

ASPECTOS METODOLOGICOS EHPM 2008-2012

1. OBJETIVOS:

1.1 OBJETIVO GENERAL

Generar información estadística actualizada, tanto cualitativa como cuantitativa, relacionada con las condiciones socioeconómicas y demográficas de la población salvadoreña, para facilitar el diseño o rediseño de políticas, planes, programas y proyectos que desarrollan las Instituciones Públicas, que contribuyan a elevar el bienestar de la población y que a la vez sea de utilidad a otros organismos nacionales e internacionales para los mismos propósitos.

1.2 OBJETIVOS ESPECIFICOS

- Proveer información estadística de las principales variables que caracterizan la situación demográfica de la población salvadoreña.
- Generar estadísticas actualizadas sobre las principales variables que caracterizan la situación y tendencia del empleo en el país.
- Estimar el monto y fuente de los ingresos en los hogares.
- Presentar indicadores sobre características y situación de tenencia de la vivienda en el país y de los servicios básicos que dispone.
- Proporcionar datos que permitan estimar indicadores sobre el nivel educativo de la población.
- Proporcionar cifras que faciliten el cálculo de indicadores sobre la cobertura de los servicios de salud.

2. COBERTURA GEOGRAFICA

La EHPM se realizó a nivel nacional en los 14 departamentos del país, por lo tanto tiene una cobertura a nivel nacional, por área urbana y rural, área Metropolitana de San Salvador (AMSS); además es posible obtener información agregada por departamento y a nivel de los 50 municipios autorepresentados.

3. COBERTURA TEMPORAL

La encuesta se realizó durante el período comprendido del mes de enero a diciembre de 2008.

4. VARIABLES INVESTIGADAS

La Encuesta de Hogares de Propósitos Múltiples (EHPM) se diseñó como un sistema continuo de encuestas, basadas en submuestras mensuales representativas del país, con aplicación de ocho módulos permanentes.

4.1 MODULOS DE LA ENCUESTA

Se investigaron durante todo el período de la encuesta, rubros básicos sobre población (características personales de los miembros de los hogares, edad, sexo, parentesco, etc.), educación, empleo e ingreso, salud, remesas familiares y gastos del hogar.

Se incluyeron los siguientes módulos en la encuesta:

SECCION	AREAS TEMATICAS
1	Características Sociodemográficas
2	Características de Educación
2A	Tecnología de información y comunicación
3	Características Generales de la Vivienda
4	Empleo e Ingreso
5	Actividad del Productor Agropecuario
6	Salud
7	Remesas Familiares
8	Gasto del hogar

SECCION 1) CARACTERISTICAS SOCIO DEMOGRAFICAS

A través de este módulo se recabó información estadística sobre la composición y estructura poblacional de los hogares, de tal forma que permita identificar su comportamiento demográfico.

Este módulo se aplicó al jefe del hogar, que es la persona que los miembros del mismo reconocen como tal y en ausencia de este a un informante adecuado (cónyuge o hijos mayores de 18 años). Constituye un registro básico que contiene información sobre cada uno de los miembros del hogar, tales como: parentesco con el jefe del hogar, sexo, edad y estado familiar.

SECCION 2) CARACTERÍSTICAS DE EDUCACIÓN

Las variables captadas en este módulo permiten caracterizar las principales condiciones de educación de la población encuestada, en relación con la cobertura del sistema educativo.

Se aplicó a todos los miembros del hogar de 3 años y más de edad, captándose información referida a la cobertura educativa, asistencia a centros educativos, nivel y grado que estudia, tipo de centros, gastos en educación, último nivel y grado aprobado y título o certificado obtenido. Este módulo también, permite elaborar los principales indicadores educativos como: tasa de analfabetismo de la población, escolaridad promedio, tasa bruta y neta de escolaridad.

SECCION 2A) TECNOLOGÍA DE INFORMACIÓN Y COMUNICACIÓN

Con la información obtenida en este módulo se han obtenido indicadores sobre las características de la población usuaria de Internet, captando el lugar de uso, frecuencia y las actividades en las cuales se utiliza el servicio.

SECCION 3) CARACTERÍSTICAS GENERALES DE LA VIVIENDA

A partir de este módulo se han obtenido indicadores sobre las características físicas, tenencia y servicios básicos de la infraestructura utilizada como vivienda por los hogares encuestados.

Se aplica de preferencia a jefes de hogar.

Los temas a investigar incluyen: tipo y tenencia de la vivienda, materiales, número de habitaciones, servicios básicos y gastos en vivienda (agua, alumbrado, tipo de servicio sanitario), equipamiento, datos financieros (cuota mensual, adquisición) y para completar las características de la vivienda, dentro de las condiciones ambientales, se incluye el tratamiento que se le da a los desechos sólidos.

SECCION 4) EMPLEO E INGRESO

La información que se obtiene en este módulo ha permitido producir estadísticas actualizadas sobre las principales variables que caracterizan la situación y tendencia del empleo en el país. Con el propósito de identificar aspectos que permitan visualizar la situación ocupacional en el período de referencia, el cual para el caso del empleo, es la semana anterior a la fecha de la encuesta.

Este módulo se aplica a cada una de las personas miembros del hogar de 5 años y más de edad aunque los datos se tabulen de 16 años y más. Este módulo es uno de los más importantes de la EHPM; ya que, a partir de éste se obtiene información sobre los niveles de empleo e ingreso de la población, los cuales influyen directamente en los niveles de vida de la misma.

El módulo consta de dos partes. En la primera se capta la información sobre la condición de actividad de la población (PEA, PEI, ocupados y desocupados, etc.) En la segunda se investiga el ingreso que percibe en la actividad principal en concepto de salario, cuenta propia o patrono; así mismo, el ingreso por empleo secundario, y otros ingresos no laborales.

SECCION 5) ACTIVIDAD DEL PRODUCTOR AGROPECUARIO

Con la información de este módulo se obtuvo el ingreso proveniente de la actividad agropecuaria para quienes declararon haber realizado algún tipo de ésta actividad como patrono o cuenta propia, captando el régimen de tenencia de la parcela, la superficie total, la actividad productiva a la que se dedica actualmente en la unidad agropecuaria, y los gastos incurridos en la actividad agropecuaria.

SECCION 6) SALUD

A través de este módulo se conocieron las condiciones de salud de la población, así como el uso de los servicios de salud al cual recurrió la población.

Además, se captó información relacionada a la morbilidad, paralelamente se investigan datos de asistencia a los centros de salud (a quien consultó, lugar de consulta), dificultades que enfrenta al buscar la atención médica y, los gastos que tuvo para atender el quebranto en la salud. Este módulo se administra a todos los miembros del hogar, que en la sección 1 (Características Sociodemográficas) en pregunta 108 responden ítems de 1 a 4.

SECCION 7) REMESAS FAMILIARES

La información de éste módulo permitió mejorar la captación de datos relacionados con el ingreso en el hogar y la medición del peso que las remesas tuvieron en la economía de las familias salvadoreñas.

Se aplicó a los jefes del hogar y permite extraer información en forma continua sobre el total de hogares y personas que recibieron remesas del exterior, así como el monto recibido anualmente en efectivo y en especie por este mismo concepto.

El monto global de remesas tiene una profunda repercusión social en la mayoría de la población. El módulo permitió conocer la distribución de las remesas en efectivo en la adquisición de bienes de consumo (vivienda y mejoramiento de las condiciones de la vida del grupo familiar), subsanando en gran medida las necesidades de muchos hogares y dinamizando la economía salvadoreña.

SECCION 8) GASTOS DEL HOGAR

Con relación a los datos que se obtuvieron con este módulo, su captación permite estimar el gasto promedio del hogar, así como establecer el destino del gasto del mismo, según niveles de ingresos.

Se aplicó al jefe o bien a la persona encargada o responsable de la administración del hogar (abuelo o abuela, el padre o madre, hermano mayor, etc.) y se investigó, además del concepto del bien, la frecuencia y forma de como lo obtuvo y el valor del mismo.

Esta sección se dividió en cuatro sub-secciones: la primera, informó sobre el gasto en alimentos que el hogar consumió la semana pasada. La segunda, se refirió a gastos mensuales del hogar en términos de artículos personales y del hogar de consumo frecuente. La tercera, investigó los gastos semestrales en concepto de bienes duraderos (prendas de vestir, calzado, artículos del hogar, reparación y mejoras de la vivienda, etc) y, el cuarto los gastos de los últimos 12 meses en concepto de pago de cuotas alimenticias, donaciones, regalos, aportes económicos y otros.

5. DISEÑO MUESTRAL

5.1 Cobertura de la Muestra

La población representada en la muestra correspondió a las viviendas particulares, a los hogares y personas que allí habitaban, con una cobertura nacional

5.2 Niveles de Estimación y de Desagregación de Resultados

A través del nuevo diseño muestral se obtuvieron precisiones confiables para el mayor nivel de desagregación, en el caso de los municipios autorepresentados se obtuvieron buenas precisiones con las encuestas recolectadas a lo largo del año.

NIVEL DE DESAGREGACIÓN	MUESTRA REQUERIDA
Nivel nacional-total	Trimestral
Nivel nacional-urbano	Semestral
Nivel nacional-rural	Semestral
Departamento	Semestral
Departamento-áreas urbanas	Anual
Departamento-áreas rurales	Anual
Municipio auto-representado	Anual

5.3 Marco Muestral y su Actualización

El marco muestral se basa en el material cartográfico del último Censo de Población y Vivienda (2007); estableciendo un nuevo marco muestral final de 1,676,002 viviendas.

El Salvador tiene una extensión territorial de **21,040.79** kilómetros cuadrados y su división política administrativa consiste en 14 departamentos, los que se subdividen en un total de 262 municipios. Cada municipio, a su vez, se divide en un área urbana (cabecera municipal) y área rural compuesta por los lugares conocidos como cantones.

CONCEPTUALIZACIÓN DE AREA URBANA

La delimitación de un área urbana se definió como el área geográfica constituida por la agrupación de los segmentos censales resultante de la aplicación de los siguientes criterios:

1. El segmento censal donde está ubicada la alcaldía municipal; o sea, el segmento origen del casco urbano - OCU.
2. El conglomerado de segmentos que cumplen las características de núcleo poblacional (NP), núcleo habitacional (NH) ó conglomerar viviendas en manzanas (CVM) y que colindan o encierran al segmento origen del casco urbano.
3. Dos o más segmentos que cumplen las características de núcleo poblacional (NP), núcleo habitacional (NH) ó por contener viviendas conglomeradas en manzanas (CVM), que sean colindantes entre si, y que suman 500 viviendas o más agrupadas continuamente.
4. Los segmentos clasificados NP, NH ó CVM que colinden con el área urbana de otro municipio.
5. Los segmentos que no cumplen ninguno de los criterios pero que queda circunscrito completamente dentro de una mancha urbana conformada.

6. El área rural de un municipio será la conformada por el conjunto de segmentos restantes a los clasificados como área urbana.

7. Aplicados los criterios anteriores, el municipio cuya población rural sea igual o menor al 5% se consideró totalmente URBANO.

DEFINICIONES UTILIZADAS EN CONCEPTUALIZACION ANTERIOR

Segmento Censal

Es la unidad básica estadística establecida como área de empadronamiento, que comprende un área geográfica conformada por 1 o más manzanas u otras áreas dispersas, y que cumple con el criterio de estar conformadas por una cantidad específica de viviendas.

Núcleo Poblacional (NP)

Segmento censal con densidad poblacional mayor o igual de 1,000 habitantes/km².

Centro Poblado Urbano

Asentamiento poblacional con un mínimo de 500 viviendas agrupadas formando manzanas contiguas. Puede estar conformado por dos ó más núcleos poblacionales.

Casco Urbano (CU)

Centro poblado urbano donde están alojadas las autoridades edilicias.

Origen del Casco Urbano (OCU)

Segmento censal donde está ubicado el edificio de la Alcaldía Municipal.

Núcleo Habitacional (NH)

Segmento censal con una superficie territorial menor o igual a 0.6 km².

Conglomerado de Viviendas en Manzanas (CVM)

Segmento que en donde el 60% o más de las viviendas están agrupadas conformando manzanas con una superficie menor o igual a 20,000 m².

Para una mejor cobertura geográfica el país se divide en 5 Regiones, estando constituidas por los departamentos que a continuación se detallan:

REGION I (Occidental)

Santa Ana, Ahuachapán y Sonsonate.

REGION II (Central I)

La Libertad y San Salvador, excluyendo los municipios que conforman el Área Metropolitana de San Salvador (AMSS); Chalatenango y Cuscatlán.

REGION III (Central II)

San Vicente, La Paz y Cabañas.

REGION IV (Oriental)

Usulután, San Miguel, Morazán y La Unión.

REGION V (AREA METROPOLITANA DE SAN SALVADOR)

Constituida por los municipios de: San Salvador, Mejicanos, San Marcos, Ayutuxtepeque, Cuscatancingo, Delgado, Ilopango, Soyapango, Antiguo Cuscatlán, Nueva San Salvador, San Martín, Nejapa, Apopa y Tonacatepeque.

5.4 Muestra Maestra

El enfoque más rentable y eficiente para proporcionar muestras para las encuestas de hogares a través de un período largo de tiempo es por medio de una muestra maestra. Si continuamente se actualiza la cartografía y los datos de las áreas en muestra en la muestra maestra, la muestra para cada encuesta de hogares permanece representativa a través del tiempo. El diseño de una muestra maestra se lleva a cabo normalmente después de haberse levantado un censo de población y vivienda.

El diseño de una muestra maestra involucra la selección de una muestra de Unidades Primarias de Muestreo (UPMs) seguido una muestra de Unidades Secundarias de Muestreo (USMS) dentro de cada UPM. La muestra tiene que ser de tal tamaño que pueda proporcionar unidades de vivienda muestrales para cada encuesta de hogares del período intercensal. Las UPMs y las USMS se definen en la Sección 5.6 La muestra maestra utilizará un diseño multietápico estratificado. Una encuesta individual puede usar todas las USM en la muestra maestra o sólo utilizar una submuestra. Si la USM en la muestra maestra es muy grande, se seleccionarán segmentos de menor tamaño en una tercera etapa de selección para una encuesta dada. Estos segmentos más pequeños pueden conformar Áreas de Empadronamiento (AEs) completas o partes de ellas, que fueron segmentadas para mantener los costos correspondientes al listado dentro de límites presupuestarios. Los segmentos muestrales pueden ser parcialmente rotados cada año para reducir el sesgo de condicionamiento y las altas tasas de no entrevista que puedan deberse al hecho de mantener los mismos hogares en muestra por mucho tiempo.

5.5. Universo y Unidad de Análisis

El universo para las encuestas de hogares incluye los hogares y la población que vive en la república de El Salvador. Sin embargo, la población que vive en instituciones tales como hospitales, bases militares, prisiones, conventos, etc., será excluida del marco ya que representa un porcentaje muy pequeño del total de la población y es muy difícil y caro de enumerar. Si se necesitara información para estas subpoblaciones, deberán prepararse diseños muestrales específicos para dichas subpoblaciones. Las unidades de análisis para las encuestas de hogares son los hogares y las personas que viven en los hogares. Mientras que los segmentos en el marco cubran el área completa del país (urbano y rural), el marco representará adecuadamente el universo especificado.

5.6 Marco Muestral y Unidades Muestrales

El marco muestral para un programa intercensal de encuestas se basa generalmente en los datos censales más recientes y en la cartografía más actualizada. Ya se tiene disponible del censo de mayo 2007 algunos archivos de datos y algunos mapas. Toda esta información se usa para construir la muestra maestra.

La **Tabla 1** muestra la distribución de los hogares, usando datos del censo 2007, por Departamento y por área de residencia urbano/rural, así como la distribución porcentual de los hogares. Se puede ver en la tabla que el porcentaje de hogares urbanos varía considerablemente por Departamento, desde un 1.22 por ciento en el Departamento de Morazán hasta un 44.90 por ciento en el Departamento de San Salvador.

Tabla 1

Composición del Marco Muestral de Hogares en El Salvador:
Número de Hogares por Departamento (Urbano/Rural) y Distribución Porcentual

Departamento	Viviendas	%	Urbano	%	Rural	%
Ahuachapán	70,664	5.12	28,337	3.48	42,327	7.46
Santa Ana	128,362	9.29	77,235	9.49	51,127	9.01
Sonsonate	101,507	7.35	46,410	5.70	55,097	9.71
Chalatenango	44,120	3.19	13,407	1.65	30,713	5.41
La Libertad	158,931	11.51	98,438	12.09	60,493	10.66
San Salvador	402,143	29.11	365,796	44.94	36,347	6.41
Cuscatlán	51,166	3.70	14,561	1.79	36,605	6.45
La Paz	73,572	5.33	28,332	3.48	45,240	7.97
Cabañas	30,655	2.22	10,757	1.32	19,898	3.51
San Vicente	36,950	2.68	17,394	2.14	19,556	3.45
Usulután	82,939	6.00	33,656	4.14	49,283	8.69
San Miguel	105,159	7.61	54,667	6.72	50,492	8.90
Morazán	38,706	2.80	9,933	1.22	28,773	5.07
La Unión	56,430	4.09	14,983	1.84	41,447	7.30
Total	1,381,304	100.0	813,906	100.0	567,398	100.0

La **Tabla 1A** presenta el cambio porcentual en el número de hogares a nivel nacional, urbano y rural entre el censo de septiembre de 1992 y el Censo de Población y Vivienda de mayo 2007.

Tabla 1A

Cambio Porcentual en el Número de Hogares entre el Censo de Mayo 2007 y el de Septiembre de 1992

	Número de Hogares Nivel Nacional	Número de Hogares Nivel Urbano	Número de Hogares Nivel Rural
Mayo 2007	1,381,304	813,906	567,398
Septiembre 1992	1,100,714	598,842	501,872
Cambio Porcentual	25.49%	35.91%	13.06%

Para el censo de 2007, El Salvador se dividió en AEs con un promedio de viviendas entre 110 a120 por AE. Sin embargo, los tamaños de las AEs varían, como puede observarse en la Tabla 2. Las unidades muestrales definidas en el marco de la muestra maestra se basan en un diseño que involucra varias etapas. Las UPMs dentro de cada estrato se definen normalmente en base a las subdivisiones administrativas del país y en base a consideraciones de conveniencia operacional y control. En un marco de muestra maestra es importante que las USMs sean lo suficientemente grandes para poder proporcionar hogares muestrales para cada encuesta durante el período intercensal. En encuestas de hogares anteriores, las AEs fueron utilizadas para este propósito. Su tamaño es particularmente conveniente para encuestas únicas y pueden ser seleccionadas en una etapa.

Tabla 2

Composición del Marco Muestral de Hogares de El Salvador: Número de AEs y Promedio de Hogares por AE por Departamento y Area Urbano/Rural

Depto	Total		Urbano		Rural	
	# AEs	# HH/AE	#AEs	#	#AEs	#H
Ahuachapán	668	106	228	125	440	97
Santa Ana	1124	114	594	130	530	97
Sonsonate	901	113	347	134	554	100
Chalatenango	514	86	115	117	399	77
La Libertad	1419	112	801	123	618	98
San Salvador	3148	128	2812	130	336	108
Cuscatlán	430	119	103	142	327	112
La Paz	747	99	230	124	517	88
Cabañas	302	102	88	123	214	93
San Vicente	412	90	152	115	260	76
Usulután	778	107	270	125	508	97
San Miguel	958	110	468	117	490	103
Morazán	389	100	85	117	304	95
La Unión	633	90	140	107	493	84
TOTAL	12,423	112	6,433	127	5,990	95

Como existen varias AEs que contienen pocas viviendas, se recomienda que las mismas se consoliden con AEs aledañas en la misma área geográfica (incluyendo zona urbano/rural) antes de la creación de las UPMs. Esta operación redujo el número de AEs de 12,423 a 12,065.

Seguidamente, crear UPMs que contengan, en promedio, 3 AEs o segmentos **contiguos**. En algunos casos, podrá ser necesario que ciertas UPMs contengan 4 AEs contiguas y en otros casos 2 áreas contiguas, dependiendo de las condiciones en el terreno y del tamaño de las AEs.

Las AEs que tienen menos de 51 hogares deberán consolidarse con AEs adyacentes antes de crear las UPMs para que, de esta manera, no haya ninguna AE que tenga menos de 51 hogares. Todas las UPMs deben encontrarse complementamente dentro de la misma Municipalidad. ***Ninguna UPMs puede contener AEs de Municipalidades diferentes o de una área de residencia diferente (urbano/rural)***. El país estará dividido en aproximadamente 4,020 UPMs de tamaño promedio igual a 3 AEs. Esta metodología proporcionará UPMs de tamaño suficiente para garantizar suficiente muestra para cualquier encuesta en el periodo intercensal. Los tamaños de las AEs son también apropiados porque muchas encuestas requieren un número grande de hogares por AE en el estrato rural. Otra consideración es que las encuestas demográficas y de migraciones normalmente requieren muestras más grandes dentro de cada USM (el AE o un segmento de la AE), ya que las características demográficas tienen una correlación intraclase más baja que las variables económicas. Por lo tanto, se recomienda definir USMs con un tamaño mínimo de 51 viviendas.

Se eligió una muestra de 1,908 UPMs. Este tamaño de muestra satisfará los requisitos de muestra de la mayoría de las encuestas que se levanten durante el período intercensal. La DIGESTYC puede usar todas las 1,908 UPMs para las encuestas de gran envergadura o puede usar una submuestra de las UPMs para encuestas más pequeñas. La submuestra de UPMs para las encuestas pequeñas deberá elegirse en el mismo orden en que se eligió la muestra original de las 1,908 UPMs. La distribución de las UPMs muestrales por área urbano/rural se presenta en la Tabla 3.

Tabla 3

Composición de los Hogares en el marco muestral:
Estimación del número de UPMs y USMs por region y subestrato (Urbano/Rural)

Región (1)	# of AEs en el Marco			# of UPMs en la Muestra		
	Urbano	Rural	Total	Urba	Rural	Total
Ahuachapán	228	440	668	38	58	96
Santa Ana	594	530	1,124	72	84	156
Sonsonate	347	554	901	62	70	132
Chalatenango	115	399	514	40	44	84
La Libertad	801	618	1,419	120	72	192
San Salvador	2,812	336	3,148	360	132	492
Cuscatlán	103	327	430	42	42	84
La Paz	230	517	747	40	56	96
Cabañas	88	214	302	38	58	96
San Vicente	152	260	412	38	34	72
Usulután	270	508	778	44	52	96
San Miguel	468	490	958	58	62	120
Morazán	85	304	389	44	52	96
La Unión	140	493	633	36	60	96
Total	6,433	5,990	12,423	1,032	876	1,908

Para el AMSS, la DIGESTYC llevó a cabo una estratificación socioeconómica de las AEs en tres categorías: estrato alto, estrato medio y estrato bajo (ver más adelante en 5.7.11). Por lo tanto, se pudieron formar las UPMs combinando AEs que pertenecieron al mismo estrato socioeconómico dentro de las fronteras de las municipalidades que conformaron el AMSS.

Una vez que las UPMs recientemente creadas fueron asignadas a los estratos socioeconómicos correspondientes, la DIGESTYC procedió a seleccionar la muestra de UPMs.

Un ejemplo permitirá ilustrar la selección de las UPMs muestrales.

Supongamos que se determina la siguiente asignación de UPMs por estrato socioeconómico para un cierto dominio de estimación, digamos el AMSS. La Tabla 4 más abajo presenta la asignación de $n = 92$ UPMs urbanas que deben seleccionarse en muestra. Supongamos que el dominio de estimación tiene el número de UPMs en el marco que se muestra en la columna 3 de la Tabla 4. Supongamos también que 192 UPMs en el marco pertenecen al estrato

socioeconómico alto, que 354 pertenecen al estrato socioeconómico medio y que 846 pertenecen al estrato socioeconómico bajo. Como necesitamos elegir 92 UPMs en la muestra, procedemos como se muestra en la Tabla 4.

Ahora, como los hogares en el estrato socioeconómico alto tienen tendencia a tener tasas altas de no respuesta, será más conveniente elegir un número más alto de UPMs del estrato socioeconómico alto para compensar por la tasa alta de no respuesta. Por lo tanto, la asignación que se muestra en la Tabla 4 tiene que considerarse a la luz de número real de UPMs deseado del estrato socioeconómico alto.

Tabla 4

Asignación de la Muestra de UPMs dentro de cada Dominio de Estimación (AMSS)

Dominio (1)	Subestrato Socioeconómico (2)	Número de UPMs en el Marco (3)	Número de UPMs en la Muestra (4)
AMSS	1 = Alto 2 = Medio 3 = Bajo	192 354 <u>846</u> 1,392	$(92 \times 192) / 1392 = 13$ $(92 \times 354) / 1392 = 24$ $(92 \times 846) / 1392 = \underline{55}$ 92

Una vez que las UPMs se seleccionan en la muestra, se debe proceder a seleccionar las USMs. Para hacer esto, de cada UPM en muestra, se seleccionará una USM con probabilidad proporcional al tamaño (PPT). En algunos casos, puede ser necesario tener que segmentar la USM si la misma contiene un gran número de hogares. Si una USM contiene 150 hogares o menos, no se debe segmentar. Sin embargo, si la USM contiene más de 150 hogares, debe segmentarse en partes que no excedan las 100-150 viviendas por parte. Una vez que la USM se ha dividido en segmentos, se debe seleccionar uno de los segmentos con PPT. Se hace esto para tener segmentos de tamaño 150 viviendas o menos antes de llevar a cabo la operación de listado en los segmentos seleccionados. Si la USM es muy grande, la operación de listado puede encarecer los costos de la encuesta.

Se debe notar que la selección de la USM con PPT seguido de la selección de una AE con PTT dentro de cada USM es equivalente a seleccionar el AE con PPT en una sola etapa. Sin embargo, las AEs adicionales en las USMs muestrales estarán disponibles para otras encuestas de hogares en el período intercensal. Desde el punto de vista operacional, esta etapa adicional de selección es bien fácil de implementar.

Para encuestas de hogares individuales, es deseable tener una muestra de áreas pequeñas en las que se pueda llevar a cabo la operación de listado de hogares de manera oportuna y económica antes de pasar a la selección de los hogares que se incluirán en la muestra. Se ha encontrado en encuestas de hogares anteriores que muchas AEs eran demasiado grandes para permitir un listado que sea económico. Por lo tanto, las AEs grandes se dividieron en segmentos más pequeños con límites bien definidos. Es muy importante asegurarse de que cada segmento tenga límites bien definidos en el terreno, sin importar el tamaño del segmento. En caso de no poder tener un límite bien definido e identificable en el terreno, es mejor usar segmentos más grandes que arriesgar tener problemas de cobertura con segmentos que no están bien delineados en el campo.

Se llevará a cabo un listado de hogares dentro de cada segmento muestral antes de cada encuesta, o al menos se hará el listado una vez al año cuando se tenga que compartir el mismo listado en varias encuestas. En muchos países, las viviendas se seleccionan en la última etapa ya que son estructuras permanentes con mucha menor variabilidad en el tiempo que los hogares. Sin embargo, en El Salvador, una estructura o una unidad de vivienda puede tener más de un hogar (esto es más pronunciado en la parte rural del país) y esta situación puede presentar problemas para controlar la carga de trabajo de los entrevistadores.

Por lo tanto, sería operacionalmente más conveniente seleccionar hogares directamente a partir del listado. Esta es una opción viable si el listado se lleva a cabo en un período no mayor de tres meses anteriores al levantamiento de la encuesta. Si no, se tendrá errores de cobertura debido a cambios en los hogares a través del tiempo.

Cuando se seleccionan UPMs, USMs, AEs y segmentos (si fuera necesario) con PPT en las etapas respectivas, y luego se selecciona un número constante de hogares por segmento en la última etapa, la muestra será aproximadamente auto-ponderada dentro del estrato. La variabilidad en las ponderaciones depende en la exactitud de las medidas de tamaño (número de hogares) en el marco. Por lo tanto, la palabra segmento significará ya sea una AE entera o uno de los segmentos derivados a partir de la segmentación de una AE grande. Por lo tanto, los segmentos representan las Unidades Terciarias de Muestreo (UTMs).

5.7 Estratificación

Uno de los rasgos más importantes de un diseño muestral eficiente es la estratificación del marco muestran en áreas homogéneas. La selección de la muestra se lleva a cabo independientemente dentro de cada estrato, aunque se desea ordenar las UPMs bajo ciertos criterios dentro de cada estrato para proporcionar, de esa manera, una estratificación adicional implícita cuando se usa selección sistemática. La naturaleza de la estratificación depende de las

características más importantes que se quieran medir en la encuesta, así como de los dominios de estimación o de análisis.

El primer nivel de estratificación corresponde a los dominios geográficos más importantes que se definen en las encuestas de hogares: Departamento y 50 Municipalidades dentro de los Departamentos. Las unidades muestrales se estratificarán aún más por área urbano/rural para propósitos de muestreo únicamente. No habrá una estratificación explícita a nivel urbano/rural.

Para mejorar la eficiencia del diseño dentro de cada estrato, es importante definir subestratos que sean homogéneos para ciertas características clave que se medirán en la encuesta. Aunque los objetivos de cada encuesta son diferentes, es posible proporcionar un cierto nivel de estratificación geográfica y socioeconómica que mejore la eficiencia de la mayoría de las encuestas.

Dado que la muestra maestra se usará en muchas encuestas de hogares durante el período intercensal, es más efectivo y eficiente invertir, al momento del diseño, en el desarrollo de una estratificación eficiente del marco. A la vez, la oportunidad de la disponibilidad de la muestra maestra es también importante. A tal efecto, la DIGESTYC ha llevado a cabo una estratificación visual del AMSS, el único centro urbano en el marco para el cual se debe llevar a cabo una estratificación socioeconómica.

Se presenta a continuación varios métodos de estratificación que requieren diferentes niveles de investigación y tiempo de implementación. Se documentan aquí para que la DIGESTYC pueda beneficiarse en censos venideros. La estratificación final adoptada depende de los recursos disponibles y de la oportunidad del diseño, así como de la planificación a largo plazo relacionada con la eficiencia muestral y con economías de escala. Las alternativas de estratificación se presentan más abajo separadamente para áreas urbanas y áreas rurales.

5.7.1 Alternativas de Estratificación Urbana

Dada la alta variabilidad en las características socioeconómicas de las grandes ciudades, es crítico llevar a cabo una estratificación en este estrato. La mayoría de las variables importantes estudiadas en las encuestas de hogares como el empleo, ingreso, educación, y salud están correlacionadas con el estado socioeconómico del hogar y de las personas dentro de los hogares. Por lo tanto, será ideal definir subestratos socioeconómicos dentro de las grandes ciudades. Como muchas características, incluyendo el estado socioeconómico, varían geográficamente dentro de una gran ciudad, es también posible mejorar la eficiencia del diseño ordenando las UPMs en el marco de la gran ciudad de manera geográfica para obtener una estratificación implícita cuando se usa el método de selección sistemática de unidades. Se presentan aquí tres alternativas para la estratificación urbana de las grandes ciudades.

Para propósitos de estratificación, el AMSS se define como una gran ciudad. En censos venideros, la DIGESTYC puede considerar la estratificación de otras ciudades (San Miguel y Santa Ana) y decidir si dicha estratificación es viable. Para invertir recursos en la estratificación de una gran área urbana debe existir en la misma una diversidad socioeconómica tal que justifique la inversión de tiempo y recursos en la estratificación de dicha área.

Al momento del diseño de la muestra maestra, la DIGESTYC había llevado a cabo el método de estratificación detallado en la Alternativa 2 de la Sección 6.1.2. de este informe. No habrá una estratificación socioeconómica para San Miguel y para Santa Ana. Los tipos de estratificación más comunes se detallan a continuación para propósitos de documentación.

5.7.1.1 Alternativa 1 para la Estratificación de una Gran Ciudad

Será eficiente y operacionalmente viable definir tres subestratos socioeconómicos dentro de cada ciudad grande: alto, medio y bajo. Cada UPM dentro de una gran ciudad puede clasificarse en una de las tres categorías antes de llevar a cabo la selección de las UPMs. Como el estrato alto es generalmente pequeño (es decir, puede no contener muchas UPMs), los criterios que definen este subestrato pueden incluir la clase media alta. El censo de mayo 2007 proporciona una riqueza de información reciente con respecto a las características socioeconómicas de la población y de los hogares. Al mismo tiempo, no es deseable introducir procedimientos de estratificación que sean muy complejos o que lleven mucho tiempo de implementación.

Para llevar a cabo este tipo de estratificación, la DIGESTYC debe identificar variables que sirvan para discriminar hogares y poder clasificarlos en diferentes estratos socioeconómicos. Como ejemplos de estas variables tenemos:

<u>Variable</u>	<u>Categorías</u>
1.Educación del jefe del hogar	Educación secundaria o mayor
2.Profesión del jefe del hogar	Identificar categorías profesionales con altos ingresos
3.Disponibilidad de inodoro	Presencia o ausencia o compartir
4.Cocina dentro de la vivienda	Presencia o ausencia o compartir
5.Agua por tubería dentro de la casa.	Presencia o ausencia

5. Número de dormitorios

Número de dormitorios

6. Número de personas por hogar

Número de personas por hogar

Una vez que se han identificado todas estas variables, se debe usar el siguiente criterio para decidir inclusión dentro de un subestrato socioeconómico: si más del 50 por ciento de los hogares en la UPM tiene una tasa de inclusión de al menos 75 por ciento para las variables definidas, la UPM pertenece al estrato socioeconómico alto; entre 50 y 75 la UPM pertenece al estrato medio, y menos de 50 por ciento, la UPM pertenece al estrato bajo. Aun esta simple estratificación puede mejorar la eficiencia de la muestra maestra. Una variable muy importante es el “Número de personas por dormitorio” como indicador de hacinamiento.

5.7.1.2 Alternativa 2 para la Estratificación de una Gran Ciudad

La segunda alternativa se basa en definir tres estratos socioeconómicos en los mapas de la gran ciudad. Un cartógrafo puede trabajar con personal que esté familiarizado con las diferentes partes de la gran ciudad y delinear límites alrededor de áreas que pertenezcan a uno de los tres estratos socioeconómicos. Se deben establecer lineamientos para la estratificación de áreas de acuerdo a su estado socioeconómico. Aunque algunas áreas de la ciudad puedan estar mezcladas, una estratificación aunque sea un poco tosca puede mejorar la eficiencia del diseño muestral.

Una vez que se han definido los límites de los estratos socioeconómicos en los mapas de la ciudad, será necesario examinar las UPMs en los mapas censales. Se identificará el subestrato socioeconómico para la UPM localizando la UPM en el mapa de la ciudad y determinando que subestrato contiene la mayoría del área de la UPM. Seguidamente, se le asignará un código de subestrato a la UPM. Estos códigos serán introducidos en la computadora conjuntamente con los datos censales correspondientes a la UPM en cuestión. Para cada estrato en la ciudad se preparará un archivo que contenga un registro para cada UPM, incluyendo datos que identifiquen a la UPM, el número total de hogares, el número total de viviendas, el total de la población, y el código de subestrato. Dentro de cada subestrato, se ordenarán las UPMs por código geográfico. Este archivo será usado para la selección sistemática PPT de las UPMs dentro de cada subestrato de la ciudad.

Se seleccionará una muestra de UPMs de cada subestrato para verificación en el terreno. Un estadístico o un supervisor de campo con mucha experiencia deberá visitar cada UPM seleccionada y verificar visualmente el código de subestrato asignado para determinar la consistencia entre el código asignado y la realidad en el terreno.

Esta alternativa debe considerarse si no hay tiempo suficiente para investigar e implementar la primera alternativa. Aunque este procedimiento de estratificación sea algo subjetivo, es mucho más eficiente que la tercer alternativa que se describe más abajo.

5.7.1.3 Alternativa 3 para la Estratificación de una Gran Ciudad

La tercera alternativa es la más simple de implementar ya que sólo involucra ordenar las UPMs geográficamente de manera serpentina dentro de cada ciudad. Como los códigos geográficos asignados a las UPMs fueron asignados de manera contigua, los códigos correspondientes para las UPMs pueden ser usados para este ordenamiento. Esta metodología proporciona una estratificación geográfica implícita que se basa en la selección sistemática de UPMs con PPT dentro de cada ciudad. Aunque la muestra que se obtiene con este método no es tan eficiente como las obtenidas con los dos métodos anteriores, hay una cierta correlación entre la ubicación geográfica y el estado socioeconómico que dará una modesta ganancia en la eficiencia de la muestra. Esta alternativa debe considerarse sólo si hay recursos insuficientes y poco tiempo para implementar el diseño.

Al decidir entre las tres alternativas, es importante tener en cuenta los planes a largo plazo del programa de encuestas durante el período intercensal. Nuevamente, una inversión en una estratificación del marco que sea eficiente proporcionará grandes ahorros a la larga ya que se podrán usar muestras más pequeñas para obtener la misma precisión que las muestras grandes provenientes de diseños menos eficientes. La estratificación socioeconómica es especialmente importante para encuestas socioeconómicas que tengan muestras relativamente pequeñas debido a la gran cantidad de información que se recoleccione.

5.7.1.4 Estratificación Urbana de Ciudades Pequeñas (Capitales Administrativas de los Departamentos)

Como hay menos variabilidad socioeconómica dentro de las ciudades pequeñas, es posible usar una estratificación implícita ordenando las UPMs de manera geográfica dentro de cada región o Departamento.

5.8 Estratificación Rural

Dada la importancia de las actividades agrícolas en la economía de las áreas rurales, sería ideal usar criterios de estratificación que estén correlacionados con los patrones agrícolas y otras características agrícolas/agropecuarias. Por ejemplo, sería eficiente estratificar u ordenar las UPMs rurales por zona ecológica que tiene una gran influencia en las actividades agrícolas. Un estudio de cultivos agrícolas en términos de zonas ecológicas puede llevar a obtener

una estratificación rural muy eficiente, de la misma manera que las variables socioeconómicas la tuvieron en la estratificación de grandes ciudades.

Una alternativa más simple que proporcionará parte de la ganancia de una estratificación por zona ecológica será el ordenamiento geográfico de las UPMs rurales dentro de cada región. Dado que hay menos variabilidad para características socioeconómicas en las áreas rurales, esta es una alternativa razonable, aunque valdría la pena llevar a cabo un poco de investigación con respecto a los otros tipos de estratificación.

5.9 Tamaño de Muestra y Asignación

El tamaño de muestra para una determinada encuesta depende de la exactitud requerida para las estimaciones en cada dominio, así como de los recursos y de las limitaciones operacionales. La exactitud de los resultados de una encuesta depende tanto del error muestral como del error no muestral. El error muestral puede medirse por media de la varianza mientras que el no muestral sólo puede parcialmente medirse a través de re-entrevistas muy caras o de estudios de validación. El error muestral es inversamente proporcional al tamaño de muestra. Por otro lado, los errores no muestrales pueden aumentar con un incremento en el tamaño de muestra ya que es más difícil controlar la calidad de una operación de gran envergadura. Por lo tanto es importante que el tamaño de muestra global sea manejable desde el punto de vista de control de calidad y control operacional.

El diseño de la muestra maestra será flexible en términos de satisfacer requisitos de tamaños de muestra diferentes para las distintas encuestas durante el período intercensal. El número de USMs será suficiente para las encuestas de hogares que requieren el tamaño de muestra más grande. Las encuestas más pequeñas usarán una submuestra de UPMs de la muestra maestra. El número de hogares que se elegirá por segmento muestral también variará de encuesta a encuesta, dependiendo de los objetivos de la encuesta y de la correlación intraclase de las principales variables que se midan. Por ejemplo, las características socioeconómicas tienen generalmente una correlación intraclase más alta que las características demográficas. Como la correlación intraclase aumenta el efecto del diseño (la varianza obtenida con el diseño actual dividido por la varianza correspondiente proveniente de una muestra aleatoria simple del mismo tamaño), es generalmente más eficiente usar tamaños de segmentos más pequeños para encuestas socioeconómicas (aunque también es necesario tener en cuenta las limitaciones de recursos y logísticas).

En base a los objetivos de las diversas encuestas de hogares que se tienen planificadas y dada la mayor variabilidad y menor costo de recolección de datos en las áreas urbanas, se recomienda elegir una muestra de 1,908 UPMs para la muestra maestra, que proporcionará 1,032 UPMs urbanas y 876 UPMs rurales en base a una asignación más o menos proporcional. Esta asignación asegura estimaciones a nivel nacional eficientes y un tamaño de muestra adecuado en

cada Departamento del país y en cada una de las 50 Municipalidades más importantes del país.

La eficiencia muestral para una encuesta particular depende en parte del número de hogares seleccionados dentro de cada área muestral (como lo es el segmento o el AE) en la última etapa.

Para encuestas socioeconómicas en otros países en vías de desarrollo se ha determinado que el límite superior del número óptimo de hogares por segmento es de 10 para áreas urbanas y 15 para las rurales. Aunque sea necesario usar un número más grande de hogares por segmento para algunas de las encuestas debido a limitaciones logísticas o de recursos, es deseable más adelante en el período intercensal tratar de reducir el tamaño del segmento para encuestas socioeconómicas.

5.10 Procedimientos de Selección de Muestra

En base a la metodología descrita anteriormente para el diseño de la muestra maestra, se describe a continuación los pasos que se han seguido para desarrollar el marco de la muestra maestra y la subsiguiente selección de unidades muestrales en cada etapa de selección.

5.10.1 Selección de la Muestra Maestra de UPMs para el AMSS

- (1) Cada AE en el AMSS ha recibido un código de estrato socioeconómico: alto (código 1); medio (código 2); y bajo (código 3). Crear las UPMs como se ha descrito anteriormente en este informe. Cuando se termine con este proceso, se tendrán UPMs con tres códigos de estratificación socioeconómica: 1, 2, y 3.
- (2) A este punto, cada UPM en el AMSS tendrá un código identificando cada uno de los 3 estratos socioeconómicos. El AMSS está compuesto de áreas urbanas que se encuentran en las siguientes Municipalidades:

Santa Tecla ¹	Antiguo Cuscatlán ¹
San Salvador	Mejicanos
San Marcos	Soyapango
Ayutuxtepeque	Nejapa
Cuscatancingo	San Martín
Delgado	Apopa
Ilopango	Tonacatepeque

¹ Pertenecen al departamento de La Libertad, los demás a San Salvador

- (3) Distribuir el número de UPMs de la columna (5) de la Tabla 3 a los tres estratos mencionados en el párrafo anterior.

- (4) Hacer la selección de las UPMs usando selección sistemática con PPT.
- (5) Seleccionar de manera sistemática y con PPT una USM de cada UPM.
- (6) Si la USM contiene más de 150 hogares, deberá segmentarse en partes que no excedan las 100-150 viviendas por parte. Una vez que se haya dividido la USM en segmentos, se procederá a seleccionar uno de los segmentos con PPT. Esto se hace para tener segmentos de unas 150 viviendas o menos antes de llevar a cabo el listado en los segmentos muestrales. De esta manera, el trabajo de subdividir el AE en áreas más pequeñas se limita a las AE muestrales para las primeras encuestas que usen la muestra maestra. Las AES adicionales en la muestra de USMS estarán disponibles para encuestas de hogares que se lleven a cabo más adelante durante el período intercensal. Desde el punto de vista operacional, esta etapa adicional de selección es bien fácil de implementar.
- (7) Si una AE va a ser subdividida en segmentos más pequeños, tener en cuenta que es de suma importancia obtener segmentos que tengan límites bien definidos e identificables en el terreno. El tamaño de segmento que se tiene como objetivo es de 150 hogares o menos, aunque este número variará de segmento a segmento. Es más importante que el segmento tenga límites bien definidos en el terreno que el tamaño del segmento. En los casos en que no se tenga buenos límites naturales, es mejor usar segmentos más grandes que arriesgar tener problemas de cobertura con segmentos pobremente definidos. Un segmento se elegirá entonces con PPT dentro de cada AE que ha debido ser segmentada.
- (8) Antes de cada encuesta, hacer el listado de hogares correspondiente. Si esto no fuera posible, asegurarse de que el listado de hogares en los segmentos muestrales se haga al menos una vez al año.
- (9) De cada segmento muestral, seleccionar el número deseado de hogares para las encuestas individuales.
- (10) Asignar códigos geográficos en el mismo orden de selección y obtener mapas de las UPMs. Estos mapas deben contener las USMs de las cuales se crearán los segmentos.

5.10.2 Selección de la Muestra Maestra de UPMs para el Resto del Estrato Urbano y el Estrato Rural

- (1) Seleccionar el número apropiado de UPMs (ver la Columna 5 y 6 de la Tabla 3) usando una estratificación implícita (geográfica) de las UPMs.
- (2) Como se hizo anteriormente, una vez que las UPMs hayan sido elegidas, procurar los mapas de dichas UPMs para poder obtener las USMs y obtener los segmentos muestrales.

5.10.3 Última Etapa de Selección

Para la última etapa de selección, se actualizará el mapa del segmento y se llevará a cabo un listado de hogares. El listado de hogares será usado para la selección sistemática con arranque aleatorio de hogares, donde el intervalo será igual al número de hogares listados dividido por el número seleccionado para la encuesta. Si se levantan dos encuestas más o menos al mismo tiempo, es posible compartir el listado de hogares y seleccionar hogares diferentes para la segunda encuesta. Sin embargo, dado que puede haber cambios potenciales en los hogares a través del tiempo, se recomienda actualizar el listado al menos una vez al año. Como se mencionó anteriormente, se desea definir la unidad de la última etapa como la unidad de vivienda que puede definirse como el espacio ocupado por un solo hogar.

5.10.4 Cálculo de las Probabilidades de Selección de los Hogares Muestrales

La probabilidad de selección de un hogar que se selecciona en tres etapas es:

$$(1) \quad p_{hijk} = \frac{n_h M_{hi}}{M_h} p_{2hij} \frac{m_{hijk}}{M'_{hij}},$$

Donde:

p_{hijk} = probabilidad de selección del k -ésimo hogar en el j -ésimo segmento dentro de la i -ésima UPM del estrato h .

n_h = número de UPMs seleccionadas en el estrato h .

M_h = Número de hogares en el marco dentro del estrato h .

M_{hi} = Número de hogares en el marco dentro de la i -ésima UPM del estrato h .

p_{2hij} = Probabilidad de selección del j -ésimo segmento dentro de la i -

ésima UPM del estrato h .

m_{hij} = Número de hogares seleccionados en el segmento j -ésimo de la i -ésima UPM del estrato h .

M'_{hij} = Número de hogares listados en el j -ésimo segmento dentro de la i -ésima UPM del estrato h .

El componente p_{2hij} en la fórmula (1) se calcula:

$$(2) \quad p_{2hij} = \frac{M_{hij}}{M_{hi}}$$

Donde:

M_{hij} = Número de hogares en el j -ésimo segmento dentro de la i -ésima UPM del estrato h .

M_{hi} = Número de hogares dentro de la i -ésima UPM del estrato h .

Si los segmentos dentro de cada UPM contienen más o menos el mismo número de hogares, se puede entonces seleccionar el segmento con probabilidad igual y, en ese caso, la fórmula para p_{2hij} viene dada por:

$$(3) \quad p_{2hij} = \frac{1}{S_{hi}}$$

Donde:

S_{hi} = Número total de segmentos en la i -ésima UPM del estrato h .

El factor de ponderación básico o inicial es la inversa de la probabilidad inicial de selección del hogar y viene dado por:

$$(4) \quad W_{hijk} = \frac{M_h M_{hi} M'_{hij}}{n_h M_{hi} M_{hij} m_{hijk}}$$

Donde:

W_{hijk} = Factor de ponderación básico para los hogares en el j -ésimo segmento muestral dentro de la i -ésima UPM del estrato h .

Si m_{hijk} es constante dentro de cada estrato (por ejemplo, 10 hogares muestrales) y p_{2hij} se basa en una selección con probabilidad proporcional al tamaño, la muestra será aproximadamente auto-ponderada dentro de cada estrato. Estos factores de expansión variarán un poco debido a la diferencia

entre el número de hogares en el marco y el número de hogares listados dentro del segmento.

Es importante también ajustar los factores de expansión para tomar en cuenta la tasa de no entrevista. Como los factores de expansión serán calculados a nivel de segmento muestral, es conveniente ajustar los factores a este nivel. El factor de expansión final (W'_{hijk}) para los hogares muestral dentro del j -ésimo segmento de la i -ésima UPM en el estrato h se puede expresar de la siguiente manera:

$$(5) \quad W'_{hijk} = W_{hijk} \frac{m_{hijk}}{m'_{hijk}}$$

Donde:

m_{hijk} = Número total de hogares válidos (es decir ocupados) seleccionados en el j -ésimo segmento muestral de la i -ésima UPM del estrato h (es decir, la suma de las entrevistas y las no entrevistas dentro del segmento muestral).

m'_{hijk} = Número total de hogares entrevistados dentro del j -ésimo segmento de la i -ésima UPM en el estrato h .

5.10.5 Estimaciones

Las estimaciones más comunes que se calculan en las encuestas son totales y razones. La estimación de un total viene dada por:

$$(6) \quad \hat{Y} = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{k=1}^{m'_{hijk}} W'_{hijk} y_{hijk}$$

Donde:

L = Número de estratos

y_{hijk} = Valor de la variable y para el k -ésimo hogar muestral en el j -ésimo segmento dentro de la i -ésima UPM del estrato h .

La estimación de una razón viene dada por:

$$(7) \quad \hat{R} = \frac{\hat{Y}}{\hat{X}}$$

Donde \hat{Y} y \hat{X} son estimaciones de totales para las variables “Y” y “X”, respectivamente, que se calculan como se muestra en la fórmula (6) más arriba.

Cuando se utilizan diseños con conglomerados las medias y las proporciones son casos particulares de razones. En el caso de la media, la variable X en el denominador de la razón es igual a 1 para cada unidad, de tal manera que el denominador es la suma de los factores de expansión. En el caso de una proporción, la variable X en el denominador también es igual a 1 para todas las unidades, pero la variable Y en el numerador es binomial y es igual a 0 o a 1, dependiendo de la ausencia o presencia, respectivamente, del atributo especificado para la unidad observada.

5.10.6 Procedimientos de Estimación de Varianzas

En la publicación de los resultados de una encuesta es importante incluir información sobre la confiabilidad de las estimaciones. Además de presentar cuadros que muestren errores muestrales, es importante también incluir algunas fuentes de errores no muestrales para que el usuario de datos pueda tener una idea clara de la validez de los datos.

El error muestral se usa para determinar el nivel de confiabilidad de una estimación y contiene una parte del error no muestral. El estimador de la varianza debe tomar en cuenta los diferentes aspectos del diseño muestral, como la estratificación y el uso de conglomerados. Hay muchos programas que se pueden utilizar para calcular errores muestrales (o varianzas). Algunos de éstos son gratuitos (CENVAR) y otros, como SPSS o STATA, contienen módulos que permiten el cálculo de varianzas para diseños complejos. Todos estos paquetes usan el estimador que en inglés se conoce con el nombre de “ultimate cluster variance estimator” para calcular varianzas para estimadores lineales. La fórmula para estimar la varianza de un valor agregado o total viene dada por:

$$(8) \quad V(\hat{Y}) = \sum_{h=1}^L \left[\frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\hat{Y}_{hi} - \frac{\hat{Y}_h}{n_h} \right)^2 \right],$$

Donde:

$$(9) \quad \hat{Y}_{hi} = \sum_{k=1}^{m_{hijk}} W_{hijk} y_{hijk}$$

$$(10) \quad \hat{Y}_h = \sum_{i=1}^{n_h} \hat{Y}_{hi}$$

La fórmula para estimar una razón viene dada por:

$$(11) \quad V(\hat{R}) = \frac{1}{\hat{X}^2} \left[V(\hat{Y}) + \hat{R}^2 V(\hat{X}) - 2\hat{R} \text{Cov}(\hat{X}, \hat{Y}) \right]$$

Donde:

$$(12) \quad \text{Cov}(\hat{X}, \hat{Y}) = \sum_{h=1}^L \left[\frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\hat{X}_{hi} - \frac{\hat{X}_h}{n_h} \right) \left(\hat{Y}_{hi} - \frac{\hat{Y}_h}{n_h} \right) \right]$$

donde $V(\hat{X})$ y $V(\hat{Y})$ se calculan como se muestra en la fórmula (8) más arriba.

5.10.7 Información que se Debe Mantener para la Muestra Maestra

Es importante mantener una hoja electrónica o un archivo en formato EXCEL que contenga un registro por cada UPM, por cada USM, y por cada segmento indicando las medidas de tamaño usadas en cada etapa de selección para poder calcular las probabilidades de selección y las ponderaciones correspondientes. Este mismo archivo puede usarse para introducir el número de viviendas y hogares listados en cada segmento muestral durante cada encuesta o período de encuesta, para facilitar la verificación de la actualización de las áreas muestrales e identificar problemas potenciales.

6. Consideraciones Metodológicas en una Encuesta de Hogares Continua

En las encuestas de hogares que miden características afectadas por la estacionalidad, es ideal tener submuestras representativas distribuidas a través del tiempo para representar a nivel nacional la estacionalidad a través del año.

Algunos países están usando este tipo de encuesta continua para medir el desempleo a nivel mensual o trimestral. Las submuestras representativas se distribuyen a través del tiempo y del espacio. Con este tipo de metodología, se puede mantener un conjunto de entrevistadores permanentes que entrevistan una submuestra de segmentos cada mes. Una ventaja de este enfoque es que es posible usar personal permanente y bien entrenado. Otra ventaja es que la recolección y el procesamiento de los datos se distribuyen equitativamente a

través del año. Esto facilita también la implementación de controles de calidad y operacionales.

La asignación propuesta de la muestra en la Tabla 3 permite la posibilidad de definir submuestras urbanas y rurales, ya sean trimestrales o mensuales, que pueden distribuirse equitativamente a través del año usando todas las USMs disponibles en la muestra maestra.

Una manera simple de distribuir la muestra de manera representativa a través del año es asignar un código de mes (del 1 al 12) a las USMs de la muestra maestra dentro de cada estrato en el mismo orden en que las mismas fueron elegidas, comenzando con un número aleatorio entre 1 y 12. Todas las USMs muestrales (y segmentos muestrales correspondientes) con código de mes 1 serán asignadas al mes de enero, con código de mes 2, al mes de febrero y así sucesivamente.

7. Consideraciones para la Rotación de la Muestra

No es deseable mantener los mismos hogares en la muestra durante muchos años porque la tasa de no respuesta aumentará a través del tiempo, y las respuestas pueden estar sujetas a un sesgo de condicionamiento después de varias entrevistas. Por esta razón la muestra deberá rotarse periódicamente en encuestas continuas.

Otra consideración es que un traslapo en la muestra entre dos rondas consecutivas mejora la precisión de las estimaciones de las diferencias (como por ejemplo la diferencia entre dos mediciones anuales del desempleo). En el caso de una encuesta continua se recomienda que la mitad de los segmentos sea rotado cada año para poder tener un traslapo en la muestra de un 50 por ciento entre dos rondas anuales.

En los primeros dos años de la encuesta, se deben rotar hogares en los mismos segmentos muestrales. Al principio del tercer año es una buena idea empezar a rotar un cierto porcentaje de segmentos (25 por ciento). De esta manera se renueva el stock de segmentos y mejora la actualización del marco muestral.

8. Nuevo Procedimiento Para El Trabajo de Campo

Una vez seleccionados los bloques de una UPM, su identificación (mapa y listado) se entregó al Supervisor. Los encuestadores se dirigirán no sólo a las viviendas que figuran como ocupadas, sino que investigaron en campo la situación de todas las construcciones del bloque, y realizaron las encuestas en todas las viviendas que se encontraron ocupadas en el área delimitada por el bloque. Esto significa que al momento de la Encuesta, algunas viviendas que figuraban como desocupadas hayan pasado a ser ocupadas, y en ese caso, el encuestador realizó más de cuatro entrevistas por bloque.

También podría ocurrir el caso contrario que una vivienda ocupada se hubiera desocupada entre el momento del listado y el momento de la Encuesta. Este procedimiento no es arbitrario sino que responde a la necesidad de calcular los coeficientes de expansión corrigiendo las cifras del marco mediante la información que los propios encuestadores obtienen en campo en ocasión de la realización de las encuestas.

El trabajo de cada equipo de encuestadores y supervisores se resume en las dos tareas siguientes:

Entrevistar a todos los hogares que residen en las viviendas ocupadas en los bloques seleccionados en la muestra y registrar el número de viviendas encontradas en cada lote, cuántas de ellas son viviendas ocupadas, y cuántas de éstas pudieron ser entrevistadas (algunas podrían no ser entrevistadas por rechazo, moradores ausentes u otras formas de no contacto).

9. Distribución de la Muestra en el Tiempo

Considerando que el período de la encuesta fue de Enero a Diciembre, los 1,448 segmentos seleccionados fueron distribuidos a lo largo del año en cada uno de los meses del año, conteniendo cada uno 120 segmentos aproximadamente a ser entrevistadas en campo. Los municipios y los cantones se distribuyeron a lo largo de los 12 meses, de forma que, en lo posible, un número similar de segmentos se visite para un departamento en particular a lo largo de todo el país urbano y rural, y así poder determinar la carga de trabajo a los 24 entrevistadores y 8 supervisores..

La asignación de un cantón a un mes en particular, estuvo determinada por la cercanía a los municipios que se visitarían en la semana de trabajo.

10. Consideraciones sobre la no respuesta

En toda investigación de naturaleza estadística se presenta el fenómeno que la información obtenida no corresponda exactamente a todas las unidades muestrales seleccionadas, conociendo este comportamiento como "no respuesta".

Una definición precisa por tipo se presenta a continuación.

TIPO A:

Se presenta aquellos casos cuando en las viviendas ocupadas no se obtiene respuesta, ya sea porque al visitar la vivienda en varias ocasiones nunca se

encontró a nadie o los miembros del hogar están ausentes temporalmente o porque éstos se negaron a proporcionar la información solicitada.

TIPO B:

Este tipo de casos está referido cuando la vivienda seleccionada se encuentra desocupada, ya sea en forma regular o permanente, así mismo cuando la vivienda la están construyendo o que ésta es utilizada temporalmente para negocio.

TIPO C:

Se refiere a los casos cuando al momento de la entrevista la vivienda seleccionada ya no existe, debido a que la han demolido o funciona como tienda, almacén, taller o cualquier otro tipo de negocio, ya no existe, etc.

Las no respuestas de los tipos B y C se deben a problemas en el marco muestral.

11. CONSISTENCIA Y PROCESAMIENTO DE DATOS

Una de las etapas posterior a la recolección de los datos es la de crítica y codificación de cada uno de los cuestionarios diligenciados, acciones que se efectúan en una sola operación. Entendiéndose como crítica el conjunto de reglas que permiten verificar la coherencia de la información con el fin de obtener datos depurados y consistentes, para un mejor rendimiento de costos y de tiempo; por otra parte, la codificación es la asignación de códigos o claves numéricas a las respuestas provenientes de preguntas abiertas, a efectos de facilitar la captura de datos en la computadora.

A partir del 2006, se ha utilizado, para la captura de datos, el programa CSPro que permite agilizar el ingreso de información, crítica, corrección y elaboración de tabulaciones, lo que hace posible que la información se trabaje de manera más eficiente.

Los módulos usados de CSPro para los procesos de la EHPM son:

Entrada de Datos.

Los siguientes 3 archivos son indispensables en la captura de datos:

- Diccionario. Este es un archivo en el cual se definen todas las secciones, variables, rangos de variables, nombres y títulos de variables y módulos, etc.
- Formulario. Este es un archivo de diseño que prácticamente se elabora lo más parecido a la boleta o formulario en ejecución para que el digitador se familiarice muy bien con la pantalla de la computadora y el formulario.
- Programa. Este es un archivo el cual coordina el flujo del cursor sobre los datos y las diferentes inconsistencias que se especifiquen con el digitador.

Prácticamente existe un control total del flujo de la información que se programan en este modulo.

Inconsistencias en la entrada de datos

Las inconsistencias de entrada de datos se controlan con 2 módulos de los anteriores que son el diccionario y programa; el formulario es solamente la parte visual.

Inconsistencias en la salida de datos

En CSPro se pueden generar programas específicos de inconsistencias sobre archivos de datos, para este caso se necesita el diccionario y se edita un programa. Además, se usa IMPS como un complemento para buscar inconsistencias. Actualmente también se usan módulos auxiliares de CSPro de Exportación hacia archivos en EXCEL para buscar inconsistencias.

Generación de variables auxiliares que no están en la boleta

Estas se generan por medio de un programa Batch de recodificación en CSPro conocido como RECODI.

Una vez concluida la digitación de toda la encuesta y sus datos han sido consistenciados y depurados, se separa la información por secciones por medio de CsPro, luego se procede a ejecutar una macro que esta hecha en UltraEdit y para finalizar se vuelven a unir las secciones en un archivo y se les aplican unos utilitarios de IMPS que son: Sort y Zerofill, que son necesarios para el funcionamiento de la nueva base de datos en IMPS.

12. CONCEPTUALIZACIONES BASICAS ADOPTADAS

→ **POBLACION TOTAL:** El conjunto de habitantes que residen en los hogares particulares de las diferentes regiones del país.

→ **HOGAR:** Personas o grupo de personas unidos por lazos familiares o amistosos que comparten una misma vivienda y que tienen una administración presupuestaria de alimentación común.

→ **GRUPOS DE EDAD:** Intervalos utilizados para estratificar la población de acuerdo a la edad cumplida en años.

→ **POBLACION EN EDAD DE TRABAJAR (PET):** También llamada población en edad activa, está integrada por las personas de 16 y más años de edad, que de acuerdo a la realidad socioeconómica del país se encuentran aptas para trabajar.

→ **CONDICION DE ACTIVIDAD:** Se refiere a la situación de las personas en cuanto a su vinculación con el mercado laboral, es decir, Población Económicamente Activa (PEA), y Población Económicamente Inactiva (PEI).

→ **POBLACION ECONOMICAMENTE ACTIVA (PEA):** Grupo poblacional constituido por las personas que estando en edad de trabajar, efectivamente forman parte de la fuerza de trabajo al mantenerse en una ocupación o buscarla activamente.

→ **OCUPADOS:** Personas económicamente activas que tienen un trabajo del cual obtienen una remuneración o ganancia, o trabajan sin pago en dinero en un establecimiento de tipo familiar.

i- Ocupados Plenos: ocupados que trabajan 40 horas o más a la semana y que perciben un ingreso mayor o igual al salario mínimo establecido.

ii- Sub-empleados Visibles: ocupados que trabajan involuntariamente menos de 40 horas a la semana.

iii- Sub-empleados Invisibles: ocupados que trabajan 40 horas o más a la semana y que perciben un ingreso menor al salario mínimo establecido.

→ **DESOCUPADOS:** Personas que no trabajan pero que buscan activamente un trabajo.

i- Cesantes: son los desocupados que han trabajado anteriormente, pero que buscan colocarse en un trabajo.

ii- Busca trabajo por primera vez: desocupados que buscan por primera vez un empleo, sin conseguirlo. Generalmente, corresponde a los jóvenes que salen del sistema educativo y buscan incorporarse por primera vez al mercado de trabajo.

→ **POBLACION ECONOMICAMENTE INACTIVA (PEI):** Grupo de personas que estando en edad de trabajar no trabajan ni buscan activamente empleo. Ejemplos de este grupo son los estudiantes, amas de casa, jubilados etc.

→ **JEFE DE HOGAR:** Es el miembro de un hogar que ejerce la autoridad para tomar las decisiones que involucra al grupo familiar y que es reconocido como tal por los miembros que lo conforman.

→ **SECTORES DE OCUPACION:** Se refiere particularmente a los segmentos básicos del mercado de trabajo urbano, vale decir, sector informal y sector formal o moderno.

i- Sector Informal: segmento del mercado de trabajo compuesto por: Asalariados y trabajadores familiares ocupados en establecimientos de menos de cinco trabajadores; y, trabajadores por cuenta propia y patronos de empresas con

menos de cinco trabajadores en ocupaciones no profesionales, técnicos, gerenciales o administrativos.

ii- Sector Formal: segmento del mercado de trabajo compuesto por: Asalariados y trabajadores en establecimientos de 5 ó más trabajadores; y, trabajadores por cuenta propia y patronos de empresas de cinco o menos trabajadores que desempeñan ocupaciones de tipo profesional, técnico, administrativo y gerencial.

→ **RAMA DE ACTIVIDAD:** Sector de la economía dentro del cual la persona o empresa ejerce o ha ejercido la ocupación. Se utiliza la Clasificación Industrial Internacional Uniforme de todas las actividades económicas (CIIU, revisión 3) del Departamento de Asuntos Económicos y Sociales Internacionales de las Naciones Unidas(1989).

→ **GRUPO OCUPACIONAL:** Conjunto de ocupaciones específicas relacionadas por la naturaleza del trabajo realizado. Se utiliza la clasificación Internacional Uniforme de las Ocupaciones (CIUO88 que sustituye al CIUO68) de la Organización Internacional del Trabajo (OIT).

→ **CATEGORIA OCUPACIONAL:** Muestra la relación entre una persona económicamente activa y su empleo, es decir, si la persona es:

i- Patrono: es la persona, natural o jurídica (Sociedad Anónima o el Estado) que posee uno o varios establecimientos donde se realiza una actividad económica, y tiene uno o más trabajadores a su cargo, a los que les paga una remuneración en dinero o en especie.

ii- Trabajador por Cuenta Propia: persona que desarrolla una actividad económica en forma independiente que no tiene empleados remunerados y que no es empleado de nadie (puede tener familiares no remunerados).

iii- Asalariado: persona que trabaja para un empleador y recibe una remuneración en forma de salario, sueldo más comisiones.

iv - Trabajador Familiar No Remunerado: persona que trabaja para un familiar sin percibir remuneración monetaria alguna.

v - Aprendiz: persona que al momento de la entrevista se encuentra en proceso de aprendizaje de un arte u oficio y que puede recibir o no, algún ingreso por dicha actividad.

vi - Servicio Doméstico: persona que trabaja para un grupo familiar, desempeñando labores propias del hogar.

→ **SECTOR INSTITUCIONAL:** Denota la situación del empleado en relación a la propiedad de los medios de producción de la empresa para la cual presta su servicio, es decir, sector público y sector privado.

→ **MIGRANTES:** Toda persona que adoptó como lugar de residencia habitual cualquier lugar diferente al de origen, durante el período comprendido en los últimos 5 años a la fecha de la entrevista, por razones de trabajo, educación, salud u otros motivos.

→ **SALARIO:** Retribución monetaria y/o en especie que el trabajador recibe en forma semanal, quincenal o mensual por la realización de servicios o por su trabajo. Formará parte del salario el aguinaldo proporcional y las comisiones.

→ **INGRESO:** Monto total percibido durante un período establecido (Período de referencia) en dinero y/o en especies por medio de su trabajo y otras fuentes tales como: comisiones, bonificaciones, renta, remesas y ayuda familiar, pensiones y jubilaciones, intereses, alquileres, donaciones, utilidades, etc.

→ **INGRESO FAMILIAR:** Monto total de las entradas monetarias percibidas en dinero o especie, por el conjunto de miembros que conforman un hogar.

→ **AGUINALDO:** Parte del salario ganado por el trabajador y pagado por el patrono al final de año.

→ **COMISION:** Pago monetario acordado entre la empresa y el trabajador según el volumen de ventas realizadas.

→ **BONIFICACION:** Pago monetario que el patrón otorga al trabajador en forma voluntaria de las ganancias generadas en la empresa.

CANASTA BASICA ALIMENTARIA (CBA): El conjunto de productos considerados básicos en la dieta de la población residente en el país, en cantidades suficientes para cubrir adecuadamente, por lo menos, las necesidades energéticas y proteicas del individuo promedio.

NIVEL DE POBREZA: Límite establecido al comparar el nivel de ingresos de los hogares con el costo de la Canasta Básica Alimentaria (CBA) y la Canasta Ampliada (CA, dos veces el costo de la canasta básica).

i) Pobreza Extrema o Crítica Absoluta: Se encuentran ubicados en esta situación, aquellas personas u hogares cuyos ingresos son menores que el costo de la CBA, la cual es diferenciada por área de residencia.

ii) Pobreza Relativa: Aquellas personas u hogares cuyos ingresos son mayores que el costo de la CBA, pero son menores que el costo de la CA, el cual, de acuerdo con la metodología del Banco Mundial, resulta ser el doble del costo de la CBA.

→ **ANALFABETOS:** Personas de 10 años y más que no saben leer y escribir.

→ **NIVEL EDUCATIVO:** Es la estratificación de acuerdo al grado o año escolar más alto aprobado en el sistema de enseñanza regular ; ejemplo Parvularia (uno a tres años de estudio), Básica (de primero a noveno grado), Media (de primero a segundo o tercer año de bachillerato, según el nuevo sistema educativo), etc.

→ **CASA IMPROVISADA:** Es la vivienda privada o independiente que ha sido construida con materiales de desecho (pedazos de madera, láminas viejas, cartón, varas de mangle, bambú, etc.), sin cumplir las normas de construcción. Para considerarla vivienda improvisada deberán tener paredes o techo con materiales de desecho, los servicios sanitario y baño podrán ser privado o común

12. INDICADORES UTILIZADOS

a) **TASA BRUTA DE PARTICIPACION:** Razón entre la PEA y la población total. Indica la proporción de la población total del país que se encuentra activa económicamente.

b) **TASA GLOBAL DE PARTICIPACION:** Razón entre la PEA y la población en edad de trabajar. Indica la proporción de personas en edad de trabajar del país, que se encuentra económicamente activa. Denota precisamente el nivel de actividad económica de la población.

c) **TASA DE OCUPACION:** Razón entre los ocupados y la PEA. Representa el grado de aprovechamiento efectivo de los recursos humanos disponibles para el trabajo.

d) **TASA DE CESANTIA:** Es la relación entre los cesantes y la PEA y denota la expulsión de mano de obra por el conjunto de unidades productivas, la cual puede deberse a una contracción económica, traslado o movilidad de la mano de obra, etc.

e) **TASA DE DESEMPLEO:** Relación que mide el volumen de desempleo o desocupados entre la población económicamente activa durante un período determinado, representando el porcentaje de la fuerza laboral que no es absorbida por el sistema económico.

f) **TASA DE INACTIVIDAD:** Razón entre los inactivos y la población en edad de trabajar. Es la proporción de la población en edad de trabajar que se encuentra económicamente inactiva.

g) **TASA DE DEPENDENCIA ECONOMICA:** Es la población total menos la población ocupada dividida entre la población ocupada. Este cociente es un indicador de la carga económica que soporta cada trabajador, es decir, cuantas personas que no trabajan deben ser sostenidos por cada persona ocupada.