

Diseño Muestral

El presente documento detalla el diseño de la muestra de la Encuesta Trimestral de Empleo (ETE), la cual llega a una pequeña porción de la población total con la que se pueden inferir los resultados de la totalidad de forma trimestral y anual.

1. Universo de Estudio

La Encuesta está dirigida al conjunto de hogares establecidos en viviendas particulares de las ciudades capitales de todo el país, más la ciudad de El Alto. Excluye a las personas que residen en viviendas colectivas, como ser: hospitales, cuarteles, hoteles, conventos, etc., pero incluye a las personas que residen en viviendas particulares que se encuentran dentro de las viviendas colectivas, como ser: serenos, porteros y otros.

2. Cobertura

El alcance de la ETE está determinado para las nueve ciudades capitales del país más la ciudad de El Alto sin ser capital, por el tamaño de población que tiene.

3. Marco Muestral

De acuerdo con los objetivos de la encuesta es necesario contar con un marco muestral que permita:

- Determinar el tamaño de la muestra.
- Seleccionar la muestra de acuerdo con las características del diseño propuesto.
- Determinar los errores y probabilidades de selección.

Por lo tanto el marco muestral para la Encuesta Trimestral de Empleo se construyó en base al Censo de Nacional de Población y Vivienda de 5 de septiembre de 2001, que facilitó la construcción de un marco de áreas, integrado por mapas y planos a diferentes escalas y una base de datos con la división geográfica, política, administrativa y estadística (ver cuadro 1), teniendo un total de 9.275 áreas geográficas, denominadas unidades primarias de muestreo (UPM).

El marco muestral está conformado por un listado de viviendas particulares: ocupadas y desocupadas del área urbana, las cuales están clasificadas mediante el orden de jerarquía de los identificadores censales.

La división censal se construyó con las siguientes características:

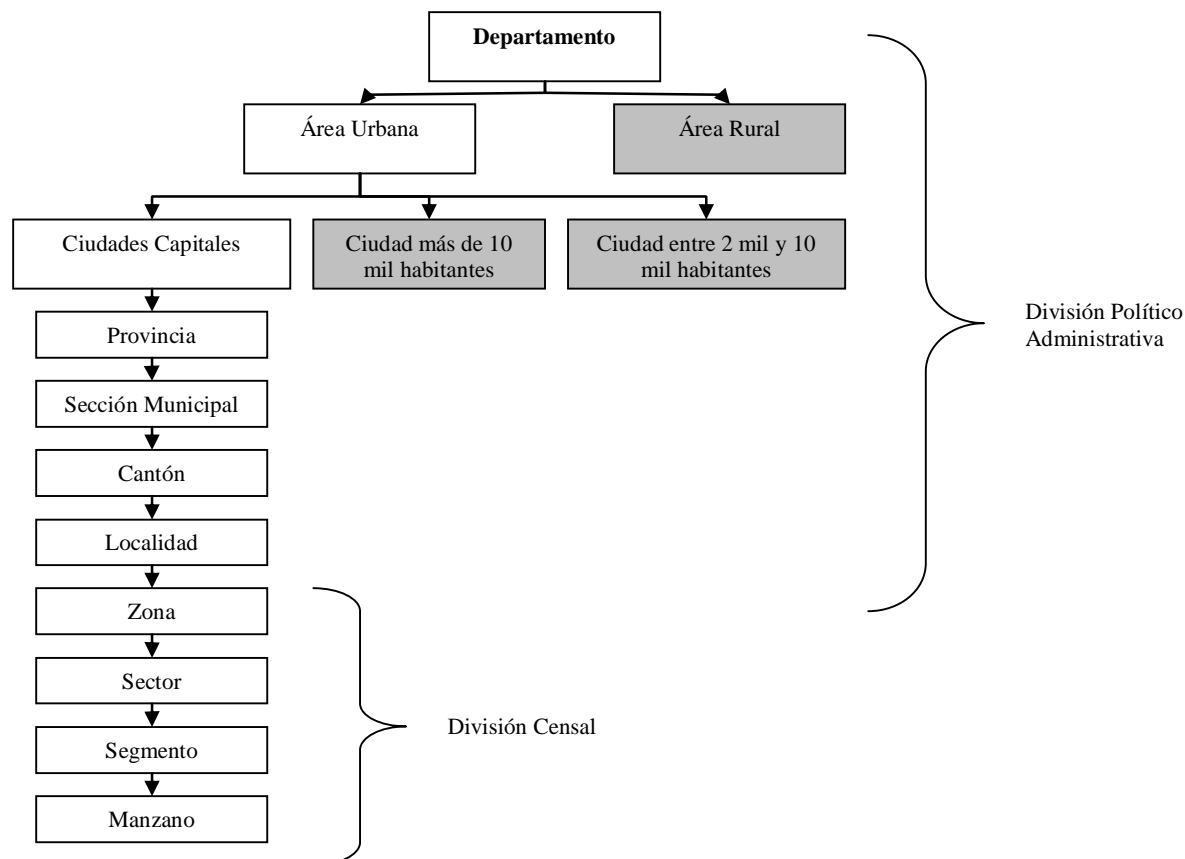
1. Las UPMs son perfectamente reconocibles en el terreno al estar delimitado por rasgos topográficos identificables y perdurables.
2. Por lo general son homogéneas en cuanto a las características sociales, económicas y geográficas.

Esta información está respaldada en un sistema automatizado, formado por una base de datos relacional, que permite la consulta directa al usuario y por la cartografía digitalizada.

Cuadro 1
Distribución del marco muestral según el Censo 2001

Departamento	Ciudad Capital	Nro. de UPMs	Nro. de Viviendas	Población Total	Población total de hombres	Población total de mujeres
Chuquisaca	Sucre	447	47.697	185.034	87.183	97.851
La Paz	La Paz	2.023	219.661	770.333	363.319	407.014
	El Alto	1.612	179.518	631.919	309.849	322.070
Cochabamba	Cochabamba	1.337	142.978	539.209	255.915	283.294
Oruro	Oruro	500	53.627	196.682	94.386	102.296
Potosí	Potosí	312	33.410	130.568	62.130	68.438
Tarija	Tarija	322	34.401	132.841	63.139	69.702
Santa Cruz	Santa Cruz	2.482	267.428	1.093.148	529.717	563.431
Beni	Trinidad	176	15.442	72.463	35.426	37.037
Pando	Cobija	58	5.044	19.393	9.692	9.701
TOTAL		9.275	999.206	3.771.590	1.810.756	1.960.834

Gráfico 1
División Geográfica, Política, Administrativa y Estadística



Las Unidades Primarias de Muestreo (UPM) están conformadas, bajo los siguientes criterios:

- Constituida generalmente por un sector censal, pero existen UPMs con dos o tres sectores censales.
- Tienen un mínimo de 40 viviendas y un máximo de 284 viviendas.
- Construidas en promedio, por 130 viviendas habiendo algunas con más viviendas y otras con menos.

Por experiencias de encuestas anteriores se sabe que en UPMs del área periurbana se tiene un crecimiento muy significativo, superando a las 284 viviendas.

4. Unidades de Observación de Muestreo

De acuerdo a los objetivos de la investigación:

- **La unidad de muestreo** en su última etapa es la vivienda.
- **La unidad de observación** es el hogar (puede haber más de un hogar bajo un mismo techo).
- **La unidad de análisis** es la población residente en las viviendas seleccionadas.

5. Tipo de Muestreo

El tipo de muestreo es probabilístico, estratificado, bietápico y por conglomerado¹. En cada estrato se obtendrá una muestra independiente que lo representará.

Probabilístico

Porque la probabilidad de inclusión de las unidades de muestreo, es decir las viviendas, tienen una probabilidad conocida y distinta de cero y por lo tanto la probabilidad de selección de cada miembro de la población es también distinta de cero.

Estratificado

Porque las unidades primarias de muestreo (UPM) con características geográficas y poblacionales similares son agrupadas en estratos.

Bietápico

Esto es porque la vivienda (unidad última de muestreo) es seleccionada de una UPM, la cual fue anteriormente seleccionada de algún estrato. Por lo tanto se llega a la unidad última de muestreo en dos etapas.

PRIMERA ETAPA: La unidad primaria de muestreo es la UPM

SEGUNDA ETAPA: La unidad secundaria de muestreo (USM) es la vivienda

¹ En algunas UPMs se introdujo una etapa más en la selección debido a la extensión geográfica.

Un proceso adicional que se realizó en la selección de la muestra es la inclusión de una etapa de selección para las UPMs que se encuentran en zonas periurbanas, puesto que la extensiones geográficas de estas UPMs son amplias y difícil de cubrir en el tiempo estipulado y con la carga de trabajo dada. Por esta razón es que en esas UPMs el muestreo es trietápico, el cual consiste en realizar una delimitación de la UPM tomando en cuenta los segmentos censales como unidad secundaria de muestreo (USM). Posteriormente de la USM se selecciona las unidades terciarias de muestreo que serían las viviendas.

Por conglomerado

Ya que las unidades primarias de muestreo son conjuntos de otras unidades muestrales (viviendas) delimitadas.

Cabe aclarar que en general, la formación de conglomerados tiende a incrementar la varianza muestral cuando no cumple la heterogeneidad dentro del conglomerado. Sin embargo, su utilización en el diseño muestral permite flexibilidad en el número de etapas de selección y reduce substancialmente los costos asociados a la construcción del marco y a la realización del trabajo de campo.

6. Estratificación

La estratificación divide las unidades de la población en grupos mutuamente excluyentes y colectivamente exhaustivos llamados también estratos, de donde se seleccionan las muestras separadas para cada uno de los ellos.

Se debe tener en cuenta que el marco muestral que se dispone, es un marco de áreas geográficas, con la información del Censo de Población y Vivienda 2001, el cual se encuentra desactualizado a la fecha.

En la Encuesta Trimestral de Empleo se consideran dos tipos de estratos, uno explícito y el otro implícito.

Estratificación explícita

En ocasiones, una sencilla estratificación geográfica representa un medio eficaz para mejorar la muestra², pero la ETE asumió dos etapas de estratificación.

En una primera etapa se tiene una estratificación geográfica (Sucre, La Paz, Cochabamba, Oruro, Potosí, Tarija, Santa Cruz, Trinidad y Cobija) alcanzando un total de 9 estratos geográficos en el marco muestral.

En la segunda etapa se asume que el desempleo está correlacionado con el nivel de educación³ de los individuos, entonces las UPMs se someten a un nuevo procedimiento de estratificación propia que emplea la variable “nivel de educación” de los jefes de hogar que habitan en las viviendas dentro de la UPM, la que se clasifica en dos grupos: con

² Manual de la OIT sobre conceptos y métodos.

³ Se trabajó un documento de estratificación donde se detalla el procedimiento.

educación superior y sin educación superior. Ambos criterios se combinan y logran un total de 18 estratos explícitos.

Se hizo un análisis para ver si esta estratificación mejoraba los errores muestrales de las variables. Con los resultados que se obtuvieron se llegó a la conclusión de que esta estratificación lograba errores muestrales e intervalos de confianza más pequeños que con una estratificación solamente geográfica.

Estratificación implícita

Dentro de cada estrato explícito se utiliza una técnica conocida como la estratificación implícita, en las UPMs antes de la selección de la muestra, donde se ordenan las UPMs al interior de los estratos explícitos con respecto a uno o más variables que están altamente correlacionados con las variables de interés y que estén disponibles para cada UPM en el estrato.

Para el caso de la Encuesta Trimestral de Empleo, se utilizó como estratificación implícita a los estratos por pobreza que se usaron en la Encuesta a Hogares 2008 donde se divide el marco muestral en cinco estratos de pobreza.

La ECH 2008 abandona el procedimiento de estratificación, basada en las necesidades básicas insatisfechas (NBI), debido a las deficiencias de dicho indicador para expresar el bienestar, toda vez que los componentes de este índice atribuyen un peso arbitrario a los componentes como vivienda, educación, atención de salud y disponibilidad de servicios básicos⁴. En reemplazo de dicho índice, el diseño construyó una variable de bienestar con la información proveniente de encuestas de hogares anteriores, específicamente de las del 2000 a 2002.

Se considera que la distribución del bienestar de los hogares está expresada en el agregado de consumo de los hogares. Éste se define como el gasto y las adquisiciones de alimentos⁵, gastos en vestimenta, transporte, servicios básicos, educación, salud, recreación y otros. También incluye los servicios de la vivienda, el alquiler imputado de los propietarios por el uso de la vivienda propia y una estimación del flujo de gastos de bienes durables⁶. La observación del agregado del consumo familiar per capita permite la construcción de un modelo predictivo, basado en variables que son observables tanto en las encuestas de hogares como también en el Censo de Población y Vivienda de 2001.⁷

Una vez aplicado este método de estratificación se logran cinco estratos: h=1,2,3,4,5 donde el más alto nivel de pobreza corresponde al estrato uno y va disminuyendo respectivamente. Pero en las ciudades capitales solamente se encuentran cuatro de los cinco estratos, ya que el estrato categorizado como el más pobre está concentrado en el área rural y no es parte de la cobertura de la encuesta.

⁴ Está compuesto por variables asociadas a la vivienda, disponibilidad de servicios e insumos de la vivienda, logro educativo y acceso a servicios de salud. La metodología completa está descrita en www.ine.gov.bo

⁵ El consumo de alimentos proviene no solo de la compra sino también del consumo de la producción propia, ingreso en especie y regalos o donaciones recibidas de otras entidades.

⁶ Esta metodología proviene de las recomendaciones del Banco Mundial (Deaton, xxx, Zaidi, xxx)

⁷ (Encuesta a Hogares 2008)

7. Dominios de estudio

De acuerdo con los objetivos de la encuesta y la necesidad de contar con información para cada uno de los departamentos se plantean los siguientes dominios de estudio:

Resultados trimestrales

- La Paz
- Cochabamba
- Santa Cruz
- Nacional

Resultados anuales

- Chuquisaca
- La Paz
- El Alto
- Cochabamba
- Oruro
- Potosí
- Tarija
- Santa Cruz
- Beni – Pando (agregado)

Lograr resultados trimestrales para cada uno de los departamentos demandaría una muestra mucho mayor a la que se tiene, lo que implica también un presupuesto mayor al que se dispone. Por tal motivo se tendrán resultados anuales para los departamentos más pequeños y resultados trimestrales para el eje troncal (La Paz, Cochabamba y Santa Cruz).

Asimismo, con la muestra planteada en este documento, desagregaciones trimestrales para los demás departamentos, generan resultados no confiables ya que la muestra es reducida para estos casos. Si el usuario decide tener resultados trimestrales para cada uno de los departamentos, se recomienda contrastar con los errores muestrales de cada indicador. Estos errores muestrales proporcionan información de la confiabilidad del resultado.

8. Tamaño de la Muestra

El tamaño de la muestra para la Encuesta Trimestral de Empleo es de 8.532 viviendas. Los tamaños de muestra fueron calculados en base a la variable: tasa de desempleo⁸; datos que fueron tomados de la Encuesta de Hogares 2003 y 2004.

Se tomó como variable auxiliar la tasa de desempleo porque este es un indicador muy importante para la encuesta y además porque es la variable que más muestra demanda. De modo que cuando se garantice el resultado para esta variable, se esta asegurando el

⁸ Se calculó con esta variable puesto que es la que más muestra demanda, de modo que se garantiza el tamaño de muestra para los demás indicadores.

resultado para las demás variables que no necesitan un tamaño de muestra mayor al propuesto.

Obviamente se puede tomar alguna otra variable auxiliar para el cálculo del tamaño de muestra, dependiendo de los intereses de cada encuesta.

La expresión matemática del cálculo del tamaño de la muestra, es la siguiente:

$$n_0 = \frac{Z^2(1-p)}{pe^2}$$
$$n = \frac{n_0 * deff}{TH * TNP * (1 - TNR)}$$

Donde:

- Z: Valor de la distribución normal estándar del nivel de confianza dado.
- P: Tasa de desempleo abierto en la Encuesta de Hogares 2003 – 2004
- Deff : Efecto del diseño
- TNP : Tasa neta de participación
- TH: Promedio de habitantes de 10 y más años de edad por vivienda
- e : Error relativo permisible
- TNR: Tasa de no respuesta

Sin duda es muy importante conocer el número de observaciones necesarias que se deben tener para poder dar estimaciones sobre una población o variable. Por lo tanto el tamaño de muestra calculado está en base de los resultados que se desean obtener, los cuales están planteados en la parte de dominios de estudio.

“Con frecuencia se planea presentar los estimadores no solo para la población total, sino para ciertas subdivisiones, entonces se realiza un cálculo separado del tamaño de la muestra para cada región y la suma de todas las muestras separadas será el tamaño de muestra total.”⁹

Considerando una confianza de 95%, un efecto de diseño de 1.72 (observado en las experiencias anteriores), una tasa neta de participación de 58.63, un promedio de 4.3 habitantes por vivienda, un error relativo máximo esperado de 7.4% para una tasa de desocupación de 8.1% y una tasa de no respuesta de 13% se determinó una muestra a nivel nacional de 8.532 viviendas.

Las 8.532 viviendas planificadas para la muestra, son agrupadas en 948 UPMs considerando 9 viviendas por UPM.

⁹ “Técnicas de Muestreo”. G Cochran.

Afijación de la Muestra

El tamaño de muestra para cada uno de los dominios de estudio se calculó de manera separada.

La muestra de viviendas de tamaño 8.532 se la distribuye empleando el **Método de Asignación Potencial** con $\alpha = 0.5$ en cada uno de los estratos dentro de cada departamento, usando la siguiente relación:

$$n_{hi} = n_i * \frac{N_{hi}^\alpha}{\sum N_{hi}^\alpha}$$

- N_{hi}: Total de viviendas en el estrato h del departamento i
- n_i: Tamaño de la muestra en el departamento i (Número de viviendas de la muestra)
- n_{hi}: Tamaño de muestra en el estrato h del departamento i

$$h = 1, 2, \dots, 18$$

Para las desagregaciones de los resultados trimestrales y anuales que se desean obtener, la distribución de la muestra por ciudad capital es la siguiente:

Cuadro 2
Distribución de las UPM y viviendas, por ciudad

CIUDAD CAPITAL	UPMS TRIMESTRALES	VIVIENDAS TRIMESTRALES
CHUQUISACA	72	648
LA PAZ	240	2160
COCHABAMBA	180	1620
ORURO	72	648
POTOSI	60	540
TARIJA	60	540
SANTA CRUZ	192	1728
BENI	48	432
PANDO	24	216
TOTAL	948	8532

9. Selección de la muestra

Selección de las unidades primarias de muestreo (UPM)

El marco muestral que se tiene proporciona información cartográfica a nivel de zonas, sectores y segmentos censales de cada una de las UPMs, las mismas que se seleccionan de manera independiente de cada uno de los estratos definidos anteriormente, utilizando la selección proporcional al tamaño (PPT) más un algoritmo de selección sistemática para ejecutar la estratificación implícita.

Selección de la Unidades Secundarias de Muestreo (USM)

Listado de las USM

Siguiendo el proceso de selección de la muestra, se debe construir el marco secundario de muestreo mediante una actualización y listado de las viviendas dentro de cada UPM seleccionada. Este trabajo lo realiza el personal de campo, para pasar a la selección correspondiente.

Selección de las USM

Ya contando con el marco secundario de muestreo, se tienen los insumos para hacer la selección de las viviendas, para lo que se construye una tabla (ver cuadro 3) aplicando un muestreo sistemático con arranque aleatorio e intervalo variable.

La manera en que se utiliza esta tabla es sumando el total de las viviendas ocupadas, objeto de estudio de la UPM y buscar esa cifra en la primera columna de la tabla. Posteriormente, las 9 viviendas seleccionadas son aquellas que tienen la numeración correspondiente a los números de la cada fila.

Cuadro 3
Tabla de selección de las Viviendas (TSV – 01)

Nro. Total de Viviendas	VIVIENDAS SELECCIONADAS								
	Viv. 1	Viv. 2	Viv. 3	Viv. 4	Viv. 5	Viv. 6	Viv. 7	Viv. 8	Viv. 9
88	1	12	23	33	44	55	66	76	87
89	7	18	29	30	41	52	62	73	84
90	1	12	23	34	45	56	67	78	89
91	4	15	26	37	48	60	71	72	83
92	5	16	27	39	50	61	62	73	85
93	8	19	31	32	43	54	66	77	88
94	8	19	31	32	43	55	66	78	89
95	3	15	26	38	49	61	72	84	85
96	2	14	25	37	49	60	72	84	95
97	10	11	23	35	46	58	70	82	93
98	6	18	30	42	54	55	66	78	90
99	10	22	23	35	47	59	71	83	95
100	4	16	28	40	52	65	77	89	90
101	2	14	26	39	51	63	75	88	100

10. Probabilidades de Selección y Factores de Expansión

Dentro de cada UPM cualquier vivienda particular tiene la misma probabilidad de ser seleccionada, es decir que se tienen muestras autoponderadas al interior de cada UPM, sin embargo, las UPM se seleccionan con probabilidad proporcional a su tamaño.

Considerando dichos aspectos se tienen definidas las siguientes probabilidades:

$$P(\text{UPM}_{hi}) = n_h \frac{M_{hi}}{M_h}$$

Donde:

$P(\text{UPM}_{hi})$: Probabilidad de seleccionar la i – ésima UPM en el h – ésimo estrato.

n_h : Número de UPM seleccionadas en el estrato h – ésimo estrato.

M_{hi} : Número de viviendas en la i – ésima UPM en el h – ésimo estrato, según el Censo 2001

M_h : Número de viviendas en el h – ésimo estrato, según el Censo 2001

$$h = 1, 2, \dots, 18$$

La probabilidad de seleccionar la vivienda dentro de una UPM es:

$$P(\text{Viv}_{hij}) = n_h \frac{M_{hi} * VS_{ih}}{M_h * VL_{ih}} = n_h \frac{M_{hi} * 9}{M_h * VL_{ih}}$$

$P(\text{Viv}_{hij})$: Probabilidad de seleccionar la j –ésima vivienda en la i – ésima UPM en el h – ésimo estrato.

n_h : Número de UPM seleccionadas en el en el h – ésimo estrato.

M_{hi} : Número de viviendas en la i – ésima UPM en el h – ésimo estrato, según el Censo 2001

M_h : Número de viviendas en el h – ésimo estrato, según Censo 2001

VS_{ih} : Número de viviendas seleccionadas en la i –ésima UPM del h ésimo estrato

VL_{ih} : Número de viviendas de la i –ésima UPM, del h – ésimo estrato, según el listado de viviendas actualizado

El inverso de la probabilidad de selección de la vivienda es el factor de expansión preliminar. El factor de expansión final incluye el ajuste por no respuesta y el ajuste a la proyección de la población¹⁰.

$$F_{ih} = \frac{1}{P(\text{Viv}_{hij})} = \frac{M_h * VS_{ih}}{n_h * M_{hi} * VL_{ih}}$$

Donde:

F_{ih} : Factor de expansión para la i –ésima vivienda del estrato h – ésimo estrato.

n_h : Número de UPM seleccionadas en el estrato h – ésimo estrato.

¹⁰ El ajuste a la proyección de la población se realiza en función a los datos que se tienen en el INE.

M_{hi} :	Número de viviendas en la i – ésima UPM en el h – ésimo estrato, según el Censo 2001
M_h :	Número de viviendas en el h – ésimo estrato, según el Censo 2001
VE_{ih} :	Número de viviendas encuestadas en la i -ésima UPM del h – ésimo estrato.
VL_{ih} :	Número de viviendas listadas en la i -ésima UPM del h – ésimo estrato.

Ajuste por No Respuesta

Una vez obtenidos los factores de expansión preliminares, se realiza un ajuste por no respuesta, el cual está dado por la siguiente expresión:

$$F_{AJUS} = F_{ih} * \frac{VS_{ih}}{VE_{ih}}$$

Donde:

- VE_{ih} : Número de viviendas encuestadas en la i -ésima UPM del h – ésimo estrato.
- VS_{ih} : Número de viviendas seleccionadas en la i -ésima UPM en el h - ésimo estrato.

Ajuste por Proyección de la Población

Los factores ajustados por no respuesta se corrigen, a fin de asegurar que en el total de UPMs se obtenga la población determinada por la proyección de población generada por el INE.

Este ajuste por proyección viene dado por la siguiente expresión:

$$F_{\text{proy}} = F_{AJUS} \frac{P_h}{\hat{P}_h}$$

Donde:

- P_h : Población en la h -ésima ciudad capital, según la proyección de la población.
- \hat{P}_h : Población en la h -ésima ciudad capital expandida por la encuesta.

11. Esquema de Rotación de la Muestra

La tasa de no respuesta de las viviendas puede reducirse mediante la rotación periódica de la muestra. Sin embargo, las unidades de rotación incrementan el costo de la encuesta debido al mantenimiento de la muestra adicional, posible capacitación adicional al personal de campo, costos extras de la recolección inicial de la información básica y dificultades en el entrenamiento de nuevas unidades para proporcionar los datos.

La rotación parcial de unidades seleccionadas a alguna tasa fija se toma como un compromiso entre la rotación total, es decir 100% de las unidades que son muy caras y dan estimaciones pobres de cambio y ninguna rotación en absoluto (panel), producirían una distribución inaceptable de tasa de respuesta.

La ETE es una encuesta por panel con una rotación parcial, la cual se refiere a un procedimiento donde las viviendas como unidad de selección (las unidades de estudio son los individuos) en la muestra son incluidas en la encuesta varias veces, pero limitado número de ocasiones.

Se tiene la información de países como México y Argentina quienes tienen un modelo de rotación igual al nuestro, que la tasa de no respuesta aumenta cuando se trabaja con periodos más largos de permanencia en el panel.

Por este motivo la construcción del esquema de rotación de la Encuesta Trimestral de Empleo se definió en 2 – (2) – 2, esto significa que 3/4 de la muestra se traslapa y 1/4 de la muestra se reemplaza con nuevas viviendas. Esto significa que cada vivienda de la muestra es incluida en la encuesta un total de cuatro veces.

Existen muchos modelos de rotación, pero se adoptó el esquema 2 – (2) – 2 porque se pretende lograr estimaciones de cambio entre trimestres y años adyacentes más precisas, con un traslape del 50% en ambos casos.

Las ventajas más importantes de una muestra de rotación son:

- a) Posibilidad de estimar cambios (entre rondas de la encuesta) con mayor precisión, utilizando la correlación entre rondas.
- b) Reducción del trabajo del campo, puesto que la actualización y listado de viviendas reduce a partir de la segunda ronda.

Las desventajas son:

- a) La estimación de agregados (por ejemplo un promedio o un total anual) es menos precisa a comparación de obtenerla con muestras independientes.
- b) Los participantes se cansan y no permanecen en la muestra todo el tiempo previsto, lo que significa que se pierde una parte de las observaciones a medida que pasa el tiempo.

Este esquema de rotación permite obtener lo siguiente:

- Correlaciones positivas entre los cuatro trimestres (12 meses) de la encuesta separadamente. Esto es posible a un costo dado para obtener estimaciones de diferencias (estimadores trimestrales, semestral y los cambios anuales) con la mejor precisión que habría sido en otro caso sin un sistema rotación de la muestra. El diseño muestral de ETE prioriza las estimaciones de promedios trimestrales y estimación de cambio entre los trimestres adyacentes.

- Comparaciones de valores de variables de las diferentes encuestas por ocasiones, para cada individuo en la muestra común y así la posibilidad de estimar los cambios brutos entre dos períodos puntuales.

Para este propósito fue aplicado el modelo de rotación 2 – (2) – 2. Las viviendas serán entrevistados dos trimestres consecutivos, será omitido por dos trimestres e incluido otra vez por un período de 2 trimestres.

El siguiente cuadro describe la distribución del panel según el modelo de rotación asumido en la Encuesta, para el periodo 2009 – 2012:

Cuadro 4
Modelo de Rotación 2- (2) -2
(2009 – 2012)

PERIODO	VISITA			
	1º	2º	3º	4º
	PANEL			
1º – 2009	P11	P12	P13	P14
2º – 2009	P15	P11	P16	P13
3º – 2009	P17	P15	P18	P16
4º – 2009	P19	P17	P12	P18
1º – 2010	P20	P19	P11	P12
2º – 2010	P21	P20	P15	P11
3º – 2010	P22	P21	P17	P15
4º – 2010	P23	P22	P19	P17
1º – 2011	P24	P23	P20	P19
2º – 2011	P25	P24	P21	P20
3º – 2011	P26	P25	P22	P21
4º – 2011	P27	P26	P23	P22
1º – 2012	P28	P27	P24	P23
2º – 2012	P29	P28	P25	P24
3º – 2012	P30	P29	P26	P25
4º – 2012	P31	P30	P27	P26
1º – 2013	P32	P31	P28	P27
2º – 2013	P33	P32	P29	P28
3º – 2013	P34	P33	P30	P29
4º – 2013	P35	P34	P31	P30

Asignación de la Muestra a Paneles

Paneles Balanceados

Para la primera ronda se selecciona una muestra de UPMs como cualquier muestra independiente. Inicialmente se seleccionan las UPMs con probabilidad proporcional al tamaño en cada estrato (ppt; el tamaño es el número de viviendas en cada UPM).

Asumiendo que se busca una tasa de rotación de 25%, la muestra debe ser dividida en cuatro paneles. Por eso, cada UPM seleccionada será asignada a uno de los paneles P11,..., P14. Esto se realiza con una lista de las UPMs seleccionadas en orden de los estratos. Empezando con un panel aleatorio (por ejemplo P3), la primera UPM de la lista es asignada a P3, la segunda UPM a P4, la tercera a P1, etc. En un estrato nuevo, se continúa con la secuencia. El número del panel debe ser incluido en la lista de las UPMs de la muestra.

Para la segunda ronda, se precisa sólo el panel P15, que debe sustituir al panel P14. Los otros paneles, con todas las UPMs y viviendas que contengan, permanecen en la muestra. El número de UPMs en P15 y la distribución por estratos es completamente determinado por P14. Basta seleccionar siempre con ppt el mismo número de UPMs, en los mismos estratos.

Se precisó establecer algunas reglas de reelección. Por ejemplo, se puede estipular que la reelección de una UPM que sale con P11, sea permisible solamente después de un año.

No se debe permitir que una UPM entre en dos o más paneles que son usados simultáneamente. Entonces, en la selección de P15, las UPMs de P12-P14 no serían permisibles.

En la Encuesta Trimestral de Empleo será conveniente seleccionar todos los paneles necesarios durante un año, es decir cuatro paneles simultáneamente en una sola operación. En este caso se propone que todas las UPMs seleccionadas sean diferentes (sin reposición). Con esta regla, la reelección de una UPM será posible solo después de un año.

Distribución de la muestra en el tiempo

- Las áreas seleccionadas se distribuyen a lo largo de 12 semanas del trimestre de manera que haya una cantidad similar de áreas por semana.
- Cada área tiene asignada una semana de referencia dentro del trimestre. Esta semana es la misma para todos los trimestres y también para los años.
- En cada trimestre se toma solamente 12 semanas de referencia, de modo que la semana 13 quede libre para poder realizar algunos ajustes o algunas encuestas pendientes de la semana 12.

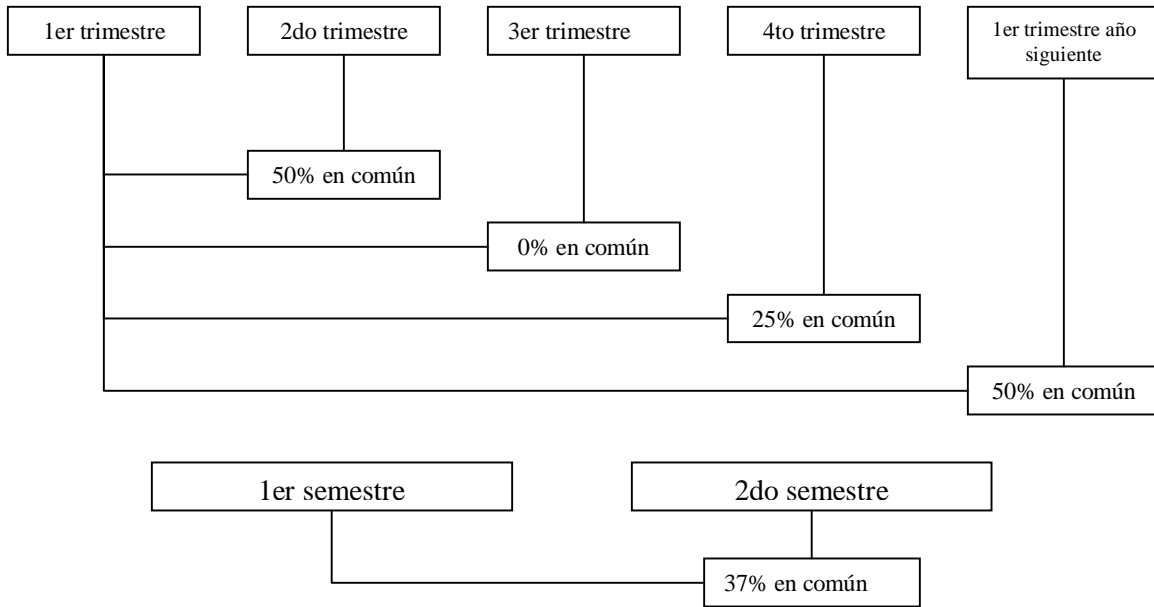
Solapamiento

El solapamiento en una encuesta continua sigue el siguiente esquema:

- Entre un trimestre y el siguiente existe un 50% de muestra en común, encuestas en los mismos meses y semanas de esos trimestres.
- Entre un trimestre y el mismo del año siguiente existe un 50% de muestra en común.
- Entre un trimestre y otro separados por un trimestre intermedio, no existe muestra en común.
- Entre un trimestre y otro separados por dos trimestres intermedios, existe un 25% muestra en común.

- Entre un semestre y el siguiente existe un 37% de muestra en común.

**Gráfico 2
Solapamiento**



12. Estimadores

Se obtienen estimadores lineales tales como la media, el total, proporción y tasas. Para los dos últimos se utilizarán estimadores de razón.

Para cada uno de los casos, se debe determinar su variabilidad y el correspondiente error de muestreo para facilitar la determinación del grado de significancia del indicador. El error muestral calculado es el coeficiente de variación, que permite obtener un indicador del error medido en porcentaje.

Por lo tanto, una vez que los datos estén expandidos se procederá al cálculo de los estimadores y sus respectivos errores muestrales.

El estimador del total de la característica X es:

$$\hat{X} = \sum_h \sum_i \sum_j F_{hij} X_{hij}$$

Donde:

F_{hij} : Factor de expansión de la j-ésima vivienda, en la i – ésima UPM del estrato h. (después de los ajustes)

X_{hij} : Valor observado de la característica de interés X en la j – ésima vivienda (o persona) en la i – ésima UPM del estrato h.

Para la estimación de tasas y proporciones se utiliza el estimador:

$$R = \frac{\hat{X}}{\hat{Y}}$$

Donde:

\hat{Y} se define en forma análoga a \hat{X} .

Estimación del error muestral

La técnica para la estimación de la varianza que se utilizó es la de replicaciones Jackknife, la que consiste en replicar k submuestras hasta obtener la distribución de la varianza.

La estimación de la desviación estándar ($\hat{\sigma}_{\hat{\theta}}$), efecto de diseño ($deff$), coeficiente de variación (CV) y el intervalo de confianza (IC) se calculan con las siguientes expresiones:

$$\hat{\sigma}_{\hat{\theta}} = \sqrt{\hat{V}(\hat{\theta})}$$

$$deff = \frac{\hat{V}(\hat{\theta})}{\hat{V}_{MAS}(\hat{\theta})}$$

$$CV = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}}$$

$$IC = [\hat{\theta} - 1.96\sqrt{\hat{V}(\hat{\theta})}, \hat{\theta} + 1.96\sqrt{\hat{V}(\hat{\theta})}]$$

Donde:

$\hat{\theta}$: Estimador del parámetro poblacional θ

$\hat{V}_{MAS}(\hat{\theta})$: Estimador de la varianza bajo el muestreo aleatorio simple.

13. Actualización del Marco de Viviendas

La actualización del marco muestral de la Encuesta Trimestral de Empleo se realiza constantemente debido a la continuidad de la encuesta y a que los datos del marco muestral pueden estar desactualizados por tratarse del Censo de Población y Vivienda 2001.

La actualización se lleva a cabo de la siguiente manera:

Actualización debido a la rotación de la muestra

- Se lleva cabo cada trimestre y se actualiza solo la cuarta parte de las UPMs que corresponden a las UPMs nuevas en la muestra.
- La actualización de las UPMs consiste en listar las viviendas, actualmente existentes, en cada una de ellas, para luego reemplazar este dato en los factores de expansión.

Actualización de la cartografía

La cartografía se actualiza cuando:

- Se fusionan manzanas
- Existe creación de nuevas manzanas
- Aparece apertura o cierre de calles
- Se presentan cambios de rasgos físicos

Bibliografía Consultada

- J.L. Sanchez Crespo. MUESTREO DE POBLACIONES FINITAS APLICADO AL DISEÑO DE ENCUESTAS. I.N.E. España 1973.
- Department of Economic and Social Affairs Statistics Division Studies in Methods 2005.
- HOUSEHOLD SAMPLE SURVEYS IN DEVELOPING AND TRANSITION COUNTRIES.
- Dr. Daniel Kasprzyk; Dr. Grez Duncan; Dr. Graham Kalton; Dr. M. P. Singh PANEL SURVEY.