

DEPARTAMENTO DE ASUNTOS ECONÓMICOS Y SOCIALES
OFICINA DE ESTADÍSTICA

ESTUDIOS DE METODOS

Serie **F** No. **19**

**ORGANIZACION Y REALIZACION
DE ENCUESTAS
SOBRE EL COMERCIO
DE DISTRIBUCION**



NACIONES UNIDAS
Nueva York, 1978

NOTA

Las firmas de los documentos de las Naciones Unidas se componen de letras mayúsculas y cifras. La mención de una de tales firmas indica que se hace referencia a un documento de las Naciones Unidas.

ST/ESA/STAT/SER.F/19

PUBLICACION DE LAS NACIONES UNIDAS

No. de venta: S.77.XVII.3

Precio: \$ 10,00 (EE. UU.)
(o su equivalente en la moneda del país)

INDICE

	<u>Página</u>
INTRODUCCION	
Observaciones generales	1
Organización del manual	2
 <u>Capítulo</u>	
I. PASOS PRELIMINARES Y CONSIDERACIONES PRINCIPALES	
Estudio de los programas existentes	3
Estudio de las recomendaciones internacionales	4
Consideraciones principales	4
 II. NECESIDADES Y USOS DE LAS ESTADISTICAS DEL COMERCIO DE DISTRIBUCION	
Observaciones generales	8
Necesidades y usos de la administración pública	8
Necesidades y usos de las empresas	9
Otros usos y necesidades	10
 III. PLANIFICACION Y ORGANIZACION GENERALES	
Preparación de un resumen para la aprobación del programa	11
Requisitos legislativos	11
Determinación de la estructura orgánica	11
Reclutamiento del personal de planificación	12
Organización de grupos de asesoramiento	13
Decisiones fundamentales	14
Método para reunir datos	15
Método para elaborar datos	15
Necesidades de equipo y personal	15
Asignación de responsabilidades	16
Coordinación con otras encuestas	16
Planificación de un ensayo previo	17
Planificación de la publicación	18
Diferencias entre planificación de programas censales y planificación de programas de encuesta	19
 IV. DIRECCION Y PRESUPUESTO	
Uso de manuales	21
Uso de organigramas	21
Calendarios	22
Programación de personal y equipo	23
Sistema de información sobre la marcha de los trabajos	23
Establecimiento de un sistema de control de calidad	24
Maximización del uso de los recursos	25
Diferencias entre censos y encuestas corrientes	25

V.	CAMPO DE APLICACION, COBERTURA Y CLASIFICACION	
	Introducción	28
	Campo de aplicación	28
	Cobertura	30
	Clasificación	33
VI.	CUESTIONARIOS Y OTROS FORMULARIOS PARA LA RECOGIDA DE DATOS	
	Información necesaria para control, correspondencia y evaluación	34
	Cuestiones sobre publicación y preparación	34
	Distinción entre anotaciones de no información y anotaciones nulas	35
	Diseño de la pregunta sobre tipo de actividad	36
	Localización de información consolidada	36
	Preparación del formulario	37
	Importancia de la codificación previa	37
	Tratamiento de las instrucciones	37
	Problemas derivados de la diversidad de actividades	38
	Diseño de los formularios	39
	Secuencia y redacción de las preguntas	40
VII.	RECOGIDA DE DATOS	
	Observaciones generales	42
	Métodos de enumeración	43
	Enumeración sobre el terreno	44
	Establecimiento y uso de un directorio	47
	Recogida y control de la información	52
	Uso del muestreo	55
VIII.	ELABORACION DE DATOS	
	Observaciones generales	59
	Operaciones administrativas	60
	Preparación del cuestionario para su tabulación	62
	Revisión y codificación con computadora	64
	Datos de registros administrativos	67
	Preparación de datos en encuestas repetitivas por muestreo . . .	68
IX.	TABULACION, REVISION Y CORRECCION	
	Observaciones generales	69
	Máquinas de sumar	69
	Elaboración electromecánica	70
	Elaboración con computadora	71
	Revisión y corrección de las tabulaciones	72

Capítulo

Página

X. PUBLICACION

Examen final	77
Análisis último sobre publicación indebida de información	77
Preparación y reproducción de ejemplares	78
Publicación de informes preliminares	78
Descripciones de la encuesta para su publicación	79
Contenido de los cuadros	80
Encuestas de frecuencia más que anual	84

ANEXOS

I. TECNICAS DE MUESTREO EN ENCUESTAS SOBRE EL COMERCIO DE DISTRIBUCION Y LOS SERVICIOS	85
II. BIBLIOGRAFIA	112

INTRODUCCION

Observaciones generales

1. La Comisión de Estadística de las Naciones Unidas en su 18^o período de sesiones de 1974, entre otras cosas, aprobó unas recomendaciones sobre estadísticas del comercio de la distribución y los servicios, suscribió la propuesta del Secretario General y recomendó que "se comenzase a preparar un manual sobre la organización y realización de encuestas, que sería publicado como suplemento metodológico de las recomendaciones internacionales" 1/.
2. Esta publicación se ha preparado para atender a esa recomendación y su finalidad es servir de guía práctica en la planificación y dirección de censos y encuestas sobre el comercio de distribución y los servicios, con especial atención a los problemas que afrontan los países que se encuentran en las primeras etapas del desarrollo estadístico.
3. El marco teórico de las encuestas cuya ejecución en sus aspectos prácticos se examina en este manual figura en sus líneas generales en las recomendaciones internacionales aprobadas por la Comisión de Estadística 2/.
4. En vista de las semejanzas entre las encuestas industriales y las de comercio de distribución, en este manual se analizan algunos métodos y procedimientos que también son aplicables a encuestas industriales. No obstante, como en la actualidad casi todos los países hacen las encuestas industriales y las del comercio de distribución por separado, se ha pensado que unos manuales distintos, al brindar orientaciones más especializadas, serían de uso más fácil por aquellos países que piensan realizar tales encuestas.
5. Para los fines del presente manual, un censo se concibe como una encuesta completa realizada a intervalos poco frecuentes, de ordinario cada cinco o diez años. Las encuestas corrientes son las que se hacen a intervalos frecuentes -mensual, trimestral o anualmente. Censos y encuestas comprenden las actividades del comercio al por mayor y por menor, restaurantes y hoteles y algunos servicios que se definen en la Clasificación industrial internacional uniforme de todas las actividades económicas (CIU) 3/, concretamente en su gran división 6 y en las

1/ Documentos oficiales del Consejo Económico y Social, 58^o período de sesiones, Suplemento N^o 2 (E/5603), párr. 97.

2/ Recomendaciones internacionales sobre estadísticas del comercio de distribución y los servicios, Informes Estadísticos, serie M, N^o 57 (publicación de las Naciones Unidas, N^o de venta S.75.XVII.9).

3/ Clasificación industrial internacional uniforme de todas las actividades económicas, Informes Estadísticos, serie M, N^o 4, Rev. 2 (publicación de las Naciones Unidas, N^o de venta S.58.XVII.8).

subdivisiones 83, 94 y 95. Las unidades sobre las que se recogen estadísticas son establecimientos (almacenes, tiendas, oficinas u otros locales caracterizados) dentro de los límites territoriales del país. La empresa no constituye la unidad estadística básica, aunque en algunos casos puede ser la unidad de notificación.

Organización del manual

6. En este manual se examinan los pasos que deberían darse para planeary realizar una encuesta del comercio de distribución. Su contenido se ha ordenado, siempre que ha sido posible, con arreglo a la secuencia en la que las operaciones de la encuesta deberían planificarse. En muchos aspectos, los métodos que se usan para los censos poco frecuentes y para las encuestas corrientes son muy similares. No obstante, cuando se requiere un tratamiento distinto, se indica.
7. El capítulo I describe los pasos preliminares y las consideraciones principales de la organización de encuestas, el capítulo II indica las necesidades y usos de los datos del comercio de distribución, los capítulos III y IV se dedican a los aspectos generales de planeamiento y organización, dirección y presupuesto. En los capítulos V y VI se examinan puntos relacionados con el alcance, ámbito, clasificación y diseño de formularios y cuestionarios. El capítulo VII trata de los métodos y problemas de la recopilación de datos. En los capítulos VIII y IX se describen las medidas y problemas asociados con la elaboración, tabulación, examen y corrección de datos. El manual termina con el capítulo X donde se analiza la publicación de los resultados. En el anexo I se habla en detalle de las técnicas concretas de muestreo en las encuestas sobre la distribución.
8. El manual debe considerarse sólo como un conjunto de sugerencias prácticas acerca de los puntos más importantes en la organización y ejecución de encuestas del comercio de distribución. En el reducido número de páginas de este documento no ha sido posible abordar todos los temas con el debido detalle y las recomendaciones que se hacen constituyen sólo una guía sobre ciertos aspectos de la ejecución de encuestas. Todo país que se proponga realizar tales encuestas tendrá que estudiar a fondo estas recomendaciones para adaptarlas a sus condiciones particulares.
9. Aunque las recomendaciones internacionales buscan reflejar prácticas comunes a la mayoría de los países, quienes se interesen en un análisis más detallado de ciertos temas o en la experiencia concreta de determinados países deberían ponerse en comunicación con éstos para conocer los problemas que puedan haber encontrado y beneficiarse de sus soluciones. Cabe obtener asesoramientos sobre problemas particulares escribiendo al Director de la Oficina de Estadística, Naciones Unidas, Nueva York, N.Y. 10017.

I. PASOS PRELIMINARES Y CONSIDERACIONES PRINCIPALES

Estudio de los programas existentes

10. Si bien las recomendaciones internacionales han sugerido el tipo de datos que conviene reunir en materia de distribución y servicios, es menester que cada país examine por su cuenta los datos que se hayan ya recopilado y los objetivos que deberían fijarse. Esto requiere que se estudien los informes administrativos que realizan los diversos organismos gubernamentales y proceder a una detenida investigación de las necesidades de datos del sector privado y del público.
11. Un examen de las actividades de la administración pública puede mostrar que ya se dispone de abundantes datos derivados del funcionamiento de los servicios impositivos, reglamentación de las organizaciones financieras, aplicación de las disposiciones laborales o actividades de la seguridad social. Tales datos deben ser examinados para saber si cabe usarlos en programas estadísticos, al margen del cometido reglamentario o administrativo para los que se reunieron originariamente.
12. Para determinar las necesidades de datos, será necesario ponerse en contacto con dos clases principales de usuarios, a saber, los públicos y los privados. Conviene señalar que las dependencias de la administración que toman decisiones y establecen normas frecuentemente actúan sin una adecuada información estadística. Si se les brindase la oportunidad de exponer sus necesidades, sus peticiones de datos podrían ser abrumadoras. El problema consistiría entonces en limitar el programa a dimensiones viables.
13. Una manera de estudiar las necesidades de datos de los usuarios privados sería ver los programas de datos que las organizaciones privadas realizan. De ordinario se encontrará que tales programas son poco fidedignos desde el punto de vista estadístico por causa de su incapacidad para obtener índices de respuesta elevados y representativos, aunque se lleven a cabo por no existir otro procedimiento. El hecho de que la organización, pese a todo, prefiere hacer el esfuerzo es una buena indicación de la utilidad de los datos.
14. Al examinar las necesidades de datos, hay que tener presente que, sin perjuicio de que estén muy integrados, son dos los grandes tipos de programas que deben tomarse en cuenta. El primero consiste en un censo general que intente desarrollar las estadísticas con bastante detalle como base para comprender la estructura de todo el sector. Un censo de esa índole se suele ejecutar a intervalos infrecuentes, tal vez cada cinco o diez años, y requieren un largo período de tiempo para su planeamiento y elaboración. El segundo tipo de programa se realiza a intervalos más frecuentes -cada mes, trimestre o año-, es mucho más restringido en cuanto a los puntos que comprende y puede hacerse recurriendo al muestreo. Si bien las muestras que se usan en las encuestas corrientes no tienen necesariamente que seleccionarse de los recuentos completos de los censos, serán más fiables y útiles si se procede de ese modo. Tal cosa permite vincular los dos conjuntos de datos y sirve para comprobar cómo funcionan las muestras entre los períodos censales.

Estudio de las recomendaciones internacionales

15. Como se ha dicho antes, las recomendaciones internacionales aprobadas por la Comisión de Estadística y publicadas después por las Naciones Unidas 4/ sirven para ayudar a los países a que desarrollen sus estadísticas de la distribución. En la publicación citada figuran análisis sobre los temas siguientes: campo de aplicación, cobertura y unidades usadas; características de la unidad estadística y su clasificación; clases de datos que han de reunirse y estadísticas que han de publicarse; y las definiciones de las diversas clases de datos. En el programa que se indica figuran recomendaciones sobre encuestas infrecuentes (censos), encuestas anuales y encuestas de frecuencia más que anual. Para cada categoría, se hacen recomendaciones sobre qué tipos de datos deberían incluirse. También se examinan los grandes conjuntos de problemas, especialmente las definiciones de unidades estadísticas y de notificación y las definiciones de las características de los establecimientos, tales como índole de actividad, tipo de operación, clase de organización económica, clase de organización jurídica y tipo de propiedad, ubicación y tamaño. No se aspira a que tales recomendaciones se encuentren aceptables en todos sus extremos por todos los países. Habrán de introducirse las modificaciones oportunas según las circunstancias especiales de cada país. No obstante, debería intentarse que el programa nacional encajara dentro del marco de las recomendaciones internacionales para que pudieran hacerse comparaciones coherentes entre unos países y otros y para que las estadísticas mundiales pudieran desarrollarse mejor sobre la base de conceptos uniformes. Conforme las economías del mundo se vuelven más integradas e interdependientes, la normalización puede conducir a una mejor comprensión de las relaciones recíprocas de las fuerzas económicas y a una solución más inteligente de los problemas económicos.

Consideraciones principales

16. A causa de la gran complejidad tanto de los censos infrecuentes como de las encuestas corrientes, es importante que todos los aspectos del trabajo que se vaya a hacer se comprendan y se planifiquen. Es menester considerar cuidadosamente cada operación y los efectos recíprocos para estar seguros de que el producto final resultará aceptable. Esto es especialmente cierto para los censos infrecuentes en los que lleva de cinco a diez años para hacer las modificaciones necesarias en las especificaciones. Para las encuestas más frecuentes, las correcciones cabe hacerlas más a menudo y las consecuencias de un error no serán tan desastrosas. Preparar un programa bien meditado y adaptado específicamente a las necesidades del país de que se trate es importante. Los sistemas estadísticos por bien que funcionen en un país, pueden resultar nada apropiados en otro.

4/ Recomendaciones internacionales sobre estadísticas del comercio de distribución y los servicios, Informes Estadísticos, serie M, N° 57 (publicación de las Naciones Unidas, N° de venta S.75.XVII.9).

17. Un presupuesto adecuado y realista es la consecuencia de una detenida planificación previa y de un buen entendimiento de la asignación de los costos. Muchas veces, la persona responsable de los aspectos técnicos del programa se ocupa también del presupuesto, con el resultado de que los costos se subestiman. Es importante que un analista de presupuestos competente examine todas las operaciones para comprobar que el presupuesto es adecuado y suficiente. Para casos de emergencia deben figurar reservas como partidas de imprevistos. La operación tiene que analizarse de modo detenido para evitar que se omita algo importante. Por causa de una planificación y presupuestación deficientes, algunos países han visto que sus fondos se gastaban sobre todo en la etapa de recogida de datos y no quedaban recursos para completar la elaboración de los cuestionarios y la publicación de los resultados.

18. Una deficiencia muy común, especialmente en los censos infrecuentes, consiste en recargar el programa y complicar el cuestionario hasta el punto de que todo el sistema se viene abajo por razón de su misma complejidad. Si bien es importante incluir información vital, tal información puede perderse bien porque la persona que responde se vea frenada por el gran volumen de datos pedidos o no tenga a mano todos ellos, o bien porque resulte imposible elaborar los datos a causa de su complejidad. Es muy importante que al iniciar un programa nuevo se mantenga el contenido del cuestionario lo más sencillo posible para asegurar una realización positiva del censo.

19. Para hacer una encuesta con éxito, es esencial obtener la cooperación de quienes han de contestar. Un modo de obtener tal cooperación es mediante obligaciones legales, y la fijación de sanciones para quien no responda es probablemente algo necesario en el caso de los censos generales. En las encuestas frecuentes basadas en muestras, tal cosa no está tan justificada, ya que en esas encuestas se le pide a la persona que figura en la muestra que asuma una carga que no se aplica con carácter universal. En todo caso, un colaborador dispuesto resulta más apto para proporcionar una información correcta y a su debido tiempo que alguien recalcitrante. Por tanto, corresponde al director de la encuesta trabajar con organizaciones empresariales y representantes de grandes firmas con el objeto de que se cree un clima propicio a la realización de la encuesta. A veces esto obliga a incluir datos de algún tipo especial que desee un sector industrial concreto o en ocasiones a excluir otros que no se vean con buenos ojos. Significa también que los datos pedidos han de obtenerse con facilidad de los registros contables de quienes responden. Un tipo de queja bastante frecuente que hacen las empresas es que los datos solicitados ya fueron suministrados a otros servicios del gobierno. Aunque tal cosa sea cierta, muchas veces es necesario pedir de nuevo la misma información por causa del calendario o por la complejidad que reviste transferir los datos desde los otros servicios.

20. Como ya se mencionó antes, cabe usar algunos servicios de la administración para lo fines de la encuesta. En primer lugar, muchas veces hay listas de direcciones en los registros de esos servicios o pueden obtenerse de ellos. Esas listas contienen frecuentemente direcciones completas además de una clasificación de industrias y tal vez algún tipo de datos como volumen de ventas. Tal información puede ser muy útil al preparar una guía para su uso en el censo o para

confeccionar las muestras de las encuestas corrientes. En segundo término, si un servicio de la administración está ya reuniendo datos, tal vez sea posible pedirle que modifique sus peticiones para ajustarlas a las necesidades de la encuesta.

21. El empleo de los registros de otros servicios frecuentemente requerirá negociaciones habilidosas, ya que tales servicios no se ocupan primordialmente de estadística. También necesitará una cuidadosa planificación para estar seguros de que se dispondrá de esos registros a tiempo y en la forma adecuada para su uso en la encuesta. Puede surgir un problema ya que los datos de un servicio de la administración pueden haberse reunido con la promesa de mantenerlos confidenciales. Como las encuestas estarán también reuniendo datos en virtud de una ley que asegura su carácter confidencial, el uso de datos de servicios de la administración puede justificarse en tanto en cuanto el personal de los servicios de estadística jure respetar ese carácter confidencial y las publicaciones no contengan datos que revelen operaciones de ninguna empresa. Lo contrario, sin embargo, no es cierto; esto es, la información sobre empresas individuales contenida en la encuesta estadística no puede transferirse a otro servicio de la administración, ya que tal información se obtuvo con la promesa de que no podría usarse para fines fiscales o administrativos.

22. Varios de los tipos de datos de las encuestas sobre el comercio de distribución y los servicios son similares a los que figuran en otras encuestas que realiza el organismo estadístico u otros servicios de la administración. Es importante lograr en todo lo posible que los mismos datos se definan del mismo modo, a fin de que los que procedan de encuestas diferentes puedan compararse o agregarse. Las modificaciones que requieran las condiciones especiales de un sector determinado deberían prepararse de tal modo que los datos resultantes sean comparables. Esto es especialmente importante con respecto a los mismos datos que se reúnen en un censo general y en una encuesta corriente. Por ejemplo, a causa de la necesidad de reunir y publicar rápidamente los datos de una encuesta mensual, la persona que responda puede tener alguna dificultad para ajustar su información sobre ventas a extremos tales como impuesto sobre el valor agregado, entregas a cuenta, bonificaciones, etc. Tal cosa podría conducir a la no comparabilidad de datos entre los censos y las encuestas corrientes.

23. En casi todos los países, existe ya de ordinario una organización para reunir datos. Habrá que decidir hasta qué punto las diferentes encuestas pueden o deben integrarse, ya que suelen tener bastantes elementos comunes. Si bien cabe integrar muchas de las operaciones de elaboración, las funciones de planificación, especialmente cuando se relacionan con el contenido de los datos y el conocimiento de la industria, tendrán que establecerse en una organización separada. Incluso en el caso de las operaciones de elaboración, hay que procurar no perder de vista el hecho de que el conocimiento de industrias específicas es importante para evaluar el carácter de los problemas y cómo deberían resolverse.

24. Muchas encuestas no se han visto coronadas por el éxito a pesar de los muchos esfuerzos desplegados antes de su realización, porque la prueba final sólo puede venir con la ejecución efectiva de la encuesta misma. En todos los casos, la actividad mínima anterior a tal ejecución debería incluir visitas a las empresas para examinar el contenido del cuestionario. No obstante, sería mucho mejor si

podiera hacerse un ensayo previo a escala completa para comprobar todos los aspectos de la encuesta desde la planificación hasta la publicación. Esto se examina con más detalle en el párrafo 65.

25. Las recomendaciones internacionales publicadas por las Naciones Unidas 5/ contienen un breve examen de la necesidad de estadísticas del comercio de distribución. En el capítulo siguiente se intenta describir cuáles son los usuarios más importantes de la información que se busca con las encuestas. Se espera que conocer quiénes son los usuarios y también los fines para los que se precisa la información ayudará a los planificadores de una encuesta no sólo a seleccionar los temas que hay que abarcar y a pronunciarse sobre su importancia sino también a preparar la publicidad y a responder a las preguntas sobre la utilidad del programa.

5/ Recomendaciones internacionales sobre estadísticas del comercio de distribución y los servicios, Informes Estadísticos, serie M, N° 57 (publicación de las Naciones Unidas, N° de venta S.75.XVII.9).

II. NECESIDADES Y USOS DE LAS ESTADÍSTICAS DEL COMERCIO DE DISTRIBUCION

Observaciones generales

26. Hay una tendencia a que la recopilación de estadísticas del comercio de distribución y los servicios corresponda a una etapa del desarrollo económico muy posterior al de las estadísticas de agricultura, industria y sector extractivo. Esto es así porque las actividades primarias del hombre tratan inicialmente de atender las necesidades básicas, esto es, alimentación, abrigo y ropa. Sólo cuando la sociedad ha sido capaz de acumular excedentes el comercio adquiere importancia. La especialización que consiste en la aparición del sector de los servicios es un hecho aun posterior. Igual que se encontró útil reunir información sobre agricultura e industria como ayuda para comprender y mejorar esas funciones, adquiere utilidad recoger datos sobre el comercio de distribución y los servicios cuando estas actividades adquieren importancia económica.

Necesidades y usos de la administración pública

27. Existen muchos modos por los que los diferentes departamentos de la administración pueden sacar provecho de los datos sobre distribución y los servicios. Uno de los más importantes es como base para evaluar las tendencias de la economía. Para las tendencias a corto plazo, las estimaciones anuales y más frecuentes de ese sector de la economía nacional desempeñan un papel importante. Esto es especialmente importante en el comercio de detalle y en los servicios personales, ya que tales actividades reflejan el consumo final, meta de toda la actividad económica. En economías orientadas hacia el consumidor, los altibajos en el consumo final afectan necesariamente a todas las demás actividades económicas y se convierten así en el más importante indicador de los que cabe considerar por separado- que deben vigilar los servicios del gobierno responsables de la política económica. Cuanto menos unidas estén la producción y el consumo por medio del comercio, menor importancia tendrá ese indicador. Además de medir el consumo final, conviene calcular otros aspectos de la distribución y los servicios, tales como sueldos y salarios del sector, variaciones en existencias y formación bruta de capital fijo.

28. Además de aquilatar las tendencias de la economía, los datos sobre el comercio de distribución y los servicios pueden usarse para establecer normas (y medir los efectos de las ya establecidas) en muchas esferas de la actividad de la administración, por ejemplo, para fijar el nivel de ingreso público que cabe esperar de gravámenes tales como impuestos sobre las ventas, sobre el valor agregado y sobre el trabajo y determinar si se ha alcanzado el nivel previsto. Los datos también son necesarios para el cálculo de las estadísticas del ingreso y producto nacionales. Entre los demás programas que pueden beneficiarse figuran los estudios para la modificación del sistema de transporte, la medición de las necesidades de mano de obra y su aprovechamiento, la planificación de planes de desarrollo regional, la preparación de programas de asistencia social tales como seguridad social, salario mínimo y seguro de desempleo, la evaluación de programas antitrust, etc.

29. A pesar de la utilidad de los datos del comercio de distribución y los servicios, para la administración, hay que tener siempre presente que el suministro de esa información constituye un gasto para la empresa que, a nivel agregado, puede ser muy superior que para el gobierno. Las peticiones de datos complicados en este sector de la economía pueden resultar particularmente onerosas, ya que la empresa más corriente suele ser pequeña y no llevar registros adecuados. Esos factores deberían sopesarse al determinar el contenido de datos del programa.

Necesidades y usos de las empresas

30. Si bien la comunidad mercantil es el proveedor principal de los datos sobre la distribución y los servicios, también es un usuario muy importante de los mismos. Además puede beneficiarse de su uso adecuado, ya que las decisiones que tome el gobierno y que afectan al comercio podrán hacerse con más inteligencia. Los representantes de la comunidad mercantil podrán apoyar u objetar tales decisiones, con mayor fundamento si disponen de datos. Si bien es cierto que sólo las empresas mayores pueden permitirse un personal que analice los datos, las de menores dimensiones podrán influir en la política económica por conducto de las asociaciones que actúan de representantes de la empresa pequeña.

31. Uno de los principales usos de los datos por las empresas es determinar la localización de las unidades de negocios. Claro está que los datos sobre el carácter de la población y la distribución del ingreso tienen que usarse en conexión con aquellos sobre las actividades mercantiles. Los analistas industriales pueden hacerse una buena idea de todo un sector en términos de tamaño de las unidades, proximidad a los centros de población, proporción entre nómina y ventas e inventario y ventas, penetración de mercados, tendencias del sector a lo largo del tiempo, salidas comerciales para productos específicos y varios otros factores muy útiles, según sea la clase de datos reunidos y el tiempo y energía invertidos en el análisis de la información recogida. La información más compleja no puede obtenerse, como es lógico, en las encuestas anuales y de frecuencia superior sino sólo en los censos generales infrecuentes.

32. Además de las empresas de un sector concreto que pueden aprovechar los datos, las firmas que realicen servicios para ese sector también encontrarán útiles tales informaciones. Por ejemplo, los servicios públicos de agua, gas y electricidad, y las compañías de transporte pueden hacer una labor de planificación mucho mejor si cuentan con un conocimiento detallado de la estructura de las muchas unidades que trabajan en los sectores del comercio minorista, y mayorista y de servicios. Igualmente, todos los fabricantes de equipo podrán estudiar el mercado de sus bienes. Quienes fabriquen artículos que se vendan por los canales del comercio al por mayor o por menor y los mayoristas que quieran conocer cómo se venden al por menor sus productos pueden sacar mucho provecho de esos datos. Las instituciones financieras son otro tipo de empresa con gran interés en las estadísticas.

Otros usos y necesidades

33. Un empleo de los datos que no debe soslayarse es el que se relaciona con el campo académico. Una aplicación primordial sería la de escuelas de administración de empresas como instrumento básico en la formación de personal que llegará a dirigir organizaciones empresariales grandes y medianas. Otra aplicación sería el adiestramiento de economistas, analistas de mercado y estadísticos de encuestas. Personas capacitándose como economistas laborales, consultores financieros, o economistas del gobiernos también encontrarán los datos útiles.

34. Las organizaciones sindicales podrán aprovechar los datos para las negociaciones salariales. Pueden relacionar los datos sobre sus afiliados con información sobre el emplazamiento y número de trabajadores en los sectores en que las organizaciones se desenvuelven para determinar el modo efectivo de desplegar sus esfuerzos. Asimismo, el conocimiento que adquieren de la estructura y funcionamiento de la industria puede servir para establecer demandas justificadas y negociar contratos efectivos.

35. Un aprovechamiento fundamental de los resultados de los censos generales infrecuentes es el diseño de muestras para usarse en otras encuestas, especialmente las que se hacen anualmente o con mayor frecuencia. El censo general es el marco ideal para seleccionar la muestra más eficiente, ya que contiene aquellas características de las unidades del muestreo tales como tipo, tamaño y ubicación de la empresa y consiste en un recuento completo que comprende el nombre y dirección de cada unidad. Si bien las organizaciones privadas de muestreo no pueden tener acceso a la información detallada de que dispone el gobierno, les será posible, sin embargo, aprovechar los datos publicados, tanto para el diseño de muestras eficientes como para contrastar los resultados de su labor con los totales de la administración.

36. Por último, el desarrollo de programas estadísticos que contengan información sobre el comercio de distribución y los servicios añadirá otra faceta al acervo general de conocimientos y mejorará el entendimiento público de los hechos de la vida mercantil, lo que permitirá una influencia informada de los particulares en la política pública. Quienes moldean la opinión pública por conducto de los diversos medios de comunicación -periódicos, revistas, televisión- tendrán acceso a unos datos que les permitirán una mejor evaluación de los programas públicos y privados.

III. PLANIFICACION Y ORGANIZACION GENERALES

Preparación de un resumen para la aprobación del programa

37. Para los propósitos de este manual, se supone en este punto que algunas personas responsables con autoridad han decidido provisionalmente que sería útil emprender un programa de estadísticas en la esfera de la distribución y los servicios. Se supone también que ya existe una organización estadística que trabaja en otras encuestas. Como paso inicial para obtener la aprobación del programa, sería aconsejable preparar un resumen o justificación del programa propuesto. Buena parte de lo que se señala a continuación puede servir para preparar tal resumen. Este debería contener algunas ideas generales sobre costo, calendario y necesidades de personal. Tal cosa no se hará con mucho detalle pero debería dar a la dirección del organismo alguna indicación sobre los órdenes de magnitud para que pueda tomar una decisión acerca de si debe recomendar la aprobación del programa o una modificación del mismo. Una buena labor de presentación en esta etapa hará mucho en favor de la aprobación del proyecto.

Requisitos legislativos

38. Una vez que el organismo de estadística haya decidido auspiciar el programa, suele ser necesario aprobar una ley (o modificar una ya existente) facultando la ejecución del programa. Esa disposición debe redactarse de tal modo que haga referencia al campo de aplicación (que debería ser lo bastante amplio para abarcar todo el comercio de detalle, el mayorista y los servicios), la frecuencia de los censos generales, el carácter de las encuestas anuales y más frecuentes, las normas sobre el carácter confidencial y la publicación de la información suministrada por las empresas y la imposición de sanciones por no responder o por falsificar voluntariamente los datos pedidos. Esto fijará el marco dentro del cual pueden planificarse y realizarse las diversas encuestas. Cuando un servicio de la administración tenga ya amplias facultades para recabar información económica, bastará un decreto o disposición de poder ejecutivo. En el decreto deberían figurar los aspectos que se han señalado.

Determinación de la estructura orgánica

39. Toda decisión sobre la estructura orgánica vendrá condicionada por los servicios de estadística que existan. Cuando haya ya una estructura funcionalizada con actividades separadas, personal de planificación, una organización de campo, una división de publicaciones, etc., se necesitará simplemente agregar un personal de planeamiento para el comercio de distribución y los servicios que prepararía las especificaciones destinadas a las diversas divisiones operativas, que, a su vez, establecerían unidades separadas para ocuparse de esas encuestas. En los casos, sin embargo, en que no se hayan establecido funciones de elaboración como operaciones independientes, será menester contar con una organización en que se combinen la planificación y la elaboración en un solo grupo.

40. Una función que normalmente se relega a un personal separado es el del trabajo de campo que con toda probabilidad se realizará para varios programas diferentes. El personal de campo reviste especial importancia cuando el programa prevea la recogida de datos por muestreo de área en una serie de encuestas corrientes. Cuando un país no cuente ya con una organización para trabajar en el campo, el comienzo de una operación de esa índole resulta una tarea importante y costosa.

41. Al ponderar los méritos de un sistema que combine las funciones de elaboración y planeamiento en una sola plantilla frente a aquellos de una división del trabajo basada en la función que hay que realizar, dos factores importantes son las características del personal disponible y el volumen de la labor que hay que ejecutar. La función de planeamiento requiere personas muy enteradas de los problemas de cada sector. La función de elaboración también requiere conocimientos y experiencias especializadas. Desde el punto de vista ideal, el personal debería estar al tanto de ambos tipos de actividades, pero tal cosa no suele ser corriente. Además, cada actividad requerirá probablemente dedicación plena del personal asignado a la misma.

42. Es muy importante comprender el efecto que la elaboración puede tener en la calidad de los datos que haya que producir. Con demasiada frecuencia, un personal cuya función principal es elaborar los datos recurrirá a simplificaciones y hará modificaciones de procedimiento que si bien pueden acelerar la elaboración y reducir el costo, afectarán también a los datos y podrá hacerlos inútiles. Por otra parte, la combinación de ambas funciones puede redundar en un personal especializado sumamente competente que prepare una realización muy ineficiente y despilfarre los fondos con la consecuencia final de que haya que abandonar el programa. Es extremadamente importante que ambas funciones se tengan en cuenta en todo momento y que quienes hagan la planificación consideren siempre de qué modo sus planes afectarán a la elaboración, y viceversa. El personal de elaboración de datos debería seguir un cursillo sobre planificación y realización de encuestas y recíprocamente, el personal que se ocupe de las encuestas debería seguir, a su vez, un cursillo sobre elaboración de datos.

Reclutamiento del personal de planificación

43. Las funciones que tiene que desempeñar el personal de planificación consisten en determinar el contenido de los diversos temas; consultar a los usuarios de datos, asociaciones comerciales y personas encuestadas; diseñar y cotejar cuestionarios; participar en la planificación de la recogida de datos; y en los métodos de elaboración; preparar instrucciones para los entrevistados; redactar especificaciones para elaborar los datos comunicados; resolver los problemas sobre la recopilación de datos; planificar las publicaciones y los textos escritos; y revisar profesionalmente los datos tabulados. El personal que ha de realizar esas funciones tiene que ser obviamente de una gran capacidad. Los conocimientos fundamentales que se requieren son de estadística y economía con alguna formación matemática.

44. Un programa completo tanto de censos infrecuentes como de encuestas frecuentes requeriría preferentemente el establecimiento de dos grupos de planificación, uno para cada programa. Esto se basa en el hecho de que un programa repetitivo de encuestas frecuentes, tales como series mensuales, necesita la atención completa del personal asignado a esa labor, y la asunción por ese personal de los deberes adicionales relacionados con un censo general haría que ambos programas se resintiesen.

45. El personal asignado a las fases específicas de planificación del censo debería después correr con la responsabilidad de su ejecución. Tal cosa contribuye a garantizar la continuidad del planeamiento y la realización de las operaciones previstas. Permite que la formación y experiencia adquiridas en la etapa de planeamiento se utilicen plenamente en las operaciones posteriores del censo.

Organización de grupos de asesoramiento

46. En una de las primeras etapas, el personal de planificación debería organizar dos grandes grupos asesores de usuarios de datos. Pueden ser muy útiles para que la opinión colectiva y los conocimientos de funcionarios y público coadyuven a establecer los objetivos del censo. También pueden contribuir a que se inicie la labor censal consiguiendo la legislación, la cooperación pública y los fondos necesarios.

47. El primer grupo que debería formarse estaría constituido por asesores pertenecientes a la administración. Consistiría en funcionarios experimentados que fuesen usuarios efectivos o potenciales de los datos. Estos funcionarios deberían ocupar cargos elevados, para que pudiesen hablar en nombre de los servicios donde trabajen. También deberían ser técnicos que comprendiesen el carácter de los programas estadísticos y sus limitaciones. Los posibles candidatos serían representantes de los ministerios de trabajo, seguridad social, comercio y otros organismos que se ocupen de regular o promover el comercio y los servicios o la planificación y el análisis económico.

48. La segunda gran categoría de usuarios de datos con los que debería constituirse un grupo asesor es la de personas que no pertenezcan a la administración. Se trata de representantes de grupos de análisis de mercado (privados y académicos), representantes de medios de comunicación (prensa, radio, televisión), organizaciones de investigación estadística, economistas y organizaciones de consumidores. Aquí también habría que procurar seleccionar miembros con algún conocimiento de los problemas de las encuestas y que puedan aportar su contribución a la labor de planificación mejor que personas que no tengan tal capacidad. Si se encuentra que el comité asesor funciona con dificultad por causa de su tamaño, será necesario dividirlo en dos o más subgrupos.

49. Dos tipos de organizaciones han sido deliberadamente excluidas de los grupos asesores, no porque su consejo no se precise, sino porque la consulta con ellos reviste carácter diferente. El primero de ellos es el de las asociaciones de comerciantes que tienen que ser consultadas con mayor detalle y con referencia

específica a las necesidades y problemas relacionados con el sector de que se trate. Se recomienda así que cada asociación que represente a un sector incluido en la encuesta se contacte por separado para examinar las repercusiones de ésta en dicho sector. El segundo grupo excluido es el de los contadores. Los representantes de este grupo tienen que ser consultados sobre la disponibilidad de datos y el tiempo necesario para su notificación. Una vez más, tal cosa debería hacerse como una actividad por separado, y los contadores no deberían figurar en los comités asesores.

50. Cuando el personal de planificación haya llegado a una conclusión preliminar sobre el contenido de datos de la encuesta, sería aconsejable establecer un sistema para llevar nota oficial de todos los pasos dados para llegar al contenido de la encuesta. En esas notas debería figurar relación puntual de los contactos, personas consultadas, método usado (correo, entrevista personal, teléfono), observaciones de la persona que haya realizado el contacto y de un representante del servicio de planificación. Tales notas serán muy útiles para referencias futuras.

Decisiones fundamentales

51. Una vez que la encuesta haya sido aprobada y se disponga del personal de planificación, hay varias cuestiones que ese personal debe decidir, al menos con carácter provisional, en una de las primeras etapas. Esas cuestiones se examinan brevemente en los párrafos siguientes.

Campo de aplicación

52. Aun cuando la decisión general sobre los sectores que se incluirán en la encuesta haya sido ya tomada, no por ello dejará de ser necesario que los planificadores examinen esa decisión con referencia concreta a cada uno de los sectores incluidos y consideren qué modificaciones serían aconsejables. Por ejemplo, con algún ligero costo adicional la encuesta podría tal vez ampliarse para abarcar un sector en mayor detalle o para incluir otro en el que previamente no se haya pensado. O podría convenir, por el contrario, que se suprimiera un determinado sector si resulta imposible hacer en él la encuesta.

Cobertura

53. Una decisión primera y muy importante que hay que tomar y que tiene una gran repercusión en el presupuesto es la de la cobertura. Especialmente importante es decidir si ha de procurarse una cobertura completa u omitir categorías tales como establecimientos caseros, aquellos que no emplean personal, o los que quedan por debajo de una determinada dimensión o volumen de actividad, etc.

Disponibilidad de fondos

54. El personal de planificación necesitará saber si la decisión es firme para seguir adelante con la encuesta y en qué nivel debe planificar. El organismo estadístico debería poder hacer una declaración bastante concreta para que los objetivos del programa se ajusten a una estimación razonable de los fondos de que se dispondrá.

Método para reunir datos

55. Una vez que estén tomadas las decisiones sobre campo de aplicación y cobertura y se tenga una seguridad razonable de que se contará con los fondos adecuados, habrá que decidir el método que se seguirá para recoger los datos. Tendrá que darse respuesta a preguntas como en qué medida se obtendrán datos de otros servicios de la administración, hasta qué punto se recurrirá al muestreo, la proporción del universo que se cubrirá por correo, la extensión en que los datos se obtendrán por recuentos sobre el terreno, de qué modo se determinará la ubicación material de los que no respondan y cómo se obtendrán direcciones para la correspondencia. Representantes de los servicios tanto de planificación como de elaboración tendrán que trabajar de consuno en esos temas.

Método para elaborar datos

56. Con los continuos avances en el diseño de equipos de elaboración de datos, el número de opciones disponibles para la tramitación y tabulación de datos sigue aumentando. Conviene decir unas palabras de advertencia en este punto. Con demasiada frecuencia se decide usar equipo electrónico sofisticado cuando en realidad no sólo es innecesario sino también inconveniente. Incluso los países estadísticamente más avanzados tienen gran dificultad para usar tal equipo, y a menos que se disponga de personal que domine los problemas de redacción de los programas, funcionamiento de las máquinas y mantenimiento del equipo, es mejor recurrir a sistemas más sencillos. Las primeras consideraciones deberían referirse al tipo de equipo disponible y a la viabilidad de usarlo para elaborar los resultados de la encuesta.

57. En cualquier caso, ha de tomarse una decisión sobre el tipo de equipo que se empleará y en qué medida se necesitará personal administrativo y profesional para preparar los informes y redactar y revisar las tabulaciones. Un requisito conexo es la evaluación de las necesidades de equipo y personal para la correspondencia, operaciones de cotejo y actividades complementarias y para la preparación y mantenimiento de una guía de establecimientos mercantiles.

Necesidades de equipo y personal

58. Una vez adoptada la decisión general sobre cómo se procederá a la elaboración, será menester enumerar todas las operaciones de esa elaboración e indicar cuándo se llevarán a cabo. Tendrán que calcularse los distintos trabajos que habrá que hacer y los requisitos correspondientes de equipo y personal, por tipos. Tal

cosa debería coordinarse con las necesidades de otros organismos y con el equipo nuevo que se pida y el personal que habrá que contratar en caso necesario. Al estimar las necesidades de equipo, debería tenerse en cuenta el hecho de que muchas veces lleva tiempo su entrega. Además, debería tenerse presente el tiempo que se perderá por averías de material o por retrasos, tales como los derivados de que una operación anterior no entregue una cantidad de trabajo suficiente para su elaboración.

59. Cuando el organismo de estadística cuente ya con un sistema centralizado de diseño de sistemas, podrá sacarse partido de un personal especialmente capacitado en la optimización de las operaciones de elaboración. No obstante, será menester que los servicios de planificación examinen los procedimientos propuestos por el personal de diseño de sistemas, ya que no siempre son obvios los efectos que una operación concreta puede surtir en la calidad de los datos que se producen.

Asignación de responsabilidades

50. Por el carácter complejo de estas encuestas y porque la ejecución del plan afectará a muchas y diferentes dependencias del organismo de estadística, es fácil que algún aspecto del programa pueda dejarse de lado. Enumerar todas las actividades principales y asegurarse de que todos y cada uno de los servicios que intervienen conocen sus responsabilidades específicas respecto de dichas actividades es, por tanto, algo necesario. Cada dependencia debería examinar sus operaciones para determinar el mayor o menor grado en que su propio trabajo depende de las actividades (tales como redacción de especificaciones, preparación de planes tabulados, elección del formato de la producción de las máquinas, evaluación de la disponibilidad de los resultados de los ensayos, etc.) de otros servicios, asegurándose de que todos ellos saben perfectamente lo que se requiere de cada uno y cuándo.

Coordinación con otras encuestas

61. Si bien este manual se ha preparado para que se use en la organización y realización de encuestas del comercio de distribución y los servicios, no hay que olvidar que tales encuestas no se hacen aisladamente. Muchas veces se llevan a cabo conjuntamente con otras sobre industrias, extracción de minerales, construcción, etc. Como todas estas encuestas contienen muchos elementos comunes y como una sola empresa con establecimientos múltiples puede trabajar a menudo en diversos sectores, es importante comprobar que no existen contradicciones en los formularios donde se recaba información y en las instrucciones que preparan los diferentes servicios de planificación que trabajan en las encuestas. Además, al preparar listas de direcciones y en las operaciones derivadas de los casos de no respuesta, una operación planeada de modo conjunto no sólo resultará más eficiente sino que también ayudará a resolver problemas sobre el alcance y definición de la unidad estadística y de información. Es, por tanto, conveniente que se establezca un "comité de elementos comunes" en una de las primeras etapas de la planificación donde figuren representantes de cada división para que se ocupe de redactar las preguntas que serán comunes a todos los cuestionarios. El mismo comité debería ocuparse de preparar una lista de direcciones conjunta y el plan complementario para los casos de no respuesta.

62. Si bien la redacción de las preguntas comunes no tiene por qué ser exactamente igual en los diferentes formularios, no debe soslayarse el efecto que una redacción diferente puede tener sobre otras encuestas. Por ejemplo, al establecer la unidad estadística para una instalación industrial, puede ser necesario hacer alguna referencia a localizaciones separadas de las ventas que estarían clasificadas adecuadamente en el sector de ventas y figurar en el comercio al por mayor. El uso de un comité de elementos comunes reviste particular importancia al iniciarse las encuestas de la distribución y los servicios, ya que por lo general éstas se empiezan después de haberse realizado las encuestas sobre el sector industrial. Se encontrará, por consiguiente, que hay tendencia a dejar de lado problemas como la identificación del sector de ventas, lo que lleva a inflar la actividad manufacturera.

63. Las atribuciones del comité de elementos comunes deberían abarcar puntos tales como la redacción de las preguntas sobre identificación de establecimientos, propiedad, forma jurídica de organización, período de operación y ubicación física. También debería velar por que se usen definiciones normalizadas para términos tales como "empleo", "nómina de pago" y "gasto de capital". El Comité debería asumir la función de planificar y redactar las especificaciones para un primer estudio de las empresas con más de un establecimiento a fin de determinar su emplazamiento, la clasificación provisional sectorial y las dimensiones aproximadas de cada establecimiento de todas las empresas de esa índole, a fin de que pueda enviarse el formulario adecuado para obtener información. Otra función es determinar el contenido del membrete que se use en la correspondencia. El comité también puede usarse de forma enlazada con otros servicios de la administración en cuestiones que son comunes a todas las divisiones, tales como obtener listas de direcciones o datos.

64. De lo contrario debería quedar claro que los miembros de ese comité tienen que ser técnicos muy competentes. También deben figurar en él representantes de las divisiones de elaboración para evaluar la viabilidad y estimar el costo de los planes que haga el comité. Si bien el número de elementos comunes que intervienen en la realización de encuestas corrientes frecuentes es mucho menor que en los censos generales, sigue siendo lo bastante importante como para estar seguros de qué preguntas y definiciones son coherentes entre sí en ambos casos. Los problemas surgen porque la necesidad de una información rápida en las encuestas corrientes hace difícil que quienes responden perfeccionen los datos; en los censos generales, quienes contestan tienen más tiempo para preparar registros precisos. Tal cosa puede conducir a falta de uniformidad entre ambos tipos de actividades. El servicio de planificación de las encuestas corrientes debería, por tanto, revisar cuidadosamente la labor del comité de elementos comunes.

Planificación de un ensayo previo

65. Un procedimiento muy útil para determinar la viabilidad y costo de un censo es el ensayo general previo que consistiría en levantar un censo real de una porción representativa del país -tal vez una provincia entera- que contenga zonas tanto urbanas como rurales. La finalidad última de ese ensayo previo

debería ser la publicación efectiva de los datos. Esto necesitará una planificación y elaboración cuidadosas. Sólo de ese modo se podrá estar seguro de que se conocen la mayoría de los problemas que es probable se encuentren al hacer la encuesta o el censo en todo el país. Inevitablemente, habrá algunos problemas propios de un sector o una zona particular que no se encontrarán incluso en un ensayo general de esa índole. Si bien puede parecer que gastar tiempo y dinero en un ensayo previo de carácter general no vale la pena, el resultado en términos de la capacidad para perfeccionar procedimientos, formar personal y sentar las bases adecuadas para estimar los costos se traducirá en una realización del censo mucho más eficiente y en una seguridad mayor de que toda la labor se verá coronada por el éxito.

66. Otra forma de ensayo previo resulta útil en relación con problemas relacionados con un sector concreto -por ejemplo, al determinar la redacción correcta de formularios y la viabilidad de la recogida de datos sobre porcentajes de ocupación de hoteles o número de entradas vendidas en cines y teatros. En tales casos, el procedimiento mejor podría ser preparar un formulario para obtener la información donde figuren los puntos concretos que interesa ensayar y que se enviaría para que lo cumplimentasen un grupo representativo de empresas del sector. La respuesta sería entonces analizada por el servicio de planificación antes de que se redactaran las preguntas definitivas. Este tipo de ensayo previo puede usarse con igual efectividad tanto en los censos generales como en las encuestas corrientes. En el caso de encuestas frecuentes, los primeros períodos de recogida de información deberían considerarse como ensayos previos. Deberían hacerse planes para un rápido examen intensivo de las encuestas iniciales a fin de que cualesquiera cambios que se necesiten en procedimientos o cuestionarios puedan hacerse rápidamente.

Planificación de la publicación

67. La medida del éxito de una encuesta viene dada por su capacidad para producir la información deseada dentro de los límites existentes de costo y tiempo. Antes de emprender una encuesta, el director tiene que conocer en detalle cuáles son sus objetivos en materia de datos. Los usuarios de esos datos deben tener la posibilidad de examinar el resultado que se persigue, y el personal de elaboración ha de conocer qué objetivos está intentando conseguir. El modo mejor de alcanzar esas metas, por lo que hace a los planificadores, es preparar un conjunto de tabulaciones posibles para que lo examinen y comenten los usuarios de los datos y el personal de elaboración.

68. Un plan provisional de publicaciones, si se prepara con atención y cuidado, obligará al personal de planificación a hacerse algunas preguntas difíciles pero importantes tales como si el cuestionario está resultando demasiado largo y complicado por intereses especiales de ciertos usuarios, o si las preguntas están redactadas de tal modo que las respuestas se excluyen unas a otras, o si todas las preguntas pueden tabularse, o si el recurrir a entrevistar a un pequeño número de personas con el fin de obtener unos datos concretos no conducirá a demasiadas supresiones por problemas derivados del carácter confidencial de esos datos, o si la tabulación propuesta conducirá a un número imposible de páginas

para recoger los datos que hay que publicar, o si los datos de dos o más cuadros no pueden combinarse en uno solo, ahorrando dinero, o si la cantidad del tiempo de máquina necesario para producir un cuadro propuesto es excesiva, o si el programa de publicación total es tan complicado que resultará demasiado costoso o llevará demasiado tiempo, o cuál será el efecto de variar la muestra en la capacidad de publicar información. Estas y otras preguntas que necesitan respuesta vendrán sobre el tapete al examinar el plan de publicación.

69. El personal de elaboración ha de estar al tanto de todas las tabulaciones propuestas para que pueda evaluarlas en términos del costo y tiempo que llevará su preparación. Esto debería hacerse en un primer momento para que puedan descubrirse los principales problemas y estudiarse soluciones alternativas. Una vez que se convenga el conjunto final de tabulaciones, el trabajo puede avanzar simultáneamente en varios frentes. Entre ellos, la fijación del contenido del formulario para recabar la información, la asignación de códigos a aspectos pertinentes de los datos, las claves de la información que se perforará, la preparación de procedimientos administrativos de revisión, el diseño de métodos de elaboración por máquina para la recogida y análisis de datos y la preparación de planes de tabulación. En esa etapa será posible reconsiderar las estimaciones originales de costo, personal y tiempo e introducir los oportunos ajustes en el programa como consecuencia de esa evaluación.

Diferencias entre planificación de programas censales y planificación de programas de encuesta

70. Es posible o bien usar el mismo personal de planificación tanto para los censos generales como para las encuestas corrientes, o bien tener dos plantillas independientes. Ambas cosas tienen sus ventajas e inconvenientes. Los inconvenientes principales de contar con dos plantillas separadas consiste en la falta de uniformidad que se creará entre ambos programas al trabajar en ellos personas diferentes, lo que redundará en problemas de comunicación, o en intentar mantener un personal formado en labores censales durante todo el ciclo del censo, con las consiguientes dificultades presupuestarias. En cambio, tener una sola plantilla hace que por lo general el censo no disponga de personal suficiente, ya que la necesidad de publicar muchos informes periódicos hace que el esfuerzo principal se concentre inevitablemente en esa actividad.

71. Al planificar encuestas frecuentes, cabe establecer un ciclo regular de operaciones con un plan de trabajo conocido y previsible. Tal cosa hace que pueda fijarse el tamaño del personal y el costo de la operación y mantener bastante bien el nivel de personal capacitado disponible. Esto no se aplica, sin embargo, cuando es menester revisar la muestra, lo que suele hacerse después de terminar un censo general, cuyos registros sirven para obtener información al día sobre clasificación, tamaño y composición. Para esta operación será posible usar parte del personal de planificación del censo que por entonces estará disponible. No obstante, habrá costos adicionales muy grandes para actividades sobre el terreno y elaboración.

72. El tamaño del personal necesario para elaborar el censo infrecuente estará sujeto a grandes variaciones a lo largo del ciclo del censo, particularmente en el caso de personal de elaboración y de equipo. Esto puede conducir a ineficiencias en el empleo del equipo y a problemas para obtener el oportuno personal administrativo durante el período relativamente breve que se necesita; por tanto, reviste especial importancia que se hagan previsiones detalladas de las operaciones. Cuando en un censo general se combina el empleo del correo y la inspección sobre el terreno, surgen problemas adicionales para establecer oficinas locales y contratar y formar entrevistadores para un período relativamente corto de tiempo. Esta operación es todavía más difícil y costosa que establecer una oficina administrativa de elaboración de carácter temporal. A diferencia de las encuestas repetitivas en las que los entrevistadores, una vez formados, pueden usarse una y otra vez, un censo general requiere que el entrevistador pueda hacer su labor bien la primera vez, ya que dispondrá de muy poco tiempo para corregir sus errores.

IV. DIRECCION Y PRESUPUESTO

Uso de manuales

73. Una de las formas más útiles de fijar en las encuestas normas y procedimientos es preparar varios manuales oficiales que se distribuyan a todas las personas responsables de la planificación y la elaboración. Esos manuales deberían estar constituidos por hojas sueltas encuadradas, con una paginación flexible para poder ir aumentando cada sección. Debería preverse la sustitución de páginas cuando se cambien especificaciones o procedimientos después de que la sección correspondiente esté ya incorporada al manual. Debería fecharse cada sección cuando se redacta, y los cambios tendrían que llevar igualmente fecha. Aquellas secciones que sean sustituidas por versiones ulteriores deberían retirarse del manual y conservarse por separado como posible referencia.

74. Fundamentalmente, el personal de planificación se ocupará de redactar las especificaciones que indiquen al servicio de elaboración lo que tiene que hacerse, y este último redactará los procedimientos para alcanzar esas metas. Muchas veces, sin embargo, para asegurar el logro de los objetivos deseados, el personal de planificación tendrá que escribir las especificaciones con tal detalle que casi constituirán un procedimiento. La determinación de las técnicas más apropiadas para alcanzar un objetivo concreto debería correr a cargo principalmente del personal de elaboración, pero tanto las especificaciones finales como los procedimientos últimos deberían establecerse después de que miembros de ambos servicios examinen las versiones preliminares.

75. El manual preparado por el personal de planificación debería tener sendos capítulos sobre:

- Confección de listas de direcciones
- Envíos por correo, comprobación y actividades complementarias
- Registros administrativos
- Operaciones sobre el terreno
- Muestreo
- Sistemas de codificación
- Elaboración administrativa
- Elaboración por máquina
- Planes de publicación

Uso de organigramas

76. Otro mecanismo que cabe usar en conexión con el manual o como parte del mismo es la presentación de especificaciones y procedimientos en forma de un diagrama de movimientos. Una encuesta puede contemplarse como una serie de acciones interdependientes que conducen a la publicación de datos. Cada acción procede de otra anterior y conduce a su vez a una posterior. La relación

recíproca entre esas acciones puede presentarse como un diagrama o bien verbalmente. Ambos métodos son útiles. La presentación gráfica tiene la ventaja de ser más concisa y hace más fácil que muchas personas pueda seguir con comodidad el desarrollo de los trabajos; también hace más sencillo descubrir pasos ilógicos o la omisión de operaciones necesarias. En un diagrama de movimientos que ocupe una página es posible presentar material que en forma escrita requeriría muchas páginas.

Calendarios

77. Uno de los primeros pasos para establecer un plan realista de trabajo es fijar un calendario de las actividades fundamentales. Como el objetivo de toda encuesta es producir datos, la primera fecha que hay que fijar es la de producción de las publicaciones. Una vez establecida, será menester ir retrocediendo a través de todas las operaciones que se necesitan para llegar a dicha meta. En esta primera etapa, sólo cabe determinar fechas clave. Después, será menester ver qué se necesita para cumplir el calendario. Algunas de las operaciones fundamentales, además de la publicación de datos, para las que deben establecerse fechas son (en orden inverso) planeamiento de la tabulación, redacción por máquina de los datos, elaboración de los formularios, envío por correo de los mismos, confección de las listas de direcciones, fijación de los procedimientos de muestreo, realización de las operaciones sobre el terreno, comprobación de que se dispone de los formularios, redacción de los formularios y establecimiento del campo de aplicación y la cobertura.

78. Al fijar su fecha de terminación, también había que determinar cuándo la operación ha de iniciarse para completarla a tiempo y qué otras operaciones tendrán que acabarse antes de que aquélla pueda empezar. Entonces cabrá averiguar si se ha previsto suficiente tiempo previo para evitar que se retrasen operaciones ulteriores. Establecer fechas para la realización de operaciones fundamentales, asegurarse de que se ha previsto suficiente tiempo para completarlas en el plazo previsto y comprobar que las operaciones anteriores estarán acabadas a tiempo puede hacer mucho para el éxito de una encuesta. Después, podrán prepararse todas las demás operaciones que giran en torno a esas metas así como planificar el empleo de personal y de equipo. Si se registra algún desfase en el cumplimiento de esas fechas clave, habrá que revisar las operaciones ulteriores para modificar el calendario o bien trasvasar recursos para mantenerlo.

79. Cabe trazar un cuadro de las operaciones principales enumerándolas e indicando las fechas de comienzo y terminación en una serie de columnas que representen trimestres del calendario. Este tipo de presentación pondrá en evidencia cualesquiera operaciones que se superpongan y aquellas cuyo calendario tiene que revisarse porque no estarán terminadas cuando se necesite. También sienta las bases para establecer el plan de trabajo por períodos de tiempo y, por lo tanto, para indicar los requisitos de personal y equipo también por períodos.

Programación de personal y equipo

80. Tomando como base el gráfico descrito antes, cada unidad orgánica que corra con la responsabilidad de una función concreta debería preparar una estimación, por trimestres, de cuánto y de qué tipo de personal y equipo necesitará para alcanzar el objetivo indicado. En la preparación de esa estimación la unidad debería desglosar las operaciones principales que corren a su cargo en suboperaciones para las que también deberían indicarse los objetivos en el tiempo. Tal cosa puede revelar estrangulamientos potenciales u operaciones excesivamente costosas. Cuando todas las estimaciones se agreguen, será posible obtener el cuadro general del costo y necesidades de personal por períodos de tiempo. Después de introducir todos los ajustes que se consideren oportunos, podrá establecerse un calendario firme, y todas las unidades conocerán sus responsabilidades. Cada una, sin embargo debería tener ocasión de presentar cualesquiera objeciones o reservas que pueda tener sobre las metas que se le hayan fijado, y esos puntos deberían discutirse detenidamente antes de tomar las decisiones finales.

Sistema de información sobre la marcha de los trabajos

81. La dirección adecuada de una encuesta depende del conocimiento que se tenga acerca de si la labor está avanzando con arreglo a las estimaciones de tiempo y costos. Una vez fijados los objetivos, hay que implantar un sistema de información sobre la marcha de los trabajos.

Por desgracia, no todas las operaciones se prestan a una medición fácil. Funciones tales como perforación por teclado, examen administrativo de informes y contestación de la correspondencia son relativamente fáciles, porque la actividad es bastante uniforme para cada una de las unidades de que se trate, y el plan de trabajo, la producción y los retrasos pueden indicarse en términos de dichas unidades. Cada persona que interviene en el proceso puede ser juzgada con arreglo a su índice de producción y evaluarse en comparación con los índices de otras personas. Además, los índices generales de producción proporcionan una base para determinar si las actividades están ajustándose al calendario. En caso necesario, los procedimientos pueden cambiarse fácilmente para acrecentar la productividad, o los empleados ineficientes pueden ser despedidos o transferidos a otros puestos. En cambio, actividades tales como la redacción de especificaciones para la elaboración administrativa o por máquina no se prestan tan bien a una medición de la marcha de los trabajos. No obstante, seguirá siendo necesario establecer objetivos en el tiempo y seguir de cerca el progreso de tales actividades. Es especialmente importante vigilarlas estrechamente porque muchas veces no es posible recuperar las demoras mediante personal adicional, ya que el que se necesita ha de estar muy especializado.

82. La función más difícil en lo que a la evaluación de la marcha de los trabajos se refiere es la de escribir y comprobar programas de computadora. Esto se debe a que sólo después de que un programa demuestre ser adecuado al aplicarse a datos reales cabe estar seguro de que el trabajo está terminado. Toda dificultad que aparezca incluso en las etapas finales puede requerir mucho tiempo para su corrección.

83. La marcha de los trabajos debería medirse de dos maneras, a saber, unidades producidas (en relación con las metas finales) y costo generado. Si bien las unidades producidas pueden medirse y presentarse sobre una base temporal, la información sobre los costos se producen con mayor dificultad y de ordinario con demasiado retraso para tener valor práctico. Es posible, sin embargo, hacer una aproximación aprovechable del costo reuniendo datos sobre hombres-día de trabajo realizado, por funciones, y asignando un costo calculado por hombre-día, que debería estimarse por separado para cada función. Igualmente, cabe distribuir un costo a esas funciones en operaciones de computadora, acumulando información diaria sobre el número de horas de máquina usadas y multiplicándolas por el costo estimado por hora. El costo acumulado y completo de ambas unidades deberían analizarse frecuentemente en relación con las fechas de terminación del calendario y asignarse fondos para esa tarea, tomando acciones correctoras con antelación suficiente como para ser efectivas.

84. Como mínimo se requieren dos niveles en la información sobre la marcha de los trabajos. Para la dirección general, la información puede limitarse a las principales operaciones y elementos del costo. Para la dirección de nivel intermedio, esas operaciones principales deberían dividirse en suboperaciones significativas relacionadas tanto con los elementos del costo como con aquellas operaciones que puedan convertirse en estrangulamientos, incluso aunque no puedan entrañar grandes gastos. Hay que tener presente que los costos derivados de que una operación lleve más tiempo del inicialmente previsto afecta no sólo a ella sino que muchas veces conduce a que quede inactivo personal y equipo asignado a operaciones ulteriores, con la consiguiente alza de los costos correspondientes.

Establecimiento de un sistema de control de calidad

85. Además de llevar las operaciones al día, es muy importante mantener una calidad aceptable. Esto es especialmente importante en un censo general, ya que no hay oportunidad de corregir un error en actividades posteriores, como sucede en las encuestas que se repiten mensual o trimestralmente. Para vigilar las operaciones, por tanto, debería organizarse una dependencia de control de calidad que diseñe un sistema de muestras que habría que elegir entre las diversas operaciones para establecer si se están manteniendo niveles de calidad aceptables. En muestras relacionadas con la perforación de tarjetas y con actividades administrativas, cabe diseñar un sistema que determine también si cada uno de los empleados están produciendo niveles de cantidad aceptables. Los sistemas de control de calidad también pueden prepararse para comprobar cosas tales como listas de direcciones confeccionadas por computadora, operaciones sobre el terreno, respuesta adecuada a la correspondencia, codificación, etc. El personal de los servicios del control de calidad debería ser independiente, esto es, no tendrá que formar parte de los servicios que se supone está supervisando. Antes de que se pongan en aplicación, los procedimientos de control de calidad deberían revisarse por el personal de elaboración y planificación.

Maximización del uso de los recursos

86. Existen varios mecanismos que pueden ser útiles para aprovechar al máximo los recursos disponibles. Uno es un uso adecuado de las horas extraordinarias. A menudo un director, al intentar conseguir que se respete el calendario, tiene que sopesar los costos alternativos de formar un personal más numeroso o de emplear una plantilla menor e invertir en horas suplementarias, aunque sea a tarifas más elevadas. Esta última alternativa puede resultar menos costosa. Las horas extraordinarias, sin embargo, nunca deberían usarse tanto como para producir ineficiencia al fatigar demasiado al personal.
87. Cabe sacar partido del hecho de que muchas operaciones de un censo se siguen unas a otras para usar el mismo personal en etapas diferentes. La formación y experiencia adquiridas en una operación se transfiere a otras ulteriores y redundan en un trabajo mejor y más eficiente. Conseguir esto requiere una planificación y programación temporal muy cuidadosas así como un cumplimiento fiel de los calendarios.
88. Un organismo de estadística que se dedique a elaborar muchas y diferentes encuestas encontrará que el uso del equipo mecánico y electrónico fluctúa mucho. Una buena programación puede aprovechar los períodos de menor trabajo y lograr una disminución general de los costos. Esto es especialmente cierto en el empleo de computadoras.
89. Uno de los principales problemas que hay que afrontar en materia de costos es el control de la cantidad de documentación que circula. En relación directa con ella está la cuestión de cómo obtener un empleo óptimo del personal administrativo. Hay que llegar a un equilibrio entre el simplificar cada función de un administrativo hasta el punto que pueda realizarla con extrema rapidez y los problemas derivados del consiguiente aumento de la documentación que tiene que pasar de un empleado administrativo a otro. Por lo general, las operaciones que tienen que realizar esos empleados deberían incluir todo lo que se pueda aprender a manejar de modo efectivo en un período limitado de formación. Cuando hay que establecer muchos y complicados controles porque la documentación que pasa de una unidad operativa a otra está resultando excesiva, hay pérdida de eficiencia, frecuentemente agravada por la dificultad, o incluso la desaparición, del control.
90. Los costos pueden controlarse en alguna medida, especialmente en un censo general, mediante el buen uso del muestreo para aquellos tipos de datos cuya publicación se requiere sólo a nivel nacional o en grandes categorías regionales. Tal cosa complicará la encuesta hasta cierto punto por la necesidad de diseñar muestras y por los correspondientes problemas de elaboración, pero permitirá obtener una ganancia neta en los costos de funcionamiento.
91. Cuando se estén haciendo operaciones administrativas en masa, es importante establecer una serie de normas con las que medir la producción de cada empleado. Tales normas deberían revisarse periódicamente y colocarse en los tablones de anuncios, quizá una vez a la semana para que cada administrativo pueda conocer

su posición relativa. En conexión con el índice de producción debe figurar un índice de calidad, ya que una producción lograda a costa de la calidad resulta negativa. Un empleado cuya producción quede constantemente por debajo de niveles aceptables debería ser despedido o transferido a otras actividades si se piensa que puede hacer allí una labor aceptable. Esto cumple la doble finalidad de eliminar a trabajadores improductivos y de actuar de incentivo para resultados mejores.

92. Independientemente de lo bien que esté planificada la encuesta, el hecho de que haya tantas actividades y que la responsabilidad de las diferentes etapas de la encuesta pueda dispersarse entre distintas divisiones del organismo de estadística conduce frecuentemente a la aparición de estrangulamiento y obliga a arbitrar entre opiniones antagónicas, a desplazar personal, a cambiar prioridades, a reajustar presupuesto, etc. Para situaciones como esas conviene nombrar una persona con autoridad suficiente que coordine toda la encuesta y prevea y resuelva problemas y conflictos. Si bien gran parte de la información que necesita para funcionar de modo efectivo procederá de los diversos informes operacionales sobre la marcha de los trabajos y el presupuesto, debería gozar de libertad para inspeccionar cualquier operación personalmente e intentar encontrar soluciones a situaciones problemáticas antes de que se deterioren demasiado. También debería tener las manos libres para convocar reuniones de personal que represente a las diferentes divisiones, reajustar las prioridades y celebrar negociaciones con otras divisiones con el fin de mantener el programa dentro de los límites del calendario y del presupuesto.

Diferencias entre censos y encuestas corrientes

93. El hecho de que las encuestas corrientes se repitan a intervalos regulares y breves, mientras que un censo completo del comercio de distribución y los servicios probablemente se haga sólo cada cinco o diez años, tiene un efecto claro sobre el presupuesto y la dirección. Los costos de las encuestas repetitivas tienden a ser estables de un mes a otro, y las operaciones pueden perfeccionarse con el tiempo. En caso necesario cabe establecer programas que fijen el día de la semana e incluso la hora del día. Tal cosa no sucede en los censos, donde el grueso del trabajo y del costo corresponde al año que sigue al período de referencia del censo, variando mucho las operaciones y presupuestos de cada uno de los demás años. En las encuestas corrientes, la necesidad de una producción rápida y regular de indicadores clave significa adherirse a programaciones muy estrictas que pueden planearse por el personal de elaboración, incluso hasta el extremo de prever períodos regulares de horas extraordinarias. En un censo, todo debe apuntar a que el trabajo se haga adecuadamente la primera vez; en otro caso, tendrá que rehacerse hasta que hayan sido corregidos los errores antes de poder pasar a la operación siguiente. Cuando un error no se detecta hasta que ya han sido dados varios pasos ulteriores, todos los intermedios pueden tener que volverse a hacer. Tal cosa tendrá casi siempre un costo elevado.

94. Las variaciones en el costo de mano de obra y de otras partidas por los incrementos de precios y salarios durante el ciclo normal de cinco o diez años de un censo pueden rápidamente invalidar las estimaciones presupuestarias. Las fluctuaciones en la disponibilidad de personal durante ese período también causarán quizá dificultades. La pérdida de personal capacitado de un período de censo al siguiente es otro inconveniente grande derivado de los cambios en las necesidades de mano de obra no sólo de mes a mes sino también de censo a censo. Esto es especialmente negativo cuando transcurren diez años entre dos censos.

ii

V. CAMPO DE APLICACION, COBERTURA Y CLASIFICACION

Introducción

95. Dos publicaciones de las Naciones Unidas deberían considerarse partes integrantes de este manual; son la Clasificación Industrial Internacional Uniforme de todas las actividades económicas (Nº de venta S.68.VIII.8) y Recomendaciones internacionales sobre estadísticas del comercio de distribución y los servicios (Nº de venta S.75.XVIII.9).

96. En la Clasificación Industrial Internacional Uniforme (CIIU) figura un examen muy completo de los principios seguidos para establecer el sistema de clasificación y de la aplicación de los mismos. Sucesivamente se analizan el propósito y naturaleza del sistema, los criterios usados en su elaboración, la identificación de unidades estadísticas, la clasificación de la unidad estadística, la estructura del sistema y la adaptación del sistema internacional a los requisitos nacionales. Las recomendaciones internacionales contienen propuestas específicas sobre el campo de aplicación y cobertura de las encuestas propuestas; un examen de la unidad estadística; las características que deberían determinarse para cada una de esas unidades; las clases de datos (y sus definiciones) que han de reunirse; y un conjunto de cuadros con recomendaciones sobre qué clases de datos deberían incluirse en encuestas infrecuentes, en encuestas anuales y en encuestas de frecuencia más que anual, con recomendaciones separadas para países con estadísticas ya desarrolladas y países en vías de obtener estadísticas sobre distribución y servicios.

97. La presentación siguiente no se propone sustituir a las publicaciones mencionadas sino que aspira a señalar las dificultades principales y a sugerir métodos para abordarlas. Se reconoce que las economías de los diferentes países variarán considerablemente y que existen consideraciones nacionales que condicionarán el carácter de las acciones concretas que se emprendan. Se espera, sin embargo, que se intentará cumplir en todo lo posible esas recomendaciones, a fin de que puedan alcanzarse los objetivos internacionales. Siempre que sea posible, en el análisis siguiente se pondrán de relieve diferencias que pueden ser importantes cuando se consideran encuestas anuales y de frecuencia más que anual en contraste con los censos generales.

Campo de aplicación

Comercio al por mayor (CIIU 61)

98. Las recomendaciones internacionales sugieren que se incluyan todos los tipos de mayoristas como parte del campo de aplicación de una encuesta del comercio de distribución y los servicios. Cuando se procede a un recuento general sobre el terreno con el fin de identificar todo establecimiento mercantil, ese objetivo puede alcanzarse, aunque se tropezará con algunas dificultades para localizar aquellas actividades que se desarrollen directamente desde los hogares, a menos que

sus direcciones figuren en las listas de licencias de comercio o en otros registros administrativos. Cuando, sin embargo, el diseño de la encuesta requiera combinar una lista de direcciones de empresas mayores y un muestreo de áreas que represente a las empresas menores (como suele ocurrir cuando la frecuencia es más que anual), surgirán dificultades que obligarán a modificar el campo de aplicación de aquella parte del comercio al por mayor que representa a los mayoristas no comerciantes.

99. Existen varios problemas. Primero, el número de mayoristas no comerciantes es relativamente pequeño, y no están distribuidos geográficamente de un modo tal que pueda hacerse un muestreo efectivo de áreas. Segundo, ese tipo de mayoristas, como agentes de ventas y corredores suelen tener negocios más bien reducidos y frecuentemente operan desde sus hogares. Esto hace difícil no sólo localizarlos en un recuento sobre el terreno sino también establecer una lista de direcciones fiable. Los departamentos y oficinas de ventas de los fabricantes presentan otro tipo de problema. A menos que simultáneamente con la del comercio al por mayor se esté realizando una encuesta de las actividades industriales, la identificación y enumeración apropiadas de esos departamentos y oficinas son muy difíciles.

100. Por lo tanto, para las encuestas anuales del comercio al por mayor y para las de frecuencia más que anual, probablemente es mejor restringir el campo de aplicación a los mayoristas comerciantes, que se definen en la CIIU como "mayoristas que reciben el título de los bienes que venden -tales como revendedores o intermediarios mayoristas, distribuidores industriales, exportadores, importadores, empresarios de silos terminales y cooperativas de compras". No obstante, si pueden llevarse listas al día de todas las empresas mercantiles juntamente con la clasificación de su clase de actividad y alguna indicación del tamaño, cabe usarlas para su aprovechamiento completo por correo o para seleccionar una muestra representativa. No obstante, es raro que se dé ese conjunto de circunstancias.

Comercio al por menor (CIIU 62)

101. Este aspecto de la distribución tiende a seguir muy de cerca las pautas de la distribución de la población y así se presta mucho al uso de un diseño eficiente de muestreo por áreas. El problema de distribución que existe con el comercio al por mayor no se plantea y es, por tanto, posible incluir todos los tipos de comercio al por menor en aquellas encuestas anuales y de frecuencia más que anual que incorporan una característica de muestreo de área. Por causa del elevado porcentaje de minoristas que salen y entran del sector en un corto período de tiempo, es muy difícil preparar una lista de direcciones que pueda usarse por sí sola para encuestas anuales y de frecuencia superior. Esa lista tiende a ser representativa de sólo las empresas mayores y más estables. Resulta, por tanto, mejor recurrir a un muestreo de área que represente a todos los negocios menores y que complemente la lista de grandes establecimientos obtenida de ordinario de un censo general.

Servicios

102. El campo de aplicación de los servicios no se presta fácilmente a una definición. Además, las recomendaciones internacionales para este grupo de actividades son muy selectivas. Ese campo de aplicación se limita "a los servicios que normalmente proporcionan a los hogares y empresas los establecimientos que pueden describirse como tiendas, hoteles, restaurantes, teatros, talleres y oficinas". Existen, sin embargo, varias excepciones, y el campo de aplicación propuesto y las matizaciones correspondientes deberían estudiarse con cuidado. Una inspección de la índole de tales actividades mostrará que algunas de ellas presentan grandes problemas si se incluyen en encuestas anuales o más frecuentes, especialmente si no pueden prepararse y llevarse al día listas de direcciones actualizadas. En varios casos (hoteles de temporada, producción de películas, radio y televisión), no estaría muy indicado usar muestras de área.

Cobertura

Observaciones generales

103. Cabe suponer que, en la práctica, una cobertura completa del comercio al por mayor, comercio al por menor y servicios es virtualmente imposible. Tal cosa sucede porque muchos de los establecimientos que trabajan en esos tipos de actividad funcionan de un modo que hace muy difícil localizarlos, incluso con recuento sobre el terreno, y muchos nunca figurarán en ningún sistema de registros fiscales o administrativos. Sería conveniente introducir ciertas distinciones en la cobertura de las encuestas infrecuentes, las encuestas anuales y las encuestas de frecuencia superior al año. Si bien las primeras deberían, en principio, abarcar todos los establecimientos minoristas, mayoristas y de servicios conexos, independientemente de su tamaño, fuesen o no reconocibles desde el exterior, tuvieran o no ubicación fija, no resulta práctico intentar tal cosa en las encuestas anuales y de mayor frecuencia. Tales encuestas, para ser viables, deberían limitarse a unidades con ubicación fija, reconocibles desde el exterior, al menos en el caso del comercio al por menor y los servicios. La reconocibilidad es probablemente demasiado restrictiva en el caso del comercio al por mayor, y puede ser práctico establecer guías adecuadas sin hacer cada vez inventarios sobre el terreno. Las pequeñas unidades de minoristas y servicios requerirán, sin embargo, cada vez ese tipo de inventarios en las áreas tomadas como muestra.

104. Los principales tipos de problemas que afectan a la cobertura son: a) establecimientos no reconocibles, en especial los que funcionan en hogares; b) establecimientos itinerantes, es decir, aquellos que se desplazan, sin lugar o actividad fijos; y c) tenderetes temporales que pueden estar abiertos sólo ciertos días de la semana o épocas del año. El director de la encuesta ha de decidir si, en su caso particular, el esfuerzo necesario para identificar y contar tales establecimientos está justificado o si puede legítimamente recurrir a mecanismos tales como arbitrarias disminuciones del tamaño de las muestras de área para reducir el problema.

Reducciones del tamaño

105. Casi siempre, los tres tipos de establecimientos mencionados serán sin duda de dimensiones pequeñas. Por ello, cabe suponer que puede determinarse un tamaño por debajo del cual quedará la gran mayoría de los mismos. Si se fija la norma que todos los establecimientos por debajo de ese nivel quedarán omitidos de la encuesta y que no se intentará encontrar o contar establecimientos no reconocibles, itinerantes o temporales, entonces el universo puede definirse como todos aquéllos de tamaño superior al nivel fijado. Claro es que cuando una proporción apreciable de las actividades de un país se lleva a cabo por esos establecimientos menores, el efecto sobre las estadísticas puede ser muy pronunciado. Un país puede así decidir cubrir todos los lugares reconocibles de actividad, con independencia del tamaño. Tal cosa debería indicarse claramente en el informe, al igual que debería hacerse con cualquier supresión de tamaños que se efectuara. En ambos casos, como guía del usuario de datos, debería hacerse algún esfuerzo para evaluar la cuantía de la cobertura perdida.

Muestreo por áreas

106. En los tipos de actividades que tienden a distribuirse en proporción a la población, como el comercio al por menor, los establecimientos de comidas y bebidas y muchos servicios, es posible diseñar una muestra eficiente por áreas que reduzca el número de pequeños establecimientos con los que tendrá que establecerse contacto. Esto puede permitir dedicar más esfuerzo a identificar y enumerar los establecimientos que son más difíciles de encontrar, pero el uso del muestreo de área reducirá mucho la cantidad de los datos sobre tipo de actividad y ubicación que pueden publicarse a niveles razonables de varianza muestral.

Problemas especiales en los recuentos sobre el terreno

107. Las dificultades específicas citadas hasta ahora -lugares no reconocibles, vendedores itinerantes y puestos temporales- resultan especialmente significativos en un recuento sobre el terreno. Las siguientes dificultades adicionales pueden afectar a la cobertura cuando se usa ese tipo de procedimiento:

a) Lugares cerrados en el momento del recuento. Cuando un negocio no está abierto al hacer el recuento, el encuestador deberá averiguar si se halla cerrado temporalmente, se ha retirado de esa actividad o si está abierto solamente durante ciertas horas o ciertas épocas del año. Obtener informes para ese tipo de establecimiento puede resultar muy difícil. En algunos casos, como en zonas de recreo o en hoteles de temporada, este problema puede ser especialmente importante;

b) Empresas que funcionan en establecimientos fuera del campo de aplicación. Puede haber una tendencia a dejar al margen actividades que deben incluirse pero que están ubicadas en establecimientos que no corresponden al comercio de distribución y los servicios, tales como museos, estaciones de ferrocarril, instituciones educacionales, edificios

de apartamentos o de oficinas, etc. Las actividades pertinentes que se desarrollan en esos lugares deberían incluirse y contarse sólo si están abiertas al público en general y dirigidas por particulares;

c) Departamentos arrendados o cedidos. Es fácil dejar de lado actividades que se ejercen en departamentos arrendados o cedidos, tales como un servicio independiente de reparación de relojes en una joyería, una peluquería en unos grandes almacenes, un servicio de carnicería en una tienda de ultramarinos. Esto puede afectar apreciablemente a la cobertura en ciertos tipos de actividades e incluso conducir a una infravaloración general de las ventas totales al por menor;

d) Espacio de oficina compartido. Para reducir los costos de funcionamiento, hay empresas que alquilan a otras "espacio de oficina". Puede representar una proporción apreciable de la actividad desarrollada por agentes de exportación, negociantes de granos, agentes de importación, corredores de alimentos, etc.;

e) Establecimientos que venden por correspondencia y operadores de máquinas vendedoras. Esos tipos de establecimiento no necesitan atraer clientes al lugar de su actividad y, por lo tanto, el recopilador puede tender a dejarlas de lado. Los establecimientos de venta por correspondencia se ponen en contacto con sus clientes mediante publicidad directa o anunciándose en revistas y periódicos; los operadores de máquinas vendedoras tratan directamente con los establecimientos donde colocan sus máquinas. Se necesita un esfuerzo especial para estar seguro de que esos tipos de empresas quedan cubiertos.

Problemas especiales en la enumeración por correo

108. En la enumeración por correspondencia, el uso de una guía, especialmente cuando se obtiene de registros administrativos, evita algunos de los problemas que se encuentran en el recuento sobre el terreno pero presenta por sus características algunas deficiencias:

a) Unidades dentro del campo de aplicación de empresas de fuera de ese campo. Algunas de las empresas mayores tendrán muchas dependencias, algunas de las cuales caerán dentro del campo de la encuesta de la distribución y los servicios y otras quedarán fuera. Por ejemplo, en los registros administrativos la empresa puede clasificarse como un ferrocarril y pensarse así que no está dentro del campo de aplicación. No obstante, esa empresa puede también tener y explotar hoteles y restaurantes. Con el incremento de empresas de actividades múltiples y de las multinacionales, es importante identificar cada unidad de una compañía que quede dentro del campo de la encuesta. Antes de la enumeración efectiva, por tanto, debería hacerse un inventario previo de todas las grandes empresas para determinar si alguno de sus establecimientos cae dentro del campo de aplicación;

b) Unidades que quedan fuera del campo de aplicación pero que pertenecen a empresas que están dentro de ese campo. Igual que en el caso anterior, es importante excluir de la encuesta aquellas dependencias de empresas incluidas en el campo de aplicación que quedan fuera del campo del censo, como puede ser una instalación industrial propiedad de una organización de ventas al por menor. Aquí también, el modo de conseguir tal cosa es mediante un inventario previo de las grandes organizaciones;

c) Limitaciones de los registros administrativos. Como la jurisdicción administrativa de un sector puede limitarse a empresas de un cierto tamaño o con características especiales, la cobertura no puede ser tan completa como la que se necesita para los fines de la encuesta. Además, la clasificación por tipos de actividad puede no ajustarse con exactitud a los criterios de la CIIU. Debería estudiarse el efecto sobre las estadísticas resultantes, y si se decide usar registros administrativos en una encuesta por correspondencia, en la publicación final debería hacerse referencia a cualquier insuficiencia apreciable en la cobertura.

Actividades de la administración y de entidades sin fines de lucro

109. En muchos países buen número de las transacciones que se celebran en un tipo de actividad determinado se llevan a cabo por servicios de la administración o por entidades sin fines lucrativos. Esas organizaciones operan en el mercado de un modo similar al de las empresas privadas como, por ejemplo, en el caso de cooperativas de productores agrícolas para la comercialización de leche o en el de la venta al por menor de cereales o de licores por organismos estatales. Para obtener un cuadro completo del comercio de distribución, será menester abarcar esas actividades juntamente con las que opera el sector privado.

Clasificación

110. Como se ha dicho antes, en la CIIU se hace un examen muy completo y detallado de los principios de la clasificación. Además, contiene un análisis de los diversos tipos de unidades estadísticas y de información con las que pueden recogerse apropiadamente los datos. Sería inútil repetir ese examen. Hay que señalar, sin embargo, que los principios de clasificación y de identificación de las unidades estadísticas y de información han de aplicarse teniendo en cuenta la situación que exista en el país donde se haga la encuesta, y las preguntas que busquen la información deseada han de prepararse con sumo cuidado.

VI. CUESTIONARIOS Y OTROS FORMULARIOS PARA LA RECOGIDA DE DATOS

111. Como ya se dijo antes, en las recomendaciones internacionales sobre estadísticas del comercio de distribución y los servicios se presentan en detalle los datos que deberían incluirse y se definen sus características. La elección de características y prioridades se basa en la utilidad relativa de las estadísticas. Se reconoce que las necesidades y las circunstancias variarán de un país a otro y que en casos concretos tendrán que hacerse algunas modificaciones. El análisis que se hace a continuación se limita a algunos de los problemas que cabe encontrar al aplicar las definiciones recomendadas y al preparar el cuestionario y otros formularios que serán útiles para realizar las encuestas.

Información necesaria para control, correspondencia y evaluación

112. Todo formulario de recogida de datos debería pedir que se indicara el nombre y la dirección postal de la empresa propietaria además de la ubicación geográfica del establecimiento, para que pueda enviarse por correo a la empresa de que se trate cualesquiera preguntas que requieran una aclaración por escrito después de que se haya recibido y preparado para su publicación el cuestionario cumplimentado. (La dirección postal suele ser diferente de la ubicación). Cuando se usa una guía para preparar el registro de direcciones, suele ser sumamente útil incorporar códigos que indiquen el tipo de actividad y el tamaño en la etiqueta de la dirección. Tal información cumple dos propósitos: primero, permite la evaluación de quienes responden tarde o no lo hacen y, segundo, sirve para indicar la intensidad del esfuerzo complementario que se requiere para compensar los informes incompletos o los que fallan. (Una pequeña empresa sin empleados evidentemente necesita menos esfuerzo complementario que una empresa con 100 empleados.)

113. Además, la lista de direcciones debería contener un único número de identificación para cada establecimiento que pueda usarse como número de control para cotejo de informes, control de quienes no responden, archivo de los formularios contestados y fines generales de control.

Cuestiones sobre publicación y preparación

114. Es obvio que en el formulario que se use para recabar la información han de figurar preguntas sobre los datos que se quiere publicar. Lo que no es tan obvio es que también deben figurar otras preguntas destinadas a lograr que los datos se notifiquen y elaboren adecuadamente. A continuación se señalan algunos ejemplos:

a) Si los impuestos indirectos y sobre las ventas se recaudan por los minoristas y si se desea incluirlos como parte del total de las ventas al por menor, no siempre será posible obtener la respuesta correcta pidiendo simplemente al encuestado que incluya esos gravámenes. El método más

seguro sería decirle que los incluyera, luego preguntarle si han sido incluidos y entonces, si la respuesta es "no", preguntarle cuánto transfirió a los servicios fiscales. La persona que prepara los datos contará entonces con toda la información que necesita para llegar a la respuesta correcta;

b) Frecuentemente, más de una empresa puede estar trabajando con un solo nombre en una sola ubicación; por ejemplo, unos grandes almacenes que ceden a terceros la sección de perfumes o la de zapatos. Será necesario preparar una pregunta que amplíe ese tipo de información y obtenga la suficiente para permitir su tratamiento apropiado;

c) La inclusión de preguntas "autoaclaratorias" puede ser muy útil. Por ejemplo, cabe pedir al encuestado que coteje su nómina e ingresos para ver si son compatibles. O puede pedírsele que compruebe la coherencia entre los totales que deben figurar en más de un lugar del formulario -una vez, quizá, en la cifra total de ventas y una segunda vez en el desglose de las mismas por tipo de mercaderías. Esos contrastes deberían ser siempre sencillos para que pudieran hacerse con rapidez;

d) Para encuestas anuales e infrecuentes, las recomendaciones internacionales indican una preferencia para que se dé información correspondiente al año del calendario. No obstante, si lo normal en casi todos los establecimientos es el año fiscal, los datos pueden recopilarse con carácter uniforme sobre esa base. Si bien el cuestionario puede indicar qué método se prefiere, será necesario pedir al encuestado qué método usó en realidad y, para quienes empleen el año fiscal, también se necesitará información sobre el período a que corresponde, ya que no todas las empresas usarán el mismo. La aceptación de información para un período distinto del que se solicita debería examinarse a la luz de sus efectos sobre los resultados y, si éstos son apreciables habrá que procurar obtener información fundamental, tal como empleo o ventas, para el período requerido, información que puede entonces usarse para prorratar otras características significativas de los datos;

e) En las encuestas de frecuencia más que anual surge un problema especial cuando la empresa sólo trabaja parte del año, como en el caso de un hotel de temporada o un establecimiento situado en un lugar de veraneo. Con frecuencia no es posible distinguir un establecimiento de esa índole de otro que haya abandonado la actividad o de un centro que se traspasa, a menos que el encuestador obtenga respuestas a una serie de preguntas. La omisión de establecimientos de temporada o provisionalmente cerrados puede producir un acusado sesgo a la baja.

Distinción entre las anotaciones de no información y anotaciones nulas

115. Un problema que se da a menudo en la elaboración de los formularios cumplimentados es el que se suscita cuando el encuestado deja de contestar a un punto en el que la respuesta puede ser "cero" o "no procede". Quien elabora los

datos se encuentra así con el dilema de determinar si se trata de un caso de no respuesta o si quien contestó no hizo una anotación porque pensó que no era necesaria. Si bien conviene incluir unas instrucciones que digan a quien responde que escriba "cero", "no procede", o " - " en todos los casos apropiados, muchos no prestarán atención. Es posible reducir las consecuencias de ese problema pidiendo que se responda a la pregunta con un simple "sí-no" y usando la respuesta para decidir si procede hacer una anotación de "cero" o si se trata de un caso de no notificación de datos.

Diseño de la pregunta sobre tipo de actividad

116. La presentación de datos por tipo de actividad es uno de los objetivos básicos de las encuestas del comercio de distribución y los servicios. Para llegar a una clasificación adecuada con arreglo a las normas establecidas en la CIIU hay que aplicar ciertos criterios objetivos que muy frecuentemente necesitan que se dé respuesta a un conjunto uniforme de preguntas. Hay algunos tipos de actividad, tales como peluquerías, estaciones de radio o productores de películas, donde la autodenominación puede resultar aceptable, pero en muchos casos será menester sondear más profundamente para obtener una clasificación correcta. Por ejemplo, "actividad en el sector del vestido" es insuficiente para la clasificación, ya que puede indicar fabricación, venta al por mayor, venta al por menor o reparación de prendas de vestir. Asimismo, quien responde puede considerarse como un minorista cuando, según la CIIU, sería considerado como un mayorista. La información solicitada sobre el tipo de actividad, por lo tanto, ha de incluir preguntas que permitan clasificar el establecimiento en primer lugar dentro del comercio, la fabricación o alguno de los servicios. Si se trata de un comerciante, ¿se dedica principalmente al comercio al por menor o al por mayor? Si está en el sector de los servicios, como por ejemplo reparaciones, ¿se trata de su actividad fundamental? Si se dedica al comercio y el programa de publicación de datos requiere la tabulación con arreglo a un desglose detallado de actividades, será necesario hacer preguntas adicionales sobre la distribución de ventas, por tipo de mercancías vendidas.

117. Si se dispone de una clasificación preliminar, gracias a alguna fuente, será posible preparar preguntas específicas sobre tipo de actividad que se ajusten a la índole general de los establecimientos. Así, podrá hacerse una pregunta especial a las empresas que se dedican al negocio alimentario, o al del vestido, etc. Tal cosa permite usar preguntas y terminologías pertinentes y hace que el cuestionario cobre más sentido para quien ha de contestarlo. También permite una clasificación con una división más detallada de la actividad y recabar información concreta que se adecúe al tipo de actividad general, como, por ejemplo, la venta de alimentos congelados en tiendas de ultramarinos.

Localización de información consolidada

118. A causa de la dificultad que encierra saber, antes de hacer una encuesta por correo, cuáles son las empresas con más de un establecimiento, se necesita algún método para descubrir tales empresas en el momento de la enumeración. Con

tal fin, en el formulario donde se pide la información debería pedirse la lista de todos los lugares donde la empresa desarrolla una actividad e información básica sobre cada uno de ellos, tal como tipo de actividad (con la denominación que indique la propia empresa), volumen de ventas, nómina anual y número de empleados. Como probablemente las empresas mayores habrán sido identificadas previamente y se les habrá pedido datos por separado para cada lugar de actividad, el número y tamaño de las empresas con varios establecimientos que se agreguen en el momento de la enumeración deberían resultar pequeños.

Perforación del formulario

119. A cada característica de los datos o a cualquier otro tipo de información debería asignarse un código en forma de clave que se tabulará o usará para la clasificación y que es también necesario para la identificación del establecimiento. El tipo de código que conviene emplear puede variar con la índole del equipo que se emplee (tal como máquina de calcular, equipo de fichas perforadas o computadora). Una asignación adecuada de los códigos puede suponer una diferencia apreciable en el costo y calidad del trabajo y debería hacerse sabiendo cómo se elaborarán cada una de las características de los datos. Por ejemplo, cuando se vaya a usar equipo de ficha perforadas, habrá que tener cuidado de conseguir en todo lo posible que aquellas características que vayan a aparecer en el mismo cuadro se incluyan en la misma ficha.

Importancia de la codificación previa

120. Se ahorrará mucho tiempo y dinero si se examina cada característica que figure en el formulario para ver si es posible diseñar las preguntas de tal modo que puedan imprimirse en dicho formulario directamente códigos preasignados en lugar de que un empleado administrativo tenga que asignarlos luego. Tal cosa presenta la doble ventaja de hacer innecesario que el empleado deba dedicar tiempo a determinar y registrar el código y facilita que el perforador encuentre y lea el código impreso previamente en lugar de depender de la escritura a veces ilegible del codificador.

Tratamiento de las instrucciones

121. Cuando los cuestionarios son completados por encuestadores sobre el terreno, las instrucciones suelen figurar en un manual, y puede llevar tiempo capacitar a ese personal y someterlo a prueba para tener la seguridad de que las instrucciones han sido entendidas. No obstante, si el encuestador tiene que dejar el cuestionario a la persona que ha de cumplimentarlo o si la encuesta se hace por correo, entonces las instrucciones que figuren en el formulario o que lo acompañen deben buscar que se suministre la información requerida. Cuando se use una combinación de trabajo de campo y de correspondencia, puede introducirse un sesgo si las instrucciones usadas por el encuestador son diferentes de las que figuren en el formulario que se envía por correo. Cuando las personas que se ocupan del censo, por un lado, y de programas anuales o de frecuencia superior, por el otro, sean diferentes, habrá que tener especial cuidado en

coordinar las instrucciones. La falta de esta coordinación puede fácilmente llevar a una no comparabilidad de los datos de ambos programas.

122. Las instrucciones destinadas a las personas encuestadas deberían comprender, de forma concisa pero completa, todo lo necesario para explicar lo que se pide y evitar equívocos. Es probable que unas instrucciones largas y difusas se lean y entiendan menos que unas breves y precisas. Si se requieren instrucciones amplias, sin embargo, una práctica común es incluir sólo los puntos más esenciales en el propio cuestionario y el resto en una hoja complementaria de instrucciones. Entrevistas anteriores a la encuesta con encuestados potenciales a menudo brindan indicaciones valiosas sobre los puntos que deberían destacarse en las instrucciones.

123. La siguiente es una lista de comprobación de puntos que normalmente figuran en las instrucciones de los cuestionarios o en hojas complementarias:

1. Autoridad legal para hacer la encuesta
2. Carácter confidencial de la información
3. Tipos de establecimientos que tienen que informar
4. Aceptabilidad de estimaciones cuando no se disponga de registros contables
5. Fecha en que ha de mandarse la información completa
6. Cómo devolver el formulario cumplimentado
7. Período cubierto por la encuesta y normas sobre el tratamiento de datos del año fiscal
8. Definiciones de las diversas características de la información solicitada
9. Instrucciones sobre la recopilación de datos especiales, por ejemplo, porcentaje de ocupación de hoteles o asistencia a teatros
10. Instrucciones específicas para sectores particulares donde son de esperar problemas que afecten a muchas personas de las que tienen que responder, tales como tratamiento de departamentos explotados por terceros en grandes almacenes.

Problemas derivados de la diversidad de actividades

124. La gran diversidad de actividades incluidas en el ámbito del comercio de distribución y los servicios crea dificultades, ya que lograr una clasificación adecuada requiere responder a muchas preguntas, algunas de las cuales sólo serán pertinentes para una pequeña proporción del universo. Además, existen preguntas especiales sobre datos que se aplican a tipos de actividad específicos. Si sólo se usara un único formulario, su longitud y complejidad inevitables confundirían e irritarían a quienes han de contestarlo y probablemente reducirían mucho el número de respuestas. El modo de resolver este problema debería ser diferente en las encuestas que se hacen directamente y en las que recurren al correo. Se supone, sin embargo, que han sido diseñados cuestionarios diferentes, cada uno de los cuales será usado por los establecimientos a los que se dirigen preguntas idénticas, por ejemplo, tiendas de alimentación, tiendas de ropa de vestir, lugares de esparcimiento, mayoristas de petróleo, etc.

125. En las encuestas que vayan acompañadas de entrevistas sobre el terreno (ya se trate de censos completos o de una zona de muestreo), en la entrevista inicial debería figurar una serie de preguntas encaminadas a determinar si el establecimiento es una industria, un comercio al por mayor, un comercio de detalle o un centro de prestación de servicios y fijar dentro de cada uno de esos grandes sectores qué formulario debería completarse para obtener la información. El encuestador puede dejar entonces ese formulario para que se rellene y devuelva a la oficina de estadística, o la hoja con la lista de establecimientos puede usarse como base para que dicha oficina envíe el formulario oportuno a cada establecimiento.

126. En las encuestas que se hacen por correo, será menester contar con información suficiente en los registros de correspondencia para enviar el cuestionario adecuado. Los formularios deberían diseñarse para que se ajusten al tipo de información disponible. Por ejemplo, si los registros existentes no permiten diferenciar entre mayoristas y minoristas, el cuestionario tendrá que abarcar ambas actividades. Así, el número de formularios y su complejidad serán función de la información que cabe usar en la correspondencia. Se suscitará un problema especial cuando el registro usado para los formularios enviados por correo comprenda algunas empresas para las que no se disponga de clasificación. En tales casos, se necesitará un formulario especial diseñado para cubrir no sólo establecimientos que queden dentro del campo de aplicación sino también para identificar a los que no les sucede tal cosa. No obstante, sólo debería pedirse datos al primer tipo de establecimiento. También puede ser necesario hacer un envío por correo complementario para aquellos establecimientos donde los datos obtenidos en la primera encuesta sean inadecuados. Esto es lo que ocurrirá si el formulario diseñado para el contacto inicial es demasiado largo y complicado por figurar en él encuestas especiales destinadas a todos los tipos de negocio. El segundo contacto puede limitarse a aquellos establecimientos lo bastante grandes como para tener un efecto apreciable en los datos. Para los de dimensiones menores, la información que falte puede evaluarse juntamente con los demás datos no recibidos.

Diseño de los formularios

127. Los formularios para recabar información son los documentos fundamentales de las encuestas donde se registran todos los datos que serán elaborados para preparar la tabulación y publicación de las conclusiones de la encuesta. Su formato, organización y contenido tendrán gran influencia en la calidad de los resultados estadísticos, el costo de la recogida de datos, la preparación y tabulación y la rapidez de su publicación. Frases ambiguas, erratas e incluso la falta de espacios suficientes han hecho que algunos países abandonen encuestas importantes o gasten recursos escasos en un esfuerzo para corregir los efectos de esos errores. En la autoenumeración, cuando quien contesta cumplimenta un formulario sin la asistencia de un entrevistador, incluso una mala elección del tipo de letra ha sido reconocida como causa de interpretaciones defectuosas muy generalizadas. Esto puede también ocurrir cuando un título clave, importante para comprender lo que se dice debajo, no se destaca suficientemente y por lo tanto es

dejado de lado por el encuestado. Los ensayos previos de los formularios pueden a menudo poner de manifiesto tales deficiencias a tiempo para corregirlas. Esas pruebas previas no tienen que ser complejas o estar diseñadas científicamente. Si se proponen formatos alternativos, tales ensayos pueden establecer cuál resulta mejor.

128. Un uso adecuado del espacio también es importante para un buen diseño de los formularios. Con demasiada frecuencia, el espacio que se deja para dar respuesta a una pregunta es inadecuado, lo que hace que quien contesta piense que el organismo encargado de la encuesta no sabe lo que hace. Por otra parte, demasiado espacio hace al cuestionario más largo de lo necesario, y la persona que responde puede creer que lleva más trabajo rellenarlo de lo que en realidad es. Si se espera que muchos cuestionarios se rellenen a máquina, los espacios entre las líneas deben corresponder al espaciado de la mecanografía. Además, el cuestionario debería ser examinado teniendo en cuenta el espacio necesario para las anotaciones de la oficina. Las que se hacen "para uso oficial solamente", son de tres tipos generales:

- a) Datos de control, tales como número del establecimiento, identificación de quienes elaboran los informes y fechas en que se dan los diversos pasos;
- b) Números de código, tales como los de ubicación, tipo de actividad, tamaño del establecimiento y tipo de propiedad. Cuando las preguntas estén preparadas para ser codificadas previamente, tal cosa no será menester;
- c) Cifras calculadas, tales como número medio de personas ocupadas, promedio de ingresos, etc.

129. A causa de que frecuentemente hay que cambiar nombres o direcciones, el espacio de estas últimas debe dejar margen suficiente para ello. Además, en cada cuestionario debería figurar espacio adecuado para que quien contesta haga las observaciones que estime oportunas con el fin de explicar cualquiera de sus respuestas.

Secuencia y redacción de las preguntas

130. Siempre que sea posible, conviene situar las preguntas fáciles de contestar al comienzo del formulario. Si quien contesta se siente desalentado tan pronto como empieza a responder el cuestionario se sentirá más inclinado a dejarlo de lado. Igualmente, las preguntas deberían disponerse de tal modo que aquellas que se sigan lógicamente unas a otras figuren juntas. Cuando un cuestionario comprende algunas preguntas especiales para sólo una parte de quienes lo reciben, tales preguntas deberían desplazarse hacia el final. Cada parte del cuestionario debería numerarse para facilitar la referencia a la misma en las instrucciones, en la correspondencia con los entrevistados y en la redacción de los procedimientos del servicio para tramitar los informes. Cuando se use más de un tipo de formulario, debería hacerse todo lo posible para conservar el mismo número en cada partida (por ejemplo, nómina y empleo), incluso si ello obliga a que en algunos

formularios tenga que figurar el número de una partida donde no se pidan datos y lleve sólo la observación "no se aplica a este formulario". Tal cosa simplificará mucho los procedimientos administrativos y la preparación de las publicaciones.

131. Aunque las instrucciones son importantes y deberían escribirse con cuidado, todavía mayor atención debe prestarse a la redacción de las preguntas. Un factor importante para lograr respuestas prontas, completas y fiables es que las preguntas estén redactadas de tal modo que sean fácilmente entendidas por quienes han de responderlas y por los entrevistadores sobre el terreno. Parte de quienes contestan dejarán de leer las instrucciones separadas, pero probablemente todos ellos leerán las preguntas antes de anotar las respuestas. Por lo tanto, debiera hacerse todo lo posible para conseguir una redacción clara y concisa de las preguntas y evitar ambigüedades. Si dos interpretaciones de una pregunta son igualmente probables y sólo una es correcta, la mitad de las preguntas serán inaceptables.

132. Para algunas partidas, el uso de una respuesta de elección múltiple donde quien contesta simplemente marca el cuadro correcta es preferible a una pregunta que obligue a que el encuestado tenga que hacer una anotación. Esas preguntas pueden usarse para recabar información sobre clase de clientes, método de distribución de los productos, forma jurídica de la empresa, etc. Cabe también emplearlas para obtener datos cuantitativos si las respuestas en términos de clases de tamaño son aceptables. Para estar seguro de que no se omite ninguna, conviene incluir una categoría de "todos los demás", lo que hace que quien contesta haga una anotación cuando opine que no le corresponde marcar uno de los cuadros disponibles. Esas anotaciones pueden revisarse por un empleado para determinar qué cuadro debería haberse usado. El valor de este tipo de pregunta donde la respuesta es un cuadro que hay que marcar aumenta por el hecho de que al lado de cada uno de ellos pueden figurar códigos preasignados, evitando así la operación administrativa de codificación.

133. Por último, hay que mencionar que el uso de colores y sombras puede mejorar las respuestas, la perforación y la elaboración general al hacer a la persona que lee el informe muy consciente de donde tienen que hacerse las anotaciones y también de que un espacio en blanco significa que una pregunta no ha sido contestada. Sistemas tales como flechas e indicadores también son útiles en este sentido. La fijación de un espacio separado para cada dígito en una encuesta de valores puede ser igualmente útil, porque requiere que quien contesta haga una anotación más cuidadosa; esto hace que el formulario sea más sencillo de elaborar por el perforista o el empleado administrativo. El diseño adecuado de todo ello puede también conducir a un redondeo de cifras previo, lo que no sólo elimina una operación administrativa sino que también reduce las posibilidades de error cuando el empleado realice ese cometido.

VII. RECOGIDA DE DATOS

Observaciones generales

134. Al decidir qué sistema se va a usar para reunir los datos, es preciso analizar los problemas especiales relacionados con el universo que hay que cubrir, los tipos de información que se quiere obtener y la disponibilidad de fondos, personal capacitado y equipo especial. Para el comercio de distribución y los servicios hay varias características que deben tenerse presentes al diseñar las encuestas. En muchas de esas actividades hay gran número de pequeños establecimientos. Frecuentemente, sus propietarios son analfabetos y tienen gran dificultad para entender conceptos, incluso cuando se use una enumeración personal. Además, tienen tendencia a no llevar registros y, por tanto, los datos del período de que se trate -año, trimestre o mes- son estimaciones tanto peores cuanto mayor sea el período de tiempo cubierto. A menudo, los establecimientos de esa índole funcionan en los domicilios de sus propietarios o en lugares que no son reconocibles o bien las actividades son realizadas por personal ambulante. Tienen también tendencia a ignorar las peticiones que reciban por correo. En razón del gran número de establecimientos que hay, la enumeración personal resulta costosa.

135. También importa considerar, en el diseño de las encuestas, el hecho de que, especialmente en el comercio al por menor, gran parte de la actividad puede llevarse a cabo por establecimientos de empresas que cuentan con muchas sucursales. Como tales empresas suelen tener una oficina central donde se reúnen datos de cada uno de sus establecimientos, es mejor obtener los datos necesarios en esa oficina que enviar entrevistadores a cada sucursal (donde incluso parte de la información necesaria puede no estar disponible) o formularios separados por correo a cada establecimiento. Si una lista de casi todas las empresas de esa índole puede hacerse en los primeros pasos de la planificación de los censos infrecuentes, cabe efectuar un inventario previo especial para obtener una lista completa de todas las ubicaciones (incluidas las auxiliares) que comprendería suficiente información clasificada para permitir el uso del formulario correcto en el momento en que se haga el inventario efectivo. Todos los formularios recabando la información pueden entonces enviarse a la oficina central. Tal cosa presenta la ventaja de que permite a la empresa coordinar toda la información que hay que suministrar sobre sus establecimientos y resolver cualesquiera contradicciones en la comunicación de datos sobre diversas sucursales. En todo caso, la oficina central preferirá frecuentemente supervisar la contestación de los cuestionarios y transmitirlos cuando estén completados. Cuando se use este procedimiento, sin embargo, habrá que prestar mucha atención a que se establezcan controles que aseguren que la información no se duplica por los encuestadores sobre el terreno o por el envío por correo de formularios directamente a los establecimientos. Los entrevistadores pueden recibir una lista de empresas o de establecimientos específicos sobre los que no hay que recoger información, y las listas utilizadas en la correspondencia han de ser adecuadamente purgadas de toda duplicación. Cuando las encuestas se hacen después de un censo general, las listas de las empresas con muchos establecimientos pueden obtenerse de los registros censales y usarse como una lista patrón por los encuestadores que hacen el trabajo de campo.

136. Otro factor que afecta al diseño de encuestas es la gran diversidad de actividades y tipos de establecimientos que hay que en la distribución y los servicios. Tanto los que se dedican al comercio al por mayor como los minoristas se caracterizan por transferir bienes, pero es esencial disponer de información sobre la clase de cliente para determinar el tipo de actividad. Además, el desglose detallado de esta última vendrá determinado por diferentes combinaciones de los tipos de bienes vendidos. En los servicios, la gama de actividades es extremadamente amplia y comprende cosas tan diversas como servir una comida, redactar un anuncio, producir una película, alquilar una máquina, reparar un reloj, etc. Esto requiere que el sistema que hay que diseñar permita seleccionar el formulario adecuado en las encuestas por correo o que el entrevistador sea formado cuidadosamente para que sepa determinar bien el tipo de actividad. También debe hacerse lo necesario para detectar y corregir errores cuando se han usado formularios equivocados.

Métodos de enumeración

137. Hay tres métodos básicos que pueden usarse para obtener los datos necesarios en las encuestas. El primero es el uso de entrevistadores sobre el terreno que visitan cada lugar de trabajo. Si el establecimiento es pequeño y el cuestionario lo bastante simple para que las respuestas puedan darse durante la visita, el recopilador puede completar el cuestionario en ese momento. Si por algún motivo no puede rellenarse en el momento de la visita, porque no está la persona que puede facilitar la información o porque los datos solicitados son demasiado complejos, el cuestionario puede dejarse para que se envíe por correo o ser recogido por el entrevistador en una visita posterior.

138. El segundo método básico es de cuestionarios enviados directamente por correo a quienes los han de contestar. Este método sólo puede usarse cuando es posible preparar una guía fidedigna de los establecimientos mercantiles y cuando el servicio postal es eficiente. Los cuestionarios pueden ser devueltos también por correo una vez cumplimentados, o cabe recurrir a recopiladores que visiten los establecimientos para recoger los formularios.

139. Un tercer método básico es el que cabe usar cuando un país cuenta ya con un sistema de registros administrativos bastante perfeccionado en el que los datos que se necesitan para la encuesta se recogen en el transcurso de la realización de diversos programas nacionales, tales como impuestos, seguridad social, paro y otros análogos, y existe en el sistema administrativo un método para identificar inmediatamente al establecimiento.

140. En la práctica, cabe combinar los tres métodos, según las condiciones particulares que existan en el momento de la encuesta en el país de que se trate. Así, es posible preparar una guía de grandes empresas a las que pueden enviarse directamente por correo los formularios solicitando información y usar esa misma lista como guía para los recopiladores sobre el terreno que se encargarían de recoger todos los demás cuestionarios.

Enumeración sobre el terreno

141. En los países que se proponen por primera vez hacer un censo general del comercio de distribución y los servicios, sería muy útil proceder a un inventario completo sobre el terreno donde se confeccione un registro en forma de lista con todos los establecimientos mercantiles localizados, ya figuren o no dentro del campo de aplicación. Tal lista puede ser muy valiosa para fijar los marcos del muestreo. También es necesaria una completa enumeración sobre el terreno cuando no sea posible obtener una guía fiable que pueda usarse para el envío por correo de los cuestionarios. En tal caso, cabe también usar un muestreo de área para partes del país que sólo tienen una pequeña actividad comercial, siempre que la fijación de tales zonas sea lo bastante detallada como para permitir una identificación adecuada. En todo caso, el empleo de la enumeración sobre el terreno requiere un esfuerzo intensivo para establecer un conjunto detallado de mapas que se usarán para controlar las asignaciones a los encuestadores.

142. Si bien una enumeración sobre el terreno de la distribución y los servicios presenta muchas características similares con un censo de población, hay diferencias que hacen que las encuestas del comercio de distribución y los servicios resulten mucho más difíciles y requieran un tipo de encuestador más capacitado. Dicho sea de paso, es posible combinar un censo de población con uno de la distribución y los servicios y ahorrar así en principio gran cantidad de dinero en mapas, adiestramiento de recopiladores y costos generales. No obstante, la complejidad de tal operación y los diferentes requisitos de tiempo para la terminación conducirían a complicar mucho la operación y llevar al colapso de ambos censos.

143. La enumeración sobre el terreno de los establecimientos del comercio de distribución y de servicios es muy compleja y requiere controles muy estrictos y una minuciosa capacitación de supervisores y encuestadores. Si bien no es esencial contar con personal que tenga experiencia mercantil, la categoría de las personas contratadas para esta labor tiene que ser tal que pueda comprender los problemas de clasificación. Si no pueden captar el concepto de campo de aplicación o si tienen un sentido escaso de los números, tales personas pueden crear grandes dificultades. Hay que tomarse tiempo para formar adecuadamente encuestadores tanto en la mecánica de la enumeración como en los objetivos de la encuesta. Tienen que tener conciencia de las consecuencias de las acciones que tomen o no tomen. Una vez tras otra ha habido necesidad de volver a recontar partes sustanciales de un censo porque las instrucciones se habían entendido mal. Si hay restricciones en los fondos disponibles, sería mejor reducir el alcance de la encuesta (y, por tanto, el número de encuestadores necesarios) más bien que la calidad del recuento.

144. Al preparar manuales de formación y al llevar a cabo la labor de adiestramiento, conviene concentrar la mayor parte posible del trabajo en manos de unas pocas personas de los servicios de planificación. En la práctica, sin embargo, por causa del gran número de gente que tiene que ser formada y la frecuente extensión de las zonas geográficas que hay que abarcar, es menester formar grupos intermedios que capacitarán a su vez a los encuestadores propiamente dichos. Tal cosa conduce a una dilución y distorsión del material original y

debería reducirse al mínimo. Un buen procedimiento es formar primero personas que después se convertirán en supervisores sobre el terreno y hacer entrevistas reales supervisadas por el personal de planificación para comprobar hasta qué punto esas personas han asimilado el material de formación. Luego ellas, a su vez, pueden usar ese mismo procedimiento para probar a los recopiladores sobre el terreno. Las clases teóricas no pueden sustituir el verdadero trabajo práctico para determinar cómo se ha aprovechado la formación. Incluso aunque el material de capacitación se haya preparado bien, nunca debe darse por supuesto que basta la simple entrega de tal material a los encuestadores para lograr que se haga un buen trabajo. Debe hacerse lo antes posible una revisión del trabajo efectivo del recopilador para que puedan corregirse los errores antes de que el daño sea grande. Como ideal, ese examen consistiría en que el supervisor volviera a recopilar la información en el distrito correspondiente.

145. Desde el punto de vista de la formación, se obtiene alguna ventaja cuando se usan los mismos recopiladores en encuestas recurrentes, como las mensuales y trimestrales. No obstante, incluso en tales casos, conviene preparar un programa ordinario de supervisión de cuando en cuando de las prácticas efectivas de enumeración. Puede resultar difícil obtener una buena plantilla de encuestadores para un censo general, ya que el período de empleo es muy breve (aunque más largo que en el caso de un censo de población). Como esa fase de la labor puede absorber el grueso de los fondos, tal vez haya tendencia a pagar salarios inferiores, lo que hace más difícil atraer personal competente. Cuando exista un programa anual pero no mensual o trimestral, la inversión en el adiestramiento de encuestadores se convierte en una partida muy costosa y puede ser muy pequeño el remanente de personal que queda desde una encuesta hasta la siguiente, ya que transcurre un largo período entre las encuestas sucesivas y los recopiladores tienden a buscar otros empleos.

146. Además de los cursos de instrucción, el encuestador debería disponer en el momento del recuento de un manual que contenga todo el material básico que necesite para saber cómo proceder a una enumeración adecuada. En ese manual deberían figurar todas las cuestiones de procedimiento que afecten a los encuestadores y los problemas principales con que puedan encontrarse. Los temas que deben figurar en el modelo de manual típico son los siguientes:

- a) Normas y reglamentos. Obligación legal de los encuestadores de proporcionar información verídica y de considerar confidencial la información suministrada por quienes respondan. Acciones prohibidas a los encuestadores, tales como peticiones o ventas;
- b) Extractos de la ley o reglamento autorizando el censo;
- c) Procedimientos. Organización del personal sobre el terreno. Cómo obtener ejemplares y transmitir los cuestionarios cumplimentados. Qué hacer si el entrevistado desea que se le deje el cuestionario para rellenarlo o para enviarlo por correo al organismo de estadística. Qué hacer si alguien rehúsa proporcionar la información. Cómo informar sobre establecimientos que cambiaron de propietario durante el año del censo. Cómo preparar informes sobre la marcha de los trabajos;

d) La lista de omisiones, esto es, la lista de establecimientos que el recopilador no tiene que visitar. Puede comprender los establecimientos con sucursales con los que se pondrá en contacto directamente el organismo central de estadística;

e) Campo de aplicación del censo, definición de establecimiento y tipos que hay que incluir. Tipos de establecimientos corrientes que no hay que incluir;

f) Lista de cuestionarios y otros formularios usados en el censo y descripción sucinta del contenido, propósito y alcance de cada uno;

g) Instrucciones detalladas para cada pregunta del cuestionario;

h) Instrucciones especiales para tipos particulares de establecimientos o industrias. Cómo manejar actividades mixtas que se encuentran con frecuencia (por ejemplo, talleres industriales y comercio al por mayor, minoristas y servicios, departamentos arrendados);

i) Empresas con sucursales que no figuren en la lista de omisiones. Cómo identificar establecimientos u oficinas centrales de tales empresas;

j) Índice alfabético del contenido del manual.

147. Al preparar el recuento sobre el terreno habrá que dividir con tal fin cada sección del país en distritos. Para encuestas basadas en el uso de muestras de área, las seleccionadas tendrán que subdividirse igualmente en distritos. Con posterioridad, se preparará un mapa para cada distrito de enumeración para su uso por el recopilador al inventariar el distrito que se le haya asignado.

148. El encuestador tendrá que recoger los datos de cada establecimiento incluyéndolos en un libro que contenga toda la información necesaria para controlar el procedimiento de recuento. Para cada establecimiento, tendrá que tomarse una determinación acerca de si entra o no en el campo de aplicación de la encuesta. Para aquellos que sí entran, se requiere ir al establecimiento y hacer una entrevista. Para reducir el costo de este sistema de recogida de información, convendría dar instrucciones al encuestador de que no enumere lugares que quedan obviamente fuera del campo, tales como consultas de médicos o despachos de abogados, explotaciones agrícolas, etc. No obstante, hay que insistir en la importancia de preguntar directamente en cualquier establecimiento sobre el que haya dudas y hacer la anotación correspondiente.

149. Para controlar el recuento, determinar si el establecimiento entra dentro del campo de aplicación, decidir qué formulario ha de rellenarse y poder identificar al establecimiento adecuadamente, será menester obtener la información siguiente:

1. Nombre y dirección del establecimiento
2. Suficiente información sobre clasificación para determinar si el establecimiento queda dentro del campo
3. Si el establecimiento forma parte de una empresa con sucursales cuya información se obtendrá desde una oficina central
4. Si una vez obtenido el cuestionario, se envía directamente al organismo central de estadística o deberá ser recogido por el recopilador en fecha ulterior
5. Fecha en que el cuestionario se obtuvo
6. Nombre de la persona entrevistada
7. Número de identificación del establecimiento, para fines de control.

150. Como los distritos de la enumeración están designados previamente, será posible establecerlos de tal modo que sus límites nunca se mezclen con divisiones geográficas sobre las que se quiera tabular o publicar datos. Así, si se conoce el número del distrito de recuento, todos los establecimientos a los que se ha asignado ese número pueden geográficamente codificarse en provincias, ciudades u otras zonas.

Establecimiento y empleo de una guía

151. Resultaría ideal si, antes de hacer una encuesta, se dispusiera de una guía de todas las unidades estadísticas que figurarán en la misma, juntamente con información sobre emplazamiento, tipo de actividad y tamaño. Tal cosa definiría el campo que habría que contar; cabría usar información sobre tipos de actividad para preparar cuestionarios adecuados, podría emplearse la información sobre la ubicación material para la asignación geográfica y la información sobre el tamaño se emplearía en las decisiones sobre muestreo y cobertura. La guía permitiría controlar la enumeración, ya se hiciese por correo o mediante entrevistadores sobre el terreno, y serviría de medio para estimar los datos de quienes no responden. Pero como es imposible preparar una guía de esa índole hasta después de que se haya terminado la encuesta, el objetivo del director de ésta será obtener antes la mayor aproximación posible.

152. Como mínimo, en la guía debería figurar la información siguiente:

1. Nombre y ubicación de cada establecimiento
2. Dirección postal, que puede ser diferente de la ubicación geográfica
3. Nombre y dirección de la oficina central o sede de los establecimientos que formen parte de empresas con sucursales
4. Tipo de actividad, descripción o código
5. Código de tamaño, que normalmente debería estar basado en el número de personas empleadas pero que también puede fundarse en el volumen de ventas o ingresos
6. Fuente y fecha de la información

153. A causa del número casi siempre grandes de pequeños establecimientos en la distribución y los servicios y por los muchos cambios en la propiedad, sería muy difícil y costoso preparar y mantener una guía completa. Por ello, algunos países establecen un mínimo de tamaño e incluyen sólo en la guía aquellos establecimientos por encima de determinadas dimensiones. Cuando para cubrir el resto del universo se usan otros mecanismos, tales como una lista completa sobre el terreno u otras fuentes de datos, como registros administrativos o fiscales, tal procedimiento es aceptable. No obstante, si la encuesta se va a basar en un tamaño mínimo que incluya sólo aquellos establecimientos por encima de un nivel, siempre habrá incertidumbre acerca de si figuran todas las unidades. Claro está que si un país ha desarrollado un método para mantener un registro al día de todas las empresas mercantiles con fines administrativos o impositivos, ese registro puede usarse para preparar la información que se necesita para la guía.

154. La utilidad de esta guía depende de la índole de los datos que contenga, su carácter más o menos completo y su precisión. Existen muchas fuentes diferentes para redactar una guía, casi todas ellas con algunas deficiencias. Toda fuente ha de examinarse con cuidado antes de usarse procurando subsanar cualesquiera defectos. Al mismo tiempo, tal vez sea menester sacrificar algo a la totalidad o precisión a fin de mantener los costos de la confección de esa guía en un nivel razonable, pero tales decisiones deberían hacerse conscientemente intentando medir y describir las deficiencias. A continuación se señalan algunas de las fuentes de las guías del comercio de distribución y los servicios:

a) Una lista completa sobre el terreno. Este es probablemente el método más costoso para reunir la información que se necesita en la guía. No obstante, especialmente cuando un país está iniciando un programa de estadísticas económicas, es sin duda el más útil. Recopiladores capacitados pueden buscar cada lugar físicamente reconocible de actividad mercantil y reunir la información necesaria mediante entrevistas directas y observación. Además de su elevado costo, los defectos se derivan del hecho de que se omitirán los centros de actividad no reconocibles;

b) Registros de la administración. Pueden servir para establecer la guía los registros con fines fiscales o los del seguro de paro, seguridad social u otros programas estatales. Tales registros, sin embargo, necesitan un examen cuidadoso para saber si son completos, si se prestan a los fines del censo o si son precisos, ya que no están destinados en principio a cubrir las necesidades de la encuesta económica. Los registros pueden no contener indicación alguna del tamaño y, así, no se podrían utilizar para la selección de muestras o para la aplicación de un nivel mínimo de tamaño determinado previamente. Pueden indicar empresas más bien que establecimientos, exigiendo así una encuesta preliminar (a menudo denominada inventario previo) para reunir la información que se necesita de cada establecimiento. Las direcciones pueden ser postales y no coincidir con la ubicación real. Los registros tal vez no estén actualizados; pueden comprender empresas que hayan cambiado de propiedad o abandonado su actividad y tal vez no incluyan las que hayan entrado recientemente en el sector. La asignación de la clase de actividad (suponiendo que tal

información figure en el registro) puede basarse en normas de clasificación diferentes de las que se empleen en la encuesta. La gran ventaja de los registros de la administración es que en ellos figurarán listas de lugares de actividad no reconocibles y también de negocios que los encuestadores pueden olvidar porque sean difíciles de localizar o porque estén omitidos erróneamente en la lista al no haberse entendido bien las instrucciones;

c) Directorios de publicaciones o asociaciones de comercio. Como los directorios establecidos por asociaciones de comercio o publicaciones mercantiles pretenden unos fines concretos, tienen ciertos defectos que sus usuarios han de conocer. Además de los mismos problemas que los registros de la administración, los de las asociaciones de comercio tienden a limitarse únicamente a sus miembros. Muchas veces, sin embargo, incluyen las empresas que prestan servicios al sector de que se trate. Así, una lista de salones de belleza puede abarcar a mayoristas y fabricantes de productos de belleza. Igualmente, como las publicaciones mercantiles buscan la mayor clientela posible, en sus listas figurarán muchas empresas que sólo tengan relaciones marginales con el sector de que se trate;

d) Otras fuentes potenciales. Comprenden guías de teléfonos o listas especiales preparadas por las compañías telefónicas y también listas preparadas por compañías que se dedican a publicar directorios privados, cámaras de comercio, etc. Cada tipo tiene sus propias características especiales que han de estudiarse con cuidado antes de que se tome una decisión sobre cómo usarlo.

155. Uno de los principales problemas al establecer una guía para su uso en una encuesta de la distribución y los servicios estriba en el hecho de que las empresas que trabajan en ese sector tienen una tendencia frecuente a operar en varios sitios diferentes. Además, existe el problema de las empresas que se dedican a más de un tipo de actividad. Para hacer que la guía resulte lo más útil posible, es necesario intentar identificar a esas empresas y ponerse en contacto con ellas antes de que se haga la encuesta propiamente dicha a fin de poder establecer un registro separado para cada ubicación, que comprenderá su emplazamiento físico, la clasificación por tipo de actividad y algún indicador del tamaño. Tal contacto o investigación previa puede basarse en una selección de empresas de las que se sepa presentan tales características (cuando se dispone de tal información).

156. Por lo general, la guía se prepara usando un registro para cada establecimiento. Para las empresas de muchos establecimientos, sin embargo, debería también haber un registro para la oficina central, y cada establecimiento figuraría con referencia a tal oficina. Los registros pueden hacerse en forma de una sencilla ficha para cada establecimiento o empresa, o cabe preparar una serie de fichas perforadas que pueden convertirse en cinta de computadora en aquellos países que tengan acceso a ese sistema. Un registro separado para cada unidad permite una flexibilidad máxima y la eliminación de los registros de los

establecimientos o empresas que abandonen la actividad mercantil, la adición de otros nuevos para los que inician tal actividad, la introducción de cambios y la clasificación de los registros por orden alfabético, numéricamente, por tipo de actividad y por zonas. Para países que tengan una guía pequeña, puede ser adecuado mantener un conjunto de fichas corrientes; no obstante, el uso de fichas perforadas permite una clasificación rápida y mayor flexibilidad. Las posibilidades del cálculo electrónico resultan aún más rápidas y flexibles pero no deberían usarse a menos que el volumen y la complejidad del trabajo lo justificase.

157. Un problema común a todas las fuentes que se usen para establecer guías es el intervalo que transcurre entre el tiempo en que la información se obtiene y el momento en que se usa la guía. Por causa de ello, la información es muchas veces incompleta y no traduce los cambios que han ocurrido entre tanto. Si la guía se va a usar en encuestas ulteriores, deberá disponerse lo necesario para completarla con nuevas empresas y con correcciones que reflejen los establecimientos que han cerrado y los cambios en la propiedad. Será menester, por tanto, introducir en el sistema un método para incorporar tales cambios. Hay que estar seguros de que tal forma no conduzca a una duplicación en aquellos casos de empresas que simplemente han cambiado sus nombres o han adquirido otras.

158. Como se dijo antes, la mejor fuente para una guía completa es la que se tiene después de haberse terminado un censo general. En éste, se habrá dedicado un gran esfuerzo a identificar y clasificar cada unidad estadística y de formación, y se habrán reunido datos sobre tamaño, ubicación y otras características. Una guía derivada de un censo puede servir para establecer series que se usarán en las encuestas anuales y de frecuencia superior basadas en la obtención de encuestas estratificadas y eficientes. Estas últimas pueden emplearse al mismo tiempo que una muestra de área. Una guía de esa índole probablemente estaría limitada a las empresas y establecimientos mayores, con todas las demás representadas mediante la muestra de área; esta muestra podrá entonces usarse para representar a los pequeños establecimientos así como a los que se creen por vez primera. La lista de grandes empresas que se incluirán en la encuesta no debería limitarse a aquellas que caen dentro del sector de la muestra de área, y los datos de tales establecimientos deberían recogerse directamente por correo independientemente de donde estén situados. La determinación de qué nivel debería usarse para decidir cuándo un establecimiento es grande debería basarse en el equilibrio óptimo entre el costo y los problemas que entraña el recuento sobre el terreno y el empleo del correo. Debería prepararse y clasificarse por zonas una lista de grandes establecimiento (y de empresas de muchas unidades) que se transmitiría a las oficinas sobre el terreno y se usaría como lista de omisiones obligadas en el recuento de las zonas seleccionadas en el muestreo de área.

159. Como una guía representa un conjunto de empresas que existían cuando se preparó y vale lo que valga la fuente de donde se obtuvo y la calidad de su preparación, existe una constante necesidad de ponerla al día o de completarla para incluir nuevas empresas y establecimientos. El uso de una muestra de área en relación con esa guía cumple tal función. Si se dispone de otras fuentes para

tener actualizada la guía, tales como registros fiscales y administrativos, cabe usarlas. No obstante, la elaboración y transferencia de tales registros en tiempo oportuno resulta extremadamente difícil.

160. Una vez que se haya establecido una guía, algunos países preparan un programa de recogida de datos a partir de una serie constituida por las unidades mayores incluidas en la guía. Cuando se usa esa serie, el requisito mínimo es establecer un sistema en que las tendencias de período a período se basen en una información de unidades iguales. Entre los defectos obvios de un programa de esa índole figura el hecho de que lo que ocurre a las grandes unidades puede no ser representativo de lo que sucede en la economía total y las unidades que dejan de informar pueden tener tendencias diferentes de aquellas que sí lo hacen. Si bien este tipo de programa es menos costoso de operar, los datos que reporta pueden resultar muy equivocados.

161. Como se indicó con anterioridad, un primer censo general que comprenda listas sobre el terreno constituye la mejor fuente para establecer una guía. En censos posteriores, la lista de establecimientos identificados en el censo anterior se convierte en la base de una nueva guía. El uso de ésta en un censo es, sin embargo, algo diferente de su empleo en las encuestas anuales y de frecuencia más que anual. La guía obtenida del censo anterior se convierte en un mecanismo de control que se usa fundamentalmente para tener la seguridad de que los establecimientos y empresas mayores no se han omitido del censo que se está haciendo o que no se han clasificado diferentemente o se han asignado a un sector distinto. Antes del recuento, los datos del censo anterior, en el caso de los grandes establecimientos, pueden clasificarse por zonas y transmitirse a las oportunas oficinas sobre el terreno para que los usen como cotejo de la recopilación que estén haciendo. Debería pedirse a los entrevistadores sobre el terreno que indicaran los establecimientos no localizados. Una vez que el recuento se haya completado, los registros deberían clasificarse por ubicación, tipo y tamaño de actividad, y deberían compararse los datos de ambos censos. Habría que investigar las diferencias y omisiones importantes. Hay que prestar atención, sin embargo, a que las diferencias no se deban a cambios de nombre o propiedad.

162. Cuando se establece una guía para su uso por correo más bien que en una recopilación sobre el terreno, su función es diferente. En un recuento por correspondencia, la guía ha de ser completa en sí misma, y tal cosa requiere que contenga alguna anotación para cada unidad sobre la que se tabularán datos. El primer requisito, por tanto, es que las fuentes usadas sean completas y recojan los datos necesarios para el censo o permitan establecer contacto con la empresa para recabar tales datos. Como el censo se limita a ciertos tipos de actividad, la fuente puede comprender información sobre ese tipo así como nombre y dirección. También sería conveniente tener información sobre tamaño. Para ser totalmente útil, la guía debería incluir asimismo una anotación separada para cada establecimiento de las empresas con sucursales. Cuando una guía contenga todos los elementos anteriores, podrá usarse no sólo como medio para controlar la recogida de información sino también como base para evaluar los datos de quienes contestan mal o no lo hacen. A partir de esa guía, la información sobre

los establecimientos por tipo de actividad, ubicación geográfica y dimensiones generales puede tabularse. Si bien esa información no será tan precisa como la que cabe obtener con un formulario debidamente completado, resultará útil para varios propósitos. Cuando la recopilación se lleve a cabo, será posible, en efecto, preparar un programa para quienes contestan defectuosamente o se niegan a hacerlo, adscribiendo a tales establecimientos iguales características que las obtenidas al tabular los datos notificados de los establecimientos del mismo tamaño y tipo de actividad, situados en la misma zona general.

163. Cuando en los diferentes registros de la administración se haya establecido un mismo sistema de identificación de las empresas, será también posible transferir los datos reales (aunque casi siempre muy limitados) a la guía. Esto permitiría que toda una clase de pequeñas empresas no tuvieran que informar directamente para el censo. Al recurrir así a los registros administrativos hay varios elementos que deben estudiarse para estar seguros de la validez del procedimiento. En primer lugar, tiene que ser posible identificar la unidad estadística o de información. En segundo término, la asignación de la clase de actividad ha de basarse en el mismo sistema de clasificación; en tercer lugar, los datos que hay que transferir tienen que definirse del mismo modo como figurarán en la encuesta; en cuarto lugar, el registro administrativo debe estar puesto al día.

Recogida y control de la información

164. En una encuesta que requiera visitas directas a los establecimientos que caen dentro del campo de aplicación, cabe ampliar la información obtenida durante la confección de listas para que puedan confundirse en uno solo los procedimientos de recogida de información y de preparación de listas. Como se dijo antes, puesto que el comercio de distribución y los servicios se caracterizan por el gran número de pequeños establecimientos, un leve aumento de la información obtenida en el momento de la preparación de las listas puede hacer que éstas sirvan de conducto para obtener datos de esos pequeños establecimientos de los que sólo se requieren datos básicos. Para todo lo demás, el entrevistador puede entregar el cuestionario apropiado y fijar una fecha para pasar a recogerlo o convenir en que se envíe por correo a la oficina correspondiente. Una vez terminada su labor de "recopilación de la lista", el entrevistador devuelve esta última a su oficina que entonces la comprueba con los informes recibidos, esto es, los entregados en mano por el entrevistador y los enviados por correo a la oficina.

165. Como parte de la operación relativa a las listas, debería obtenerse para cada establecimiento algún indicador del tamaño. Ese indicador puede usarse como base para estimar los datos de quienes no contestan. También es útil para conocer qué esfuerzo vale la pena hacer al intentar recoger la información de quienes no devuelven los cuestionarios.

166. Cuando el entrevistador encuentre establecimientos que pertenezcan a empresas con sucursales, debería obtenerse el nombre y la dirección de la oficina central desde la que los datos tienen que comunicarse. Convendría preparar fichas con referencias cruzadas para tales establecimientos y enviarse a la oficina central

de estadística que puede entonces cotejarlas con fichas similares del resto del país y usar todo ello como comprobación de los datos de las empresas de muchos establecimientos.

167. Si la preparación de las listas se ha hecho y controlado adecuadamente, los establecimientos que figuren en ellas y que se ha decidido figuren dentro del campo de la encuesta se convierten en el universo sobre el que se tabularán datos. Las operaciones ulteriores deberán estar vinculadas con ese universo si se quiere que la encuesta esté bajo control.

168. Por razones prácticas, conviene detectar y corregir los errores en los datos comunicados lo antes posible una vez que se hayan cumplimentado los cuestionarios o las listas. Si hay demoras en volver a entrevistar o preguntar a quienes contestaron, la persona que suministró inicialmente los datos quizá tal vez ya no trabaje en la empresa o ésta puede haber abandonado su actividad o los registros pueden haberse tirado. Por otra parte, de ordinario no resulta viable ni conveniente hacer en las oficinas locales las operaciones completas de preparación de los datos que haya que publicar. No obstante, el personal de esas oficinas debería hacer una operación de cribado que entraña una cantidad limitada de cotejo de datos. Tal cosa podría consistir en comprobar que el cuestionario está debidamente identificado, que no se han omitido cifras y otras sencillas verificaciones similares. Una comprobación algo más complicada pero fundamental consistiría en sumar los costos notificados (sueldos y salarios, materiales, etc.) para ver que no exceden del valor notificado de las ventas o servicios. De producirse tal cosa, la información recibida debería devolverse al entrevistador o darse a otro recopilador para su verificación o corrección.

Ese examen a escala local también proporciona algún control sobre los resultados de cada entrevistador. Cuando se encuentra que los cuestionarios o las listas son defectuosos, debería pedirse al entrevistador responsable que los corrigiera, si el tiempo lo permite. Un número excesivo de listas o cuestionarios defectuosos puede significar que el entrevistador es incompetente y debería prescindirse de sus servicios.

169. Una guía completa de establecimientos y otras unidades de información, después de que se haya actualizado, podrá usarse para seleccionar entre dichas unidades las que se incluirán en la encuesta, tanto si ésta se basa en un recuento completo o en una muestra. Las unidades se elegirán con arreglo al tipo de actividad y se enumerarán usando cuestionarios específicamente diseñados para cada uno de los grandes grupos de actividad. Si la encuesta va a basarse en una información obtenida por correo, se usará la dirección que figure en la guía para tal fin y se pedirá a quienes han de contestar que devuelvan el cuestionario cumplimentado antes de una fecha límite a la oficina central de estadística, donde se comprobará al recibirlo. De cuando en cuando será menester hacer operaciones complementarias para aquellas empresas que no envían el cuestionario. Una realización rápida y eficiente de esas operaciones requiere que se disponga de un conjunto rígido de controles y que se respete cuidadosamente. De otro modo, hay muchas posibilidades de que se omita alguna información y se duplique otra.

170. Además del cuestionario adecuado que se base en el tipo de actividad de la empresa, en el conjunto que se envía por correspondencia deben figurar varias otras cosas. Entre ellas, un sobre con la dirección y que no tenga que ser franqueado para que quien contesta pueda usarlo únicamente al devolver el cuestionario cumplimentado, una copia del cuestionario para que la persona que lo contesta pueda conservarla en sus propios archivos y cualesquiera instrucciones especiales que se necesiten para rellenar y devolver el formulario. También debería incluirse una declaración sobre el carácter confidencial de los datos que se comuniquen y sobre las sanciones de no hacerlo o de dar datos falsos. Unos sobres con indicaciones visibles señalarán a la atención de quienes lo reciban la importancia del contenido. El uso de sobres con una ventanilla transparente a través de la cual se ve la dirección impresa que figura en el cuestionario evita tener que repetir esa dirección en el sobre. Conviene indicar claramente la fecha en que los formularios cumplimentados han de devolverse. Es importante fijar una fecha límite que permita un tiempo razonable para contestar el cuestionario pero que no sea tan excesivo como para que el encuestado lo deje de lado y se olvide de él. Por lo general, un mes puede considerarse un período razonable. No será posible establecer un día y un mes específico para su devolución a menos que todos los cuestionarios estén listos para ser enviados a sus destinatarios aproximadamente en la misma fecha. Si esto no es posible, se pedirá que se devuelvan los cuestionarios 30 días después de hacerlos recibido. Igualmente, países que cuentan con un sistema postal poco fiable tendrán dificultades para fijar una fecha concreta en la devolución y pueden tener que pedir que se cumplimente y envíe 30 días después de su recepción. En esas circunstancias, detectar y completar los datos de quienes no responden resulta muy difícil.

171. Para estar seguro de que cada uno de los establecimientos que figuran en la lista postal ha quedado incluido y que las tabulaciones finales representan así a todo el universo, hay que fijar un rígido sistema de verificaciones. Tal cosa comienza con la creación de un registro, ya sea en fichas corrientes, fichas perforadas o cinta de computadora, para cada unidad a la que se ha enviado un cuestionario. La asignación de un solo número de identificación a cada una de tales unidades resultará útil, ya que la comprobación mediante el nombre comercial puede ser complicado porque esos nombres cambian, se repiten, se escriben con errores, se producen cambios en la propiedad, etc. Ese número de identificación puede también usarse para el archivo, en su caso, de los informes y simplificará encontrar los formularios que se necesiten durante la labor de tabulación. Dos tipos de listas resultarán muy útiles durante la elaboración de la encuesta: una en orden numérico y otra en orden alfabético. Si bien su uso proporciona mayor flexibilidad, las fichas por desgracia a veces se pierden o traspapelan. Las listas no tienen ese defecto, aunque debería procurarse juntarlas de modo que no se deterioren, preferentemente en formato encuadernado.

172. Una vez que se vayan recibiendo los formularios cumplimentados, se harán anotaciones que permitan cribar el registro completo cuando se decida que es necesario subsanar los daños de quienes no devolvieron el cuestionario. Este esfuerzo complementario que se requiere para obtener índices de respuesta aceptable es grande, incluso en los países más avanzados. Al cabo de un mes del envío de

los cuestionarios por correo, es de esperar que no se habrán devuelto más de la mitad. Como no sería posible tabular los datos sin incluir a las unidades mayores, es aconsejable establecer procedimientos especiales en su caso. Tales unidades deberían ser visitadas antes de la encuesta para establecer cualesquiera procedimientos específicos que puedan requerirse. Durante la encuesta, deberían desplegarse esfuerzos especiales para seguir de cerca los formularios correspondientes y estar seguros de que se reciben a tiempo y de que se comunican datos aceptables. La preparación de los cuestionarios de esas unidades puede llevar más tiempo, y en su caso la fecha exigida tal vez tenga que ampliarse. También debe pensarse en una prolongación razonable de las fechas en unidades con problemas especiales. Cuando una empresa que presta servicios contables tenga la responsabilidad de llenar gran número de contestaciones, también debería dejársele más tiempo.

173. La recogida de la información por correo generará un gran volumen de correspondencia, ya que los destinatarios frecuentemente harán preguntas que necesitan respuesta antes de que se rellenen los cuestionarios. Debería prepararse un procedimiento especial para clasificar tales cartas y contestarlas con rapidez.

174. Al decidir qué sistema usar para la recogida de datos, es importante tener en cuenta el hecho de que uno de los principales problemas de tales encuestas es obtener series cronológicas de datos. Las tendencias cíclicas a largo plazo se descubren con el estudio de las encuestas de frecuencia más que anual. En ambos casos, sin embargo, la validez de los cambios puede anularse cambiando el método de realización de la encuesta. Al diseñarla, por tanto, el sistema que se adopte debería poderse reproducir encuesta tras encuesta. Si, por ejemplo, se decide usar registros administrativos de unos determinados servicios del gobierno para establecer una guía, hay que tener alguna seguridad de que esos servicios seguirán manteniendo el material empleado como fuente o como base aceptable en un futuro previsible. De igual modo, al decidir el alcance y carácter de la muestra, se debe tener alguna certeza de la continuidad del proceso. Toda "mejora" ha de ser examinada para ver si redundará en una discontinuidad de los datos. No obstante, si se decide que es preciso adoptar un procedimiento que producirá una discontinuidad, habría que hacer todo lo posible para tender un "puente" entre las series antiguas y las nuevas a fin de que los usuarios de los datos puedan conocer los efectos del cambio e introducir los ajustes oportunos.

Uso del muestreo^{6/}

175. Al hacer una encuesta, ya se trate de un censo general con gran cantidad de datos o de un recuento anual o de frecuencia superior con miras a producir un menor número de datos lo más rápidamente posible, el director de la encuesta intentará diseñarla de tal modo que los datos requeridos se obtengan con un costo

^{6/} En el anexo I se hace un examen detallado de las técnicas de muestreo.

mínimo. Como el muestreo es un método con el que cabe reducir esos costos, es aconsejable examinar todo programa para determinar hasta qué punto y de qué manera dicha técnica puede usarse. Al tomar una decisión, el director ha de sopesar las ventajas derivadas del muestreo, tales como una reducción en número de contactos, número de cuestionarios y necesidades de mano de obra, y una mayor rapidez en la publicación, con los inconvenientes, tales como las dificultades inherentes al control de las muestras, la mayor calidad del personal requerido y la pérdida de datos. Al diseñar encuestas en el comercio de distribución y los servicios, la fijación del alcance y carácter de las muestras que conviene usar viene condicionada por la heterogeneidad de los tipos de actividad que hay que cubrir. Por ejemplo, una muestra que podría ser adecuada para medir la actividad del comercio al por menor probablemente sería inapropiada para medir la actividad de las agencias de publicidad. Un diseño de la muestra que sería aceptable para un censo general puede no dar resultado en una encuesta mensual.

176. Cuando la encuesta recurra a un trabajo de campo mediante listas, será posible fijar un tamaño y aplicar el criterio de que todos los establecimientos que queden por encima de él se recuenten y los que queden por debajo se seleccionen por un muestreo que permita elegir cada n ésimo establecimiento. Esta selección puede hacerse en las listas con arreglo a pautas previamente fijadas o después de que dichas listas hayan sido elaboradas en la oficina central de estadística. Multiplicados adecuadamente, esos casos representarán entonces todos los establecimientos por debajo del tamaño fijado como divisoria. Tales datos, claro es, contendrán un error propio del muestreo. En lugar de seleccionar sólo un caso de cada n por debajo del tamaño fijado a fines de enumeración, un procedimiento alternativo sería reunir información completa de tales casos así seleccionados y hacer que todos los demás que quedan igualmente por debajo también contesten a un cuestionario muy simplificado donde figuren sólo datos básicos. Todos los establecimientos por encima del tamaño tendrían que suministrar información completa.

177. Cuando un procedimiento de recuento por correo emplea una guía que contiene información sobre tamaño y tipo de actividad, será posible seleccionar una muestra estratificada basada en tales características. Al diseñar cualquier muestra, sin embargo, es importante que se incluyan siempre los grandes establecimientos. No obstante, como hay muchos pequeños, cabe diseñar una muestra que reducirá considerablemente el número de unidades que hay que entrevistar, pero el uso de tal muestra disminuirá mucho las posibilidades de publicar información por zonas y tipo de actividad.

178. Los objetivos en materia de datos de las encuestas anuales y de frecuencia superior suelen ser muy limitados. Además, la producción rápida de esos datos reviste gran importancia, especialmente en las encuestas de frecuencia más que anual, ya que el objetivo es reunir información sobre tendencias actuales que servirán de orientación para tomar medidas inmediatas. Esto es especialmente cierto en sectores como las ventas de los minoristas. Por consiguiente, el diseño de la encuesta debería ser lo más sencillo posible y susceptible de una rápida elaboración. Para tener validez, debería contarse con una muestra con una probabilidad conocida de selección para que pueda calcularse la variabilidad del

muestreo. El diseño más corriente consiste en preparar una lista de grandes empresas de las que se obtendrá información en cada encuesta y del que cabe obtener un sistema para seleccionar una muestra de todas las demás. Esta muestra puede ser de área, o si se dispone de una lista completa al día de todas las empresas, seleccionarse de dicha lista. Como es casi imposible tener actualizada una lista de esa índole, ese diseño debería completarse con el uso de una muestra de área que puede ser bastante menor que la que se requerirá de no tener tal lista. Si bien la función básica de la muestra de área es representar a los pequeños establecimientos, también cumple la importante finalidad de encontrar y encuestar nuevos establecimientos independientes de su tamaño, que hayan surgido después de que se eligieran a los "grandes" comercios en la muestra.

179. Existen algunos tipos de actividad donde el uso de una muestra de área es muy ineficiente y conduce a variaciones tan grandes como para limitar o incluso excluir su empleo. Tal cosa se debe al hecho de que el uso principal de la encuesta es obtener datos sobre tendencias del comercio al por menor, ya que éste es uno de los indicadores económicos más sensibles. Por ello, la selección de segmentos de la muestra se basa en las concentraciones de población que guardan mucha relación con la actividad de los detallistas. El hecho de que actividades tales como hoteles de temporada, estaciones de petróleo a granel, mayoristas de chatarra, productores de películas y emisoras de radio, que también están incluidas en el campo de aplicación, no acostumbren a estar ubicadas de ordinario en zonas que coinciden con concentraciones de población significa que la proporción de datos reunidos en los segmentos de la muestra contendrá varianzas tan elevadas de muestreo como para hacerlos inútiles. Para esas actividades, por tanto, es menester basarse fundamentalmente en una guía bastante completa que puede tenerse al día. De no ser ello posible, entonces tales actividades no deberían figurar en las encuestas corrientes.

180. Desde el punto de vista teórico, como el tamaño de una muestra significa que tiene que reunirse y elaborarse menos información, su uso presente varias ventajas. Entre ellas figura el no hacer recaer sobre muchas personas la carga de dar la información y reduce en consecuencia el trabajo del organismo de estadística. Esto último permite que se ejerza un mayor control sobre el acopio y tramitación de los cuestionarios y, por ende, redundan en una mejora de la calidad de las estadísticas producidas. Además, el calendario de la producción de datos podrá mejorarse, y será posible ajustar el tamaño de la muestra a los fondos disponibles. Cuando estos fondos no sirvan para un censo completo, el muestreo puede ser el único método posible.

181. El sistema de muestreo usado debería ser una muestra de probabilidad que dé resultados de calidad que puedan medirse. El uso de una muestra de esa índole introduce pasos adicionales que tienen que darse debidamente. Todo ello entraña un costo y crea la posibilidad de errores que no hay con una cobertura completa. A menos que la muestra esté diseñada para ajustarse a las condiciones en que hay que trabajar y esté satisfactoriamente controlada, su uso puede reportar más pérdidas que ganancias. El diseño de una muestra muy elaborada y teóricamente conveniente que requiera personal más capacitado, más registros o más equipo de lo que se dispone sería inútil o incluso peor. Otro inconveniente del muestreo es que la

cantidad de detalles fiables que cabe publicar se verá inevitablemente reducida, ya que el menor número de observaciones disponibles de cada célula de datos conduce a un incremento en la varianza del muestreo, en muchos casos hasta un nivel que haga a las estadísticas resultantes nada fidedignas. Además, incluso aunque la recogida, preparación y tabulación se controlarán mejor cuando haya que elaborar un menor número de cuestionarios, los errores en esas funciones no pueden nunca eliminarse totalmente.

VIII. ELABORACION DE DATOS

Observaciones generales

182. Una vez recibidos en la oficina los cuestionarios, será necesario someterlos a un examen que garantice que contienen toda la información necesaria y que se pueden elaborar en operaciones ulteriores. En ese examen, entre otras cosas, se comprobará que se han notificado los datos necesarios, que el trabajo mecánico de redondear cifras y codificar se ha hecho debidamente, que las anotaciones son legibles, que de las explicaciones que hayan dado por escrito las personas que han respondido los cuestionarios se toma nota para proceder en consecuencia y que se han realizado las comprobaciones administrativas para detectar datos erróneos. Ese examen descubrirá un gran número de problemas que han de resolverse antes de que el cuestionario esté listo para una elaboración posterior. Un elevado porcentaje de las personas que hayan respondido tendrán que ser contactadas para verificación o corrección de datos. Como esta labor lleva mucho tiempo, debería iniciarse pronto y hacerse con rapidez para no demorar las operaciones posteriores.

183. La función más difícil es la detección de los errores en la información. Estos errores pueden deberse a varias causas: personas que no comprenden los conceptos que están implícitos en una información correcta; personas que rellenan los cuestionarios sin la atención debida; datos que no están disponibles en las unidades a las que se piden; datos notificados sobre empleo y sobre nómina de pago por ejemplo, con diferentes períodos de referencia, etc. Estos problemas han de ser detectados y zanjarse antes de la tabulación. La cantidad de trabajo que tendrá que realizar el personal administrativo y profesional variará con la calidad de éste y el equipo de que disponga. Con el uso de computadoras, cálculos complejos para detectar errores pueden realizarse con rapidez. Cuanto menos perfeccionado sea el equipo, más trabajo tendrá que hacer el personal administrativo y profesional.

184. Es despilfarrar recursos tratar igual todas las deficiencias en los cuestionarios, por lo que hay que establecer en su revisión algunas normas y criterios. Por lo general, la cantidad de esfuerzo dedicado a detectar y corregir errores y omisiones debería guardar proporción con sus efectos probables sobre los datos publicados. Hay una cierta seguridad en los números; si los errores son de tipo aleatorio, tenderán a anularse unos con otros, siempre que individualmente sean pequeños. Esto conduce a la conclusión de que no es necesario un alto grado de precisión ni normas rígidas cuando se trata de pequeños establecimientos, típicos de muchas ventas al por menor y servicios. La cantidad de atención que se preste a cada uno de los cuestionarios durante su revisión y codificación debería variar en proporción directa con el tamaño de los establecimientos representados. De esto se sigue que los informes de los grandes establecimientos deberían normalmente revisarse con más cuidado que los de los pequeños, excepto cuando la cobertura de estos últimos se base en una muestra.

Operaciones administrativas

185. La asignación de la labor, el mantenimiento de los registros y la información sobre la marcha de los trabajos se verán simplificados si los cuestionarios se clasifican por tipos y se agrupan en unidades de trabajo de tamaño conveniente. De ordinario no más de 100 deberían reunirse en una de esas unidades; su número, en realidad, debería basarse en la cantidad de trabajo que ha de hacerse, pero no tendría que ser tan elevado como para tener ocupada a ninguna unidad de trabajo durante un largo período de tiempo. Las unidades deberían numerarse en serie y mantenerse un registro centralizado de control que lleve nota de la situación en todo momento de cada una, como fecha asignada, nombre del empleado al que se haya asignado, número de cuestionarios que requieren un nuevo contacto con quienes los contestaron, número de cuestionarios que hay que remitir para su examen al personal profesional, y fecha de terminación. Una unidad de trabajo contiene normalmente un solo tipo de cuestionarios. Quizás sea conveniente clasificar los cuestionarios de otras maneras, separando, por ejemplo, los de los pequeños establecimientos de los grandes, ya que ambos pueden tener normas de revisión distintas. De igual modo, la información sobre empresas con muchos establecimientos pueden agruparse en series separadas de las unidades de trabajo.

186. Una de las principales decisiones que hay que hacer en esta etapa de la labor se refiere al mayor o menor grado en que la revisión y preparación de los cuestionarios deben dividirse en operaciones muy especializadas; cada empleado puede o bien repetir unas cuantas cosas muy sencillas de rutina que se dominan rápidamente o aprender y realizar todos los pasos necesarios para la preparación completa del cuestionario con miras a la perforación. La primera alternativa tiene el inconveniente de requerir una operación de control más bien compleja y mucha manipulación de los informes. La segunda necesita tiempo para capacitar a las personas a fin de que puedan seguir procedimientos más complejos y difíciles. La decisión dependerá del sentido común y del tipo de personal disponible. Quizá sea conveniente llegar a un equilibrio entre ambos extremos.

187. La operación administrativa se presta muy bien al uso de un sistema de control de calidad destinado a aplicar un programa de alerta rápida que identificaría a aquellos empleados que no están siguiendo debidamente los procedimientos. El sistema debería también permitir que se midiera el índice de producción en relación con la calidad. A los empleados se les debería notificar inmediatamente sus errores y aquellos que estuviesen constantemente por debajo del nivel aceptable deberían ser despedidos. Frecuentemente, esa clase de comprobación puede también revelar un error en los procedimientos que afectaría a todos los que están revisando los cuestionarios.

188. Incumbirá al personal de planificación preparar especificaciones que usará el servicio de elaboración para redactar las instrucciones a los empleados sobre qué pruebas deberían hacerse para determinar si la información es aceptable o defectuosa. Ese servicio debería también suministrar el material de referencia que permita evaluar toda información omitida y los sistemas de codificación que tienen que aplicarse a los datos. Como con todos los procedimientos, el personal de planificación debería revisar los resultados obtenidos. También han de hacer

lo necesario para supervisar la labor de algunas de las primeras unidades de trabajo inmediatamente después de que hayan terminado e introducir lo antes posible cualesquiera modificaciones necesarias en el procedimiento. También es conveniente que compruebe personalmente las operaciones administrativas de cuando en cuando para ver si están surgiendo nuevos problemas.

189. Como parte de las operaciones de revisión y preparación de los cuestionarios debería establecerse una unidad de correspondencia. Se encontrará que muchas de las respuestas que se necesitan para que un formulario resulte completo o aceptable pueden obtenerse sobre la base de una contestación a una carta circular. El empleado puede indicar la carta que hay que enviar (a veces una sola carta circular puede incluir varios puntos diferentes) y enviar una breve nota de referencia con esa información a la unidad de correspondencia que puede mandar entonces la carta y establecer controles para ocuparse de las respuestas.

190. En algunos casos no cabe introducir un procedimiento sencillo en las operaciones administrativas de tramitación; habrá ciertas situaciones donde se necesitará algo más complicado. En esos casos debería constituirse un segundo grupo de empleados bastante más capacitados que hayan sido especialmente formados para ocuparse de esas operaciones. Los problemas que sean singulares y no puedan tratarse por ninguno de los empleados administrativos deberían remitirse al personal profesional. Estos tendrán la opción de buscar ellos mismos una solución o de volverse a poner en contacto con las personas encuestadas de que se trate para resolver el problema.

191. Entre las normas administrativas de procedimiento que presentan ventajas y que usan muchos países figuran las siguientes:

- a) Las anotaciones en el cuestionario de quienes lo revisan deberían hacerse con lápiz de color o bolígrafo. El color escogido debería contrastar con el que se haya usado para responder al cuestionario o con el empleado por otras personas que lo hayan anotado después;
- b) Las anotaciones hechas en la revisión y basadas en las respuestas de las personas encuestadas deberían hacerse también en un color distinto;
- c) Cuando se haga una corrección a una respuesta, la anotación original no debería borrarse ni tacharse. El empleado debería trazar una línea sobre esa anotación de tal manera que quede legible y poner la cifra correcta sobre ella. Si no hay espacio, la cifra corregida podría ponerse en otra parte e indicarse con una flecha su posición adecuada;
- d) En el formulario debería dejarse espacio para las iniciales de las personas (revisores, codificadores, miembros de la plantilla profesional) que revisen los formularios y la fecha correspondiente.

Preparación del cuestionario para su tabulación

Comprobación y cotejo

192. El primer paso es determinar si el cuestionario se refiere a un establecimiento que queda dentro del campo de aplicación de la distribución y los servicios. Tal cosa es especialmente necesaria cuando el establecimiento ha sido contactado por correo y su tipo de actividad ofrece dudas. Si se decide que queda fuera del campo, no será menester ninguna tramitación más. Como la comprobación del campo de aplicación entraña una inspección del tipo de actividad a que se dedica el establecimiento, podría parecer lógico que en este punto se asignara también un código detallado de tipos de actividad. No obstante, la aplicación de las normas para la asignación de tales códigos puede ser muy complicada, y quizá fuese conveniente reservar esa operación a un grupo de empleados especialmente formados para tal función. Para asignar códigos de actividad, será necesario usar información tal como ventas por clase de cliente, volumen de ingresos por la prestación de servicios en contraste con el derivado de la venta de mercancías, ventas por clase de artículo, servicios por tipo de prestación, tipo de actividad que se asigna a sí mismo la empresa y código histórico (cuando exista). Las normas sobre cómo usar esa información para llegar a la clasificación por tipo de actividad serán suministradas por el servicio de planificación que debería también facilitar su personal para resolver los muchos problemas que puedan surgir.

193. Hay varios puntos que deben estar presentes en todo establecimiento incluido en el universo cubierto por la encuesta. Entre ellos figura de ordinario la cifra total de ventas, el empleo total y la nómina anual total. (En las encuestas de frecuencia más que anual, los datos necesarios pueden limitarse al volumen de ventas e incluir inventarios en el caso del comercio mayorista y minorista.) Cuando grandes establecimientos no hayan facilitado los datos necesarios, será menester volverse a poner en contacto con ellos. Para los pequeños, sin embargo, sería más conveniente evaluar los datos que falten. Para ello debería establecerse una serie de cuadros que sentara las bases de tal evaluación.

194. Se dará una situación especialmente perturbadora cuando se pida una respuesta a una pregunta sobre una actividad a la que no se dedican todos los que contestan. La falta de respuesta a esa pregunta es ambigua, ya que puede significar que quien responde no se dedica a tal actividad o simplemente que no se ha molestado en suministrar la información. La tabulación de las respuestas a ese punto podría prestarse a confusión si la experiencia de quienes contestaron es diferente de quienes no lo hicieron, especialmente si la falta de respuesta supone no actividad en ese sector. Dado que es especialmente importante distinguir la no contestación de la no aplicabilidad, la pregunta debería incluir un desarrollo "sí-no", del modo siguiente:

¿Tuvo usted ingresos de _____ durante el año del censo? Sí No

Si la respuesta es "sí", ¿cuánto le reportó esa actividad durante el año del censo? _____

Si se ha contestado "sí" pero se ha omitido la cantidad, quedaría claro que el encuestado ha dejado de indicar una cifra que debería haber señalado.

Preparación y codificación

195. Frecuentemente las anotaciones de datos resultan incorrectas, y se requiere una revisión crítica de los datos notificados para detectar problemas concretos. Hay dos tipos de revisión que deberían hacerse. El primero coteja la coherencia interna de los datos indicados en diferentes partidas para el mismo establecimiento, y el segundo comprueba si la información sobre un punto concreto resulta por sí sola razonable. El cotejo de la coherencia interna también consiste en determinar si los componentes señalados son iguales al total señalado, si las respuestas a la misma pregunta en diferentes partes del cuestionario son las mismas y si ciertos cocientes, tales como paga y empleados o paga y ventas, quedan dentro de unos límites aceptables. La determinación de la aceptabilidad de los límites para esos cocientes puede basarse en promedios derivados de los datos comunicados por establecimientos similares del mismo sector y de la misma zona geográfica. Si en períodos anteriores se han hecho encuestas similares, los datos correspondientes pueden usarse como base para establecer tales límites. Los nuevos datos, sin embargo, deberían ajustarse a los cambios conocidos en el nivel de precios y salarios. En encuestas iniciales, será menester usar como guía datos de otros países. Un ejemplo de incoherencia interna es una respuesta "cero" para ingresos procedentes de la venta de alimentos en un establecimiento que se titula de alimentación.

196. Se necesitan también otros tipos de comprobación para verificar si la respuesta es razonable. Habría que poner en tela de juicio una respuesta de "120%" a una pregunta sobre porcentaje de ocupación de hoteles, una tienda de reparación de televisores sin ingresos por servicios, salas de cine con un número de asientos de 10, lavanderías sin máquinas para lavar, etc. En el transcurso de la revisión, la efectividad de las diversas comprobaciones debería examinarse en términos del número de rechazos que surja. Si la mayoría de los cuestionarios se rechazan de resultas de la aplicación de una determinada prueba de revisión, las tolerancias son casi con seguridad demasiado estrictas. Si, en cambio, prácticamente no se rechaza ninguna información, las tolerancias son probablemente demasiado grandes y deberían introducirse los cambios oportunos.

197. Cuando el cuestionario haya sido diseñado para incluir códigos impresos previamente, el empleado tendrá que examinar la información notificada y asignar un código numérico que pueda perforarse y usarse en la clasificación y tabulación. Incluso cuando los códigos han sido predesignados en el formulario, será menester revisar ciertos casos tales como una respuesta escrita en su totalidad que tendrá que codificarse o la indicación de dos o más alternativas mutuamente exclusivas.

198. Las dos operaciones más difíciles de codificación son las del tipo de actividad y la de ubicación. Determinar el código correcto de tipo de actividad requiere que se solicite y suministre información apropiada en el cuestionario y que se apliquen normas muchas veces complicadas. La codificación de ubicación puede parecer una simple cuestión de aplicar un código numérico al nombre de una

localidad, pero la experiencia muestra que es más difícil que todo eso, especialmente al codificar ciudades. Los límites administrativos son difíciles de identificar, y la gente a veces usa nombres de ciudades en sus direcciones cuando el establecimiento en realidad está situado fuera de sus límites. En un recuento sobre el terreno, tal dificultad puede subsanarse haciendo que el encuestador determine la ubicación, lo cual no se puede hacer por correo. Otros puntos también que codificar son organización jurídica, forma de propiedad, método de venta, tipo de mercancía, diversos códigos de tamaño, etc.

199. Los valores suelen casi siempre notificarse en términos de la unidad monetaria menor, pero a menudo no conviene perforar o tabular los datos de ese modo. La perforación de la cifra completa llevaría más tiempo y se prestaría a más errores; la tabulación también requeriría más tiempo y espacio. Igual ocurre cuando las cantidades se notifican hasta la unidad. Parte de la preparación de los datos para su tabulación consiste en reducir el número de dígitos que habrá que elaborar redondeando las cifras. La capacitación para esa labor debería hacerse con cuidado; mucha gente no parece capaz de captar los principios en que se basa. Es más fácil, claro es, usar un procedimiento que permita eliminar dígitos en lugar de redondearlos, pero los datos de que se trate han de examinarse para comprobar que tal cosa no tiene un efecto apreciable en las estadísticas resultantes.

200. El formulario debería examinarse para buscar cualesquiera notas u observaciones hechas por el encuestado que puedan requerir respuesta. Como paso final, el empleado tiene que examinar el cuestionario para estar seguro de que los datos que haya que perforar son legibles y que el perforador no tendrá dificultad para localizarlos rápidamente.

201. Una encuesta que recurra mucho al correo para distribuir y obtener los formularios generará una gran cantidad de correspondencia. Buena parte de ella es de rutina y podrá manejarse con el empleo de circulares. No obstante, siempre quedará por hacer mucho trabajo para manejar las cuestiones técnicas que tendrán que resolverse antes de que los encuestados puedan completar el cuestionario. Si no se asigna a esa labor un personal debidamente formado para manejar el asunto, las respuestas pueden fácilmente demorarse durante semanas y semanas. Tal cosa no sólo retrasará toda la encuesta sino que frecuentemente desalentará a las personas para que cumplimenten el cuestionario.

Revisión y codificación con computadora

202. Si el volumen de la labor que hay que realizar lo justifica y si se dispone del tipo adecuado de personal y equipo, buena parte del trabajo de revisión puede hacerse con computadora. No obstante, seguirá siendo necesario que empleados administrativos revisen la información para tener la seguridad de que se han comunicado los datos fundamentales, que no existen errores obvios y que toda la labor necesaria que no pueda hacer la computadora se realiza antes de que la información se envíe para la perforación. Una vez que las tarjetas perforadas han sido pasadas a cinta magnética, la computadora puede llevar a cabo muchas operaciones, tales como relaciones, límites de comprobación, evaluación de datos y asignación de códigos. También puede aplicar una serie de pruebas para determinar

la aceptabilidad e identificar casos que requerirán una contrastación o, en circunstancias apropiadas, sustituir un valor estimado a uno obviamente erróneo. Por lo general, la computadora puede realizar tales operaciones con mayor rapidez y precisión que una persona.

203. Se estima, por lo general, que el uso de computadoras abrevia el tiempo necesario para producir las publicaciones finales. No obstante, la experiencia no justifica tal opinión. Otra idea extendida que no responde a la realidad es que el empleo del cálculo electrónico redundaría automáticamente en una enorme reducción de personal. Esto, también, no tiene por qué ocurrir necesariamente. Hay muchas cosas que considerar antes de pronunciarse por la computadora para la preparación de publicaciones:

- a) Conversión de las especificaciones del planificador en el lenguaje de la máquina usado por el programador establece una barrera adicional en las comunicaciones. A menos que el planificador pueda leer el lenguaje del programador para estar seguro de que sus ideas han sido correctamente interpretadas, habrá la posibilidad de que se introduzcan puntos erróneos. Los planificadores tendrán que recibir suficiente formación para poder analizar inteligentemente la labor del programador;
- b) Tendrán que corregirse muchos de los informes que son señalados por la computadora para examen manual. Resulta costoso corregir tarjetas perforadas y cintas de computadora. De ahí que si muchos cuestionarios, llegado el caso requieren una tramitación manual y si hay que volver a perforar las tarjetas y otras tareas, la elaboración de los datos por computadora puede redundar en un costo general muy superior al de la elaboración corriente;
- c) La computadora en concreto que se use puede tener una capacidad de almacenamiento insuficiente para realizar todas las revisiones eficientemente además de otras operaciones computarizadas, tales como codificación y tabulación. El programa debería comprobarse todo lo que sea menester para determinar el efecto de las limitaciones de almacenamiento;
- d) Una computadora con capacidad suficiente es capaz de hacer muchas más comprobaciones que en una operación manual. Puede multiplicar el número de comprobaciones de coherencia interna, por ejemplo. Algunos planificadores, entusiasmados con esas extraordinarias posibilidades, han especificado tantas comprobaciones en la elaboración que prácticamente todo el cuestionario tiene que pasar a examen manual. Al igual que en la revisión manual, la utilidad de toda comprobación individual debería considerarse en relación con su costo y también el efecto cumulativo general de gran número de comprobaciones;

- e) La preparación por computadora se hace por fuerza en una etapa posterior a la manual. Como se dijo antes, conviene dirigirse a las personas que hayan contestado los formularios lo antes posible después de haber recibido éstos. La demora en volver a ponerse en contacto con tales personas como consecuencia de la elaboración de los datos con computadora ha de tenerse en cuenta cuando se recurra a ese sistema;
- f) Una diferencia importante entre la elaboración por cálculo electrónico y los métodos más antiguos, tales como los sistemas manuales y mecánicos o electromecánicos, estriba en que con la primera, las interpretaciones equivocadas que conducen a errores reiterativos o sistemáticos son a menudo bastante más difíciles de corregir. Muchas tabulaciones costosas han ido a parar a la papelera porque el programador de la computadora no comprendió bien lo que el planificador deseaba. En los métodos más tradicionales, esos errores se descubrían y rectificaban por lo general antes de que hubiesen hecho mucho daño. En cambio, la enorme velocidad de la elaboración con computadora hace que todo el proyecto esté muchas veces terminado antes de que se descubran los errores sistemáticos.

204. Es extremadamente importante que se deje tiempo suficiente para una detenida comprobación de los programas de computadora. Un procedimiento muy útil es pasar a cinta magnética un "lote de prueba" de tarjetas perforadas similares a las que eventualmente se elaborarán, pero con errores y otro material deliberadamente introducido para poner en funcionamiento todos los rasgos del programa. Con ello, éste puede probarse con un costo relativamente pequeño y los resultados usarse para corregir el programa si fuese menester. En la práctica, sin embargo, la producción de un lote de prueba que contenga todas las posibilidades no suele ser posible.

205. El uso de la computadora ha reducido mucho la cantidad de codificación requerida, especialmente la que ha de hacerse directamente por el personal administrativo. En primer lugar, ya no son necesarios los códigos cuyo propósito primordial es facilitar la clasificación de las fichas perforadas. La computadora puede operar igualmente bien sobre una clase de datos -número de empleados, por ejemplo- como con un código de tamaño de un dígito basado en la cifra original. Además, si se requiere este último código de un dígito por otras razones, la computadora puede fácilmente asignar uno. En segundo término, los empleados no suelen hacerlo muy bien cuando se trata de asignar códigos basados en un cálculo o comparación de magnitudes relativas, mientras que la computadora siempre es precisa en este punto, si ha sido programada debidamente. Por lo general, hay que preferir la computadora siempre que se vaya a asignar un número de código basado en un dato numérico, especialmente si ello entraña cálculos complejos. Los sistemas de codificación también pueden simplificarse; mientras que la tabulación mecánica puede facilitarse con un número de dígitos apreciables donde la posición de cada uno es significativa (por ejemplo, los primeros dos dígitos representan la provincia; los dos siguientes, el municipio; y los dos últimos, la ciudad o pueblo), las operaciones con computadora pueden ser igualmente eficientes usando un número de tres dígitos

asignado de modo consecutivo a las ciudades, manteniéndose en la memoria de la computadora la referencia al condado y la provincia. Muchas de las prácticas que se usan en los procedimientos manuales o mecánicos tendrán que revisarse si se va a emplear el cálculo electrónico.

206. La asignación de un código de tipo de actividad y ubicación por la computadora, sin embargo, entraña alguna programación extremadamente difícil y resulta también muy costosa, ya que requiere el uso de mucho tiempo de máquina y una computadora que tenga suficiente capacidad para almacenar gran cantidad de material de referencia. Estos y otros puntos que no suponen información numérica pueden muchas veces ser manejados mejor mediante el uso de personal.

Datos de registros administrativos

207. La disponibilidad de la información sobre empresas mercantiles reunida en el curso de la administración de diversos programas oficiales brinda la oportunidad de usar tal información como fuente original de datos y como instrumento para la preparación de los mismos. No obstante, es fundamental para ello un método sencillo que permita transferir datos de una fuente a otra. Uno de esos métodos consiste en proporcionar un número de identificación que sea común tanto a la fuente de registros administrativos como a la encuesta. Si bien es posible transferir los datos de los registros administrativos sin equipo mecánico o electrónico, tal cosa no resulta muy práctica. El tipo de equipo que resulta más idóneo es la computadora electrónica que puede clasificar, comparar y transferir datos con precisión y a gran velocidad.

208. Como se indicó con anterioridad, para aprovechar datos de registros administrativos como sustituto de la respuesta a un cuestionario, el organismo correspondiente ha de usar la misma clasificación de tipos de actividad, y los registros han de incluir las clases de información básica que necesita la encuesta, tal como empleo, nómina y volumen de ventas y cualesquiera otras características que no sean datos y que también se requieran, tales como forma jurídica de organización y emplazamiento. No obstante, un uso secundario, aunque también importante, de los registros administrativos es el de instrumento de preparación de los datos como fuente de información para el caso de las personas que no devuelven el cuestionario. Como tal instrumento, los datos de los registros administrativos proporcionan un método independiente de contrastar la precisión de los datos recogidos en la encuesta. Para los establecimientos que dejan de responder al cuestionario, el registro administrativo sirve de base para estimar la información necesaria cuando otros métodos resulten inadecuados.

209. Como organismos diferentes pueden emplear técnicas distintas de elaboración de datos, frecuentemente será necesario ajustar los ficheros de cinta a los diversos requisitos. Esto puede resultar bastante costoso y llevar mucho tiempo. Como en todo trabajo de computadora, hay que evaluar el tiempo y el costo antes de tomar una decisión.

Preparación de datos en encuestas repetitivas por muestreo

210. Como la realización de encuestas por muestreo de frecuencia más que anual obliga casi siempre a que se creen oficinas locales estratégicamente distribuidas por el país, la recogida y preparación de datos puede hacerse mejor por el personal sobre el terreno que por una oficina central de estadística. Es más fácil para resolver los problemas ponerse en contacto con las empresas directamente o por teléfono o correo. Esto tiene su importancia, ya que en esas encuestas suele hacerse hincapié en una publicación rápida de los datos, y las respuestas a las preguntas han de obtenerse en poco tiempo.

211. En aquellos casos que entrañen el uso de un encuestador para hacer una entrevista y obtener los datos en ese momento, esa persona puede ser formada para que efectúe al mismo tiempo algunas sencillas comprobaciones de revisión de los datos y anote las observaciones que le parezcan oportunas para explicar aparentes discrepancias. Estas pueden examinarse cuando entregue su trabajo a la oficina local. Cuando la encuesta suponga devolver el cuestionario por correo a esta oficina, cabe hacer en ella una serie de comprobaciones que incluirían no sólo la coherencia interna de los datos sino también una comparación de los mismos con los que notificó la misma empresa en períodos anteriores. Este es un rasgo muy útil de las encuestas repetitivas, ya que una misma empresa estará informando regularmente durante un largo período, lo que permite establecer unas pautas que faciliten la detección de errores.

212. El propósito primordial de la revisión de los datos sobre el terreno en las encuestas de frecuencia más que anual es asegurar que los cuestionarios se han respondido completa y correctamente antes de que se transmitan a la oficina central de estadística que ha de elaborar los cuestionarios de todo el país. En esa etapa, los formularios son objeto sólo de un examen limitado, ya que las tabulaciones se producen con un calendario de tiempo estricto. Todos los errores que se descubran después de la tabulación son difíciles de corregir.

IX. TABULACION, REVISION Y CORRECCION

Observaciones generales

213. Los tipos de equipo de tabulación de que se dispone normalmente oscilan en complejidad desde sencillas máquinas de sumar hasta costosas calculadoras electrónicas de gran velocidad y capacidad con su equipo auxiliar especializado. Cuando existan opciones para decidir qué equipo conviene adquirir, será menester ponderar muchos factores, como número de cuestionarios que hay que elaborar, cantidad de datos, alcance y tipo de la revisión que haya que hacer, número y complejidad de las operaciones de clasificación y tabulación, limitaciones de tiempo y costo de los sistemas alternativos. La decisión debería tomarse por personas que estén muy familiarizadas con todos los tipos de equipo disponible y que también comprendan bien el trabajo que se quiere hacer. Las ventajas teóricas de las computadoras deberían examinarse con cuidado, y los países que tengan experiencia en su uso para trabajos similares deberían ser consultados antes de tomar una decisión. Seguidamente se discuten los tipos principales de equipo de elaboración.

Máquinas de sumar

214. Si el número de cuestionarios es pequeño y el propósito es producir totales parciales para sólo una o dos características de los establecimientos, la tabulación con máquina de calcular puede ser más rápida y más eficiente que con otros métodos. Para tomar un ejemplo, supongamos que todas las partidas tienen que agregarse por a) tipo de actividad y b) zona geográfica. Después de la revisión y codificación, los cuestionarios se clasifican por zona geográfica y luego por tipo de actividad. Los datos de cada grupo de cuestionarios que representen un sector dentro de una zona se suman entonces con las máquinas. Si hay muchas partidas, la labor se facilitará usando una máquina que registre al mismo tiempo varias columnas. Este tipo de máquinas se llama a veces máquina contable. Debería ser del tipo que imprime los datos para verificación. Una vez comprobada la precisión, los totales obtenidos pueden recogerse en borradores sumarios, que pueden volverse a sumar para producir a) totales nacionales por tipo de actividad y b) totales por zonas.

215. Si se requiere una nueva tabulación -por ejemplo, un cuadro que indique algunas partidas o todas ellas clasificadas por tamaño de establecimiento- la información deberá clasificarse de nuevo y repetirse las operaciones. La duración de la clasificación a mano suele ser demasiado larga si los datos han de tabularse por más de dos características y por lo tanto deberían considerarse con tal fin otros medios. Existen sistemas que facilitan la clasificación, tales como tarjetas con orificios perforados marginalmente donde los datos se escriben a mano o se mecanografían. Tales sistemas son a veces muy útiles en las encuestas repetitivas pero no es probable que en las infrecuentes ofrezcan ventajas sobre la tabulación con máquina de sumar.

Elaboración electromecánica

216. El término "electromecánico" se usa para describir el tipo de equipo que supone usar tarjetas perforadas en la tabulación mecánica de los datos. Los tipos básicos de equipo que se necesitan en este sistema comprenden máquinas de perforación, máquinas de verificación de las tarjetas, clasificadores, intercaladoras, perforadoras de tarjetas reproductoras y máquinas de contabilidad (tabuladoras). Con este tipo de equipo, los datos de los cuestionarios se transcriben primero en las tarjetas con ayuda de la máquina perforadora, y la perforación se comprueba por otro operador con la máquina verificadora. Las tarjetas se clasifican seguidamente y se intercalan en caso necesario para la tabulación o enumeración que se desee; las tabulaciones se preparan para su utilización por la máquina contable.

217. El equipo de tabulación electromecánico varía en velocidad y capacidad. La primera depende en parte del número de totales parciales y generales que se requiera. Algunos modelos son capaces de calcular e imprimir totales parciales, intermedios y generales para el mismo campo de datos -por ejemplo, datos de provincias, ciudades y la nación en su conjunto. Algunos pueden producir nuevas tarjetas perforadas que muestren los totales resumidos para las categorías que se indiquen. Esas tarjetas con resúmenes pueden luego usarse en tabulaciones ulteriores.

218. Después de que los datos se hayan transcrito en las tarjetas perforadas, éstas pueden usarse en conexión con tabuladoras electromecánicas o como entrada de las computadoras electrónicas, ya sea directamente cuando se usen calculadoras pequeñas, o indirectamente, después de su conversión en cinta de computadora. El uso de tarjetas perforadas simultáneamente con clasificadoras y tabuladoras electromecánicas resulta especialmente útil en los censos económicos donde los datos básicos pueden clasificarse y relacionarse entre sí mediante diversas características de los establecimientos, tales como tipo de actividad, tamaño, ubicación y método de operación. Cuando se emplean en conexión con computadoras electrónicas, es posible incluso una mayor flexibilidad, velocidad y precisión.

219. Algunas máquinas perforadoras sólo sirven para perforar datos numéricos, mientras que otras también tratan información alfabética. Algunos tipos imprimen automáticamente cada carácter en la parte superior de la tarjeta cuando se está perforando, y esto puede ayudar en la verificación.

220. La perforación de tarjetas es una operación que no puede hacerse sin errores. Estos son frecuentemente superiores a los que surgen en la transcripción manual. Al no dejar los espacios debidos, por ejemplo, el perforador puede fácilmente transformar el número 123 en el número 123.000. Estos errores pueden resultar muy graves. Resulta así esencial que toda la perforación se someta a una labor de verificación que consiste en que otro operador perfore los datos en una máquina verificadora. Casi siempre, sin embargo, la verificación de la perforación resulta ser sólo parcialmente efectiva. En una operación de verificación de eficiencia media, cabe esperar que se detecten del orden del 95% de los errores.

221. La perforación y verificación normalmente se harán con mayor rapidez y eficiencia si las fichas están diseñadas de tal modo que las cifras se perforen en el orden en que aparecen en el cuestionario. En la elaboración electromecánica, es conveniente agrupar en una ficha las partidas que tengan que tabularse juntas. Si, para facilitar la perforación, las partidas no se han perforado en la secuencia adecuada para la tabulación, pueden crearse nuevas fichas con una máquina perforadora reproductora, pero tal cosa aumentará el costo. Si hay dos o más fichas por cuestionario, la cantidad de información de identificación y clasificación necesaria en cada una se efectúa también por el método de tabulación. Con la elaboración electromecánica suele reproducirse en cada ficha el número de identificación del establecimiento, los códigos de tipo de actividad y geográfico y todos los demás que se usen para la clasificación de los datos, tales como códigos de tamaño.

222. Si se emplean computadoras, podrán reagruparse muchas veces los campos de datos con un costo pequeño o nulo cuando se transfieren de las tarjetas a una cinta magnética. Cuando se usa con la computadora equipo de teclado directo a cinta, la información completa sólo tiene que ponerse en clave una vez, lo que no ocurre con el equipo electromecánico. La computadora puede entonces reunir electrónicamente el cuestionario completo de un establecimiento y volverlo a disponer en cualquier formato que se quiera.

223. Los adelantos de la tecnología han hecho posible evitar enteramente la producción de fichas perforadas como medio para la entrada de datos en la computadora. Los dos grandes tipos de sistemas son los que usan cinta como medio (teclas a cinta) y aquellos que tienen acceso directo a la computadora (teclas a disco). Existe también un sistema que usa un "lápiz electrónico" en lugar de una máquina perforadora para transferir datos directamente a la computadora. Si bien no es necesario crear primero una serie de fichas perforadas o una cinta para el acceso a la computadora, será posible, una vez que los datos se han introducido en ésta, preparar una cinta con tales datos, y también crear una serie de tarjetas perforadas a partir de la cinta de computadora.

Elaboración con computadora

224. El desarrollo y uso de calculadoras electrónicas en los últimos 20 años se ha sucedido a un ritmo vertiginoso. El resultado ha sido una tecnología en rápida expansión donde nuevas ideas han conducido a continuos cambios en el equipo y los procedimientos técnicos, de grandes consecuencias en la elaboración de datos. No obstante, tales cambios también han introducido mucha confusión conforme aparecían en escena varias "generaciones" sucesivas de computadoras y equipo auxiliar y surgían varios idiomas de máquina. En diversos lugares de este manual, se han hecho advertencias sobre los riesgos potenciales de lanzarse de cabeza a este método de elaboración de datos. No se trata de olvidar las posibles ventajas en velocidad, precisión y flexibilidad, sino simplemente advertir que la computadora sólo puede seguir instrucciones y que la redacción de las mismas ha de ser completa y precisa en todos sus detalles. Según sea el tamaño de su memoria y el tipo de equipo auxiliar que use, la computadora electrónica puede manejar cantidades enormes de datos, hacer muchas suertes de cálculos en respuesta a instrucciones

complicadas e imprimir los resultados en el formato deseado. La tabulación de los datos de censos industriales con máquinas de sumar puede teóricamente iniciarse con sólo una notificación previa de unas pocas horas, siempre que se disponga de los operadores de las máquinas. El uso de equipo electro-mecánico requerirá más tiempo para la planificación y para la conexión de los tableros de control. En la tabulación computarizada, sin embargo, se necesita mucho más tiempo previo. En primer lugar, tendrá que encontrarse un programador capacitado y le llevará tiempo familiarizarse con los objetivos del censo, especialmente con los planes y especificaciones para preparar y tabular datos. Tal cosa puede obligar al diseño de sistemas intensivos y extensivos antes de iniciar la programación. Seguidamente, tendrá que planear las líneas generales del programa de la computadora y trazar sus rasgos esenciales en forma de un diagrama de flujos. Antes de seguir adelante, el programador y el personal de planificación deberían discutirlo en detalle, tomándose todo el tiempo necesario para estar seguros de que se comprende completamente por todos los interesados. El objetivo de todo ello es poner de manifiesto conceptos e interpretaciones equivocadas que, de no detectarse, puede resultar sumamente costoso corregir después. Nunca se insistirá demasiado en la necesidad de una educación mutua (es decir, el programador ha de conocer los problemas del estadístico y viceversa) y de una clara comunicación, que es el principal problema de la computarización.

225. Una vez aprobado el plan general, el programador tendrá que escribir el programa propiamente dicho. El tiempo necesario dependerá de la complejidad del programa y del tipo de computadora que se use. Unos programas sencillos pueden escribirse en sólo unos pocos días; otros pueden llevar meses. Después de que se escriba, el programa ha de comprobarse. Esto se hace de ordinario con un "lote de prueba" de tarjetas perforadas del tipo que haya que tabular, que se diseña para poner de relieve los rasgos del programa y representar los problemas que se prevé surjan al abordar los datos reales. Una vez ensayado y corregido, el programa está listo para la preparación o tabulación de los datos recibidos.

226. La verificación automática de la perforación de la información codificada resulta posible mediante el dígito de comprobación -un número determinado rápidamente por la computadora a partir del número mismo del código y luego añadido al último dígito. Con un código de seis dígitos, por ejemplo, el dígito de prueba se convierte en el séptimo dígito. Los dígitos de prueba pueden calcularse de muchos modos. En una operación computarizada de preparación de datos, el dígito de prueba se calcula una segunda vez, y si el resultado no se ajusta al obtenido la primera vez, el código se señala como incorrecto.

Revisión y corrección de las tabulaciones

227. La eficacia del trabajo que se haya hecho para diseñar el sistema de recogida de datos y elaboración de cada uno de los cuestionarios puede comprobarse bastante bien cuando los datos se reúnen y presentan en forma de tabulaciones. Es sólo en esa etapa cuando cabe hacerse una idea sobre la mayor o menor validez de los resultados. Para ello, será menester disponer de varios instrumentos diferentes para que los aplique a la evaluación de las tabulaciones un personal que se haya formado en su empleo.

228. Si bien el uso de equipo complejo como computadoras electrónicas permite realizar pruebas más complejas de los datos, todavía no existe lo que sustituya a la persona que puede mirar una serie de números y, a la luz de su formación y experiencia, decir si tienen o no sentido. No obstante, el analista nunca debería hacer ningún ajuste arbitrario sobre la base de que, porque ciertos datos resulten incorrectos, están equivocados. A veces surgen situaciones insólitas, y los datos deben de contrastarse con el establecimiento que los proporcionó. La revisión de las tabulaciones finales no debería ser la ocasión de repetir la labor realizada en las etapas anteriores de la elaboración. En ese momento, la revisión de datos consiste en examinar resúmenes para determinar su carácter más o menos razonable. Sólo cuando el bloque de datos no supera la prueba debería hacerse un intento para determinar si el problema viene motivado por operaciones anteriores. Tendrá que establecerse un presupuesto adecuado para esta fase de la labor, ya que el no identificar y corregir los errores que produzcan datos inútiles significará que todo el esfuerzo aplicado a las operaciones anteriores se habrá malogrado. La tramitación equivocada de una información puede invalidar los datos de todo un sector de actividades.

229. Existen dos tipos principales de error que pueden ocasionar dificultades. El primero consiste en la inclusión de datos incorrectos (sea cual fuere el motivo) para uno o un pequeño número de establecimientos grandes, lo que se pondrá de manifiesto por el hecho no sistemático de que un bloque de datos no supere la prueba de revisión. El segundo tipo de error es sistemático y afecta a gran número de bloques para una partida determinada de datos. Este último tipo de error puede aparecer incluso cuando se están revisando datos para sólo pequeños establecimientos y creará muchas dificultades ya que tal vez requiera rehacer toda una operación.

230. Las tabulaciones preparadas en esta etapa, si bien están diseñadas para contener toda la información que se ofrecerá en la publicación final, no tienen que aparecer en la publicación misma. En cambio, deberían estructurarse básicamente para facilitar el análisis. La disposición y el espaciado han de hacerse de tal modo que faciliten la labor de quien revisa la información para aislar bloques de datos con problemas y proporcionar toda la asistencia posible para identificar las causas de esos problemas.

231. Existen dos tipos de comprobaciones al revisar los totales de los bloques: comprobaciones de coherencia interna y comprobaciones con fuentes de información externa disponible. Las primeras buscan aislar aquellos bloques de datos que quedan fuera de unos límites aceptables cuando un campo de datos se pone en relación con otro. En el análisis de las tabulaciones, sin embargo, los límites aceptables deberían ser considerablemente más estrechos de lo que eran para establecimientos individuales, ya que ahora estamos tratando con una acumulación de datos para un grupo de establecimientos donde el efecto de casos extremos aislados queda muy atenuado. Como corolario, cuanto mayor sea el número de observaciones en un bloque, menor será el límite aceptable y viceversa. Debería prepararse un conjunto de cuadros señalando los límites aceptables en términos del número de observaciones en un bloque. Al determinar las relaciones aceptables, deberían tenerse presentes factores tales como variaciones debidas a localización geográfica y tipo de actividad. Si se usa una computadora en la preparación de las

tabulaciones para su revisión, ese cuadro puede programarse en la computadora como un fichero de referencia, y los bloques que queden fuera de los límites establecidos pueden ser identificados por la computadora y señalados para recibir atención especial.

232. La tabulación que se encontrará más útil para detectar problemas es la que presenta los datos con mayor detalle. Usualmente consiste en un cuadro que proporciona datos con un desglose detallado de actividades para la zona identificable más pequeña y que presenta información básica sobre número de establecimientos, ventas o ingresos, empleo y nómina. La detección y corrección de errores significativos en este nivel de detalle conducirá, cuando se lleve a niveles superiores, a la producción de tabulaciones aceptables en todos los niveles en la medida en que se incluyan en ellos las mismas partidas de datos. Pero si hay otras características (tales como, por ejemplo, números de habitaciones de un hotel) que no figuran en otras tabulaciones, tendrán, claro es, que revisarse independientemente.

233. Además de los resultados de las comprobaciones de coherencia interna, los datos pueden revisarse contrastándolos con información disponible de encuestas anteriores y otras fuentes. Habrá alguna dificultad en ese tipo de análisis por causa de los cambios en los niveles de precios y de las relaciones de salarios a lo largo del tiempo. Esas magnitudes pueden variar mucho de un tipo de actividad a otro, y el analista ha de conocer los diferentes ritmos de cambio.

234. Una vez que un bloque de datos se identifica como posible problema, el analista ha de encontrar el medio de determinar qué es lo que está equivocado, si es que lo está. Si el problema parece ser sistemático, tendrá que retroceder la operación hasta el punto que pueda haber generado la dificultad. Para ello necesita estar familiarizado con todo lo que se ha hecho antes de la tabulación. Si el problema no es sistemático, tendrá que ser capaz de algún modo de identificar aquellos establecimientos dentro del bloque de datos que pueden haberlo ocasionado. Tal cosa significa que tendrá que examinar la fuente original de datos -el cuestionario- o su equivalente. Para ser accesibles, los formularios necesitan estar en la misma secuencia en que la tabulación se está revisando, preferiblemente por zonas, tipo de actividad y tamaño. Como el problema casi siempre se plantea respecto de grandes establecimientos, la clasificación de la información puede limitarse a ese grupo.

235. Cuando se usen computadoras, será posible preparar listas de varios tipos que ayuden a resolver muchos problemas sin necesidad de volver a clasificar los cuestionarios. Las máquinas de elaboración de datos pueden clasificar rápidamente muchos ficheros en las secuencias más útiles y eliminar la obligación de referirse a gruesos ficheros de información donde los formularios suelen estar perdidos o mal situados. Así, las listas pueden hacerse y encuadernarse en forma de libro para facilitar su consulta y para secuencias tales como zona por tipo de actividad, tipo de actividad por tamaño basado en diversas características, sector por zona, etc. Incluso con tales listas, sin embargo, puede haber necesidad a veces de revisar los cuestionarios.

236. Una vez que se ha llegado a la convicción de que hay que introducir una corrección, surge el problema del modo mejor de hacerlo. Cuando un bloque de datos figura en sólo un lugar, tal cosa no presenta mayores dificultades, pero en tabulaciones de datos económicos es corriente que los datos de un establecimiento aparezcan en niveles diferentes de zona y tipo de actividad y en clasificaciones transversales. Esto hace que sea esencial poder introducir la corrección en muchos puntos de las tabulaciones. Tales correcciones pueden hacerse a mano o llevarse a la ficha perforada o a los ficheros de cinta, ya sea al nivel del establecimiento individual o al nivel inferior de la ficha resumen, rehaciéndose todas las tabulaciones afectadas por la corrección. Si se prevé que los registros del establecimiento de que se trate se utilizarán para preparar tabulaciones adicionales en secuencias distintas de las que vayan a publicarse en la encuesta, será menester corregir tales registros, incluso aunque ello no sea requerido por los fines de la encuesta donde sólo se usan las fichas resumidas para preparar las tabulaciones finales.

237. Debería haber un solo documento de las fuentes que indicara todas las correcciones que hay que efectuar para el establecimiento o bloque de datos. Las anotaciones de datos nuevos o corregidos en esa etapa han de ser exactas, ya que se usarán para preparar los cuadros finales de las publicaciones, y la verificación de cómo se hacen las correcciones ha de ser absoluta. En este punto no hay margen para errores.

238. Incluso si se emplea equipo electromecánico o computadoras, es necesario comprobar que la suma de las cifras que figuren en el total de la columna vertical de una tabulación es igual a la suma de las cifras que aparecen en los epígrafes de la columna. Es sorprendente cuántas discrepancias cabe detectar así. Un estudio de la causa de esas discrepancias puede llevar al descubrimiento de serios errores.

239. Al igual que ocurre en la revisión de los datos de un bloque para ver si hay problemas, el momento más oportuno para investigar aquellos casos en que la publicación dará a conocer indebidamente información sobre las operaciones de un establecimiento o empresa es cuando los datos se revisan con mayor detalle. En ese momento, cabe averiguar si un solo establecimiento domina un bloque de datos (sobre tiendas de alimentación, por ejemplo, en una ciudad) en tal medida que la publicación de tales datos permitirá a la concurrencia, servicios de la administración u otras personas hacerse una idea muy clara de los ingresos u otras características importantes de esa unidad. Es en este punto donde resulta muy útil para fines de comparación enumerar las empresas mayores en la secuencia en que figuran en las tabulaciones.

240. El método que suele seguirse en el análisis sobre publicación indebida de información es suprimir los bloques que de darse a conocer, revelarían las actividades de una sola unidad estadística. En esa situación, la unidad es normalmente un establecimiento, pero la norma se extiende muchas veces a las actividades de empresas que poseen o controlan dos o más establecimientos. La publicación de datos de un bloque que contenga sólo dos unidades estadísticas está

prohibida, porque se revelaría la actividad de cada una a la otra. Así, los bloques publicables han de incluir, como mínimo, datos de tres unidades estadísticas.

241. Algunas estadísticas se consideran más confidenciales que otras. El volumen de ingresos y la cuantía de la nómina son ejemplos de datos que se consideran por lo general especialmente sensibles. El número de establecimientos suele considerarse no confidencial a menos que la clasificación sea tan detallada que revele información confidencial.

X. PUBLICACION

Examen final

242. Antes de que los datos se mecanografien o impriman (o antes de que los cuadros generados por una computadora para su publicación se hagan por última vez), deberían ser objeto de un examen final. Para esa labor, sólo debería recurrirse a las personas más capacitadas. Si bien, conviene disponer de empleados administrativos y otro personal que realice la labor necesaria para extraer la información que necesite el analista, la determinación de qué partidas parecen sospechosas y la decisión sobre qué es lo que convendrá hacer debería hacerse sólo por él. Su misión no es detectar pequeños errores o volver a comprobar revisiones, codificaciones y correcciones anteriores sino examinar los datos a la luz de sus conocimientos, experiencia e intuición. Si en esa fase final se descubre que la calidad de las estadísticas es escasa, ya sea por las respuestas o por errores de elaboración, o, en el caso de encuestas por muestreo, que la varianza de la muestra es demasiado grande, será mejor suprimir los datos más bien que publicarlos.

243. Al hacer un examen final, el analista, en toda la medida de lo posible, debería medir la magnitud de estimación y evaluación que hay en las cifras definitivas. Además, tendría que conocer todas las dificultades técnicas que se hubiesen encontrado y cómo se superaron, a fin de que tal información le sirva no sólo para su propia comprensión de las estadísticas sino también para preparar cualesquiera notas explicativas de pie de página o aclaraciones en el texto sobre las limitaciones de los datos.

Análisis último sobre publicación indebida de información

244. Al examinar las tabulaciones se habrá hecho un análisis directo de los datos para descubrir aquellos bloques cuya publicación conduciría a suministrar información inmediata o fácilmente deducible sobre las actividades de un establecimiento. No obstante, hay otro aspecto del análisis sobre publicación indebida que debería hacerse en el momento del examen final. Se trata de la búsqueda de bloques de información complementaria indebida. En los cuadros estadísticos con totales horizontales o verticales, surge una complicación cuando se suprime uno de los componentes de un total, pues si al menos un bloque adicional más de datos no se suprime también, el bloque original que suministraba información indebida podrá obtenerse sustrayendo la suma de los bloques publicados en detalle del total. Cuanto mayor sea la cantidad de detalles para los que se esté tabulando la información, mayor será el número de bloques de información indebida directa; cuando se agregan las informaciones complementarias indebidas, el resultado puede ser un número tan elevado de supresiones como para hacer que el cuadro resulte prácticamente inútil. Al planificar el nivel de detalle de lo que se publica, por tanto, debería tenerse presente el hecho de que el deseo de publicar más detalles puede conducir en realidad a que se suministre menos información.

Preparación y reproducción de ejemplares

245. Una vez examinadas las tabulaciones y hechas todas las correcciones oportunas, el paso siguiente será preparar los cuadros en su forma final para su conocimiento por el público en general. Como ya se analizó en el apartado sobre planificación de publicaciones, la labor básica sobre el contenido y formato de los cuadros debería haberse realizado con anterioridad en la fase de planificación, ya que muchas de las operaciones de la encuesta dependerán de lo que se vaya a publicar.

246. Un factor esencial para determinar el tiempo necesario y el costo de la preparación y reproducción de las publicaciones es el método que se vaya a seguir en estas dos últimas tareas. Fundamentalmente, existen tres modos de proceder. El primero consiste en escribir a mano los cuadros, incluidos los restos, epígrafes de columnas, títulos, notas de pie de página, etc., tras lo cual el manuscrito se envía a una imprenta para que lo tipografíen, impriman y encuadernen. Si bien este sistema conducirá a una publicación muy cuidada, tiene el inconveniente de llevar mucho tiempo y resultar costoso. El segundo método básico también supone escribir a mano los cuadros que luego se mecanografían exactamente como se publicarán y se reproducen en fotolitografía. Si bien este sistema no permite una publicación tan bien presentada, tiene la ventaja de ser menos costosa y más rápida. El tercer método fundamental consiste en la preparación de la publicación final en una computadora. Es el más rápido y potencialmente el menos costoso pero requiere una programación de cálculo electrónico extensa y difícil y debería hacerse sólo cuando el número de páginas de un determinado cuadro sea lo bastante grande como para justificar el esfuerzo de programación y el costo de la computadora. En este método, la reproducción también se haría en fotolitografía, aunque se están perfeccionando técnicas para usar la cinta de la computadora como medio para fotolitografiar. Los impresos de salida de la computadora también pueden usarse para escribir a mano y mecanografiar en conexión con cualquier sistema de reproducción en tipografía o fotolitografía.

247. Ningún medio para preparar el ejemplar final es infalible. Todos requieren un examen cuidadoso para estar seguros de que no se deslizan errores durante el proceso de escribir a mano, mecanografiar, tipografiar o hacer las operaciones en computadora.

Publicación de informes preliminares

248. La oficina de estadística estará siempre sometida a una gran presión para que publique los datos lo antes posible. Esto puede llevarle a la tentación de publicar información antes de que su calidad esté debidamente verificada. Si se encuentran después errores que requieren revisiones sustanciales, la oficina será objeto de muchas críticas. Una transacción que sin ser totalmente satisfactoria resulta sin embargo útil consiste en publicar informes preliminares que contengan bastantes menos detalles de los que figurarán en los informes finales y en donde se indique claramente que están siendo objeto de revisión. De ser posible, las

cifras resumidas que se recojan en las publicaciones preliminares deberían representar totales controlados que hayan sido confirmados y se repetirán en las publicaciones finales.

Descripciones de la encuesta para su publicación

249. Las recomendaciones internacionales adoptadas por la Comisión de Estadística piden que con objeto de que los usuarios de las estadísticas que se publiquen puedan evaluar su comparabilidad y fiabilidad, debería publicarse la información siguiente:

- a) Una descripción del campo de aplicación de la encuesta, es decir, de las actividades incluidas en el campo de la encuesta -y una definición de la unidad estadística en términos del campo de la encuesta y de las clases de datos reunidos;
- b) Una descripción de la cobertura de la encuesta -es decir, si se han incluido unidades de todos los tipos (tamaño, tipo de propiedad, clase de organización jurídica, etc.);
- c) Una descripción de los métodos adoptados para cubrir el campo de la encuesta -es decir, si se utilizaron procedimientos directos de recopilación (enumeración por correspondencia o sobre el terreno), registros administrativos o muestreo (incluida una descripción del diseño de la muestra utilizado y cálculos de probables errores del muestreo). La descripción también podría incluir una evaluación del grado de completabilidad de la cobertura lograda;
- d) Las definiciones operacionales de las clases de datos reunidos, quizás en forma de copias de los cuestionarios y de las instrucciones básicas utilizadas, incluida una descripción de los métodos de valoración utilizados respecto de las distintas clases. También se habrán de dar las definiciones de las clases derivadas que se hayan elaborado a partir de las clases de datos reunidos;
 - i) El número y la importancia de las unidades estadísticas conocidas que no hayan respondido al cuestionario y, cuando sea posible, algunas de las características fundamentales de los no informantes -es decir, su clase de actividad y, sobre todo, su tamaño. Además, se indicará también si en los datos publicados se han incluido estimaciones para los no informantes;
 - ii) La magnitud de la falta de respuesta a preguntas determinadas para las cuales no se hicieron estimaciones o no fue posible hacerlas;
- e) Una descripción de la clasificación por clase de actividad y de otras clasificaciones que se hayan utilizado 7/.

7/ Recomendaciones internacionales sobre estadísticas del comercio de distribución y los servicios, párr. 100.

250. Además, convendría incluir en el texto un examen analítico de los puntos más destacados de las conclusiones de la encuesta. El examen puede basarse en la preparación de un conjunto de cuadros muy resumidos que expondrían cosas tales como cocientes (por ejemplo, de inventarios y ventas, de nómina y ventas, de nómina y plantilla) por tipo de actividad y región geográfica, distribuciones porcentuales (por ejemplo, ventas por tamaño del establecimiento basado en el volumen de las mismas, número de empleados, región geográfica). Esto puede completarse con la preparación de gráficos que hagan las estadísticas más vivas y expongan aspectos importantes de la encuesta.

Contenido de los cuadros

251. Por ser de gran utilidad, especialmente en la presentación de datos de censos periódicos, en los cuadros principales debería figurar alguna información de los períodos de censos anteriores. Tal cosa puede incluirse en los cuadros resumidos que den datos por tipo de actividad y por zona. También sería útil que esos cuadros indicaran el cambio porcentual desde el censo anterior hasta el actual.

252. La utilidad de la información sobre el comercio al por menor y los servicios personales para los establecimientos que se dedican a esas actividades guarda relación directa con el tamaño reducido de la zona para la que se publiquen datos. Esto es especialmente cierto en el comercio de detalle. Los establecimientos no están tan interesados en los datos en términos de los límites materiales de ciudades y provincias. A pesar de la presión para producir datos sobre los grandes centros del comercio al por menor y los principales distritos de actividad mercantil, ese tipo de labor no debería emprenderse sin una clara idea de la cantidad de trabajo, tiempo y dinero que requerirá.

253. Un problema especial que se suscita en los cuadros relativos a los servicios es el que se debe al carácter heterogéneo de estos últimos. La cosa se complica todavía más por el hecho de que la recomendación sobre el campo de aplicación de los servicios, si bien debe prepararse "seleccionando las actividades con más probabilidades de ser comunes a todos los países", se matiza con la afirmación de que "no todos los servicios enumerados brindan las mismas posibilidades de reunión de datos y es difícil determinar prioridades separadas para cada uno de ellos. Es posible que en algunos países sea necesario introducir ajustes en el campo de aplicación de las encuestas, por ejemplo eliminar los servicios de poca importancia o los no existentes e incluir otros que tengan cierta importancia en el país de que se trate". De estas afirmaciones se desprende claramente que para la comparabilidad internacional, los totales de todos los servicios tienen que ser cuidadosamente examinados para saber su campo de aplicación en unos países y otros. También, los tipos de actividad que se combinan bajo este epígrafe -bienes inmuebles, producción de películas, hoteles, servicios de publicidad, emisoras de radio, lavanderías, servicios de elaboración de datos, etc.- hacen que sea difícil captar el significado del total de todo ello. Fundamentalmente, los datos de muchos de los servicios cobran sólo sentido cuando se examina cada sector de que se trate y se preparen cuadros

igualmente especializados. Cuando se reúne información de esa índole en cantidad suficiente, podrá estar justificada una publicación específica para ese sector.

254. Los cuadros que contengan datos basados en todo o en parte en el uso de muestras deberían claramente identificarse como tales. Los datos deberían siempre mostrarse sobre una base ampliada. Debería explicarse el hecho de que totales basados en el uso de muestras no sean iguales que otros basados en un recuento completo, y también deberían siempre incluirse las estimaciones de la varianza del muestreo.

255. Prácticamente todos los cuadros que se vayan a publicar consistirán en datos básicos clasificados en diversas relaciones recíprocas por características de establecimientos. El número de tabulaciones que podrían prepararse de ese modo es tan elevado que los países tienen que elegir los pocos que revestirán mayor interés general. Se sugieren ejemplos de algunos de los posibles cuadros mediante los títulos y las breves descripciones que se hacen a continuación. Los cuadros considerados se basan en la lista de estadísticas que habrá que tabular con arreglo al tipo de actividad que figura en las recomendaciones internacionales. La lista se adapta mejor a la tabulación de resultados de censos, ya que algunos puntos no se recomiendan para su recogida en encuestas anuales o más frecuentes. No se pretende ofrecer una lista completa, y los títulos no figuran necesariamente por orden de importancia. La presentación de datos por zonas geográficas, además de lo que se sugiere en el cuadro 1, debería determinarse por cada país a la luz de sus necesidades y recursos.

Cuadro 1. Principales indicadores de actividad, clasificados por grupo de la CIIU y por zona

Este cuadro presentaría componentes de datos básicos, tales como número de establecimientos, número de personas contratadas, número de empleados, sueldos y salarios abonados, valor de ventas de bienes e ingresos de explotación, valor agregado y formación bruta de capital fijo. Los grupos de la CIIU podrían ser los que figuran en el párrafo 62 de las recomendaciones internacionales. La clasificación por zonas correspondería a las grandes áreas administrativas del país, tales como estados, provincias, etc. Si se requiere un mayor detalle geográfico, probablemente sería necesario reducir el nivel de la clasificación por actividades al grupo principal de tres dígitos o a la división de dos dígitos de la CIIU.

Cuadro 1a. Comercio al por mayor y al por menor: indicadores principales de actividad, clasificados por grupo de la CIIU y por zona

Este cuadro presentaría, sólo para el comercio al por mayor y al por menor, los mismos componentes de datos básicos que el cuadro 1 más margen bruto, clasificados por grupos de la CIIU y por zona. Los grupos de la CIIU podrían ser los que se sugieren en el párrafo 62

de las recomendaciones internacionales o una clasificación nacional semejante. Como con el cuadro 1, cada país debería considerar el nivel de detalle de la clasificación geográfica en comparación con el nivel de la clasificación por actividades a fin de disponer de un cuadro útil de extensión razonable.

Cuadro 2. Indicadores principales de actividad, clasificados por división de la CIIU y por tamaño de establecimiento

Este cuadro presentaría los mismos componentes de datos básicos que el cuadro 1, clasificados simultáneamente por tipo de actividad en el nivel de dos dígitos de la CIIU y por tamaño de establecimiento. Los criterios obvios para la definición del tamaño son el número de personas que trabajen y el valor total de ventas e ingresos.

Cuadro 2a. Comercio al por mayor y al por menor: indicadores principales de actividad, clasificados por grupo de la CIIU y por tamaño de establecimiento

Este cuadro presentaría, para el comercio al por mayor y al por menor solamente, los mismos componentes de datos básicos que el cuadro 1a, clasificados además por tipo de actividad en el nivel de cuatro dígitos de los grupos de la CIIU sugeridos en el párrafo 62 de las recomendaciones internacionales o una clasificación nacional semejante, y por tamaño de establecimiento.

Cuadro 3. Indicadores principales de actividad, clasificados por división de la CIIU y por tipo de establecimiento

Este cuadro presentaría los mismos componentes de datos básicos que el cuadro 1, clasificados simultáneamente por tipo de actividad en el nivel de dos dígitos de la CIIU y por clases seleccionadas de organización económica y jurídica, incluida la propiedad. Si el número de clases de estas últimas características es grande, sería preferible en la práctica desglosar ese cuadro en dos o tres partes, una para cada característica.

Cuadro 3a. Comercio al por mayor y al por menor: indicadores principales de actividad, clasificados por grupo de la CIIU y por tipo de establecimiento

Este cuadro presentaría, para el comercio al por mayor y al por menor solamente, los mismos componentes de datos básicos que el cuadro 1a, clasificados en relación con el tipo de actividad en los grupos de cuatro dígitos de la CIIU sugeridos en el párrafo 62 de las recomendaciones internacionales o una clasificación nacional semejante, y por tipo de organización económica y jurídica entre clases seleccionadas, tipo de explotación y tipo de propiedad. Si el número de clases de las últimas características es grande, sería preferible en la práctica desglosar ese cuadro en dos, tres o cuatro partes, una para cada característica.

Cuadro 4. Indicadores de empleo y remuneración, clasificados por grupo de la CIIU

Este cuadro presentaría para el país en su conjunto componentes de datos sobre empleo y remuneración, tales como número de propietarios que trabajan y miembros de la familia que lo hacen sin remuneración, número de empleados (si es posible, distinguiendo los empleados de dedicación exclusiva), sueldos y salarios abonados (de ser posible, distinguiendo los pagados a empleados de dedicación exclusiva) y complementos de sueldos y salarios. Los datos se clasificarían en el nivel de cuatro dígitos de la CIIU o equivalente nacional.

Cuadro 5a. Comercio al por mayor y al por menor: valor de ventas e ingresos, clasificados por grupo de la CIIU

Este cuadro presentaría, para el comercio al por mayor y al por menor solamente, datos sobre la producción bruta total y sus componentes, tales como valor de la venta de mercancías, clasificado en ventas al por mayor (ordenadas a su vez en cuenta propia, cuenta ajena y comisiones percibidas) y ventas al por menor, e ingresos totales de explotación. Los datos se clasificarían en el nivel de cuatro dígitos de los grupos de la CIIU sugeridos en el párrafo 62 de las recomendaciones internacionales.

Cuadro 5b. Comercio al por mayor y al por menor: valor de las ventas, clasificado por tipo de producto

Este cuadro presentaría, para el comercio al por mayor y al por menor solamente, datos sobre el valor de las ventas de mercancías clasificadas en ventas al por mayor y al por menor por tipo de producto o grupos de productos. Cada país tendrá que redactar listas especiales para reunir información de las ventas por tipo de producto. Convendría que la lista de productos vendidos al detalle estuviera vinculada con la clasificación de bienes y servicios familiares a efectos de la contabilidad nacional 8/.

Cuadro 6. Valor de las compras y otros costos, clasificados por grupo de la CIIU

Este cuadro presentaría datos sobre el valor del insumo y sus componentes, tales como valor de las compras de bienes (aquellos destinados a la venta y otros), electricidad y combustible adquirido, trabajos de reparación y mantenimiento, actividades por contrata y comisión y otros costos de funcionamiento. Los datos se clasificarían en el nivel de cuatro dígitos de la CIIU o equivalente nacional.

8/ La orientación sobre el desarrollo de una clasificación de esa índole se encontrará en los estudios preparados por la Oficina de Estadística de las Naciones Unidas en el curso de la preparación de una clasificación generalizada de productos. Véase "Clasificación internacional uniforme de todos los bienes y servicios" (E/CN.3/493).

Cuadro 7a. Comercio al por mayor y al por menor: valor de las existencias, clasificadas por grupo de la CIIU

Este cuadro presentaría, para el comercio al por mayor y al por menor solamente, el valor de las existencias a principios y finales de cada año, distinguiendo las existencias de bienes para su venta de otras existencias, de ser ello posible. Otros puntos que hay que mostrar en ese cuadro son los cambios en las existencias durante el año y rotación de las mismas. La clasificación de los datos se haría en el nivel de cuatro dígitos de los grupos de la CIIU sugeridos en el párrafo 62 de las recomendaciones internacionales.

Cuadro 8. Formación bruta de capital fijo y su composición, clasificada por grupo de la CIIU

Este cuadro presentaría datos sobre el valor de las adquisiciones de activos fijos y adiciones brutas a los mismos, con arreglo al tipo de activo (esto es, edificios y otras construcciones, terrenos, equipo de transporte, maquinaria y otro material). La clasificación de los datos se haría en el nivel de cuatro dígitos de la CIIU o equivalente nacional.

Encuestas de frecuencia más que anual

256. En la publicación de los datos basados en encuestas de frecuencia más que anual, la fecha que hay que respetar reviste mucha mayor importancia que en el caso de encuestas infrecuentes, pero el hecho de que la cantidad de datos sea muy limitada hace que la preparación y reproducción de ejemplares sea un problema mucho menor. Los datos basados en muestras reducen al mínimo el problema de publicación indebida de información pero también tienden a hacer que el examen final sea más arduo, ya que resulta más difícil averiguar la causa de los errores de elaboración. En el texto debería hacerse una descripción de la muestra y un cuadro de las varianzas del muestreo para orientación de los usuarios de los datos. Como de ordinario la prensa está sometida a limitaciones de tiempo para la publicación de datos corrientes, es normal que el texto de la publicación comprenda un análisis resumido de los datos de tal modo que pueda recogerse en su totalidad o usarse con sólo modificaciones secundarias.

Anexo I

TECNICAS DE MUESTREO EN ENCUESTAS SOBRE EL COMERCIO DE DISTRIBUCION Y LOS SERVICIOS

1. El muestreo muchas veces resulta conveniente cuando hacen falta estadísticas pero no se necesita un censo completo del universo. Trabajar con una muestra en lugar de levantar un censo completo suele costar mucho menos y reduce enormemente el número de quienes han de contestar los cuestionarios, con el esfuerzo correspondiente. Se introduce un error de muestreo, pero de resultados de la menor labor que hay que hacer, el trabajo puede desempeñarse con más cuidado, disminuyendo posiblemente el error total. El muestreo también permite publicar mucho antes los resultados de la encuesta.
2. El muestreo aleatorio significa que todo miembro del universo tiene una probabilidad de selección no nula y la muestra final viene determinada por el azar y no por la opinión de nadie. Si se procede debidamente, los errores del muestreo aleatorio pueden evaluarse a partir de las propias muestras. Estas, sin embargo, se obtienen a menudo de otros modos, tomando, por ejemplo, aquellas unidades a las que se pueda acceder con más facilidad o aquellas otras que, a juicio de alguien, sean representativas del universo. El error de muestreo de las muestras obtenidas de esa suerte no puede calcularse; así, su valor queda muy reducido. El resto de este análisis se limita, por tanto, al muestreo probabilístico.
3. Como las muestras aleatorias son consecuencia de una selección al azar, es posible un gran número de resultados que varíen en mayor o menor grado del promedio (valor esperado) de esos resultados. Esta variación alrededor del resultado medio se denomina error muestral. La varianza se define como la desviación cuadrática media de cada resultado respecto del resultado global. La raíz cuadrada de la varianza se llama error estándar. En el supuesto generalmente razonable de normalidad, del error estándar cabe inferir una distribución de errores. El error estándar dividido por la estimación misma se llama coeficiente de variación. Esta medida del error relativo permite comparar el error de la muestra entre estimaciones muy diferentes en tamaño y características. El cuadrado de tal medida se denomina la varianza relativa que equivale a la varianza dividida por el cuadrado de la estimación. El cálculo del error muestral que se describe a continuación se hace en términos de la varianza y puede referirse fácilmente a las otras medidas señaladas, si así se desea.
4. El error de muestreo puede reducirse aumentando el tamaño de la muestra (en términos generales, el error estándar varía en proporción inversa a la raíz cuadrada del tamaño de la muestra). Un procedimiento que se usa mucha veces para tantear el error de muestras de diferentes tamaños es calcular (o suponer por experiencias pasadas) una varianza relativa para una sola unidad de la muestra y luego dividir ese valor por el número propuesto de unidades de la muestra a fin de obtener la varianza relativa del diseño propuesto. Por ejemplo, si suponemos

por experiencias anteriores que la varianza relativa por establecimiento comercial en una muestra de una lista es 1, entonces la varianza relativa de una muestra que conste de 100 establecimientos sería 0,01, y el coeficiente de variación sería 0,1, ó 10%. (Adviértase que para ese cálculo no se necesita ni el número de unidades del universo ni su tamaño medio.)

5. Hay otra clase de error en las encuestas que se llama error no muestral y que no es resultado del proceso de selección al azar ni, por lo general, se ve afectado por el tamaño de la muestra. Ese error, por ejemplo, puede deberse a deficiencias en el marco que se esté usando, a errores en la notificación, al hecho de que no todas las unidades informen, a defectos de elaboración, etc. El error de esa índole resulta especialmente peligroso porque no se puede medir con facilidad, a diferencia del error muestral. Debería mantenerse bajo una vigilancia razonable mediante programas adecuados de control de calidad, incluso si es necesario sacrificar para ello el tamaño de la muestra. Al decidir sobre el equilibrio entre el esfuerzo desplegado para reducir el error del muestreo y el que se dedica a disminuir el error no muestral, hay que recordarse que el objetivo es producir el menor error total posible con el presupuesto disponible, no el menor error muestral posible.

6. De este principio se sigue que un diseño sencillo que pueda ejecutarse en la forma prevista es preferible a otro más refinado que, aunque sea teóricamente más eficiente, está sujeto a serios errores no muestrales en su realización por causa de su complejidad.

7. El presupuesto de toda encuesta (y el correspondiente tamaño de la encuesta) debería fijarse idealmente en el punto donde la suma del costo de la encuesta más las pérdidas previstas por decisiones equivocadas hechas como consecuencia de errores en las estadísticas sea mínima. Por lo general resulta imposible determinar el presupuesto incluso de modo aproximado, principalmente por la dificultad de evaluar las probables pérdidas por decisiones. Además, los presupuestos no suelen estar bajo el control de los estadísticos. No obstante, prestar atención a ese principio puede impedir error de bulto, tales como levantar un censo completo cuando serviría una encuesta por muestreo o tomar una muestra demasiado pequeña.

8. El primer paso para obtener cualquier muestra es decidir el marco dentro del cual aquélla tiene que seleccionarse. En el caso de establecimientos comerciales, el marco es bien una lista de todos los establecimientos de esa índole, bien una lista de unidades con la que los establecimientos comerciales pueden ser asociados (por ejemplo, las empresas que sean propietarias de ellos o las zonas en que están ubicados).

9. Los marcos para muestras de establecimientos comerciales pueden clasificarse en dos categorías generales:

- a) Establecimientos o empresas propietarias (cuya lista se obtiene muchas veces de registros fiscales o de los ficheros de asociaciones comerciales). Las muestras obtenidas de ese tipo de marco se denominan en lo sucesivo muestras de lista;

- b) Zonas del país. Para evitar sesgos, todas las zonas deberían tener una posibilidad de selección. Las muestras obtenidas de ese tipo de marco se llaman en adelante muestras de área.

También se usan muestras combinadas si parte del universo está representada por una muestra de lista y otra parte por una muestra de área.

10. La muestra de lista suele presentar una ventaja en términos de error muestral sobre la muestra de área por los motivos siguientes:

- a) La muestra de área requiere una inspección personal para poner al día la lista de establecimientos comerciales en las zonas de que se trate. La inspección de las muestras de lista puede hacerse muchas veces fundamentalmente por correo/teléfono, que suele costar menos;
- b) Las muestras de lista pueden de ordinario estratificarse con más efectividad que las muestras de área;
- c) El muestreo por áreas es un muestreo por conglomerados que reduce casi siempre la eficiencia por caso seleccionado. Las muestras de área no suelen ser por lo general muestras por conglomerados.

11. Si bien las muestras de lista suelen tener una ventaja clara sobre las de área desde el punto de vista del error muestral, presentan serios problemas con el error no muestral. Muchas veces resulta muy difícil dar a cada establecimiento comercial una posibilidad de selección y sólo una a partir de un marco de muestras de lista, por las razones siguientes:

- a) El marco de muestras de lista nunca está actualizado para las altas recientes;
- b) El marco de muestras de lista suele omitir por definición parte del universo de interés. Muchas veces, aunque no siempre, esas omisiones son pequeños establecimientos;
- c) El marco de muestras de lista puede también omitir algunos establecimientos que por definición deberían figurar en él. Por ejemplo, quienes evadan los impuestos no figurarán en los registros fiscales; quienes no sean miembros no estarán en los ficheros de las asociaciones comerciales;
- d) El marco de muestras de lista muchas veces contiene duplicaciones; un solo establecimiento puede tener dos o más denominaciones que no sean exactamente idénticas en nombre, dirección u otra identificación;
- e) En una muestra de lista, los establecimientos que no informen pueden considerarse en actividad cuando en realidad no lo estén, o viceversa;

- f) En una muestra de lista suele seguirse la práctica de suprimir de la lista de direcciones aquellas unidades muestrales de las que se haya notificado que están fuera de actividad. Tales notificaciones pueden ser erróneas o anticuadas; unidades que no estaban activas en el momento de la notificación pueden haber sido reactivadas. Este factor tiende a producir un sesgo acumulativo a la baja en las estimaciones derivadas de muestras de lista.

12. Por lo general, el sistema de muestras de área, al basarse en la observación real, proporciona una base firme para hacerse una idea cabal de si los establecimientos comerciales están en la actualidad funcionando o no y para darles a cada uno de ellos una posibilidad de selección y solo una. No obstante, ese sistema también tiene problemas de sesgo en la cobertura. Los principales son los siguientes:

- a) Algunos establecimientos comerciales no son reconocibles como tales por observación exterior. En general, escapan a la atención del encuestador;
- b) Este cometerá errores al hacer las listas. Generalmente, incluirá menos establecimientos de los que debiera. Tales errores pueden minimizarse de ordinario mediante un programa de control de calidad.

13. Hay que señalar que es más probable que los errores de cobertura sean acumulativos en las muestras de lista que en las de aérea. Tales errores acumulativos son especialmente perturbadores, ya que distorsionan las estimaciones del cambio así como las de nivel.

14. La elección de un marco para una determinada encuesta debería hacerse sobre la base de la fuente del mismo que arroje el menor error total para el presupuesto disponible. Esto, a su vez, dependerá de las condiciones de la encuesta de que se trate. Con todo, las pautas siguientes pueden ser útiles:

- a) En buena medida, la elección depende de la calidad de las listas disponibles. Si éstas contienen muchas omisiones o duplicaciones, las muestras de área serán probablemente la opción mejor. Adviértase que en el apartado d) infra no se recomienda combinar una muestra de una lista incompleta con una muestra de área que subsane las deficiencias de la primera;
- b) Habría que evitar los marcos de muchas listas con coberturas que se traslapan. Teóricamente, en las encuestas cabe tener en cuenta la probabilidad de la duplicación, cuando ésta se conoce, pero los problemas prácticos de su determinación son grandes;
- c) Si se dispone de una simple lista adecuada o de un grupo de listas que no se traslapan y que juntas forman una lista adecuada, la opción no es tan clara;

- i) Para una encuesta que sólo se hace una vez (por ejemplo, una muestra derivada de un censo), la muestra de lista es la opción obvia;
 - ii) El sistema de muestras de lista puede también resultar práctico para las encuestas periódicas infrecuentes en donde se deja un largo período de tiempo para la tabulación (por ejemplo, seis meses o más). Esto es así porque se dispone de tiempo para resolver los problemas de cobertura del sistema de muestras de lista;
 - iii) Los problemas de cobertura de este sistema adquieren toda su gravedad en una encuesta periódica frecuente (mensual o trimestral, por ejemplo) que cuenta sólo con un breve intervalo entre el final del período de la encuesta y la tabulación (un mes o menos quizá). Aquí el sistema de muestras de lista es casi seguro que tendrá un problema sustancial de cobertura. El que sea más importante que el error mayor de muestreo del sistema de muestras de área dependerá en parte de la cuantía del presupuesto disponible. Como el error no muestral de la muestra de lista no disminuirá con un tamaño mayor de la muestra, puede constituir una parte importante del error total de una encuesta de presupuesto elevado pero una parte mucho menos importante de ese error en una encuesta de presupuesto pequeño;
 - iv) En parte, la decisión depende de si cabe reunir la información por correo/teléfono. De no ser así, las visitas personales a unidades de una muestra de lista muy diversas podrían resultar más costosas que las visitas a una muestra de área por conglomerados;
- d) Si se usa una muestra de área, se recomienda que se complemente con una lista de grandes establecimientos o compañías redactada con seguridad. Como estas unidades seguras se inspeccionan todos los meses, será fácil estar al corriente de su situación para que puedan "saltarse" adecuadamente en la muestra de área. Además, la posibilidad de duplicación requeriría que el informante contestara por partida doble. No se recomienda combinar muestras de lista y de área cuando la lista no esté redactada con seguridad. Aquí el problema es decidir qué unidades hay que suprimir entre aquellas de la muestra de área que han tenido una posibilidad de selección. Esto suele ser tan difícil que no se recomienda una combinación.

15. En los párrafos siguientes se describe un procedimiento para obtener una muestra de área de establecimientos comerciales. El primer paso consiste en decidir el carácter del conglomerado último, definido como un grupo de m segmentos de muestra de área conteniendo un promedio de \bar{N} establecimientos comerciales cada uno e incluidos en una unidad elegida de muestreo primario (UMP). Entre las decisiones que hay que tomar sobre el conglomerado final figuran:

- a) ¿Qué tamaño debería tener \bar{N} ?
- b) ¿Qué tamaño debería tener m ?
- c) ¿Cuál debería ser el tamaño y el carácter de la unidad muestral primaria?

16. Las respuestas a estas preguntas dependerán del costo y también de las condiciones de la varianza de la encuesta de que se trate. No obstante, las siguientes normas empíricas deberían dar resultados bastante próximos al óptimo en casi todas las situaciones:

- a) Se recomienda que \bar{N} , número de establecimientos comerciales dentro del campo de aplicación en cada segmento, sea lo más próximo a 6 que se pueda (4, si sólo se consideran establecimientos minoristas). Investigaciones empíricas hechas en los Estados Unidos de América encontraron óptimo ese tamaño para cada segmento, que puede usarse como punto de partida en otros países. No obstante, como el costo y las condiciones de la varianza pueden diferir apreciablemente en otros lugares, se recomienda que se hagan investigaciones concretas para determinar qué tamaño de los segmentos brinda mayor fiabilidad por unidad de gasto;
- b) Se recomienda que el número de esos segmentos seleccionados para la muestra en cada UMP aleatoria se fije lo más cerca posible de m , número de segmentos que puede fácilmente manejar un solo encuestador;
- c) Se recomienda que el tamaño y carácter de las unidades muestrales primarias se determinen con arreglo a las consideraciones siguientes:
 - i) Para cada una de las UMP debería disponerse de una medida de tamaño, tales como ingresos de los establecimientos comerciales o estimaciones de población de censos pasados;
 - ii) La UMP no debería ser tan grande como para requerir que el encuestador pase la mayor parte de su tiempo viajando de su domicilio al segmento o entre segmentos (tal vez el viaje medio domicilio-segmento no debería durar más de tres cuartos de hora, de ser posible);
 - iii) La UMP no debería ser demasiado pequeña o habría pérdidas considerables por una elevada correlación entre clases;
 - iv) Con las limitaciones anteriores, la UMP debería ser lo más heterogénea posible (por ejemplo, todas deberían ser una mezcla de zonas urbanas y rurales, si fuese posible).

17. Las UMP serán muy probablemente subdivisiones políticas o combinaciones de ellas. Si una muestra de área de hogares es objeto de encuesta en el país, muy posiblemente el mismo conjunto de UMP puede usarse para una muestra de área de establecimiento comerciales. No se recomienda, sin embargo, que la muestra de segmentos de área usada para la encuesta de hogares se emplee para la muestra de establecimientos comerciales, ya que la correlación entre hogares y establecimientos comerciales es muy baja en ese nivel.

18. El paso siguiente es determinar ℓ , número de los conglomerados finales que se introducirán en la muestra. Esto, claro es, viene determinado por el presupuesto que se tenga para tal fin dividido por el costo estimado de cada conglomerado. La cuantía del presupuesto debería determinarse por las consideraciones analizadas supra en el párrafo 7. Seguidamente debería tantearse el error muestral probable (coeficiente de variación, o CV) para las muestras de área de los diversos tamaños propuestos. Se calcula $\sqrt{m\ell}$, raíz cuadrada del número de segmentos del diseño propuesto. Se divide ℓ por ese número para obtener una aproximación del CV de nivel, 0,2 por este número para obtener un CV estimado del cambio mes a mes y 0,4 por este número para obtener una aproximación del CV del cambio de un año a otro. Supongamos, por ejemplo, que estamos considerando una encuesta que consta de 40 conglomerados finales (ℓ) con un promedio de 10 segmentos (m) cada uno. Tal cosa proporcionaría unos 400 segmentos estimados ($m\ell$). El CV de nivel sería entonces 5% ($1 \div 20$), el CV del cambio mes a mes sería 1% ($0,2 \div 20$) y el CV del cambio de un año a otro sería 2% ($0,4 \div 20$). Estas aproximaciones primeras pueden perfeccionarse cuando se disponga de datos de ensayos previos o de las encuestas propiamente dichas. Una vez que se haya determinado ℓ , cabe calcular P, índice de la muestra general.

$P = 6 \ell m/N$, donde N es el número total estimado de establecimientos comerciales del país. Una vez que se ha fijado P y se obtengan muestras con ese índice, no debería cambiarse incluso aunque N resultase estar equivocado.

19. Procédase a identificar todas las UMP ciertas en el país. Serán aquellas donde $S_j > \frac{S}{\ell}$, siendo S_j la medida del tamaño, S la misma medida para el país y ℓ es el número de conglomerados finales de la muestra, tal como se definió antes. Por ejemplo, supongamos que la población se usa como medida del tamaño al obtener las UMP. Si la población total estimada (S) es 10.000.000 y hay que obtener 100 conglomerados finales (ℓ), entonces debería tomarse con seguridad toda UMP con una población que pasara de 100.000 ($\frac{S}{\ell}$).

20. Después hay que estratificar todas las UMP aleatorias restantes. Los estratos deberían ser de un tamaño lo más próximo posible a $2S/\ell$, en términos de la medida del tamaño (S) que se está usando. En nuestro ejemplo, procuraríamos tener estratos con poblaciones de 200.000. Las UMP situadas en un estrato dado deberían ser lo más parecidas posible.

21. Luego hay que subdividir cada estrato formado como se ha dicho en el párrafo anterior en dos substratos. También ellos deberían ser lo más iguales posibles (el tamaño ideal en el ejemplo es ahora S/ℓ , o 100.000).

22. Selecciónese una UMP de cada substrato formado con una probabilidad de selección proporcional a S. Para seguir con el ejemplo, supongamos que tenemos un substrato dado donde figura una UMP A con una medida de tamaño de 10.000 (personas, en este caso), una UMP B con una medida de tamaño de 30.000 y una UMP C con una medida de tamaño de 60.000. Estableceríamos una suma acumulativa c como sigue:

<u>UPM</u>	<u>Medida del tamaño</u>	<u>Medida acumulativa del tamaño</u>
A	10.000	10.000
B	30.000	40.000
C	60.000	100.000

Luego seleccionaríamos un número al azar entre 1 y 100.000. Todo número comprendido entre 1 y 10.000 seleccionaría la UMP A; todo número entre 10.001 y 40.000 seleccionaría la UMP B, y cualquier número entre 40.001 y 100.000 nos daría la UMP C. Las probabilidades de selección son $10.000/100.000$ (ó 0,1), $30.000/100.000$ (ó, 0,3) y $60.000/100.000$ (ó, 0,6) para las UMP A, B y C, respectivamente.

Nota: Este método de selección conduce a una UMP muestral por estrato. Por consiguiente, resulta imposible obtener una estimación insesgada de la varianza. Una sobreestimación de la varianza se obtiene por el llamado método de estratos agrupados para el cálculo de la varianza que se describe a continuación. En él se supone que las dos UMP se obtuvieron del mismo estrato.

23. Existen métodos que permiten seleccionar dos UMP por estrato grande con probabilidad proporcional al tamaño sin reposición. No obstante, son más complicados y tienen una varianza real superior a la calculada en el sistema de muestreo descrito aquí.

24. Otra alternativa posible es obtener dos UMP por estrato grande con probabilidad proporcional al tamaño con reposición. Este método brinda una estimación insesgada de la varianza, pero la varianza real es superior a la calculada con el método antes indicado.

25. Una vez identificada la UMP cierta y seleccionada la UMP aleatoria, el paso siguiente en el proceso de muestreo de área es subdividir toda la zona de las UMP seleccionadas en segmentos de muestras de área. Cada uno de éstos debería reunir los rasgos siguientes:

- a) Límites razonablemente permanentes (calles, carreteras, ferrocarriles, ríos, etc.) que sean reconocibles por los encuestadores;
- b) Seis establecimientos comerciales, en toda la medida de lo posible. Si el menor segmento posible con límites reconocibles contiene 10 o más establecimientos, debería designarse "de no inclusión de todos" con una media de tamaño igual al número estimado de establecimientos dividido por 6 (cifra redondeada en unidades).

26. Para realizar el proceso de segmentación, será menester contar con mapas que muestren aspectos reconocibles, tales como ríos, carreteras, etc., y la ubicación de establecimientos comerciales. Si se dispone ya de ellos (por ejemplo, de un censo anterior de establecimientos comerciales), deberían usarse. No es necesario que sean totalmente precisos o estén actualizados. Si no existen, los mapas deberían confeccionarse mediante inventario directo de las UMP de las muestras. Partiendo de unos mapas que sean lo más detallados posibles (ampliados, en caso necesario), los encuestadores enumerarían todos los establecimientos comerciales (determinados por observación exterior) y agregarían cualesquiera rasgos reconocibles que no figurasen ya en los mapas.
27. Con unos mapas de los que ya se dispusiera o que se hayan confeccionado del modo indicado en el párrafo anterior, la muestra UMP debería segmentarse en la oficina según las normas del párrafo 25. Toda la zona de la muestra UMP debería tenerse presente en este proceso de segmentación.
28. Los segmentos de la muestra UMP (ciertos y aleatorios) deberían numerarse sistemáticamente (por ejemplo, en forma de serpentina). En la enumeración, cada medida de un segmento de no inclusión de todos los establecimientos debería tratarse como un segmento separado.
29. Ahora debería calcularse el índice del muestreo interior. Para la UMP no probabilística ese índice es P , el general del muestreo. Para UMP aleatorias, el índice interior es $P_{ghj} = P \cdot S_{gh}/S_{ghj}$.
30. En la ecuación anterior, S es la medida del tamaño usado en el proceso de muestreo de las UMP, S_{gh} es la medida del substrato h dentro del estrato g -ésimo (grande) y S_{ghj} es la medida de la UMP j -ésima dentro de ese estrato. S_{gh}/S_{ghj} , por tanto, es el inverso de la probabilidad de elegir la UMP.
31. Para seguir con nuestro ejemplo, supongamos que se seleccionó la UMP A; en ese caso, $S_{gh}/S_{ghj} = 1/0,1$, es decir, 10. Si suponemos una fracción de muestreo general (P) de 1 cada 100, el muestreo interior en la UMP sería $(1/100) 10$, ó 1 cada 10. La muestra se selecciona en realidad usando el inverso de la probabilidad interior de selección, $1/P$ o $1/P_{ghj}$, como intervalo, con un punto de partida aleatorio comprendido entre 0 y ese número. Para elegir los segmentos, la pauta resultante se aplica a los segmentos numerados sistemáticamente de que se ha hablado en el párrafo 28. Los segmentos de cada UMP no probabilística se asignan sistemáticamente conforme se van obteniendo a cinco grupos aleatorios para facilitar el cálculo de las varianzas descrito en el párrafo 63.
32. Merece la pena casi siempre completar el muestreo por áreas con una lista de establecimientos (o empresas) comerciales muy grandes. Tal lista puede obtenerse de cualquier fuente, pero todos sus miembros deberían ser elegidos con certeza y enumerarse para cada período de encuesta. El límite inferior de esa lista podría fijarse inicialmente en $S/6m1$, ó 5 dividido por el número total estimado de establecimientos en la muestra de área. S es hasta cierto punto una medida del tamaño (ventas, empleo, nómina) disponible tanto a nivel

nacional como para cada establecimiento grande. No tiene por qué ser la misma medida usada en las UMP del muestreo. Por ejemplo, si los ingresos totales estimados del comercio al por menor en un país son de 1.000 millones de dólares y esperamos 2.000 casos para el muestreo por áreas, entonces el límite inferior por encima del cual se incluirían a todos los establecimientos sería $\$1.000.000.000/2.000$, ó 500.000 dólares, y habría que hacer lo necesario para que figurasen en la lista no probabilística todos los establecimiento de ese tamaño o mayores.

33. Esta es una aproximación muy poco precisa que supone fundamentalmente que la varianza mayor y el costo superior por establecimiento del muestreo de área (en contraste con establecimientos de un muestreo de lista) se anulan recíprocamente en la fórmula de la afijación óptima. Cabe determinar empíricamente un límite inferior más exacto experimentando con diferentes niveles después de que se disponga de algunas listas del muestreo por áreas. En esos experimentos, el costo total debería igualarse haciendo la muestra de área mayor o menor, en la medida necesaria para compensar las variaciones en el tamaño de la muestra de lista.

34. Las instrucciones al encuestador deberían consistir en enumerar todos los establecimientos comerciales que se encuentren dentro de los límites de los segmentos de la muestra, con las excepciones siguientes:

- a) Los establecimientos de la lista no probabilística deberían suprimirse de la muestra de área. Para facilitar ese proceso, el encuestador debería recibir una lista de los establecimientos (o empresas) incluidos con certeza y que se sabe operan en su UMP o podrían hacerlo.
- b) Para los establecimientos que no se incluyen todos, aquellos que figuren deberían elegirse por muestreo en la proporción de k por cada K , donde esta última es la medida del tamaño asignado al segmento y k es el número de "eliminaciones automáticas" de la muestra (usualmente una) que no quedan en el segmento. El muestreo puede hacerse sobre el terreno por el encuestador o bien en la oficina. En el primer caso, el punto de partida y el intervalo deberían designarse previamente y el encuestador recibir normas de enumeración que no dejen margen para decidir una cosa u otra. El muestreo en la oficina entraña un riesgo menor de sesgos pero puede llevar bastante tiempo y aumentar los costos.

35. Habría que comprobar para la calidad una submuestra de la lista y de la enumeración de cada encuestador. Ello resultará más efectivo si se hace independientemente, es decir, el encuestador y el revisor prepararán listas independientes y las cotejarán. La revisión resulta también más efectiva si no se hace con una pauta fija que permita al encuestador saber qué meses se comprobarán.

Nota: Los temas de ponderación, cálculo de la varianza, uso de información auxiliar en los procedimientos de estimación y rotación de muestras se aplican a las muestras tanto de lista como de áreas y se examinan en los párrafos 58 a 95.

36. Si se decide usar el marco muestral de listas, el primer paso es estratificar estas últimas. La estratificación significa colocar a los miembros de la lista en clases (estratos) de las que se obtendrán muestras independientes. El objetivo debería ser que las unidades muestrales de los estratos fuesen lo más parecidas posibles; los criterios mejores para estratificar son el tipo de actividad principal y el tamaño, si se conocen. Cinco o seis estratos de tamaño aleatorio dentro de cada tipo de actividad deberían ofrecer casi todas las ventajas, sin resultar demasiado complejas. No obstante, si la estratificación va a hacerse por un encuestador en el transcurso del trabajo de campo, se recomienda en aras de la sencillez un solo estrato aleatorio. El objetivo debería ser que los estratos aleatorios fuesen de un tamaño lo más parecido posible. No obstante, para simplificar cabe usar los mismos límites de tamaño para cada tipo de actividad. Tal cosa haría que fuesen iguales los estratos de tamaño en el nivel de todos los tipos de actividad pero desiguales dentro de cada tipo de actividad por separado.

37. El tamaño total de una muestra de lista vendrá determinado por el presupuesto total disponible (con exclusión de los costos fijos) dividido por el costo por unidad de la muestra. La cuantía total del presupuesto debería determinarse por las consideraciones expuestas en el párrafo 7. Al evaluar el error de muestreo, cabe inicialmente suponer que el coeficiente de variación de una estimación de nivel de una muestra de lista será 1 dividido por la raíz cuadrada del número de unidades de la muestra. El coeficiente de variación del cambio mes a mes podría ser del orden de 0,3 dividido por la raíz cuadrada del número de unidades de la muestra; 0,5 dividido por la raíz cuadrada del número de unidades de la muestra brindaría una primera aproximación del coeficiente de variación del cambio de un año para otro. (Véase el párrafo 18 para un ejemplo de cómo hacer esos cálculos sin usar el número de unidades que hay que obtener en la muestra de listas.) Si la clase no probabilística es importante, háganse esas aproximaciones para la clase aleatoria y multiplíquense luego los resultados por P_{NC} , proporción de los ingresos totales que corresponden a la clase aleatoria. Esas primeras aproximaciones poco precisas pueden perfeccionarse conforme se vaya disponiendo de datos de la encuesta.

38. Un método sencillo para asignar la muestra total entre los diversos tipos de actividad consiste en hacer que la afijación sea proporcional a la medida total del tamaño de cada tipo de actividad. Esto producirá un total general casi óptimo. El error de muestreo (coeficiente de variación) para cada estimación del tipo de actividad será en líneas generales inversamente proporcional a la raíz cuadrada de su tamaño. Claro está que la distribución puede alterarse si se opina que algunos tipos de actividad son más o menos importantes de lo que su afijación indicaría.

39. La afijación óptima de la muestra a un determinado tipo de actividad entre los diversos estratos de tamaño viene dado por la fórmula:

$$n_h = \frac{N_h \cdot h}{\sum_{n=1}^R N_h \cdot h} n_o$$

En esta fórmula, n_0 es la muestra total asignada al tipo de actividad (como en el párrafo 38 anterior), n_h es la muestra asignada al estrato h-ésimo de los de tamaño R dentro del tipo de actividad, N_h es el número de unidades muestrales en el estrato h-ésimo de tamaño mientras que σ es la desviación estándar (para un sola extracción) del estrato h-ésimo. Si la n_h indicada por esta fórmula es igual o mayor que las unidades N_h del estrato, todas ellas se toman sin excepción.

40. Una aproximación sencilla y bastante buena de la afijación anterior es la siguiente:

- a) Fíjese el límite inferior de los establecimientos que se incluyen todos en X/n_0 , donde X es la medida total del tamaño para el tipo de actividad y n_0 la muestra asignada a ese tipo;
- b) Asígnese la muestra restante en proporción a X_h (medida total de tamaño en el grupo de tamaño h-ésimo) en el tipo de actividad de que se trate. Por ejemplo, supongamos que una muestra de 100 establecimientos se ha asignado al tipo de actividad que tiene una medida total de tamaño (por ejemplo, nómina) de 100.000.000 de dólares. Entonces todo establecimiento con una nómina de 1.000.000 de dólares o más se tomaría de modo cierto. Si hubiera 10 establecimientos de esa índole a los que correspondiera 20.000.000 de dólares de nómina, seguiría habiendo entonces para distribuir una muestra aleatoria de 90 establecimientos. Si hubiera cinco estratos aleatorios, a cada uno se le asignaría una muestra de 18 establecimientos.

41. En la aproximación anterior se supone que los coeficientes de variación (para una sola extracción) son los mismos en todos los estratos. En realidad, tales coeficientes suelen ser mayores para las clases de menor tamaño. A veces se obtiene una afijación algo mejor añadiendo un pequeño incremento a cada medida del tamaño antes de proceder del modo descrito en los párrafos 39 y 40 anteriores. Por ejemplo, si el empleo se está usando como medida del tamaño, cabría agregar un empleado a la medida de tamaño de cada establecimiento (o empresa, si ésta es la que se usa como unidad muestral) antes de asignar la muestra.

42. Para extraer la muestra indicada en cada estrato de tamaño y tipo de actividad, en primer lugar hay que hacer una lista y recuento sistemáticos de todas las unidades muestrales del estrato (por tamaño, tipo de actividad, ubicación, etc.). Después hay que seleccionar una muestra aplicando la configuración que se obtiene al usar el intervalo $I_h = \frac{N_h}{n_h}$ (redondeado al entero más próximo) y un punto de partida aleatorio entre 1 é I_h . Conforme se van obteniendo, las unidades muestrales deberían asignarse sistemáticamente a los cinco grupos o réplicas aleatorias a los efectos de los cálculos de la varianza descritos antes.

43. Como ya se ha dicho, los problemas principales de muestrear una lista para una encuesta recurrente consisten en lograr que cada establecimiento tenga una posibilidad y sólo una de selección y tomar las decisiones correctas para que los

miembros de la muestra estén en activo en ese momento. En los párrafos siguientes se hacen recomendaciones para reducir al mínimo esos problemas.

44. Debería llevarse un registro con el universo de todas las unidades que hayan tenido alguna vez una posibilidad de elección. Las unidades recién aparecidas y otras unidades de las que se sospeche que sean nuevas pueden cotejarse con esa lista para ver si tuvieron una posibilidad anterior de selección.

45. Debiera disponerse de una fuente fidedigna de nuevas unidades que deberían mostrarse y agregarse a la muestra lo antes posible.

- a) Si en el momento de darse de alta nuevos establecimientos no se dispone de información fiable para la estratificación (tipo de tamaño y (o) tamaño), debería usarse un muestreo bietápico. Habría que obtener una muestra inicial grande para lograr información sobre la estratificación. Después debería extraerse una muestra menor del primer muestreo para que el producto de las probabilidades primera y segunda sea apropiado para el estrato final. El tamaño de la muestra de la primera etapa de los establecimientos recién aparecidos (es decir, el ritmo al que se extrae aquella muestra) no es susceptible de una optimización matemática explícita. Como con otros problemas de optimización de esa índole, cabe experimentar empíricamente tan pronto como se tengan datos de la encuesta. Ese experimento debería hacerse con arreglo a las líneas siguientes:
 - i) Con las unidades de la muestra, se puede volver a crear un universo ponderando cada unidad por el inverso de la probabilidad de selección. Tal universo incluiría tanto la muestra original como los establecimientos recién creados;
 - ii) Experimentétese con diversas fracciones muestrales de la primera etapa para las apariciones recientes en el universo simulado creado antes. Adviértase que fracciones mayores en la primera etapa suscitarán menos valores atípicos (causados por los casos donde la probabilidad en la primera etapa es inferior a la requerida por su estrato final de tamaño y tipo de actividad). No obstante, unas fracciones mayores en la primera etapa redundarán en un costo más elevado por causa de la muestra mayor;
 - iii) En cada nivel de la muestra simulada de la primera etapa, estímesese el costo de esta muestra. Ajustése el tamaño total de la muestra para tener en cuenta esas diferencias en los costos de la muestra de la primera etapa a fin de que se mantenga constante el costo total;
 - iv) Calcúlese la varianza general en cada nivel usando las probabilidades indicadas y selecciónese el diseño que brinda el equilibrio óptimo entre los efectos de las altas recientes atípicas y el tamaño total de la muestra.

Hasta que se disponga de datos de la encuesta, las probabilidades iniciales de la muestra de primera etapa de altas recientes deberían ser bastante elevadas para hacer improbable que surjan valores atípicos. Por ejemplo, si es posible que las altas recientes sean lo bastante numerosas como para justificar una selección sin aleatoriedad, entonces la muestra inicial de la primera etapa debería tomarse también sin aleatoriedad. Tener una muestra inicial grande de primera etapa de establecimientos que hayan aparecido recientemente también hará que la experimentación descrita antes resulte más efectiva, ya que el universo simulado de altas recientes será cercano al universo real de dichas altas;

- b) Para el muestreo de las altas recientes se usan dos sistemas principales. En uno, sólo se muestrean las altas reales (es decir, sin predecesores); los sucesores se representan sustituyéndolos a los miembros correspondientes que estén en la muestra. En el otro sistemas, se muestrean todas las altas y se eliminan todos aquellos miembros de la muestra que dejen de ejercer la actividad de que se trate, independientemente de que tengan un sucesor o no. Como no existe una definición clara de lo que es un sucesor y como no siempre es fácil obtener información fidedigna al respecto, el sistema de muestrear todos los establecimientos que aparezcan por vez primera se presta probablemente menos a un sesgo acumulativo.

46. Se recomienda que se investigue la situación de quienes no responden a la encuesta para determinar si están todavía en actividad. Cuando tal cosa se haga, cabe tomar la decisión correcta sobre si procede asignar una estimación distinta de cero a la unidad correspondiente. También se recomienda que las bajas se investiguen al menos una vez al año para determinar si han sido reactivadas.

47. Se recomienda que la muestra se vuelva a extraer a menudo por los motivos siguientes:

- a) Para reducir el efecto de sesgos cumulativos;
- b) Para extender entre más establecimientos el trabajo de tener que dar información;
- c) Para evitar incrementos en la varianza por cambios en las medidas del tamaño.

La muestra puede obtenerse independientemente o con una superposición máxima o mínima respecto de la muestra anterior. La obtención de una muestra independiente es con mucho el método más sencillo y que permite la mayor rotación posible.

48. Una nueva extracción de una muestra de lista es probable que cree una discontinuidad entre los resultados antiguos y los nuevos. Este es uno de los inconvenientes de usar muestras de lista para encuestas recurrentes. Es verdad que la muestra de área también tiene que volverse a extraer eventualmente pero ese paso hay que darlo menos a menudo que en las muestras de lista y, como se presta

menos a sesgos cumulativos, no es tan probable que conduzca a discontinuidades importantes.

49. Extraer una muestra de un censo completo representa otra aplicación de los principios del muestreo por listas a establecimientos comerciales. Debería hacerse (para ahorrar tiempo y dinero tanto para los encuestados como para el gobierno) cuando algunos de los datos no hay que detallarlos por capítulos concretos (por ejemplo, ingresos por línea de productos, inventarios, gastos de capital). Esa aplicación evita muchos de los problemas descritos en los párrafos 43 a 48 anteriores, ya que el censo completo de datos básicos para aspectos tales como número de establecimientos e ingresos totales proporciona una lista del universo en un momento que resulta bastante actualizado, completo y sin duplicaciones. No obstante, esa aplicación presenta por su parte algunos problemas especiales que dependen del método de recogida de información que se haya seguido para el censo completo y que se describen a continuación.

50. Existen tres métodos generales para levantar un censo completo de establecimientos comerciales. El primero es mediante una inspección personal donde los encuestadores confeccionan una lista de establecimientos mediante una enumeración personal. El segundo consiste en inventariar (generalmente por correo/teléfono) una lista preestablecida de establecimientos comerciales obtenida casi siempre de registros administrativos. Un tercer método es recopilar los datos básicos que se buscan en los propios registros administrativos. El censo puede hacerse combinando esos tres métodos, usando cada uno de ellos para partes diferentes del universo. Seguidamente se describen los problemas de muestreo que surgen con cada uno de esos tres procedimientos para censos completos.

51. Si un censo (o parte de un censo) se hace mediante inspecciones personales sin lista preexistente, la muestra de establecimientos debería obtenerse entonces durante la inspección o después de ella. Teóricamente, la muestra podría extraerse antes del censo diseñando una muestra por áreas, pero no parece haber muchos motivos para aceptar las ineficiencias del muestreo de área cuando se dispondrá como resultado de la labor de inspección de una lista actualizada. Si la muestra se obtiene en el transcurso de la enumeración, tal cosa debería ser extremadamente sencilla y no dejar margen para que el encuestador tuviera que tomar decisiones por su cuenta. Por lo general, conviene tener dos estratos, uno no aleatorio o cierto y otro aleatorio. Por ejemplo, podría fijarse como normas tomar en la muestra de modo cierto todos los establecimientos con 10 o más personas empleadas y una muestra de uno por cada 10 de aquellos establecimientos con menos de 10 empleados (el límite inferior para la inclusión de todos los establecimientos y el porcentaje del estrato aleatorio vendrían determinados por los principios examinados en los párrafos 36 a 42). El encuestador debería recibir unas normas inflexibles para determinar qué establecimientos corresponden al muestreo aleatorio. Por ejemplo, si la muestra está designada previamente de la lista, habría que dar al encuestador una ruta de inspección que debería seguir sin desviaciones. Alternativamente, cada cuestionario podría tener una designación de la muestra bajo sello. La norma en este caso sería no quitar el sello hasta que se haya terminado el censo de puntos básicos. Cuando el encuestador lo levanta, determina entonces si el establecimiento corresponde a la muestra o no.

52. A pesar de las precauciones anteriores, las muestras obtenidas por el encuestador durante la enumeración es casi seguro que estén sesgadas, aunque a menudo el sesgo es pequeño y no invalida los resultados. El efecto del sesgo del encuestador (o del causado por otros factores) puede muchas veces reducirse con el procedimiento de estimación que se describe más adelante en el párrafo 57.

53. La alternativa a que sea el enumerador quien obtenga la muestra durante la labor de inspección es extraerla una vez terminado el censo básico. En este sistema, el encuestador envía por correo los cuestionarios del censo básico a la oficina nacional o a las regionales, donde se muestrean (probablemente conforme se van recibiendo). Este método permite usar sistemas de estratificación más complejos y más eficientes. También reduce el riesgo de sesgos. No obstante, puede resultar imposible por dos razones:

- a) Aumentará la cantidad total de tiempo necesario para levantar el censo, ya que introduce una demora entre el censo básico y el censo inspeccionado por muestreo. Además de causar retrasos en la publicación, ese factor tal vez introduzca sesgos, ya que un número apreciable de establecimientos del sector puede darse de baja en el intervalo.
- b) Incrementará el costo total, ya que tendrá que hacerse una labor de inspección por separado (ya sea personal, por correo o por teléfono) para completar el censo por muestreo. Si el encuestador hace este muestreo al levantar el censo básico, puede coger todos o parte de los cuestionarios de la muestra o al menos dejarlos a los entrevistados.

54. Si un censo (o parte de un censo) se obtiene mediante una inspección (probablemente por correo/teléfono) de una lista preexistente, habrá entonces la posibilidad de obtener la muestra antes de que se envíe por correo el censo básico o después de que se hayan recibido en la oficina los cuestionarios con la información básica. En ambos casos, el muestreo se hace en la oficina con lo que cabe usar diseños más complicados y eficientes que apliquen los principios indicados en los párrafos 36 a 42. No obstante, si el muestreo se hace antes de que se envíen los cuestionarios del censo básico, tal vez no se dispondrá en grado adecuado de la información sobre tamaño y tipo de actividad que se necesita para la labor de estratificación. Esto introduce un problema especialmente difícil si el cuestionario de la muestra que hay que enviar depende del tipo de actividad de quien haya de responderlo.

55. Si el muestreo se efectúa después de que los cuestionarios del censo básico se hayan recibido en la oficina, se dispondrá de información mejor sobre tamaño y tipo de actividad para usarla en el proceso de estratificación. También aquí, sin embargo, el método presenta serios inconvenientes de tiempo y costo por el hecho de que al levantamiento del censo de la muestra tenga que seguir (a veces con una demora de tiempo apreciable) al del censo básico.

56. Si todo el censo básico o parte de él se obtiene recopilando datos de registros administrativos, una vez más cabe elegir entre obtener la muestra antes o después de que los datos que se vayan a usar en el censo se reciban. Muchas

veces no se dispone de una lista puesta al día de la que sea posible obtener una muestra hasta después de que se hayan recibido los datos corrientes. A menudo es tan tarde que resulta imposible obtener y encuestar una muestra de la lista. Por tanto, los datos de la muestra no podrán obtenerse para esa porción del universo, especialmente si se trata de una pequeña parte del mismo.

57. Con independencia de cómo se obtenga y encueste la muestra, se dispondrá de los datos básicos (por ejemplo, número de establecimientos e ingresos totales) tanto para el universo como para la muestra. Esos datos pueden usarse para mejorar las estimaciones de la muestra mediante una estimación proporcional de la forma $\frac{Y'}{X'} X$. En ella Y' es una estimación de alguna característica disponible solamente con el cuestionario de la muestra; el símbolo X' representa una estadística correlacionada que se obtiene del censo básico; X representa el mismo total estimado de la muestra. (Las X escogidas deberían ser aquellas que se consideren más estrechamente correlacionadas con la estimación Y). Por ejemplo, supongamos que Y' representa estimaciones de una determinada línea de productos (por ejemplo, prendas de vestir en unos grandes almacenes) y X representa las ventas totales de los grandes almacenes. Entonces, una estimación normalmente mejor (por ejemplo, Y'') de las ventas de prendas de vestir en esos almacenes puede obtenerse con la estimación $Y'' = \frac{Y'}{X'} X$. El uso de ese tipo de estimación presenta tres ventajas:

- a) Y'' (estimación proporcional) puede tener una varianza menor que Y' (estimación simple) si hay suficiente correlación entre X e Y .
- b) La estimación puede usarse como medio de calcular los casos de no respuesta con los datos de la muestra (Y). Tal cosa supone que la X (total del censo) es completa;
- c) El procedimiento puede usarse para reducir el efecto de los sesgos en el proceso de muestreo, introducido por los encuestadores o por otras fuentes. Ese método supone que la relación entre Y y X es la misma para la población real que para la muestra obtenida imperfectamente. Si bien tal cosa no es necesariamente cierta, muchas veces está más cerca de la verdad que la alternativa de suponer que la Y' simple está insesgada;
- d) Si los ingresos totales se desglosan en sus componentes (tales como ventas por línea de productos), entonces la suma de los componentes se agregará a los ingresos totales del censo básico (siempre que X' se haga igual a Y' cuando se suman todos los componentes por productos).

En relación con las ventajas b), c) y d) señaladas antes, ese procedimiento no debería usarse para ocultar defectos apreciables de un censo por muestreo. Si los valores (simples) de X' difieren de los de X (censo) más de lo que cabría razonablemente esperar por el error del muestreo, entonces los usuarios deberían ser informados de los cocientes $\frac{X}{X'}$ usados, para que sepan cuánta corrección se ha

introducido con ese procedimiento en los sesgos muestrales o en la evaluación, o en ambas cosas.

58. En los párrafos siguientes se examinan algunos principios del muestreo que se aplican a las muestras tanto por listas como por áreas. El primer punto se refiere al proceso de ponderación. Después se analizan, por orden, cálculo de las varianzas, rotación de muestras, uso de información auxiliar para reducir las varianzas y los peligros de una reducción arbitraria por ponderación.

59. La ponderación para producir estimaciones simples insesgadas es siempre la inversa de las probabilidades de selección. En el caso de las muestras de área, la probabilidad de selección es el producto de la probabilidad en las diversas etapas. En el caso de una muestra bietápica, tal cosa sería la probabilidad de obtener la unidad muestral primaria por el número de veces que represente la probabilidad de extraer los segmentos dentro de dicha unidad.

60. En el ejemplo que se ha dado en los párrafos 29 a 31, la UMP se seleccionó con una probabilidad de 1 a 10 y los segmentos dentro de esa unidad se obtuvieron con una probabilidad de 1 a 10 también. No es coincidencia que la probabilidad general de 1 a 100 sea igual al índice (P) designado previamente de 1 a 100. El coeficiente de ponderación que hay que usar en ese caso es 100. Si la muestra de área se obtiene del modo recomendado, entonces la probabilidad de selección (P) es uniforme para todas las partes de la muestra, y los resultados se ponderan por 1/P. (Una muestra donde los coeficientes de ponderación son uniformes se suele denominar muestra autoponderada.)

61. En el caso de la muestra por listas, el coeficiente de ponderación es el intervalo (I_h) usado en el proceso de selección (véase párrafo 42) ó 1, si se toma dicha muestra de modo no aleatorio.

62. Las estimaciones simples insesgadas (ponderadas) se usan siempre como estimaciones finales propiamente dichas o como bloques para construir las estimaciones más complicadas que se describen más adelante. También se usan en los cálculos de la varianza, también directamente o como bloques de construcción, si se usa uno de los procedimientos de estimación más complicados.

63. Para los diseños aquí recomendados (tanto para muestras de área como de lista), la varianza de la estimación simple insesgada puede calcularse con la fórmula siguiente que sirve para cualquier muestra estratificada:

$$\sum_{n=1}^R (1-f_h) \frac{m_h}{m_h-1} \sum_{j=1}^{n_h} (Y'_{hj} - \bar{Y}'_h)^2 \quad (1)$$

En la fórmula anterior, h es el estrato h-ésimo (de los estratos totales R). En la muestra por áreas cada una de las UMP no aleatorias constituyen un estrato separado. En la parte aleatoria de la muestra de área, cada uno de los estratos originales de los que se obtuvieron los dos últimos conglomerados constituye un estrato, independientemente de que esos conglomerados se obtuvieran creando dos substratos y obteniendo uno de cada uno de ellos u obteniendo los dos últimos conglomerados del estrato original (con o sin reposición).

En la muestra de lista, cada celda de tamaño por tipo de actividad de la que se extrajo una muestra independiente constituye un estrato, y f_h es la fracción de muestreo usada, o probabilidad (inverso del coeficiente de ponderación) empleada en el estrato h -ésimo. Cuando se recurra a la técnica del estrato agrupado (es decir, cuando sólo se haga una extracción por substrato), f_h se fija en cero. Y'_{hj} es el valor simple ponderado de la réplica j -ésima en el estrato h -ésimo. En un muestreo monoetápico, Y'_{hj} es el valor ponderado del grupo j -ésimo de las extracciones muestrales que se hicieron en el proceso de muestreo. En el polietápico, Y'_{hj} es el valor ponderado del conglomerado último j -ésimo en el estrato h -ésimo si las extracciones se hicieron en realidad de los estratos o (como en el método del estrato agrupado) extrayendo un conglomerado último de cada uno de los dos substratos.

El número de los totales de grupo descritos antes en el estrato h -ésimo es m_h . En la porción aleatoria de la muestra de área, $m_h = 2$ si la extracción se hace del modo recomendado. \bar{Y}'_h es la media simple de la inspección Y'_{hj} descrita antes, esto es:

$$\bar{Y}'_h = \frac{\sum_{j=1}^{m_h} Y'_{hj}}{m_h}$$

Si la estimación que se usa es más complicada que la simple estimación insesgada, entonces la varianza se calcula de la misma forma que antes pero sustituyendo U'_{hj} a las Y'_{hj} descritas antes. Las U'_{hj} son combinaciones lineales de las Y'_{hj} . Se indican las U'_{hj} para las estimaciones complicadas a que se hace referencia concreta más adelante. Para un método general de calcular las U'_{hj} de cualquier estimación complicada, se remite el lector a la bibliografía; el método ha sido computarizado en un programa que puede obtenerse de la Oficina del Censo de los Estados Unidos de América. Ese programa calcula varianzas de estimaciones complicadas sin ninguna entrada más que la requerida para obtener la propia estimación.

64. En las encuestas recurrentes de establecimientos comerciales suelen hacerse rotaciones de las muestras a lo largo del tiempo. Hay dos motivos para ello:

- a) Extender el trabajo de suministrar la información entre un mayor número;
- b) Reducir el error de muestreo de las estimaciones de la encuesta.

65. La renovación periódica recomendada de las muestras por listas para reducir sesgos cumulativos (párrafos 47 a 48) puede lograr la finalidad de extender el trabajo del suministro de la información sin aumentar la rotación de la muestra. Como el sesgo cumulativo en una muestra por áreas es un problema menos grave, las renovaciones de la muestra pueden ser demasiado infrecuentes como para conseguir la difusión deseada del trabajo de proporcionar la información. Para ello cabe introducir alguna rotación adicional de la muestra.

66. La rotación de la muestra para conseguir la reducción de la varianza es más bien complicada y entraña por tanto el riesgo de un sesgo adicional. Se recomienda un panel no rotatorio si existe alguna duda sobre la capacidad para manejar esas complicaciones adecuadamente. Seguidamente se indican los principios de un plan básico de rotación. Para más detalles, véase la bibliografía.

67. Se recomienda usar una vez tras otra un número limitado de paneles rotatorios (llamados un conjunto cerrado). Esto tiene dos ventajas con respecto a usar un panel nuevo para cada nuevo período de información:

- a) Cuesta una considerable cantidad de dinero preparar un panel para la enumeración, ya se trate de un panel de lista o un panel de muestra de área. Al limitar su número, se limita el costo inicial;
- b) Si usamos un conjunto cerrado de paneles, podremos aplicar el amplio procedimiento de observación descrito antes que puede ser de gran interés para reducir la varianza.

68. Un número conveniente de paneles que cabe usar en un conjunto cerrado para encuestas mensuales es 12. Esto se debe a que un panel puede ser designado como el de enero y se enumerará ese mes todos los años, otro en febrero y así hasta diciembre. Tal cosa significa que las estimaciones de cualquier mes del año se basarán en el mismo conjunto de paneles para todos los años.

69. Esto mejorará las estimaciones importantes del cambio de un año para otro porque existirá una correlación año a año. Si los costos iniciales son demasiado elevados para 12 paneles, entonces cabe usar un número menor. Si el número de paneles es 2, 3, 4 ó 6, los mismos se repetirán todos los años. Si se usan 3, el mismo conjunto se repetirá cada trimestre. Claro está que reducir el número de paneles en el conjunto reduce algo las ventajas que cabe obtener, pero incluso un conjunto tan reducido como 4 permitiría extraer casi todo el beneficio en la varianza producido por la rotación.

70. Existen dos tipos diferentes de estructuras de información que cabe usar con un conjunto de paneles rotativo. Si pueden obtenerse dos meses de datos de cada entrevistado en cada inspección, entonces todo el panel debería ser sustituido en el siguiente período de información. Para ilustrar esto, pondremos el ejemplo de un panel de enero de una serie de 12 mensuales. A principios de febrero cuando se haga la inspección, si podemos lograr que el encuestado suministre los datos de enero y diciembre (por ejemplo, ingresos de ambos meses), entonces dejaríamos el panel de enero en su totalidad para la inspección siguiente y sólo enumeraríamos el panel de febrero. En cambio, si no pudiésemos obtener información fidedigna sobre las ventas de diciembre a principios de febrero, entonces sólo pediríamos las ventas de enero a partir del panel de ese mes. No obstante, a fin de lograr ventas de establecimientos idénticos, retendríamos el panel durante un mes más y pediríamos las ventas de febrero en el momento de la inspección del panel de febrero a principios de marzo. Con ese sistema, estamos utilizando cada panel dos veces al año en lugar de una

vez. Si quisiéramos igualar la labor de inspección con el primer sistema, entonces cada panel tendría que ser sólo la mitad de grande.

71. Evidentemente, el primer sistema ofrece más información por unidad de la muestra. Al usarlo, podemos basar nuestra medida del cambio mes a mes para establecimientos idénticos de todo el panel en lugar de para la mitad. No obstante, ese primer sistema lleva consigo el riesgo de un sesgo adicional que no es trivial. La información obtenida en el mes corriente ha de tener el mismo valor esperado que la información obtenida en el mes anterior. Usemos nuevamente enero como ejemplo. Se obtendrán dos informes relativos a la actividad en ese mes. Uno (la información corriente) se obtendrá del panel de enero inmediatamente después de que haya transcurrido ese mes. Otra estimación para enero se obtendrá del panel de febrero a principios de marzo. Existen varias razones por las que el valor esperado de esas informaciones puede diferir. En primer lugar, si la información procede de registros, los de enero pueden estar más completos a principios de marzo de lo que lo estaban a principios de febrero, lo que conducirá a niveles diferentes de ventas notificadas. En segundo término, si los propietarios del establecimiento estiman la información, pueden muy bien tender a notificar las ventas de enero en marzo según las condiciones del negocio en ese momento más bien que según las vigentes en enero. Y en tercer lugar, es muy difícil tratar las altas y bajas adecuadamente cuando se obtienen dos meses en una inspección. Por ejemplo, supongamos que un establecimiento está operando a primeros de febrero pero ha dejado de hacerlo a primeros de marzo. En una enumeración por muestreo de área, el encuestador incluirá el establecimiento y registrará sus ventas en enero como parte de la inspección para el panel de ese mes. No obstante, si ese establecimiento cayera en el panel de febrero, sus ventas de enero probablemente no figurarían, ya que el establecimiento no existiría en el momento de la inspección de dicho panel.

72. A pesar de las ventajas tan atractivas del primer tipo de sistema de información, no debería usarse hasta que se hayan adquirido pruebas empíricas de que la información para el mes actual y el anterior tienen fundamentalmente el mismo valor esperado. Esto es especialmente cierto porque la estimación compuesta que se describe más adelante magnifica cualquier diferencia que exista en las estimaciones simples insesgadas.

73. Un método para reducir la varianza de las estimaciones de nivel con paneles rotativos es usar la estimación compuesta, con la cual usamos la información que ha sido notificada en los períodos anteriores para reforzar la estimación actual. La forma de la estimación compuesta depende del tipo de estructura de notificación. Para el primer tipo, la estimación compuesta que denominaremos Y_i''' (el subíndice i se refiere al mes cuyos datos se están estimando) tiene la forma:

$$Y_i''' = (1-K)Y_i' + K \frac{Y_i'}{Y_{i-1}''} Y_{i-1}''' \quad (2)$$

En la fórmula anterior de estimación supondremos una vez más que estamos obteniendo una estimación para el mes de enero. En tal caso:

Y_i''' = la estimación compuesta del mes i (enero)

Y_i' = la estimación actual simple insesgada para el mes i-ésimo (la estimación simple insesgada para enero del panel de ese mes)

Y_{i-1}'' = la estimación simple insesgada del mes anterior para el mes i-1 (la estimación simple para el mes de diciembre obtenida del panel de enero). Adviértase que las estimaciones Y_i' e Y_{i-1}'' se derivan por tanto del mismo panel

Y_{i-1}''' = la estimación compuesta para el mes anterior (diciembre) que se obtuvo de la misma fórmula

K = constante que determina el peso relativo que se da a las dos estimaciones que se combinan. Depende, a su vez, de la correlación mes a mes de las estimaciones simples derivadas del mismo panel. Si llamamos a esa correlación p el valor óptimo de K es

$$K = \frac{1 - \sqrt{1 - p^2}}{p}$$

Si K se fija en 0,72, será óptima cuando $p = 0,95$ y se acercará al óptimo cuando se aproxime a ese valor. Cuando $p = 0,95$, la varianza de la estimación compuesta (Y_i''') será del orden de un tercio de la varianza de una estimación de una muestra no rotatoria del mismo tamaño.

74. Si se usa el segundo tipo de estructura de notificación, tendremos dos estimaciones simples insesgadas disponibles para el mes de enero (una del panel de enero que llamaremos Y_i'' y una obtenida del panel de diciembre que denominaremos Y_i'). Nótese, sin embargo, que cada una de tales estimaciones se basa en paneles de la mitad de tamaño. La estimación compuesta en este caso reviste la forma siguiente:

$$Y_i''' = K \frac{Y_i'}{Y_{i-1}''} Y_{i-1}''' + LY_i' + (i-K-L)Y_i'' \quad (3)$$

Adviértase una vez más que las Y_i' e Y_{i-1}''' son del mismo panel (diciembre) mientras Y_i'' es del panel de enero. Y_{i-1}''' es la estimación compuesta tal como se calculó con una fórmula de ese tipo para el mes de diciembre.

75. Las constantes óptimas (una vez más en términos de las correlaciones mes a mes de paneles idénticos) son:

$$K = \frac{1 - \sqrt{1 - p^2}}{p}$$

$$L = \frac{(1 - \sqrt{1 - p^2})(1-p)}{p^2}$$

Si la correlación mes a mes se sitúa en el nivel de 0,95, tales constantes tienen los valores siguientes:

$$K = 0,72$$

$$L = 0,04$$

76. La estimación compuesta de nivel tendrá una varianza un poco inferior a la mitad de la de una muestra no rotatoria del mismo tamaño.

77. Si bien las varianzas de las estimaciones de nivel mejoran sustancialmente con una estimación compuesta de esa índole, las varianzas de las estimaciones del cambio mes a mes empeoran. Para el primer tipo de estructura de notificación con la correlación y las constantes supuestas, la varianza de la estimación del cambio mes a mes aumenta aproximadamente en un tercio. La varianza de esa estimación se duplica con creces cuando es necesario usar el segundo tipo de estructura de información. Hay que señalar que mientras la estimación compuesta ocasiona una pérdida en las estimaciones del cambio mes a mes, esas pérdidas pueden compensarse por las reducciones muchas veces muy sustanciales que cabe lograr mediante el uso del procedimiento amplio de observación que se describe en los párrafos 84 a 87.

78. Es evidente que el uso de paneles rotatorios y del procedimiento de estimación compuesta vale mucho más la pena si es posible obtener dos meses de información con cada inspección (es decir, primer tipo de información). De no ser así, podría muy bien ocurrir que las complicaciones de ese sistema no lo hiciesen aconsejable.

79. Las varianzas para la estimación compuesta (fórmulas 2 y 3) se calculan con arreglo a la norma general del párrafo 63, esto es, se produce una variable de sustitución U'_{hj} para cada extracción de la muestra en cada estrato y ese valor reemplaza a las Y'_{hj} en la fórmula de estimación de la varianza (fórmula 1). La forma que reviste las U'_{hj} para estimaciones del primer tipo de forma de estimación compuesta es:

$$U'_{(i)hj} = Y'_{(i)hj} + (K) \frac{Y'_j}{Y''_{i-1}} \left[U'_{(i-1)hj} - Y''_{(i-1)hj} \right] \quad (4)$$

K es la constante generalizada que sugerimos se fijase en 0,72.

80. Los símbolos tienen el mismo significado que en la fórmula 2 excepto que el subíndice hj significa que sólo se incluyen datos ponderados de la obtención j -ésima del estrato h -ésimo. Adviértase, sin embargo, que el cociente $\frac{Y'_i}{Y''_{i-1}}$ se sigue calculando de los totales generales.

$$\frac{Y'_i}{Y''_{i-1}}$$

81. Puede demostrarse que la suma $\sum_{h=1}^R \sum_{j=1}^M U'_{(i)hj}$ es exactamente igual a la Y''_i calculada en la fórmula 2. Por tanto, las $U'_{(i)hj}$ se calculan de cada extracción y se suman simplemente para obtener los totales nacionales deseados. Hay también que señalar que este es un proceso interactivo, es decir, las U'_{hj} calculadas para cada mes i se usan como parte de la entrada de datos para el cálculo de las U'_{hj} del mes $i + 1$.

82. Para estimaciones derivadas del segundo tipo de estructura (fórmula 3), las U'_{hj} toman la forma siguiente:

$$U'_{(i)hj} = (K+L)Y'_{i(hj)} + K \frac{Y'_i}{Y''_{i-1}} (U'_{(i-1)hj} - Y''_{(i-1)hj}) + (i-K-L)Y''_{(i)hj} \quad (5)$$

83. Los símbolos tienen el mismo significado que en la fórmula 3 excepto que, una vez más, los subíndices hj indican que sólo se usan datos ponderados 0,04 y 0,72, respectivamente, cuando se usan las primeras aproximaciones descritas en los párrafos 73 a 78.

84. Si se emplea un conjunto cerrado de paneles rotatorios, de 12 meses, por ejemplo, cabe usar un procedimiento amplio de observación para reducir la importancia de los valores atípicos de una forma insesgada. Tal procedimiento puede reportar algunas reducciones apreciables de la varianza, especialmente para el índice de variación mensual. Esto es importante porque la varianza de ese índice aumenta cuando se recurre al procedimiento de estimación compuesta indicado en los párrafos 73 a 78.

85. Supongamos una vez más que tenemos 12 paneles. El peso de cualquier valor atípico puede reducirse a un doceavo de su tamaño original para todo establecimiento que esté por encima del límite inferior de los establecimientos que se incluyen. Para ello, claro está, los establecimientos han de ser enumerados todos los meses, en lugar de hacerlo una vez cada 12. Este será un procedimiento insesgado, siempre que el establecimiento reciba el mismo tratamiento con independencia de en cuál de los doce paneles le corresponda caer. Para asegurar ese mismo tratamiento, usaremos el procedimiento siguiente: Fíjese un límite inferior para cada mes del año; para el mes i -ésimo, le denominaremos C_i . Ese límite puede ser el mismo para todos los meses del año, pero si existe una estructura estacional clara, sería mejor variar tales límites de modo aproximado a dicha estructura. Los límites se fijarían en términos de valores ponderados (de ingresos, por ejemplo, o de alguna otra característica). El modo mejor de determinar el límite óptimo es mediante experimentaciones empíricas con datos reales tan pronto como se disponga de ellos (en condiciones de costos fijos) pero, como aproximación inicial, se sugiere que el límite se establezca en un nivel igual a cinco veces las ventas mensuales ponderadas medias que se prevén de la muestra.

86. Una vez determinados los límites, los valores ponderados de cada establecimiento en un panel mensual dado se comparan con los límites de ese mes concreto. Si los ingresos ponderados quedan por encima de los límites para el mes apropiado, el establecimiento se coloca en una lista de "observación potencial ampliada". Después de que se haya hecho tal cosa con los 12 paneles, se buscan los datos de ingresos de los 11 meses que faltan de todo el panel de observación potencial ampliada. Si los ingresos figuran por encima de los límites inferiores para los 12 meses sin excepción, el establecimiento queda incluido en el panel de "observación ampliada permanente" y se recogen sus datos en todos los paneles, dividiéndose por 12.

87. Si el procedimiento de observación ampliada se usa en conexión con el de estimación compuesta, no hay necesidad de cambiar la fórmula de las U_{hj} usada para el cálculo de la varianza. Los valores ponderados de ese panel deberían incluirse en todos los valores simples insesgados de Y que figuren (valores tanto de Y' como de Y'').

88. Si se usa un muestreo por áreas, es poco probable que las UMP aleatorias sean objeto de una rotación frecuente. La rotación de la muestra de área es probable que suponga sólo la rotación de conjuntos de segmentos dentro de las UMP. En tal caso, se usan las mismas fórmulas y procedimiento, pero únicamente se verá afectada la componente interior de la varianza de las UMP.

89. Muchas veces información auxiliar, tal como empleo, nóminas y datos parecidos, que figuran en registros fiscales recientes, o datos sobre ingresos u otras estadísticas recogidas en un censo reciente, puede usarse para mejorar las estimaciones corrientes. Para ser útiles, ha de disponerse de esos datos para el universo y también debe ser posible evaluarlos para la muestra que se obtuvo.

90. La fórmula recomendada para la estimación utilizando esa información auxiliar es:

$$Y'_R = \sum_{h=1}^L \frac{Y'_h}{X'_h} X_h \quad (6)$$

Donde Y'_R es la estimación ajustada, Y'_h es la estimación simple corriente ponderada, X'_h es la estimación ponderada de la muestra de la variable correlacionada (siendo X_h el universo total de esa variable) y h es la clase h-ésima de las L tipos de actividad principales en el plano nacional. (Adviértase que las h clases pueden representar aquí combinaciones de los R estratos usados en el proceso de selección.)

91. Sobre esta estimación cabe hacer las observaciones siguientes:

- a) Por lo general, si la correlación entre X e Y excede de 0,5, Y'_R tendrá una varianza menor que Y' (estimación simple);

- b) Las L clases de tipo de actividad para las que se hacen las estimaciones de los cocientes deberían tener estimaciones bastante estables de X'_h (esto es, un coeficiente de variación de 15% o menos). Las clases que no satisfagan ese requisito deberían combinarse con otras;
- c) Si los datos representados por el símbolo X se emplearan en los procesos de estratificación y muestreo, se obtendrían pocas ventajas adicionales con las estimaciones de los cocientes;
- d) Si se usa esta forma de estimación proporcional, las U'_{hj} usadas en la fórmula de la varianza se convierten en $U'_{hj} = Y'_{hj} - R'_h X'_{hj}$ donde R'_h es el cociente de dividir Y'_h por X'_h en la fórmula 6;
- e) Cuando se usa el muestreo por áreas, no se suele disponer de datos para las estadísticas correlacionadas a nivel de segmento. En tal caso, las X'_{hj} serán totales de UMP ponderadas y afectarán sólo al componente entre estratos de la varianza;
- f) Si se usa la estimación compuesta para reducir la varianza de la UMP interior de un diseño de muestreo por áreas y la estimación proporcional se emplea para reducir la varianza de la UMP intercalar, entonces la estimación resultante (que llamaremos Y''_R) revestiría la forma siguiente:

$$Y''_{(i)R} = \sum_{h=1}^L \frac{Y''_{(i)h}}{X'_h} X_h \quad (7)$$

$Y''_{(i)h}$ es la estimación compuesta descrita en los párrafos 73 a 78 para el tipo de actividad principal h-ésimo, X'_h es la estimación del tipo de actividad principal h-ésimo para las estadísticas auxiliares obtenidas de los totales de las UMP ponderadas para el tipo de actividad h-ésimo y X_h es el total del universo de la estadística del tipo de actividad h-ésimo.

92. Si se usa la estimación anterior, entonces el valor U'_{hj} empleado en la fórmula de la varianza (que llamaremos U''_{hj}) adoptará el valor siguiente:

$$U''_{hj} = U'_{hj} - R'_h X'_{hj} \quad (8)$$

En la fórmula anterior, U'_{hj} tiene el mismo valor que en la fórmula 4 ó 5, según sea el tipo de rotación, R'_h es el correspondiente cociente $\frac{Y''_{(i)h}}{X'_h}$ de la extracción hj-ésima y X'_{hj} es el total de la UMP ponderada para el conglomerado último hj-ésimo.

93. Hay que decir algo advirtiendo contra una reducción arbitraria a gran escala de la importancia de los valores atípicos en cualquier tipo de muestra. Esta es una tentación grande en una muestra de establecimientos, porque un establecimiento

(o empresa) puede llegar a ser muy grande, entre el momento en que quedó incluido en la muestra de la lista y aquel en que se enumeró. También, es posible que establecimientos muy grandes aparezcan en el muestreo por áreas, ya que los establecimientos no están estratificados (excepto para una posible lista de inclusiones seguras). Por ejemplo, supongamos que unos grandes almacenes figurasen en la muestra de la ciudad A con un coeficiente de ponderación de 100. Como A es una ciudad de tamaño moderado, se sabe que no pueden existir 100 grandes almacenes en ella. Por lo tanto, la tentación es reducir arbitrariamente el coeficiente (probablemente a 1). Tal cosa sería probablemente correcta si el único propósito de la encuesta fuera estimar los grandes almacenes de la ciudad A. No obstante, el total nacional suele ser de mucha mayor importancia y en ese universo podría fácilmente haber 100 o más valores atípicos de esa índole. Eliminar ese caso o reducir su coeficiente produciría una subestimación apreciable.

94. El hecho es que no existe modo alguno para decir a partir de la muestra si los casos extremos están representados por exceso o por defecto en la muestra de que se trate. Por lo general, la información disponible fuera de la muestra es demasiado limitada como para permitir hacerse esa idea en el plano general. Por lo tanto, las reducciones arbitrarias de los coeficientes de ponderación deberían hacerse con parsimonia, ya que sólo unas pocas de tales reducciones pueden producir un sesgo considerable a la baja.

95. El procedimiento de observación ampliada descrito en el párrafo 86 es un método para reducir los coeficientes de ponderación de los valores atípicos de una forma insesgada y puede revestir considerable importancia en las encuestas donde esos valores atípicos suponen un gran problema.

Nota: El análisis anterior se dirige concretamente y se limita a las técnicas prácticas para aplicar el muestreo a las encuestas del comercio de distribución. Tendría también interés que se leyesen los documentos sobre muestreo que figuran en la bibliografía donde se explican algunos otros métodos que pueden ser útiles y se examinan algunos puntos conexos que no se han tratado aquí.

Anexo II

BIBLIOGRAFIA

Recomendaciones internacionales sobre estadísticas del comercio de distribución y los servicios, serie M, N° 57. 63 págs. Publicación de las Naciones Unidas, N° de venta S.75.XVII.9.

Clasificación industrial internacional uniforme de todas las actividades económicas (CIIU), serie M, N° 4, Rev.2. Publicación de las Naciones Unidas, N° de venta S.68.XVII.8.

Indexes to the International Standard Industrial Classification of All Economic Activities, serie M, N° 4, Rev.2, Add.1. 247 págs. Publicación de las Naciones Unidas, N° de venta E.71.XVII.8.

Clasificación uniforme para el Comercio Internacional, revisión 2 (CUCI), serie M, N° 34, Rev.2. 117 págs. Publicación de las Naciones Unidas, N° de venta S.75.XVII.6.

Clasificación por Grandes Categorías Económicas, serie M, N° 53, Rev.1. 53 págs. Publicación de las Naciones Unidas, N° de venta S.76.XVII.7.

"Clasificación Internacional Uniforme de Todos los Bienes y Servicios" (E/CN.3/493). (Aparecerá como publicación de las Naciones Unidas.)

Bibliography of Industrial and Distributive-trade Statistics, serie M, N° 36, Rev.4. 177 págs. Publicación de las Naciones Unidas, N° de venta E.75.XVII.7.

Recomendaciones para el Programa Mundial de Estadísticas Industriales de 1973, parte III. Organización y realización de censos industriales, serie M, N° 54. 111 págs. Publicación de las Naciones Unidas, N° de venta S.72.XVII.10.

Breve manual de muestreo, Vol. I. Elementos de la teoría de encuestas por muestreo, serie F, N° 9, Rev.1. 348 págs. Publicación de las Naciones Unidas, N° de venta S.72.XVII.5.

Breve manual de muestreo, Vol. II. Programas de calculadora electrónica para diseños muestrales, serie F, N° 9. 77 págs. Publicación de las Naciones Unidas, N° de venta S.71.XVII.4.

Breve manual de muestreo, Vol. II. Programas de calculadora electrónica para diseños muestrales, serie F, N° 9, Add.1. 55 págs. Publicación de las Naciones Unidas, N° de venta S.73.XVII.8.

Recomendaciones para la preparación de informes sobre encuestas a base de muestras
(edición provisional), serie C, N° 1, Rev.2. 18 págs. Publicación de las
Naciones Unidas, N° de venta S.64.XVII.7.

Encuestas por muestreo de interés actual (duodécimo informe), serie C, N° 13,
367 págs. Publicación de las Naciones Unidas, N° de venta S.73.XVII.5.

Las publicaciones de las Naciones Unidas pueden obtenerse en la Sección
de Ventas, Naciones Unidas, Nueva York o Ginebra.