje variable. Este último tiene el objetivo de representar los cambios de mediano plazo que se producen en el sector, ya que obtiene información de empresas más pequeñas y por ende más inestables, a quienes los shocks afectan con mayor fuerza y producen cambios más drásticos, como

por ejemplo el cierre o la apertura de nuevos negocios. De esos datos se obtienen pautas que pueden influir en los conceptos o preguntas que se desarrollarán en la encuesta del año posterior. En general, las preguntas del formulario y los conceptos detrás de ellas son evaluadas anualmente y se mantienen por el periodo que dure la recolección de información. Sin embargo, los procedimientos que permiten dicha recolección y la posterior depuración de los datos se encuentran en constante adaptación.

b. Pericia en las áreas relevantes.

El departamento de Comercio y Servicios mantiene una fuerte preocupación por la calidad de los trabajadores que pertenecen a él y los procesos asociados a la encuesta. Todo el personal (incluyendo a analistas, jefes de producto, personal administrativo, etc.) es profesional de áreas relacionadas y con experiencias y habilidades acorde a sus labores, que deben pasar por el proceso de selección requerido por el INE para obtener los cargos asociados.

Conjunto a esto, los analistas del departamento trabajan bajo un esquema de especialización del trabajo, donde existen equipos destinados a analizar las encuestas correspondientes a una rama de actividad económica específica.