

1. Introducción

Las encuestas del grupo trabajador son una de las operaciones más complejas que se realizan de forma continua en la práctica totalidad de los países desarrollados y en una gran parte de los países en desarrollo. La actual revisión se encuadra dentro de las revisiones que se deben realizar periódicamente a este tipo de investigación continua, especialmente cada vez que se realiza un nuevo Censo de Población y Vivienda, como recomienda la Organización Internacional del Trabajo (OIT) en su manual sobre encuestas de Población Económicamente Activa. Por lo tanto, una vez realizado el Censo de Población del año 2000, la revisión actual se hace sobre la encuesta rediseñada a partir de los datos del Censo de Población y Vivienda de 1990.

Esta encuesta es la responsable de producir los datos de empleo, desempleo y personas fuera del grupo trabajador en Puerto Rico, así como ofrecer información a nivel total y por región sobre salarios, horas trabajadas, ocupaciones, tipos de trabajos e información diversa (socio-demográfica de la población). Empezó a realizarse en 1947, por lo cual es una de las primeras encuestas de este género. Está basada en la encuesta realizada en Estados Unidos (Current Population Survey), la cual ha servido como base a las encuestas similares realizadas en un gran número de países.

La revisión que se está llevando a cabo no sólo abarca el tipo de muestreo o método de investigación a utilizar, sino que trata de

potenciar e incluso crear aspectos tan importantes de la muestra como son los sistemas de control de calidad, la incorporación de sistemas informáticos de recogida de información y la actualización del cuestionario con más de 30 años existencia.

Como parte de la revisión de esta encuesta, además de estudiar la metodología de la encuesta actual, se han analizado y tomado como referencia las recomendaciones generales que hace la Organización Internacional del Trabajo sobre diseño y rediseño de encuestas al grupo trabajador, incluidas en el *Manual de la OIT sobre Conceptos y Métodos para las Encuestas de Población Activa, Empleo y Desempleo*; y también se han analizado encuestas similares realizadas en países desarrollados, especialmente las llevadas a cabo en Estados Unidos (*Current Population Survey*, CPS) y España (*Encuesta de Población Activa*, EPA), e incluso se hizo contacto con el personal a cargo de estas encuestas.

Los métodos de diseño y de selección de la muestra propuestos en esta revisión son científicos y cumplen con los requisitos fundamentales que la OIT recomienda para este tipo de encuestas. Como son el tipo de muestreo (Probabilístico – Polietápico – Estratificado), la periodicidad de la muestra (mensual), el ámbito (nacional) y la Tamaño de la Muestra (3,500 hogares o viviendas ocupadas).

2. Objetivos de la encuesta

Esta encuesta es una investigación estadística de tipo continuo que tiene la responsabilidad de producir los datos de empleo, desempleo y personas fuera del grupo trabajador en Puerto Rico. Además ofrece información sobre salarios, horas trabajadas, ocupaciones, tipos de trabajo y datos de la población como: género, edad, estado civil, nivel educacional, total personas incapacitadas, total de personas en la escuela. Es decir, está orientada a dar datos de las principales categorías poblacionales en relación con el mercado de trabajo y a obtener clasificaciones de estas categorías según diversas variables.

Además, con los estimados obtenidos en esta investigación, se calculan las principales tasas relacionadas con el empleo, como son las tasas de participación en el grupo trabajador (Población Activa), la de empleo y desempleo.

Se obtienen resultados detallados para el conjunto de Puerto Rico y para las regiones ofrecerá información sobre las principales características de cada una de ellas.

3. Fuentes de información relacionadas

3.1. Resumen de las recomendaciones internacionales de la OIT

Como parte de la fase de estudio de la investigación, nos pusimos en contacto con la Organización Internacional del Trabajo (OIT) con sede en Ginebra (Suiza), la cual nos proveyó del *Manual sobre Conceptos y Métodos de las Encuestas de Población Activa, Empleo,*

Desempleo y Subempleo, en el cual se especifica, entre otras cosas, una serie de características generales o comunes que deben cumplir las encuestas al grupo trabajador.

Recomienda algunos aspectos prácticos para el diseño de la encuesta como:

- la utilización de un muestreo probabilístico;
- la moderación en la selección del tamaño de la muestra;
- el uso de la técnica de estratificación;
- la preferencia por muestras autoponderadas;
- evitar introducir etapas innecesarias en el diseño de la muestra (muestreo polietápico);

Este documento también establece cuáles deben ser los fundamentos de las revisiones (rediseño) periódicas a las cuales deben de estar sujetas estas encuestas. De acuerdo a éste, es indispensable la revisión y mejora periódica de las encuestas de fuerza del trabajo. Sin embargo, se requiere un compromiso entre la necesidad de mejorar constantemente el diseño y los procedimientos técnicos, por una parte, y la necesidad de mantener la comparabilidad y el valor de la información existente, por otra.

La OIT a través de este manual hace recomendaciones en cuanto al tamaño de la muestra. Suponiendo que el requisito crítico de una encuesta sea estimar una proporción (como la tasa de desempleo) que constituye aproximadamente un 10% del dominio particular

del estudio, y que dicha proporción ha de ser estimada con un margen de error de $\pm 1\%$ y un nivel de confianza alto, esto puede exigir un tamaño muestral del orden de 5,000 individuos (2,000 a 3,000 hogares), dependiendo de la eficacia del diseño.

Respecto a los **errores de muestreo**, de acuerdo a este manual, una práctica reconocida como recomendable consiste en que los resultados vayan acompañados de información detallada sobre la variabilidad muestral de las estimaciones de la encuesta, y que la interpretación de los resultados tenga en cuenta esta variabilidad. Esto es especialmente importante en el caso de las encuestas continuas a gran escala.

Para poder extraer conclusiones significativas de los resultados de la encuesta es necesario que los errores de muestreo se mantengan dentro de ciertos límites. Además, la información acerca de la magnitud de los errores de muestreo es esencial para decidir el grado de detalle con que es posible realizar la clasificación de los datos de la encuesta sin que pierdan sentido. En encuestas de fuerza de trabajo, la correcta interpretación de los resultados requiere habitualmente una clasificación muy detallada de los resultados por sexo, edad, localidad geográfica y otras características personales y del hogar.

La información sobre el error de muestreo es fundamental para el diseño, evaluación y rediseño de la muestra. La magnitud de un determinado error de muestreo contenido en la estimación de una encuesta dependerá, entre otros factores, del tamaño y diseño muestral, y

especialmente del grado en que las unidades del área muestral estén agrupadas y si son homogéneas dentro de las agrupaciones. En las encuestas continuas de fuerza de trabajo las consideraciones acerca de los solapamientos y de la rotación, la representatividad simultánea en el espacio y en el tiempo, la necesidad de revisión y rediseño periódicos de la muestra y de otros muchos factores, pueden hacer que los procedimientos de diseño y estimación sean especialmente complejos. Como es lógico, muchas de estas decisiones no pueden ser tomadas de forma científica sin una información sobre los errores de muestreo, en la medida que son determinados por el tamaño y diseño de la muestra.

De acuerdo a este manual, **los métodos prácticos para calcular los errores de muestreo han de ser suficientemente sencillos como para proporcionar fórmulas de cálculo fáciles, que a su vez permitan realizar de manera económica los cálculos detallados que requieren las numerosas estimaciones necesarias.** Al mismo tiempo, **los métodos prácticos de cálculo necesitan ser bastante generales, de forma que engloben las complejidades y variaciones que frecuentemente ocurren en el diseño muestral de las encuestas a gran escala.**

3.2. Estudio de otras investigaciones similares

También como parte de la fase de estudio, se recopiló y analizó información de investigaciones similares realizadas en otros países, prestando un interés especial a la *Current*

DEPARTAMENTO DEL TRABAJO Y RECURSOS HUMANOS

Encuesta sobre el Grupo Trabajador

Population Survey (CPS) realizada en Estados Unidos y en la Encuesta de Población Activa (EPA) realizada en España y, por tanto, la de un miembro de la Comunidad Económica Europea.

- Para el estudio de la CPS se utilizó como referencia el Documento Técnico 63 (*Design and Methodology of CPS*) y su posterior revisión, publicados ambos documentos por la Oficina del Censo y el Departamento del Trabajo de Estados Unidos, también recibimos el insumo del personal del Departamento del Trabajo sobre diferentes aspectos de la encuesta.
- En el caso de la EPA, el análisis se realizó a partir de los libros técnicos publicados por el Instituto Nacional de Estadística de España (INE), los cuales se incluyen en la parte de este documento dedicada a la bibliografía utilizada), además de contactos directos con personal de este organismo que trabaja directamente con la EPA.

4. Definiciones

Las definiciones y conceptos de este tipo de investigaciones deben tener la capacidad de medirse objetivamente y depender lo menos posible del juicio o criterio de las personas responsables de recopilar los datos y preparar la información. Por tal razón, estas definiciones deben ser simples, prácticas y fáciles de entender incluso para un ciudadano común que no esté familiarizado con el tema del grupo trabajador.

Las definiciones que a continuación se presentan están basadas en las recomendaciones de la

Organización Internacional del Trabajo (OIT) aprobadas en la Decimotercera y Decimosexta Conferencia Internacional de Estadísticos del Trabajo (Ginebra 1982 y 1998, respectivamente) y son compatibles con las utilizadas por el Departamento del Trabajo de Estados Unidos en el *Current Population Survey*.

a) Grupo Trabajador (Población Económicamente Activa)

Se puede definir como el conjunto de personas de 16 años o más que, en un periodo de referencia dado, suministran mano de obra para la producción de bienes y servicios económicos o que están disponibles y hacen gestiones para incorporarse a la mencionada producción.

Para efectos de esta encuesta, el grupo trabajador estará formado por todas las personas civiles no residentes en instituciones de larga duración ni hogares colectivos que durante la semana de referencia (la semana del día 12 de cada mes) cumplen con las condiciones expuestas en el párrafo anterior y, por tanto, se pueden encuadrar en alguno de los dos grandes subgrupos que forman el Grupo Trabajador que son: empleados y desempleados, los cuales se definen a continuación.

b) Empleados (o personas con empleo)

Lo constituyen las personas civiles de 16 años o más que durante la semana del día 12 de cada mes (semana de referencia) estaban:

- **trabajando:** Es decir, realizando algún trabajo por paga o ganancia (al menos 1

hora), o las que trabajaron sin paga por 15 horas o más durante la semana de referencia en la finca o negocio de un familiar que reside en la misma vivienda.

- **o con empleo no trabajando:** Son aquellas personas que, si bien no estaban trabajando la semana de referencia, tenían un empleo o negocio del cual se encontraban temporalmente ausentes debido a enfermedad, vacaciones, mal tiempo, conflictos laborales, cualquier tipo de licencia (estudios, maternidad, etc.), ausencia voluntaria, paro temporal con instrucciones de volver en una fecha determinada.

Dentro de los empleados podemos identificar 2 tipos:

- las personas con trabajo **por cuenta ajena o asalariadas** o aquellas que durante la semana de referencia hayan trabajado al menos una hora a cambio de un sueldo, salario u otra forma de retribución tanto en metálico como en especie.
- las personas **con actividad o negocio propio** o aquellas que durante el periodo de referencia hayan trabajado al menos 1 hora a cambio de un beneficio o de una ganancia familiar, en metálico o en especie.

También dentro de los ocupados cabe destacar la situación de los subempleados

por insuficiencia de horas. De acuerdo a las recomendaciones de la OIT, existe subempleo por insuficiencia de horas, cuando el ocupado desea trabajar más horas (en el mismo empleo, en otro empleo o encontrando un segundo empleo), está disponible para trabajar más horas y ha trabajado menos de 35 horas en la semana de referencia.

c) **Desempleados**

Son las personas de 16 años o más civiles que estaban **sin trabajo durante la semana de referencia, pero que estaban buscando (tomando medidas concretas para buscarlo) durante las 4 semanas anteriores a la de la entrevista (incluida la semana de referencia) y estaban disponibles para trabajar (se excluyen los enfermos temporalmente).**

Si una persona satisface las condiciones de la definición de desempleado se considerará como tal, aunque compatibilice su situación con otras como la de estudiante o labores de hogar. Además cabe destacar que, de acuerdo a la OIT, en este grupo también se incluyen las personas que están esperando ser llamadas de nuevo a un trabajo que ya tenían (*laid off*), es decir están sin empleo de forma temporal, pero tienen ya una fecha concreta para volver a trabajar o esperan (le han indicado o sugerido) ser llamados en los próximos 6 meses. Este grupo de personas no necesita estar buscando trabajo para incluirse en el grupo de los desempleados, pero si estar disponible para trabajar la

DEPARTAMENTO DEL TRABAJO Y RECURSOS HUMANOS

Encuesta sobre el Grupo Trabajador

semana de referencia (excepto si están enfermos temporalmente).

Por el contrario, las personas que están esperando a ser llamadas en un nuevo trabajo, si han buscado trabajo activamente en las últimas 4 semanas antes de la entrevista se consideran desempleados, pero si no lo han hecho pasan a formar parte de los que están fuera del Grupo Trabajador. Este pequeño cambio en la definición se produce para que la misma concuerde con la definición de Estados Unidos¹.

Dentro de los desempleados podemos distinguir dos grupos: los que buscan un primer empleo y los que han trabajado con anterioridad.

d) Fuera del Grupo Trabajador (**Inactivos**):

Está compuesto por el conjunto de la población no institucional de 16 años o más que no forma parte del grupo trabajador (empleados y desempleados), es decir los que no tienen empleo, ni lo han buscado en las 4 últimas semanas antes de la entrevista y/o no están disponibles para trabajar.

Dentro de este grupo poblacional se incluyen:

- Las personas que se ocupan de su hogar: son las que, sin ejercer ninguna actividad económica, se ocupan de cuidar sin ninguna remuneración sus propios hogares; como por ejemplo, las amas de casa y otros familiares que se

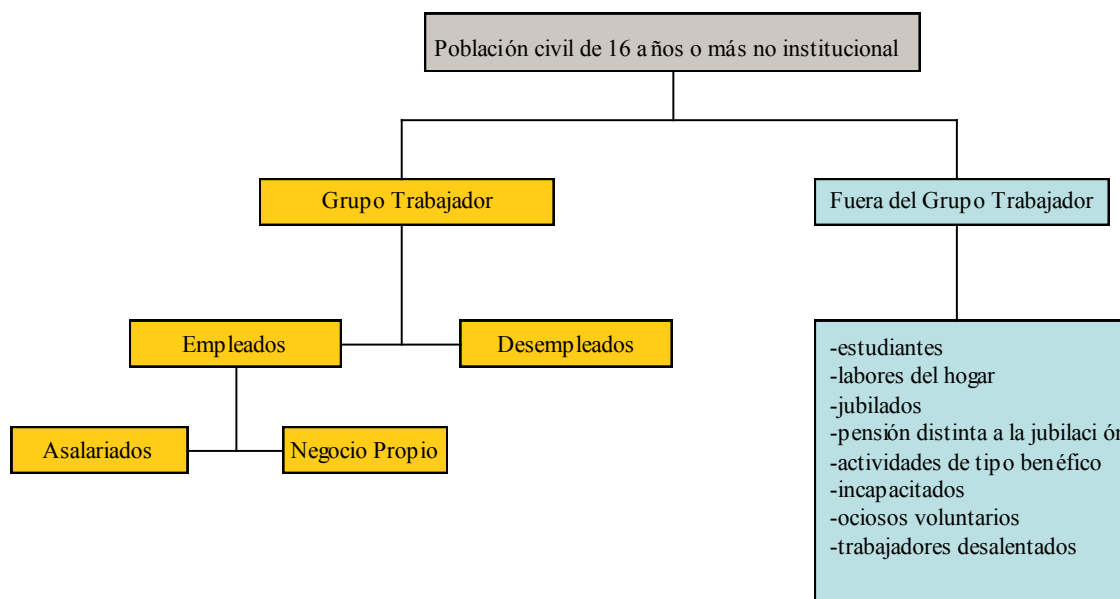
encargan del cuidado de la casa y de los niños.

- Estudiantes que no realizan ninguna actividad económica.
- Jubilados o prejubilados: personas que han desarrollado una actividad económica anterior y que por su edad u otras circunstancias la han abandonado, percibiendo una pensión (o ingresos por su jubilación) con motivo de su actividad anterior.
- Personas que reciben pensión por razones distintas a la jubilación o prejubilación.
- Personas que realizan trabajos sociales, actividades de tipo benéfico, etc. pero que no reciben remuneración por ello.
- Personas incapacitadas para trabajar.
- Ociosos voluntarios, etc.

Se debe hacer una mención especial, dentro de este grupo, a las personas denominadas como trabajadores desalentados, es decir, aquellas personas que quieren y están disponibles para trabajar, pero que no buscan porque piensan que, por diferentes razones, no hay trabajo para ellas.

A continuación, una vez establecidas las principales definiciones relacionadas con el Grupo Trabajador, la siguiente ilustración muestra de forma gráfica cómo se clasificaría la población de acuerdo a estas definiciones.

¹ Sección 5 del documento técnico *Current Population Survey, Design and Methodology, Technical Paper 63RV, U.S. Department of Labor. Bureau of Labor Statistics.*



5. Principales tasas a calcular

Una vez realizadas las estimaciones de los diferentes grupos poblaciones de acuerdo a las definiciones que se han presentado en el apartado anterior, y a partir de las mismas, se calculan las diferentes tasas asociadas al grupo trabajador. Las más importantes se detallan a continuación:

- Tasa de participación

La tasa de participación mide la población civil no institucional de 16 años o más que forma parte de la fuerza o grupo trabajador (empleados + desempleados). Es el cociente entre el grupo trabajador y el total de personas de 16 años o más civiles no institucionalizadas.

Se calcula de forma general y de forma separada por género y región.

- Tasa específica de participación

La tasa específica de actividad para un intervalo de edades determinado mide la población civil no institucional de ese intervalo de edad que forma parte de la fuerza o grupo trabajador. Es el cociente entre el grupo trabajador de ese intervalo de edades y el total de personas de ese intervalo de edad civiles no institucionalizadas.

Generalmente se calcula para ambos sexos y para cada uno de ellos por separado. Los intervalos de edad suelen ser quinquenales o decenales.

La tasa específica de participación para las personas de 16 y más años es lo que se suele denominar *tasa de participación*.

- **Tasa de desempleo**

Esta tasa es el cociente entre la población que está desempleada y el grupo trabajador (empleados + desempleados).

Se calcula para toda la población general y de forma separada por género y región.

- **Tasa específica de desempleo**

La tasa específica de desempleo para un intervalo de edades determinado es el cociente entre la población desempleada de ese intervalo de edad y el grupo trabajador de dicho intervalo.

Generalmente se calcula para ambos sexos y para cada uno de ellos por separado. Los intervalos de edad suelen ser quinquenales o decenales.

6. Ámbito de la Encuesta y Periodo de referencia

6.1. Ámbito Geográfico

La encuesta cubre todo el territorio de Puerto Rico excepto los municipios de Vieques y Culebra y los grupos de bloques que estén dentro de los límites de las bases militares.

6.2. Ámbito Poblacional

Esta investigación va dirigida a la población civil de 16 años o más que reside en viviendas familiares, es decir, las utilizadas todo el año o la mayor parte de él como vivienda habitual o permanente. Se excluyen, por tanto, los menores de 16 años, el personal militar y las personas internadas o recluidas en hogares colectivos

(hospitales, hoteles, conventos) o instituciones de larga duración como hospitales, hoteles, conventos, cárceles, asilos, orfanatos, etc.. Sí se incluyen, sin embargo, las familias que, formando un grupo independiente, residen en estos establecimientos, como puede ocurrir con los directores de los centros, conserjes y porteros o personal de mantenimiento.

6.3. Periodo de Referencia de la Información

Las entrevistas se realizan mensualmente la semana en la que está incluido el día 19 del mes. El periodo de referencia de las preguntas de la encuesta es la semana del día 12 del mes.

7. Marco Muestral

El marco muestral de esta encuesta está constituido principalmente por el Censo de Población y Vivienda del 2000 que produce para Puerto Rico la Oficina del Censo de Estados Unidos.

Para la definición del marco es necesario partir de la división administrativa de Puerto Rico, la cual consiste de 78 municipios que a la vez se dividen, de acuerdo al Censo de Población del 2000, en sectores censales, éstos en grupos de bloques y por último en bloques. Por último, para llegar a cada vivienda familiar es posible, gracias al Censo de Población del 2000, confeccionar para cada grupo de bloques una lista de segmentos o grupos de viviendas.

Por tanto, el marco de la muestra lo constituye:

- El marco de áreas, formado por los 2,429 grupos de bloques en los que, de acuerdo al Censo de Población y Vivienda del 2000, se encuentra dividido Puerto Rico exceptuando Vieques y Culebra (ya que estos municipios no forman parte de la encuesta como se indica en el apartado 6.1 referente al ámbito geográfico de la encuesta).
- El marco de listas, constituido por la relación de segmentos (grupo de 6 viviendas familiares) que se puede constituir para cada grupo de bloques seleccionado en la muestra.

Los archivos SF (Summary Files) del Censo de Población del 2000 contienen información por Sectores Censales, divididos en Grupos de Bloques. Además ofrecen para cada grupo de bloques información sobre: municipio, total de viviendas, total de viviendas ocupadas, población, pago de alquiler mediano de las viviendas, valor mediano de las viviendas ocupadas y otra serie de datos sobre población y vivienda.

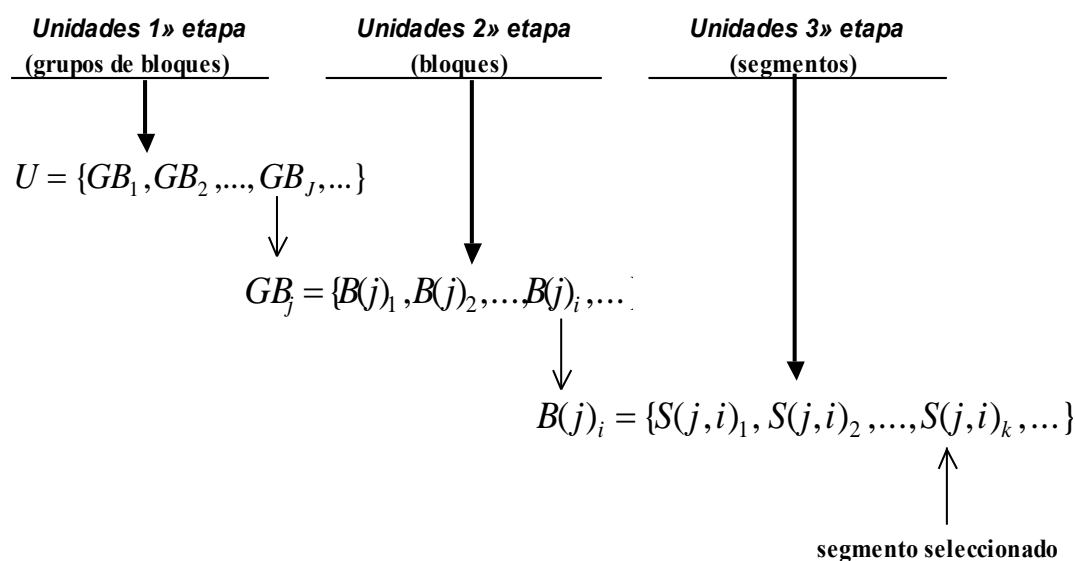
También formarán parte del marco muestral los mapas censales, los cuales identifican los límites políticos, legales, censales, físicos y naturales de Puerto Rico. Además, como en revisiones anteriores, pueden ser necesarios o, por lo menos, de utilidad la serie de mapas topográficos de Puerto Rico, fotos aéreas, mapas de zonificación actualizados y diferentes guías de mapas.

8. Tipo de Muestreo

El tipo de muestreo utilizado es probabilístico polietápico con estratificación de las unidades de primera etapa y con probabilidades proporcionales a los tamaños de las unidades de primera etapa.

Entre los métodos de muestreo probabilístico se encuentra el muestreo por etapas. Éste se puede considerar como una modificación del muestreo por conglomerados, ya que una vez seleccionados los conglomerados (en este caso los grupos de bloques) dentro de cada uno de ellos no se cogen todos los elementos o unidades que los forman, sino que se realiza otra selección o submuestreo dentro de cada uno de ellos. Y así sucesivamente de acuerdo a las etapas que se hayan determinado. Para esta encuesta se han determinado 3 etapas, por lo cual podemos hablar de un muestreo trietápico, cuyas etapas están constituidas por las siguientes unidades de muestreo:

- Las unidades de 1ª etapa están constituidas por los grupos de bloques (GB_j);
- Las unidades de 2ª etapa están constituida por los bloques [$B(j)_i$]; y
- Las unidades de 3ª etapa son los segmentos. Definiremos como segmento a un grupo de viviendas normalmente de 6 unidades, que podrá tener un mínimo de 4 y un máximo de 8 unidades de vivienda [$S(j)_k$].



Dentro de las unidades de 3ª etapa no se realiza submuestreo alguno, por lo cual en cada segmento se recoge información de todas las viviendas que forman ese segmento. A su vez, en cada vivienda se recopila información de todas las personas que tengan su residencia habitual en la misma.

9. Estratificación de las unidades de muestreo

Para la determinación de los criterios de estratificación se han tenido muy en cuenta las recomendaciones de la OIT² y las encuestas realizadas en otros países desarrollados. Según esta organización, de las encuestas de este tipo se espera que no se centren en áreas particulares o núcleos de personas con características específicas.

También cabe destacar que las encuestas al grupo trabajador son una de las operaciones más complejas que se realizan de forma continua en la práctica totalidad de los países desarrollados y en una gran parte de los países en desarrollo. La complejidad de esta muestra se hace patente en países de gran extensión y diversidad geográfica e incluso demográfica. En el caso de Puerto Rico, por su extensión, características geográficas y de densidad demográfica (donde cada vez son menos las zonas rurales), no son necesarias algunas de las complejas técnicas o recursos utilizados en otros países.

A partir de todo lo anterior, los grupos de bloques (unidades de 1ª etapa) se estratificarán atendiendo a un triple criterio:

- A. Criterio Geográfico (de estratificación). La OIT recomienda un ordenamiento geográfico que permita la selección sistemática de la muestra, a partir de lo cual Puerto Rico se divide en 6 grandes zonas o estratos. De acuerdo a las teorías de

² Encuestas de Población Económicamente Activa, Empleo, Desempleo y Subempleo. Manual de la O.I.T sobre Conceptos y Métodos. Página 414.

DEPARTAMENTO DEL TRABAJO Y RECURSOS HUMANOS
Encuesta sobre el Grupo Trabajador

muestreo, cada una de los estratos o regiones debe estar formado por unidades lo más homogéneas posibles, a partir de esto, se han creado los estratos no sólo teniendo en cuenta la cercanía geográfica (municipios limítrofes), si no también la similitud industrial y laboral de los municipios que forman cada estrato, para de esta manera

intentar minimizar la dispersión o variación de las principales tasas relacionadas con el empleo y presentadas en la sección 5 de este documento, que se estimarán a partir de la información obtenida de la muestra. En la siguiente ilustración y el mapa Anejo III a este documento se presentan las principales características de los estratos determinados:

Distribución de la Población de 16 años o más de acuerdo a las regiones/estratos del GT

Región	Población de 16 años o más	%	Viviendas	%
Metropolitana	975,401	34.4%	486,903	34.5%
Norte	418,919	14.8%	209,277	14.8%
Central	223,077	7.9%	106,722	7.6%
Oeste	341,939	12.1%	179,966	12.7%
Sur	312,608	11.0%	150,165	10.6%
Este	561,491	19.8%	280,031	19.8%
Total	2,833,435	100%	1,413,064	100%

Fuente: Censo Población 2000. No se incluye la población de Vieques y Culebra.

B. Criterio Tipo de Municipio (de subestratificación): Este criterio es muy utilizado en encuestas similares a la del Grupo Trabajador, según el cual, una vez ordenada la información a partir del criterio

geográfico, dentro de cada región se determinaron 4 estratos de acuerdo al tamaño de los municipios que forman cada región. Los municipios se agruparon en los siguientes estratos:

Tipo de Municipio	# Municipios	Población	%
menos de 25,000	20	363,907	9.6%
25,000 a 49,999	40	1,441,356	38.0%
50,000 a 99,999	9	620,718	16.3%
100,000 o más	7	1,371,655	36.1%
Total	76*	3,797,636	100%

* Para propósitos de la encuesta no se incluye ni Vieques, ni Culebra.

Fuente: Censo de Población 2000.

C. Criterio Socio-económico (de subestratificación). También dentro de cada estrato se realizó una subestratificación económica. Según la cual, las unidades primarias de muestreo (grupos de bloques) se agruparon en 4 niveles económicos, los

cuales se determinaron a partir del valor de la mediana de ingresos anuales en el hogar de acuerdo al Censo de Población del 2000. El valor de esta mediana para Puerto Rico fue de \$14,412, a partir del mismo se

crearon 4 niveles económicos que se definen a continuación;

- Nivel 1: agrupa a todos los grupos de bloques cuya mediana de ingreso anual en el hogar está 25% o más por debajo de la mediana de ingresos en Puerto Rico (grupos de bloques con mediana de ingresos en el hogar inferior a \$10,809).
- Nivel 2: formado por los grupos de bloques cuya mediana de ingresos anuales es inferior en menos de un 25% a la mediana de ingresos en el Hogar en Puerto Rico. (grupos de bloques con mediana de ingresos entre \$10,809 y \$14.412).
- Nivel 3: formado por los grupos de bloques cuya mediana de ingresos anuales es superior en menos de un 25% a la mediana de ingresos en el Hogar en Puerto Rico. (grupos de bloques con mediana de ingresos entre \$14.413 y \$18.015).
- Nivel 4: agrupa a todos los grupos de bloques cuya mediana de ingreso anual en el hogar sea superior en más de un 25% a la mediana de ingresos en Puerto Rico. (grupos de bloques con mediana de ingresos en el hogar superior a \$18,015).

Respecto a los tres criterios de estratificación que presentamos, cabe destacar, que la OIT

indica en su manual³ que “una estratificación geográfica unida a la estratificación por tipo de Municipio proporcionará un medio sumamente eficaz para mejorar la eficiencia y distribución de la muestra”.

10. Tamaño de la Muestra.

El tamaño de la muestra, es de 3,500 viviendas, es decir, se ha reducido aproximadamente en 2,700 el tamaño muestral con respecto a la revisión efectuada en 1990. Esta reducción responde a las necesidades del propio Departamento del Trabajo de Puerto Rico, las recomendaciones respecto a la moderación en la selección del tamaño muestral que hace la O.I.T⁴. y a las recomendaciones del personal del BLS (Bureau of Labor Statistics) del Departamento del Trabajo de Estados Unidos.

Cabe destacar que, de acuerdo a la OIT⁵, suponiendo que el requisito crítico de una encuesta sea estimar una proporción (como la tasa de desempleo) que constituye aproximadamente un 10% del dominio particular del estudio; y que dicha proporción ha de ser estimada con un margen de error de $\pm 1\%$ y un nivel de confianza alto, esto puede exigir un tamaño muestral del orden de 5,000 individuos (2,000 a 3,000 hogares), dependiendo de la eficacia del diseño.

³ Encuestas de Población Económicamente Activa, Empleo, Desempleo y Subempleo. Manual de la O.I.T sobre Conceptos y Métodos. Página 413.

⁴ Encuestas de Población Económicamente Activa, Empleo, Desempleo y Subempleo. Manual de la O.I.T sobre Conceptos y Métodos. Página 411.

⁵ Encuestas de Población Económicamente Activa, Empleo, Desempleo y Subempleo. Manual de la O.I.T sobre Conceptos y Métodos. Página 413.

En todo caso, independientemente del tamaño y diseño de la nueva muestra, la incorporación de la misma ha sido gradual, tal como proponen tanto el Manual de la OIT, como el de la *Current Population Survey (Technical Paper 63)*, además de diversos autores como Azorín (Francisco) y Sánchez Crespo en su libro sobre *Métodos y Aplicaciones del Muestreo* (Alianza,1994). De tal manera la nueva muestra se ha ido introduciendo paulatinamente, incorporándose mensualmente en el grupo de rotación que entraba nuevo en la muestra (8 grupos), por lo cual el 1er mes que entró, solamente constituyó 1/8 de la muestra total y así sucesivamente. De este modo, sólo hasta 16 meses después de comenzar la incorporación de la nueva muestra, la misma estuvo formada completamente por grupos de rotación seleccionados con el nuevo diseño muestral (8 grupos). Así, con esta técnica de incorporación de la muestra, las diferencias mensuales de las estimaciones se suavizaron.

11. Selección de la muestra

La selección de la muestra se realizó con la intención de que cualquier vivienda tuviera la misma probabilidad de ser seleccionada (muestra equiprobable) y siguiendo las recomendaciones de la OIT que sugieren un ordenamiento de los estratos para la aplicación de muestreo sistemático. Aplicando también un muestreo con probabilidades proporcionales a los tamaños de las unidades de primera etapa (grupos de bloques).

De acuerdo al diseño muestral, las viviendas se agruparon en segmentos (unidades de 3ª etapa), y

cada segmento esta formado en promedio por 6 viviendas, por lo que se seleccionaron unos 583 segmentos que formarán parte de otros tantos grupos de bloques (unidades de 1ª etapa).

Los grupos de bloques se clasificaron de acuerdo al criterio de estratificación (criterio geográfico), por lo cual se pueden agrupar 6 grandes subconjuntos o estratos generales. Una vez determinados estos subconjuntos (regiones), de forma independiente en cada uno de ellos se ordenaron los grupos de bloques de acuerdo a los otros dos criterios de estratificación (tipo de municipio y socio-económico). Este ordenamiento constituye el marco de selección.

Para la selección de cada una de las unidades de muestreo se siguieron las siguientes técnicas de selección:

1ª Etapa. Selección de los grupos de bloques:

Se utilizó un muestreo sistemático teniendo en cuenta el tamaño de los grupos de bloques, es decir, con probabilidades proporcionales a los tamaños de los grupos de bloques.

2ª Etapa. Selección de los bloques:

Dentro de cada Grupo de Bloques se selecciona un bloque. Para ello, se aplica un muestreo aleatorio pero aplicando un peso proporcional al número de viviendas de cada bloque.

3ª Etapa. Selección de los segmentos:

Dentro de cada bloque, se selecciona un segmento a partir de un muestreo sistemático con arranque aleatorio con

un intervalo de muestreo de 1 en 3 (se elige uno de cada 3 segmentos).

En el caso de que el bloque solamente tuviera un segmento, por defecto, se seleccionaría ese segmento, mientras que si tuviera dos segmentos se seleccionaría aleatoriamente uno de los dos segmentos.

Por último, cada segmento se visita 8 meses de acuerdo a la rotación de la muestra y se realizan entrevistas en todos los hogares que forman el mencionado segmento.

Para conocer con más detalle el proceso de selección utilizado para cada una de las unidades muestrales, se recomienda leer el documento “Revisión de la Encuesta del Grupo Trabajador. Metodología y Cuestionario” del Departamento del Trabajo y Recursos Humanos.

12. Mantenimiento de la muestra

La sustitución de las diferentes unidades muestrales (segmentos, bloques y grupos de bloques) se realizará bajo las siguientes circunstancias:

A. Sustitución de segmentos:

Cada segmento se enumera 8 meses en dos periodos de 4 meses. Se visita 4 meses consecutivos, descansa 8 y vuelve a ser visitado otros 4 meses más. Esto quiere decir que todo segmento está activo en la muestra 16 meses, siendo sustituido en la muestra por el próximo en la tercera posición del mismo bloque. Una vez que todos los

segmentos del bloque se inactivan, se procede a la sustitución del bloque.

B. Sustitución de bloques:

Una vez que su último segmento se inactiva todo bloque es sustituido, de acuerdo al ordenamiento, por el siguiente en el grupo de bloques

C. Sustitución de grupos de bloques:

Cuando el último bloque del grupo de bloques sale de la muestra, es sustituido por el próximo grupo de bloques de acuerdo al ordenamiento establecido por los criterios de estratificación.

13. Rotación de la Muestra

En las encuestas continuas, con el fin de evitar el cansancio de las viviendas seleccionadas y poder incorporar las nuevas viviendas para que aseguren la representatividad de la muestra a lo largo del tiempo, se suele utilizar la técnica de rotación de la muestra.

El método de rotación es el mismo que el realizado hasta ahora, y que es uno de los más completos y a la vez complejos que se realizan para las encuestas continuas. Está basado al utilizado en Estados Unidos para la realización de la *Current Population Survey (CPS)*.

Es el modelo de rotación con solapamiento parcial conocido como “4-8-4”. Según el mismo, la muestra se divide en 8 submuestras denominadas “grupos de rotación”. Cada grupo de rotación se incluye en la muestra 4 meses consecutivos, descansa 8 meses y vuelve a

formar parte de la muestra otros 4 meses, saliendo ya de forma definitiva.

Según este sistema, cada mes, de los 8 grupos de rotación, 6 se mantienen respecto al mes anterior, 1 es sustituido por otro nuevo y el que resta es sustituido por el grupo de rotación que termina el descanso de 8 meses. Por lo tanto, este sistema garantiza que el 75% de la muestra sea igual a la del mes anterior.

14. Actualización de la Muestra (Inclusión de las viviendas de nueva construcción)

Otro de los métodos de actualización de la muestra y por consiguiente de mantener la representatividad de la misma, es la inclusión de las viviendas de nueva construcción (aquellas viviendas construidas después de la realización del Censo 2000).

Anualmente, a partir de la información que puedan suministrar tanto agencias públicas como privadas, se incluirán los proyectos de al menos 10 unidades de vivienda, tanto públicas como privadas.

Para que las nuevas unidades se puedan integrar a la muestra se le aplicarán los diferentes criterios de estratificación. Así:

- se identificarán geográficamente de acuerdo a la región donde estén ubicadas (6 regiones de acuerdo al criterio de estratificación);
- respecto al criterio de tipo de municipio (subestratificación), estas nuevas unidades se clasificarán en los estratos determinados por

este criterio de acuerdo a la población del Municipio donde han sido construidas

- se determinarán unos niveles económicos de acuerdo al precio de venta de la vivienda si ésta es construida por el sector privado; y si es construida por el sector público, se determinará su nivel o estrato de acuerdo a la naturaleza del proyecto.

La actualización se llevará a cabo mediante la información que se puede obtener anualmente de diferentes agencias públicas y privadas como:

- Administración de Reglamentos y Permisos
- Departamento de la Vivienda
- Junta de Planificación
- Oficinas Locales del Negociado de Estadísticas
- Compañías Privadas de Construcción
- Prensa local

Es importante la realización de esta actualización de las viviendas de nueva construcción para todo Puerto Rico, ya que esta información, junto con la del Censo de Población del 2000, constituyen el marco muestral.

15. Estimación

El estimador propuesto es un estimador de razón. Este tipo de estimadores son muy utilizados para este modelo de encuestas y se caracterizan por utilizar el conocimiento de alguna información complementaria correlacionada con la característica objeto de estudio (lo que se denomina variable auxiliar). También se

DEPARTAMENTO DEL TRABAJO Y RECURSOS HUMANOS

Encuesta sobre el Grupo Trabajador

consideraron, de acuerdo a la recomendación del personal del BLS, estimadores similares a los que ellos utilizan en la CPS (Current Population Survey) con varias etapas de ajuste, pero al analizar este procedimiento se encontraron varios aspectos que limitaban el uso de este tipo de estimación. Entre estos aspectos, cabe destacar:

- La información que **mensualmente** puede proveer la Junta de Planificación sobre la población de Puerto Rico es limitada.
- La complejidad demográfica de Puerto Rico es mucho menor que la de EEUU, no teniendo sentido, por ejemplo, utilizar variables relacionadas con la raza de las personas para la determinación de los estimados.

Se optó, por tanto, por un estimador de razón con dos variables auxiliares, siendo éstas la edad y el género de las personas. Para realizar la estimación, la Junta de Planificación a través de la Oficina del Censo (Programa de Planificación Económica y Social) proveerá mensualmente una proyección por género y edad de la población no institucional de Puerto Rico de 16 años o más. Anualmente, estas proyecciones serán revisadas y contrastadas con los estimados que publica anualmente la Oficina del Censo de Estados Unidos.

A partir de todo lo anterior, se utilizará el siguiente estimador:

$$\hat{X}_t = \sum_k \frac{\hat{P}_{tk}}{P_{tk}} X_{tk}$$

donde:

\hat{X}_t = estimador de razón de la característica investigada en el periodo t.

X_{tk} = es el valor de la característica investigada de las personas del grupo de edad/género k obtenida en la muestra en el periodo t.

\hat{P}_{tk} = es el valor de la población teórica proyectada del grupo de edad/género k en el periodo de tiempo t (de acuerdo a los estimados mensuales facilitados por la Junta de Planificación).

P_{tk} = personas del grupo de edad/género k obtenidas en la muestra del periodo t.

Por otro lado, cabe destacar que todas las tasas que se calculan a partir de la información de esta muestra se deben hacer a través de los estimados poblacionales obtenidos de la estimación de razón. Según lo cual, las tasas a nivel regional se deben calcular a partir de los valores poblacionales que se estimen para cada región; y no a través de resultados directos de la muestra.

16. Error de Muestreo

Los errores de muestreo aparecen como consecuencia de la utilización de una parte de la población para estimar características de todo el conjunto poblacional. El error de muestreo está íntimamente ligado a la estimación y a los estimadores que se utilicen, ya que es una

medida de variabilidad del estimado encontrado, lo cual permite obtener un margen de error para el estimado puntual y crear intervalos de estimación.

Se analizaron diferentes métodos para el cálculo del error de muestreo utilizados para este tipo de encuestas. Teniendo en cuenta las recomendaciones del personal del BLS y que la OIT hace en su manual⁶, el método que se ha considerado más adecuado es el de semimuestras reiteradas. Este método es muy utilizado en encuestas de diseños complejos como pueden ser la encuesta al Grupo Trabajador. De hecho, es utilizado en otros países para el cálculo del error de muestreo de las encuestas al Grupo Trabajador (p.e. España y Estados Unidos), aunque adaptado a las características de cada país.

Este método consiste en la utilización de reiteraciones (de 30 a 100 reiteraciones, en esta encuesta pueden ser 40), cada una de las cuales es una semimuestra completa constituida aleatoriamente por la mitad de los grupos de bloques que forman la encuesta. Para la formación de las 40 reiteraciones se agrupan por pares los grupos de bloques de cada estrato, asignando aleatoriamente la 1ª sección de cada par a 20 reiteraciones y la 2ª a las otras 20.

La fórmula que permite calcular el error de muestreo es:

$$D(\hat{\mathbf{X}}) \approx \sqrt{\frac{1}{r} \sum_{i=1}^r (\hat{\mathbf{X}}_i - \hat{\mathbf{X}})^2}$$

⁶ Encuestas de Población Económicamente Activa, Empleo, Desempleo y Subempleo. Manual de la O.I.T sobre Conceptos y Métodos. Página 597.

siendo:

r = el número de reiteraciones utilizadas

$\hat{\mathbf{X}}_i$ = la estimación obtenida con la i -ésima reiteración aplicando el mismo proceso de estimación que para la muestra completa

$\hat{\mathbf{X}}$ = la estimación completa de la muestra.

A partir de la cual se obtiene un error de muestreo relativo en forma de coeficiente de variación en porcentaje, cuya expresión es:

$$C.V. = \frac{D(\hat{\mathbf{X}})}{\hat{\mathbf{X}}} \times 100$$

Una vez expuestos los diferentes métodos de estimación del error de muestreo, cabe destacar que el error de muestreo proporciona un intervalo numérico en el que existe una cierta confianza (medida en términos de probabilidad) de contener el verdadero valor de la característica estimada. El intervalo de confianza más utilizado es el comprendido entre la estimación más-menos dos veces (1.96) el error de muestreo, el cual tiene un nivel de confianza del 95% de encontrar dentro de sus límites el valor verdadero que se está estimando.

17. Errores ajenos al muestreo

En este grupo se incluyen todos los errores que no se producen por la utilización de muestras para la estimación de una característica determinada. Es decir, son errores independientes al tipo de muestra seleccionado. Incluyen todos los errores que se producen en las distintas fases de la realización de una encuesta, desde el diseño del cuestionario hasta la

DEPARTAMENTO DEL TRABAJO Y RECURSOS HUMANOS

Encuesta sobre el Grupo Trabajador

publicación de los resultados finales (definiciones deficientes, errores de marco, falta de respuesta en las unidades de informantes, errores de codificación y grabación, etc.).

Hay que destacar la importancia de controlar estos errores ya que los mismos pueden hacer que, a pesar de tener un buen diseño muestral, los resultados obtenidos no tengan ni la calidad ni la fiabilidad esperadas. Es muy importante tener en cuenta que el control de calidad de una muestra tiene que abarcar todo el proceso de la misma.

En la encuesta al Grupo Trabajador se realizará el estudio y tratamiento de la falta de respuesta, así como de los errores de cobertura y contenido.

- **Falta de respuesta** originada por las negativas de las unidades informantes: se puede distinguir 2 tipos: parcial y total. En el caso de la falta de respuesta parcial (faltan algunos datos de la persona encuestada) no se realizará ningún tipo de imputación. Por otro lado, la falta de respuesta total (vivienda encuestable que no ha podido ser encuestada por ausencia, negativa o inaccesibilidad en el momento de la entrevista) recibe un tratamiento distinto según se trate de la primera entrevista o sea una posterior.
 - En el caso de la primera entrevista; las viviendas ausentes e inaccesibles serán visitadas nuevamente hasta conseguir la información. Si la vivienda sigue inaccesible o los ocupantes

ausentes se volvería el mes siguiente.

- En las visitas posteriores a la primera; el tratamiento consistirá en repetir los datos de la vivienda obtenidos en el mes anterior⁷. Esta imputación sólo se realiza la primera vez que sucede esta incidencia.

Por otro lado, las negativas, tanto si es en la primera visita como en las posteriores, no se sustituyen y se vuelven a visitar los meses posteriores.

- **Errores de cobertura:** son los originados por la omisión o inclusión errónea de unidades en la encuesta.
- **Errores de contenido:** se producen cuando alguna o algunas características del entrevistado están erróneamente recogidas en el cuestionario.

Para evaluar y detectar los errores de cobertura y contenido, se propone la realización de una encuesta de evaluación consistente en seleccionar una submuestra de viviendas en las cuales se repite la entrevista, se recomienda una submuestra equivalente al 10% del total de entrevistas. Este método es conocido como el de la **entrevista repetida** y se basa en el modelo desarrollado por Hansen, Hurwitz y Bershada para la Oficina del Censo de Estados Unidos (*Measurement errors in censuses and surveys*).

⁷ Instituto Nacional de Estadística de España (2002). *Encuesta de Población Activa. Descripción de la Muestra.*

DEPARTAMENTO DEL TRABAJO Y RECURSOS HUMANOS
Encuesta sobre el Grupo Trabajador

Los datos obtenidos en la primera entrevista o entrevista original (E.O.) se cotejan con los obtenidos en la segunda entrevista o entrevista repetida (E.R.). A partir de esta comparación se obtienen indicadores de la cobertura de personas y viviendas. Mientras que para cuantificar los errores de contenido de una determinada característica C, con K modalidades, se utilizan los siguientes indicadores:

1. Porcentaje de idénticamente clasificados de la modalidad i-esima de la característica C.

$$P.I.C = \frac{n_{ij}}{n_i} \times 100$$

Es indicador de la estabilidad de respuesta.

2. Índice de cambio neto de la modalidad i-esima de la característica C.

$$I.C.N. = \frac{n_{.i} - n_i}{n_i} \times 100$$

Es un indicador del signo y de la cuantía del sesgo de la estimación proporcionada por la encuesta.

3. Índice de consistencia global de la característica C.

$$I.C.G. = \frac{\sum_{i=1}^k n_{ii}}{n} \times 100$$

donde para los tres indicadores:

i= representa la modalidad determinada de la característica C (i=1,...,K)

n_{ii} = es el número de personas que han sido clasificadas con la modalidad i en ambas entrevistas (E.O. y E.R.)

n_i = es el número de personas que han sido clasificadas con la modalidad i en la entrevista repetida (E.R.)

$n_{.i}$ = es el número de personas que han sido clasificadas con la modalidad i en la entrevista original (E.O.)

n = es el total de personas clasificadas en ambas entrevistas respecto a la característica C de referencia

$$n = \sum_{i,j} n_{ij}$$

El realizar esta submuestra supone un esfuerzo adicional tanto de recursos humanos como económicos, pero está demostrado, de acuerdo a evaluaciones de este tipo realizadas en otros países para encuestas similares, que es muy importante este tipo de control de la información ya que nos ayuda a detectar errores o patrones erróneos de respuesta, de recopilación de la información, de selección de las unidades muestrales, e incluso en el caso que la entrada de datos siga siendo manual, de errores asociados a la entrada de datos.

También se realizará una evaluación lo más personalizada posible de los resultados por encuestador (especialmente los primeros meses de entrada de la nueva encuesta) para que se puedan detectar posibles errores y sesgos producidos por la entrada de la nueva planilla. Esto se llevará a cabo tanto a través de la

submuestra anterior como de los datos recopilados mensualmente en la muestra.

18. Recogida y Tratamiento de los datos

Los métodos de recogida de la información combinan entrevistas personales con telefónicas. Si bien la 1ª y 5ª visitas se deben hacer específicamente a través de entrevistas personales, las otras 6 consultas se pueden hacer a través de entrevistas telefónicas.

La recogida de información está a cargo de personal cualificado y entrenado del Departamento de Trabajo de Puerto Rico, que se encuentra distribuido por las oficinas de área.

Actualmente, El Departamento del Trabajo y Recursos Humanos, al igual que la mayoría de los países desarrollados, está cambiando el sistema de recogida de la información, pasando del sistema clásico (conocido internacionalmente por P.A.P.I. o Paper Assisted Personal Interviewing, por sus siglas en inglés) por sistemas de recogida de la información más modernos y dinámicos, conocidos como CAPI (Computer Assisted Personal Interviewing o entrevistas personales asistidas con computadoras personales portátiles) y CATI (Computer Assisted Telephone Interviewing o entrevistas telefónicas asistidas con computadoras).

19. Bibliografía consultada

Departamento del Trabajo de Puerto Rico. Negociado de Estadísticas del Trabajo. *Diseño de la Muestra de Viviendas 1990.*

Organización Internacional del Trabajo. *Encuestas de Población Económicamente Activa, empleo, desempleo y subempleo. Manual sobre conceptos y Métodos.*

US Department of Labor. Bureau of labor Statistics (2000). *Current Population Survey. Design and Methodology. Technical Paper 63.*

US Department of Labor. Bureau of labor Statistics (2002). *Current Population Survey. Design and Methodology. Technical Paper 63RV.*

Instituto Nacional de Estadística de España (2001). *Encuesta de Población Activa 2001: Descripción de la Encuesta, Definiciones e Instrucciones para la Cumplimentación del Cuestionario.*

Instituto Nacional de Estadística de España (1999). *Encuesta de Población Activa 1999: Tratamiento de la Información.*

Instituto Nacional de Estadística de España (2002). *Encuesta de Población Activa 1999: Informe Técnico.*

Cochran, William G. (1987). *Técnicas de Muestreo.* Editorial Continental, México.

Azorín, Francisco y Sánchez Crespo, José Luis (1994). *Métodos y Aplicaciones del Muestreo.* Alianza Editorial, Madrid.

Pérez, Cesar (1999). *Técnicas de Muestreo Estadístico.* Editorial Ra-ma.

DEPARTAMENTO DEL TRABAJO Y RECURSOS HUMANOS
Encuesta sobre el Grupo Trabajador

Santos J., Muñoz A., Juez P. y Guzmán L.

(1999). *Diseño y Tratamiento Estadístico de Encuestas para Estudios de Mercado*. Editorial Centro de Estudios Ramón Areces.

Ferrán Aranaz, Magdalena (1996). *SPSS para Windows. Programación y Análisis Estadístico*. Editorial Mc Graw Hill.