



SECRETARÍA
**TÉCNICA DE PLANIFICACIÓN
DEL DESARROLLO ECONÓMICO
Y SOCIAL**



**SECRETARÍA TÉCNICA DE PLANIFICACIÓN DEL DESARROLLO ECONÓMICO Y SOCIAL
DIRECCIÓN GENERAL DE INFORMACIÓN SOCIAL**



ENCUESTA LONGITUDINAL
DE PROTECCIÓN SOCIAL

**Primera Ronda
ENCUESTA LONGITUDINAL DE PROTECCIÓN SOCIAL 2015**

**Diseño muestral y factores de expansión
Base de datos de cuestionario
(Asunción, 2016)**

19 de abril de 2016

Diseño muestral y factores de expansión
Base de datos de cuestionario
Encuesta Longitudinal de Protección Social en Paraguay - (ELPS)

Resumen

La Encuesta Longitudinal de Protección Social (ELPS) 2015 de Paraguay fue ejecutada por la Dirección General de Información Social de la Secretaría Técnica de Planificación (STP) y forma parte del Programa de Colaboración Técnica que el Banco Interamericano de Desarrollo está llevando a cabo para la implementación de este estudio en la Región.

El siguiente documento detalla el diseño muestral del estudio y la construcción de los factores de expansión para la base de datos de entrevistados. La elaboración de este documento se realizó en base a la información oficial disponible por la Dirección General de Estadísticas y Censos de Paraguay a Octubre del 2015, los informes recolectados durante el trabajo de campo y los resultados generados a partir del procesamiento de la base de datos de entrevistados del 19 de abril del 2016. Los factores de expansión fueron elaborados por Francisca Gallegos bajo la supervisión directa de la Coordinadora General de la ELPS – Paraguay Sra. Nimia Torres y su equipo.

Contenido

1. Antecedentes	5
2. Diseño de la muestra	6
2.1 Definición de la población longitudinal	6
2.2 Características del marco muestral	8
2.2.1 Cobertura del marco muestral	8
2.2.2 Estratificación del marco muestral	12
2.2.3 Conglomerados.....	14
2.3 Descripción y tamaño de la muestra.....	15
2.3.1 Tipo de muestreo	15
2.3.2 Unidades muestrales y selección	15
2.3.3 Distribución de la muestra.....	18
2.4 Errores muestrales	20
2.5 Probabilidades de selección de las unidades muestrales	22
2.6 Construcción de pesos (factores de expansión).....	23
2.6.1 Pesos transversales de la Ronda 1 ELPS - Paraguay.....	23
2.6.2 Factor de ajuste de la Ronda 1 ELPS – Paraguay información EPH.....	24
2.6.3 Factor de ajuste de la Ronda 1 ELPS – Paraguay información DGEEC.....	25
2.6.4 Factor final de expansión de entrevistado de la Ronda 1 ELPS – Paraguay	28
3. Referencias	29
4. Anexos.....	30
4.1 Actualización del número de viviendas Pre CENSO 2012	30
4.2 Análisis de la actualización de viviendas	32
4.3 Tabla de kish	33

Índice de tablas

Tabla 1: Paraguay. Población total, estimada y proyectada, por área, 2002 y 2015	6
Tabla 2: Paraguay. Población proyectada de 15 años o más según sexo, por área, 2015.....	6
Tabla 3: Paraguay. Población total, proyectada, por área y departamento, 2015	7
Tabla 4: Paraguay. Población total, proyectada, por sexo y departamento, 2015	8
Tabla 5: Paraguay. UPM, por área y departamentos, 2015 (con exclusiones)	10
Tabla 6: Paraguay. Viviendas, por área y departamentos, 2015 (con exclusiones).....	10
Tabla 7: Paraguay. Población de 15 años o más, proyectada, por área y departamento, 2015 (con exclusiones)	11
Tabla 8: Paraguay. Población de 15 años o más, proyectada, por sexo y departamentos, 2015	11
Tabla 9: Distribución de los estratos ELPS - Paraguay	13
Tabla 10: Distribución de UPM según área para los departamentos del estudio	15
Tabla 11: Distribución de UPM según departamento y área.....	16
Tabla 12: Distribución de Viviendas a visitar según departamento y zona Urbano – Rural.....	17
Tabla 13: Distribución de los estratos según UPM, viviendas y muestra esperada	18
Tabla 14: Distribución de la muestra con sobredimensión 20% según Departamento y Área.....	19
Tabla 15: Distribución de muestra esperada según Departamento y Área	19
Tabla 16: Distribución de la muestra lograda según Departamento y Área	20
Tabla 17: Distribución de la muestra esperada y error muestral según Departamento y Área (considera la fórmula del error por M.A.S.)	21
Tabla 18: Distribución de la muestra lograda y error muestral según Departamento y Área.....	22
Tabla 19: Distribución muestra con factor de selección ELPS según quinquenios	24
Tabla 20: Quinquenios de Población de 15 años y más (EPH)	25
Tabla 21: Factor de ajuste por sexo/departamento a Proyección DGEEC OCT 2015 (delta).....	26
Tabla 22: Distribución población de 15 años o más según quinquenios (Proyecciones DGEEC).....	26
Tabla 21: Distribución muestra con factor de ajuste a al total población DGEEC ELPS según quinquenios - Base de datos marzo 2016	28
Tabla 24: UPM sin encuestas logradas.....	32

1. Antecedentes

La Encuesta Longitudinal de Protección Social (ELPS) está diseñada para monitorear y evaluar los sistemas de protección social, y su interrelación con el mercado del trabajo (BID, 2015)¹ entregando información para medir el impacto de los programas de previsión y seguridad social. En tanto una encuesta longitudinal, “permite construir una base de datos que recoge a lo largo del tiempo, y en un mismo instrumento, la historia laboral y previsional de los encuestados profundizando en áreas como educación, salud, seguridad social, capacitación laboral, patrimonio y activos, historia familiar e información del hogar” (Prieto; 2015). Y en cuanto esfuerzo regional, permite comparar los distintos sistemas latinoamericanos y sus cambios en el tiempo, permitiendo al conjunto de países aprender de las lecciones de las políticas implementadas en la Región.

Una vez culminado el trabajo de campo de las rondas de ELPS se presenta el desafío de documentar y entregar la información necesaria y suficiente para que los usuarios de la base de datos conozcan las condiciones de producción y uso de la misma. Este documento entrega una especificación del diseño muestral y la construcción de los factores de expansión que se diseñaron para este estudio.

El documento se ha organizado en función de cada una de las etapas del diseño de muestral: (1) Diseño muestral, (2) Tamaño de la muestra y distribución, (3) Errores muestrales, (4) Probabilidades de selección y (5) Construcción del factor de expansión.

¹ Presentación del Proyecto Encuesta Longitudinal de Protección Social disponible en: {HYPERLINK: <http://www.observatorioregional.net/download/bid-presentacion-proyecto-regional-elps/>} Última visita: 30 de mayo 2015.

2. Diseño de la muestra

2.1 Definición de la población longitudinal

La población objetivo de este estudio corresponde a todas las personas de 15 y más años que componen la población nacional. La cobertura de este estudio se extiende a 16 Departamentos, incluyendo sus áreas urbanas y rurales, de acuerdo a la actual División Política Administrativa de la República del Paraguay. Los departamentos Boquerón y Alto Paraguay no fueron considerados dentro del estudio, pues concentran menos del 1.5% de la población.

De acuerdo a las proyecciones poblacionales basadas en el CENSO de Población realizado el año 2002², la población total de Paraguay para el año 2015 está compuesta por 6.755.756 personas, de las cuales un 69% corresponden a personas de 15 años o más (n= 4.682.717; hombres=2.352.214; mujeres= 2.330.503). Asimismo, un 60% se concentra en áreas urbanas y un 40% en área rural, la composición por Departamento y Área Urbana y Rural se detalla en las tablas n°1, n°2, n°3 y n°4.

Tabla 1: Paraguay. Población total, estimada y proyectada, por área, 2002 y 2015

	Año 2002		Año 2015	
	N	%	N	%
Total País	5.484.610	100%	6.755.756	100%
Urbano	2.997.281	55%	4.085.396	60%
Rural	2.487.329	45%	2.670.360	40%

Fuente: Elaboración propia en base “Proyección de la Población por Sexo y Edad, según Departamento, 2000-2025. Revisión 2015” DGEEC (Octubre, 2015)

Tabla 2: Paraguay. Población proyectada de 15 años o más según sexo, por área, 2015

	Total		Hombres		Mujeres	
	N	%	N	%	N	%
Total País	6.755.756	100%	3.408.566	100%	3.347.190	100%
> 15 años Total País	4.682.717	69%	2.352.214	69%	2.330.503	70%
> 15 años Urbano	2.917.930	43%	1.409.889	41%	1.508.041	45%
> 15 años Rural	1.764.787	26%	942.324	28%	822.463	25%

Fuente: Elaboración propia en base “Proyección de la Población por Sexo y Edad, según Departamento, 2000-2025. Revisión 2015” DGEEC (Octubre, 2015)

² Las proyecciones son desarrolladas por la Dirección General de Estadística, Encuestas y Censos (DGEEC) a partir de la información recopilada en el CENSO de Vivienda y Población del año 2002 del Paraguay. En Octubre del año 2015 la DGEEC publicó el documento “Proyección de la Población según departamento 2000 – 2025” en el cual entrega las proyecciones para cada departamento según sexo y edad, pero no la distribución según área urbana y rural. Según señala la DGEEC, “la proyección de la población por departamento fue realizada utilizando el método de los componentes por cohorte, y se emplearon dos de los diversos paquetes demográficos disponibles para realizar estimaciones y proyecciones: PAS1, que se aplicó principalmente para proyectar los componentes demográficos y MORTPAK, para operaciones de estimación y proyección por componentes aplicativo PROJCT” (pág.11). En el documento antes citado, la DGEEC dedica un capítulo especial para referirse a la población base e inicial, en el cual señala que para la elaboración de las proyecciones se utilizó el CENSO 2002 pues el Censo Nacional de Población y Viviendas 2012 tuvo una cobertura de personas del 74,4% resultando un excesivo nivel de omisión que le quitan precisión a las estimaciones que se esperan realizar. En los anexos de este documento se presenta en detalle el capítulo indicado.

Tabla 3: Paraguay. Población total, proyectada, por área y departamento, 2015³

Departamentos	Total		Urbana		Rural	
	N	%	N	%	N	%
Total	6.755.756	100%	3.808.661	56%	2.947.095	44%
Asunción	526.408	100%	526.408	100%	0	0%
Concepción	236.959	100%	78.986	33%	157.973	67%
San Pedro	409.381	100%	75.254	18%	334.127	82%
Cordillera	291.311	100%	95.059	33%	196.252	66%
Guairá	216.335	100%	77.493	36%	138.842	64%
Caaguazú	534.522	100%	194.372	36%	340.151	64%
Caazapá	179.576	100%	30.961	17%	148.615	83%
Itapúa	576.577	100%	214.540	37%	362.036	63%
Misiones	120.576	100%	57.417	48%	63.159	52%
Paraguarí	252.255	100%	69.070	27%	183.185	73%
Alto Paraná	773.303	100%	504.163	65%	269.140	35%
Central	1.985.384	100%	1.622.714	82%	362.670	18%
Ñeembucú	87.750	100%	45.128	51%	42.621	49%
Amambay	159.263	100%	98.008	62%	61.255	38%
Canindeyú	212.637	100%	69.771	33%	142.865	67%
Pdte. Hayes	116.536	100%	26.221	23%	90.316	75%
Boquerón	60.402	100%	18.121	30%	42.281	70%
Alto Paraguay	16.582	100%	4.975	30%	11.608	70%

Fuente: Elaboración propia en base “Proyección de la Población por Sexo y Edad, según Departamento, 2000-2025. Revisión 2015” DGEEC (Octubre, 2015) y la distribución porcentual por área según Pre CENSO 2012.

³ Dado que la DGEEC no entrega en las proyecciones para el 2015 información oficial respecto de la distribución de la población según área para cada uno de los departamentos, la construcción de esta tabla se utilizó a partir de la distribución informada por el Pre - CENSO 2012, por ello la diferencia del porcentaje de la distribución urbano rural entre la tabla n°1 y la tabla n°3.

Tabla 4: Paraguay. Población total, proyectada, por sexo y departamento, 2015

	Total País		Hombres		Mujeres	
	N	%	N	%	N	%
Total País	6.755.756	100%	3.408.566	100%	3.347.190	100%
Asunción	526.408	7,8%	247.819	7,3%	278.588	8,3%
Concepción	236.959	3,5%	121.744	3,6%	115.215	3,4%
San Pedro	409.381	6,1%	214.428	6,3%	194.953	5,8%
Cordillera	291.311	4,3%	151.228	4,4%	140.083	4,2%
Guairá	216.335	3,2%	111.506	3,3%	104.829	3,1%
Caaguazú	534.522	7,9%	275.998	8,1%	258.524	7,7%
Caazapá	179.576	2,7%	92.553	2,7%	87.023	2,6%
Itapúa	576.577	8,5%	292.614	8,6%	283.963	8,5%
Misiones	120.576	1,8%	60.882	1,8%	59.694	1,8%
Paraguarí	252.255	3,7%	130.509	3,8%	121.746	3,6%
Alto Paraná	773.303	11,4%	392.924	11,5%	380.379	11,4%
Central	1.985.384	29,4%	979.918	28,7%	1.005.466	30,0%
Ñeembucú	87.750	1,3%	44.123	1,3%	43.627	1,3%
Amambay	159.263	2,4%	79.596	2,3%	79.666	2,4%
Canindeyú	212.637	3,1%	112.021	3,3%	100.615	3,0%
Pdte. Hayes	116.536	1,7%	60.478	1,8%	56.059	1,7%
Boquerón	60.402	0,9%	31.310	0,9%	29.092	0,9%
Alto Paraguay	16.582	0,2%	8.915	0,3%	7.667	0,2%

Fuente: Elaboración propia en base “Proyección de la Población por Sexo y Edad, según Departamento, 2000-2025. Revisión 2015” DGEEC (Octubre, 2015)

2.2 Características del marco muestral

El marco a ser utilizado para la selección de la muestra fue el Pre CENSO 2012, el cual está compuesto de Unidades Primarias de Muestreo (UPM), conformadas a partir de las áreas geográficas generadas para el trabajo de los censistas del CENSO Nacional de Población y Viviendas 2012.

Este marco muestral se encuentra actualizado hasta el año 2012 y para cada UPM tiene información referida al Departamento, Distrito, Barrio, Cantidad de viviendas, entre otras variables.

El Pre CENSO 2012 se realizó entre marzo del año 2011 a marzo del 2012, en su levantamiento se realizó el conteo de la totalidad de viviendas por UPM incluyendo, además de las residenciales, aquellas desocupadas, con hogares colectivos, de uso comercial, entre otros. En el levantamiento del Pre CENSO 2012 se recopiló la información de la cantidad de integrantes de hogar y la actividad del jefe de hogar. Sin embargo, existió una alta no respuesta, por lo que el Pre CENSO 2012 se utilizó sólo como un marco de UPM cuya información del número de las viviendas fue actualizado durante el desarrollo de la encuesta a la que se refiere este documento.

2.2.1 Cobertura del marco muestral

Los errores asociados con el marco muestral se denominan de cobertura y se establecen por omisión o duplicación de elementos que deberían formar parte del mismo. Para evitar ello, las unidades primarias de muestreo deben tener cuatro características fundamentales para el diseño muestral: (1) cubrir el territorio completo del país; (2) tener sus fronteras bien delineadas; (3) que existan cifras de la población disponibles para las unidades; y (4) estar cartografiadas (Naciones Unidas, 2008).

En el caso de que alguna de estas condiciones no se cumpla se pueden generar sesgos en las estimaciones de la encuesta por dos razones:

- Si las unidades que deberían estar incluidas en el marco no lo están, la probabilidad de selección de las unidades será cero, produciendo una subestimación. Por ejemplo, si la lista de unidades de vivienda no se ha actualizado desde que se confeccionó, - en el caso de Paraguay sería desde el año 2012 - , las unidades añadidas durante los últimos tres años a los sectores urbanos o rurales no formarían parte del marco muestral. De igual manera, las viviendas abandonadas continuarían formando parte del marco muestral. En esos casos habría tanto omisión de unidades que pertenecen a la población como inclusión de unidades que no pertenecen.
- Si las unidades se duplican, existirá una sobre-cobertura, y así, una sobreestimación. Por ejemplo, problemas de demarcación de los sectores urbanos y rurales durante el proceso cartográfico del Pre CENSO que dupliquen las viviendas - también se podrían omitir-.

Para abordar ambas situaciones la actualización cartográfica del número de viviendas se considera una estrategia adecuada para minimizar el error de cobertura, la cual para esta encuesta se especifica más adelante. Es importante señalar que la magnitud y el efecto del error de cobertura son difíciles de estimar, dado que requieren información externa al marco utilizado, y que no son planeados por los diseñadores de la encuesta. En ese sentido, hay que distinguir entre la falta de cobertura – no inclusión de algunas unidades de una población en el marco muestral por errores no intencionados -, de las exclusiones deliberadas y explícitas de ciertos sectores de la población en función de los objetivos de la encuesta y las dificultades prácticas de realizar una entrevista presencial en la vivienda del entrevistado.

Considerando lo anterior, se pueden identificar dos tipos de exclusiones del segundo tipo en el proceso de selección de la muestra de la ELPS-Paraguay:

- Las viviendas clasificadas durante la actualización cartográfica como no residenciales⁴ (por ejemplo: hotel, pensión, residencial u hostel, cuartel militar, de policía o bomberos, centro de rehabilitación social o cárcel, centro de acogida y protección para niños y niñas, mujeres e indigentes, hospital, clínica, convento o institución religiosa, asilo de ancianos u orfanato, otra vivienda colectiva y sin vivienda).
- Los Departamentos de Boquerón y Alto Paraguay por concentrar menos del 1,5% de la población nacional.

De esta manera el marco muestral queda conformado por 31.329 unidades primarias de muestreo correspondiente a los departamentos en sus zonas urbanas y rurales (Tabla n°5). Para el año de aplicación del Pre CENSO 2012 en las unidades incluidas en el estudio se contabilizaban un total de 1.583.577 viviendas ocupadas en Paraguay (Tabla n°6), en las que residían 6.678.772 personas, de las cuales 4.547.742 corresponden a personas de 15 años o más (Tabla n°7 y n°8).

⁴ Cabe señalar que se consideró como criterio de exclusión a las viviendas colectivas que albergan a un conjunto de personas generalmente no ligadas por vínculos de parentesco, que comparten la vivienda por razones de trabajo, atención médica, militares, religiosas, reclusión, etc.

Tabla 5: Paraguay. UPM, por área y departamentos, 2015 (con exclusiones)

DEPARTAMENTO	Total		Urbana		Rural	
	n	%	n	%	n	%
Total país	31.329	100%	16.621	100%	14.708	100%
Asunción	2.288	7%	2.288	14%		0%
Concepción	1.099	4%	369	2%	730	5%
San pedro	2.366	8%	429	3%	1.937	13%
Cordillera	1.648	5%	543	3%	1.105	8%
Guaira	1.160	4%	412	2%	748	5%
Caaguazú	2.485	8%	900	5%	1.585	11%
Caazapá	1.017	3%	179	1%	838	6%
Itapuá	2.986	10%	1.110	7%	1.876	13%
Misiones	734	2%	346	2%	388	3%
Paraguarí	1.454	5%	396	2%	1.058	7%
Alto Paraná	3.557	11%	2.320	14%	1.237	8%
Central	7.434	24%	6.069	37%	1.365	9%
Ñeembucú	612	2%	310	2%	302	2%
Amambay	682	2%	422	3%	260	2%
Canindeyú	1.113	4%	363	2%	750	5%
Pdte. Hayes	694	2%	165	1%	529	4%

Fuente: Elaboración propia en base al Pre CENSO 2012 – Paraguay

Tabla 6: Paraguay. Viviendas, por área y departamentos, 2015 (con exclusiones)

Departamento	Total		Urbana		Rural	
	N	%	N	%	N	%
Departamento	1.583.577	100%	980.456	100%	603.121	100%
Asunción	139.597	9%	139.597	14%	0	0%
Concepción	50.394	3%	21.680	2%	28.714	5%
San pedro	98.301	6%	22.752	2%	75.549	13%
Cordillera	79.995	5%	30.828	3%	49.167	8%
Guaira	55.900	4%	24.000	2%	31.900	5%
Caaguazú	123.087	8%	53.183	5%	69.904	12%
Caazapá	44.013	3%	9.838	1%	34.175	6%
Itapuá	141.529	9%	64.344	7%	77.185	13%
Misiones	35.256	2%	20.036	2%	15.220	3%
Paraguarí	66.570	4%	21.215	2%	45.355	8%
Alto Paraná	184.545	12%	134.631	14%	49.914	8%
Central	428.280	27%	365.802	37%	62.478	10%
Ñeembucú	27.508	2%	16.828	2%	10.680	2%
Amambay	35.929	2%	25.990	3%	9.939	2%
Canindeyú	48.885	3%	20.495	2%	28.390	5%
Pdte. Hayes	23.788	2%	9.237	1%	14.551	2%

Fuente: Elaboración propia en base al Pre CENSO 2012 – Paraguay

Tabla 7: Paraguay. Población de 15 años o más, proyectada, por área y departamento, 2015 (con exclusiones)

Departamentos	Total		Urbana		Rural	
	N	%	N	%	N	%
Total	4.547.742	100%	2.728.645	60%	1.819.097	40%
Asunción	375.977	100%	375.977	100%	0	0%
Concepción	150.822	100%	49.771	33%	101.050	67%
San Pedro	264.979	100%	47.696	18%	217.283	82%
Cordillera	201.863	100%	66.615	33%	135.248	67%
Guairá	149.483	100%	53.814	36%	95.669	64%
Caaguazú	353.581	100%	127.289	36%	226.292	64%
Caazapá	115.133	100%	19.573	17%	95.560	83%
Itapúa	381.471	100%	141.144	37%	240.327	63%
Misiones	83.351	100%	40.009	48%	43.343	52%
Paraguarí	175.521	100%	47.391	27%	128.131	73%
Alto Paraná	517.566	100%	336.418	65%	181.148	35%
Central	1.392.598	100%	1.141.931	82%	250.668	18%
Ñeembucú	63.394	100%	32.331	51%	31.063	49%
Amambay	105.628	100%	65.490	62%	40.139	38%
Canindeyú	139.587	100%	46.064	33%	93.523	67%
Pdte. Hayes	76.787	100%	17.661	23%	59.894	78%

Fuente: Elaboración propia en base “Proyección de la Población por Sexo y Edad, según Departamento, 2000-2025. Revisión 2015” DGEEC (Octubre, 2015) y la distribución porcentual por área según Pre CENSO 2012.

Tabla 8: Paraguay. Población de 15 años o más, proyectada, por sexo y departamentos, 2015

	Total		Hombres		Mujeres	
	N	%	N	%	N	%
Total general	4.547.742	100%	2.290.520	100%	2.256.613	100%
Asunción	375.977	8%	173.388	8%	202.588	9%
Concepción	150.822	3%	78.103	3%	72.718	3%
San pedro	264.979	6%	141.012	6%	123.967	5%
Cordillera	201.863	4%	106.152	5%	95.712	4%
Guairá	149.483	3%	77.798	3%	71.685	3%
Caaguazú	353.581	8%	184.207	8%	169.375	8%
Caazapá	115.133	3%	59.857	3%	55.276	2%
Itapúa	381.471	8%	193.676	8%	187.795	8%
Misiones	83.351	2%	42.141	2%	41.210	2%
Paraguarí	175.521	4%	91.849	4%	83.673	4%
Alto Paraná	517.566	11%	262.570	11%	254.995	11%
Central	1.392.598	31%	681.017	30%	711.581	32%
Ñeembucú	63.394	1%	31.797	1%	31.597	1%
Amambay	105.628	2%	52.481	2%	53.147	2%
Canindeyú	139.587	3%	74.756	3%	64.831	3%
Pdte. Hayes	76.787	2%	39.715	2%	36.463	2%

Fuente: Elaboración propia en base “Proyección de la Población por Sexo y Edad, según Departamento, 2000-2025. Revisión 2015” DGEEC (Octubre, 2015).

2.2.2 Estratificación del marco muestral

En el diseño de una muestra, la estratificación corresponde al proceso de agrupar a los elementos de una población en grupos homogéneos previo a la selección de la muestra. Así, los estratos son subconjuntos de la población esencialmente independientes y mutuamente excluyentes, es decir, cada elemento de una población debe pertenecer a un estrato y sólo a uno. Su propósito es obtener un grado mayor de representatividad disminuyendo el error muestral probable. Para lograr este objetivo, cada estrato debería ser lo más diferente posible de los demás (Naciones Unidas, 2008).

Dado que las poblaciones rurales y urbanas presentan diferencias en diversas variables como empleo, ingresos, estructura del hogar, entre otras, las zonas urbanas y rurales son los estratos más utilizados en la mayoría de las encuestas sociales de cobertura nacional. La ELPS-Paraguay no es la excepción. El diseño muestral de la ELPS-Paraguay se conformará por $h=31$ estratos, a partir de la interposición de la división política-administrativa (a nivel de departamentos) y la división censal (urbano-rural). De éstos, 16 estratos se encuentran en áreas urbanas y 15 en áreas rurales. La diferencia refiere a que en el Departamento de Asunción sólo existe población urbana (Tabla 9).

Tabla 9: Distribución de los estratos ELPS - Paraguay

Nº Estrato	Departamento	Área
1	Asunción	Urbano
2	Concepción	Urbano
3	San pedro	Urbano
4	Cordillera	Urbano
5	Guairá	Urbano
6	Caaguazú	Urbano
7	Caazapá	Urbano
8	Itapuá	Urbano
9	Misiones	Urbano
10	Paraguarí	Urbano
11	Alto Paraná	Urbano
12	Central	Urbano
13	Ñeembucú	Urbano
14	Amambay	Urbano
15	Canindeyú	Urbano
16	Presidente Hayes	Urbano
17	Concepción	Rural
18	San pedro	Rural
19	Cordillera	Rural
20	Guairá	Rural
21	Caaguazú	Rural
22	Caazapá	Rural
23	Itapuá	Rural
24	Misiones	Rural
25	Paraguarí	Rural
26	Alto Paraná	Rural

27	Central	Rural
28	Ñeembucú	Rural
29	Amambay	Rural
30	Canindeyú	Rural
31	Presidente Hayes	Rural

Fuente: Elaboración propia en base al Pre CENSO 2012 – Paraguay y documento de diseño muestral STP (2014)

2.2.3 Conglomerados

En las encuestas de hogares de corte transversal el diseño de la muestra requiere finalmente utilizar alguna forma de muestreo por conglomerado para asegurar la contención de los costos de la encuesta. El muestreo por conglomerado refiere a la etapa final del muestreo cuando se encuesta a una submuestra de viviendas dentro de cada agrupación de viviendas seleccionada o conglomerado. Por ejemplo, la Encuesta Permanente de Hogares - EPH de Paraguay – selecciona 12 viviendas ocupadas por unidad censal⁵.

Para las encuestas longitudinales, debido a la movilidad residencial y la atrición de la muestra, la relación entre los conglomerados y los costos del encuestaje va cambiando en las olas posteriores de la encuesta. En efecto, a medida que la muestra se dispersa sobre el territorio de estudio, cualquier ahorro en los costos debido a la selección de una muestra por conglomerados en la ola inicial se irá perdiendo en las olas posteriores de la encuesta, como una consecuencia del “desconglomeramiento” de la muestra (Lynn et al, 2005). Considerando que una encuesta panel está concebida para durar muchos años, no es claro que los menores costos de encuestaje en las primeras olas de la encuesta sean suficientes para balancear los mayores costos en términos de pérdida de precisión debido a los conglomerados, que persisten a lo largo de toda la vida del estudio.

No obstante lo anterior, en miras de aumentar la eficiencia del trabajo de campo y en consideración a la extensión del estudio, la ELPS – Paraguay siguió el modelo de la EPH y se consideró seleccionar 12 viviendas, tanto en conglomerados urbanos como rurales. Estas fueron empadronadas a partir de la actualización cartográfica, como se señala más adelante.

⁵ Cabe señalar que uno de los problemas asociados a esta estrategia guarda relación con que la conglomeración de la muestra reduce la fiabilidad, debido a la probabilidad de que las personas que viven en un mismo conglomerado sean homogéneas o presenten características más o menos similares.

2.3 Descripción y tamaño de la muestra

2.3.1 Tipo de muestreo

En consideración a los procesos de tratamiento del marco muestral antes mencionados, el diseño de la muestra de la ELPS – Paraguay es estratificado geográficamente por zona urbana – rural, de conglomerados, trietápica, con afijación proporcional al tamaño, para todos los departamentos. Parte de un diseño basado en viviendas particulares ocupadas para finalizar con la selección de un individuo de 15 años o más por vivienda, seleccionado mediante la aplicación de la tabla de Kish.

2.3.2 Unidades muestrales y selección

Primera etapa de selección (UPE): dentro de cada estrato, se seleccionará un número determinado de UPMs en forma sistemática, con arranque aleatorio y con probabilidad proporcional al tamaño. Las que corresponden a los segmentos seleccionados dentro de los departamentos. De un total de **31.329** UPMs y los 16 departamentos a ser incluidos en la encuesta, se seleccionaron un total de 1.800 UPMs en Paraguay.

Tabla 10: Distribución de UPM según área para los departamentos del estudio

	Total de UPM	UPM a seleccionar
Total	31.329	1.800
Área urbana	16.621	955
Área rural	14.708	845

Fuente: Elaboración propia en base a datos del Pre CENSO 2012 de población y Vivienda elaborado por la DGEEC

En las áreas urbanas y rurales del país, el número promedio de viviendas por UPM es diferente.

- **Áreas urbanas:** Constituyen todas las cabeceras de distritos oficiales, definidas conforme a las leyes administrativas, presentan forma amanzanada, sin tener ninguna otra consideración especial. Cada UPM está conformado por un conjunto de 60 viviendas adyacentes en promedio.
- **Áreas rurales:** Están constituidas por todo territorio situado fuera de las cabeceras distritales, en las cuales las viviendas se encuentran dispersas y por lo general están rodeadas de cultivos y campos. Cada UPM está conformado por un conjunto de 40 viviendas adyacentes en promedio.

Las UPM que salieron sorteadas y tienen un número menor de viviendas que la mínima cantidad de encuestas esperadas fueron incluidas tanto en zonas urbanas como rurales. En el área urbana la UPM con menor cantidad de viviendas tiene 15 y en el área rural cuatro UPM tienen menos de 12. Las UPM seleccionadas por departamento y área se observan en la siguiente tabla.

Tabla 11: Distribución de UPM según departamento y área

Departamentos	Total	Urbana	Rural
Total	1.800	955	845
Asunción	131	131	
Concepción	63	21	42
San Pedro	136	25	111
Cordillera	95	31	64
Guairá	67	24	43
Caaguazú	143	52	91
Caazapá	58	10	48
Itapúa	172	64	108
Misiones	42	20	22
Paraguarí	84	23	61
Alto Paraná	204	133	71
Central	427	349	78
Ñeembucú	35	18	17
Amambay	39	24	15
Canindeyú	64	21	43
Pdte. Hayes	40	9	31

Fuente: Elaboración propia en base a la distribución porcentual datos de Proyección y Densidad poblacional 2012. DGEEC – Gobierno de Paraguay

Segunda etapa de selección (USE): dentro de cada UPM escogida en la etapa anterior, se seleccionaron las viviendas en conglomerados, siguiendo la metodología utilizada en las Encuestas Permanentes de Hogares de la DGEEC. El número de viviendas que fueron seleccionadas en cada UPM es fijo e igual a 12 para todos las UPM, fueron distribuidas en dos segmentos compactos de 6 viviendas cada uno. Al llegar a la UPM lo primero que se realizó fue la actualización cartográfica.

Esta actividad consistió en realizar un recorrido de la UPM⁶ y se actualizó el número de viviendas indicado por el Pre CENSO 2012. Luego, se identificó la vivienda seleccionada como número 1 a partir del número aleatorio indicado en la planilla de actualización de vivienda. Dicho documento trae pre impreso la variable “pre CENSO 2012” informando la cantidad de viviendas originales y el supervisor anotó en la variable actualización la cantidad de viviendas encontradas. Cabe señalar que en la selección de las viviendas a visitar se consideró todas las viviendas ocupadas para uso residencial, sea en el área urbana o rural. Como resultado de esto se obtuvo una base de datos con la cantidad de viviendas por UPM del Pre CENSO 2012 y la cantidad de viviendas actualizada en el recorrido ELPS.

La selección de las viviendas fue sin reemplazo, pero con sobredimensión de la muestra de manera de enfrentar los problemas de no respuesta. La DGEEC reporta alrededor de un 10% de no respuesta en la Encuesta Permanente de Hogares por lo que se previó para la ELPS - Paraguay una sobredimensión de un 20%, por la mayor dificultad que implica realizar un seguimiento a una persona individualizada dentro de la vivienda. Así la cantidad de viviendas a visitar serán de 21.600 distribuidas en las 1.800 UPMs esperándose entrevistar a, al menos, 10 viviendas en cada una de ellas.

⁶ Se sugiere revisar el anexo de este documento que detalla el proceso realizado para el recorrido.

Tabla 12: Distribución de Viviendas a visitar según departamento y zona Urbano – Rural

Departamentos	Total		Urbana		Rural	
	n	%	n	%	N	%
Total	21600	100%	11460	100%	10140	100%
Asunción	1572	7%	1572	14%		0%
Concepción	756	4%	252	2%	504	5%
San Pedro	1632	8%	300	3%	1332	13%
Cordillera	1140	5%	372	3%	768	7%
Guairá	804	4%	288	3%	516	5%
Caaguazú	1716	8%	624	5%	1092	11%
Caazapá	696	3%	120	1%	576	6%
Itapúa	2064	10%	768	7%	1296	13%
Misiones	504	2%	240	2%	264	3%
Paraguarí	1008	5%	276	2%	732	7%
Alto Paraná	2448	11%	1596	14%	852	8%
Central	5124	24%	4188	37%	936	9%
Ñeembucú	420	2%	216	2%	204	2%
Amambay	468	2%	288	3%	180	2%
Canindeyú	768	4%	252	2%	516	5%
Pdte. Hayes	480	2%	108	1%	372	4%

Tercera etapa de selección (USE): dentro de cada vivienda escogida, se seleccionó aleatoriamente mediante la tabla de Kish a una persona de 15 años o más para ser entrevistada⁷. Para evitar el posible sesgo que este procedimiento pudiera generar se asignó esta tarea al encuestador con el apoyo del supervisor y se revisó en el proceso de digitación y validación, siendo eliminadas todas aquellas encuestas que no cumplan con el protocolo de selección aleatorio⁸.

⁷ Es importante considerar que la segunda etapa de selección de la muestra es de viviendas y no hogares, puede existir más de un hogar en una vivienda -, en este caso se deben listar todas las personas que habitan en la vivienda, independientemente de cuál sea el hogar al que pertenecen. Se consideran hogares separados dentro de una vivienda cuando se cumplen en forma conjunta tres condiciones: acceso común, privacidad al interior de la vivienda y cocinar separadamente.

⁸ La selección de los entrevistados fue realizada por los encuestadores al momento de establecer contacto con la vivienda e incluyó recopilar información sociodemográfica de los hogares seleccionados. Las instrucciones de aplicación de la hoja de composición para la selección del entrevistado usando la tabla de Kish se encuentra entre los anexos de este documento.

Tabla 13: Distribución de los estratos según UPM, viviendas y muestra esperada

N° Estrato	Departamento	Área	UPM	Viviendas	Muestra
1	Asunción	Urbano	131	1572	1310
2	Concepción	Urbano	21	252	210
3	San Pedro	Urbano	25	300	250
4	Cordillera	Urbano	31	372	310
5	Guairá	Urbano	24	288	240
6	Caaguazú	Urbano	52	624	520
7	Caazapá	Urbano	10	120	100
8	Itapuá	Urbano	64	768	640
9	Misiones	Urbano	20	240	200
10	Paraguarí	Urbano	23	276	230
11	Alto Paraná	Urbano	133	1596	1330
12	Central	Urbano	349	4188	3490
13	Ñeembucú	Urbano	18	216	180
14	Amambay	Urbano	24	288	240
15	Canindeyú	Urbano	21	252	210
16	Presidente Hayes	Urbano	9	108	90
17	Concepción	Rural	42	504	420
18	San Pedro	Rural	111	1332	1110
19	Cordillera	Rural	64	768	640
20	Guairá	Rural	43	516	430
21	Caaguazú	Rural	91	1092	910
22	Caazapá	Rural	48	576	480
23	Itapuá	Rural	108	1296	1080
24	Misiones	Rural	22	264	220
25	Paraguarí	Rural	61	732	610
26	Alto Paraná	Rural	71	852	710
27	Central	Rural	78	936	780
28	Ñeembucú	Rural	17	204	170
29	Amambay	Rural	15	180	150
30	Canindeyú	Rural	43	516	430
31	Presidente Hayes	Rural	31	372	310

Fuente: Elaboración propia en base al Pre CENSO 2012 – Paraguay y documento de diseño muestral STP (2014)

2.3.3 Distribución de la muestra

La Encuesta Longitudinal de Protección Social – Paraguay tiene como objeto entrevistar un panel de individuos a través del tiempo. Sin embargo, a través de los entrevistados también se puede analizar tanto su hogar como su vivienda. Esto es posible porque el individuo proporciona información de sí mismo y de todos los integrantes de su hogar. Se esperaba que esta muestra alcance un tamaño efectivo de 18.000 personas mayores de 15 años en la primera ronda- 2015 -. En las siguientes tablas se presenta la distribución de la muestra con la sobredimensión (Tabla n° 14), la muestra esperada (Tabla n°15) y la muestra lograda (Tabla n°16).

Tabla 14: Distribución de la muestra con sobredimensión 20% según Departamento y Área

Departamentos	Total		Urbana		Rural	
	n	%	N	%	N	%
Total	21600	100%	11460	100%	10140	100%
Asunción	1572	7%	1572	14%		0%
Concepción	756	4%	252	2%	504	5%
San Pedro	1632	8%	300	3%	1332	13%
Cordillera	1140	5%	372	3%	768	7%
Guairá	804	4%	288	3%	516	5%
Caaguazú	1716	8%	624	5%	1092	11%
Caazapá	696	3%	120	1%	576	6%
Itapúa	2064	10%	768	7%	1296	13%
Misiones	504	2%	240	2%	264	3%
Paraguarí	1008	5%	276	2%	732	7%
Alto Paraná	2448	11%	1596	14%	852	8%
Central	5124	24%	4188	37%	936	9%
Ñeembucú	420	2%	216	2%	204	2%
Amambay	468	2%	288	3%	180	2%
Canindeyú	768	4%	252	2%	516	5%
Pdte. Hayes	480	2%	108	1%	372	4%

Fuente: Elaboración propia en base a documento de diseño muestral de ELPS Elaborados por STP – Gobierno de Paraguay

Tabla 15: Distribución de muestra esperada según Departamento y Área

Departamentos	Total		Urbana		Rural	
	n	%	n	%	n	%
Total	18.000	100%	9.550	100%	8.450	100%
Asunción	1.310	7%	1.310	14%		0%
Concepción	630	4%	210	2%	420	5%
San Pedro	1.360	8%	250	3%	1.110	13%
Cordillera	950	5%	310	3%	640	7%
Guairá	670	4%	240	3%	430	5%
Caaguazú	1.430	8%	520	5%	910	11%
Caazapá	580	3%	100	1%	480	6%
Itapúa	1.720	10%	640	7%	1.080	13%
Misiones	420	2%	200	2%	220	3%
Paraguarí	840	5%	230	2%	610	7%
Alto Paraná	2.040	11%	1.330	14%	710	8%
Central	4.270	24%	3.490	37%	780	9%
Ñeembucú	350	2%	180	2%	170	2%
Amambay	390	2%	240	3%	150	2%
Canindeyú	640	4%	210	2%	430	5%
Pdte. Hayes	400	2%	90	1%	310	4%

Tabla 16: Distribución de la muestra lograda según Departamento y Área

	Total		Urbana		Rural	
	n	%	n	%	n	%
Total	15.173	100,0%	6.973	100,0%	8200	100,0%
Asunción	655	4,3%	655	9,4%	0	0,0%
Concepción	620	4,1%	181	2,6%	439	5,4%
San pedro	1.403	9,2%	240	3,4%	1.163	14,2%
Cordillera	823	5,4%	251	3,6%	572	7,0%
Guaira	624	4,1%	200	2,9%	424	5,2%
Caaguazú	1.416	9,3%	466	6,7%	950	11,6%
Caazapá	596	3,9%	105	1,5%	491	6,0%
Itapúa	1.758	11,6%	625	9,0%	1.133	13,8%
Misiones	421	2,8%	195	2,8%	226	2,8%
Paraguarí	805	5,3%	205	2,9%	600	7,3%
Alto Paraná	1.708	11,3%	1.067	15,3%	641	7,8%
Central	2.720	17,9%	2.124	30,5%	596	7,3%
Ñeembucú	330	2,2%	168	2,4%	162	2,0%
Amambay	355	2,3%	213	3,1%	142	1,7%
Canindeyú	605	4,0%	201	2,9%	404	4,9%
Pte. Hayes	334	2,2%	77	1,1%	257	3,1%

Fuente: Elaboración propia en base a los datos de la base de datos de entrevistados ELPS – Primera Ronda Paraguay. 2015. Versión 19 de abril de 2016

2.4 Errores muestrales

Esta muestra permite acotar el error muestral absoluto en la ronda 1 a 0,9 por ciento a nivel nacional total, a 1,0 por ciento a nivel nacional urbano y a 1,1 por ciento a nivel nacional rural. Estas estimaciones se hicieron utilizando la expresión para Muestreo Aleatorio Simple (M.A.S.), bajo los siguientes supuestos:

- Uso de estimador de proporción
- Varianza máxima ($p = q = 0,5$)
- Nivel de confianza igual a 95% ($z = 1,96$)
- Efecto diseño igual⁹ a 1,4. Dado que no se contó con información para saber a priori el efecto diseño se consideró un nivel de correlación entre-clases de 0,1 para un cluster de 12 individuos por conglomerado seleccionado en la última y tercera etapa del diseño muestral para el presente estudio.

⁹ El efecto del diseño refiere a la medición del grado de distorsión que sufren las varianzas debido al diseño muestral empleado, entregando una valoración directa de la alteración que sufren los intervalos de confianza estimados cuando el diseño muestral no es aleatorio simple. “La forma de poder evaluar el efecto de un diseño muestral es comparar la varianza de un estimador obtenida mediante ese diseño con la que se obtendría mediante un diseño simple de referencia. Habitualmente se considera muestreo de referencia el obtenido empleando una selección aleatoria simple. El cociente entre ambas varianzas se conoce como el efecto del diseño y se denomina DEFF (*design effect*)” (Guillén & Ayuso; 2004). Es decir, indicaremos un DEFF igual a 1,4, lo que quiere decir que por efecto del diseño la varianza se ha multiplicado por 1,4.

La expresión del error muestral para M.A.S. es la siguiente:

$$e = z * \sqrt{\frac{p * q}{n}}$$

Para el cálculo del tamaño muestral, se toma en cuenta en forma complementaria el efecto diseño que introduce el muestreo por conglomerados, debido a que supone una pérdida de heterogeneidad de la muestra. Esto se explica porque los individuos en un mismo conglomerado son más parecidos entre sí (homogéneos) que los individuos en conglomerados diferentes.

El efecto diseño está definido como el cociente entre la varianza del estimador que se basa en el muestreo a testear respecto a la varianza del estimador obtenido bajo el supuesto de diseño aleatorio simple.

A continuación se presenta la distribución de la muestra esperada por Departamento considerando la distribución porcentual entregada en las proyecciones de la DGEEC.

Tabla 17: Distribución de la muestra esperada y error muestral según Departamento y Área (considera la fórmula del error por M.A.S.)

Departamentos	Total		Urbana		Rural	
	n	error	N	error	n	error
Total	18000	0,7%	9550	1,0%	8450	1,1%
Asunción	1310	2,7%	1310	2,7%	0	
Concepción	630	3,9%	210	6,8%	420	4,8%
San Pedro	1360	2,7%	250	6,2%	1110	2,9%
Cordillera	950	3,2%	310	5,6%	630	3,9%
Guairá	670	3,8%	240	6,3%	430	4,7%
Caaguazú	1430	2,6%	520	4,3%	910	3,2%
Caazapá	580	4,1%	100	9,8%	480	4,5%
Itapúa	1720	2,4%	640	3,9%	1080	3,0%
Misiones	420	4,8%	200	6,9%	220	6,6%
Paraguarí	840	3,4%	230	6,5%	610	4,0%
Alto Paraná	2040	2,2%	1330	2,7%	710	3,7%
Central	4270	1,5%	3490	1,7%	780	3,5%
Ñeembucú	350	5,2%	180	7,3%	170	7,5%
Amambay	390	5,0%	240	6,3%	150	8,0%
Canindeyú	640	3,9%	210	6,8%	430	4,7%
Pdte. Hayes	400	4,9%	90	10,3%	300	5,7%

Fuente: elaboración propia

La Tabla 18 entrega la distribución de la muestra y error muestral para las encuestas logradas según departamento y área.

Tabla 18: Distribución de la muestra lograda y error muestral según Departamento y Área (considera la fórmula del error por M.A.S.)

	Total		Urbana		Rural	
	n	e	n	e	n	e
Total	15173	0,796%	6.973	1,174%	8200	1,082%
Asunción	655	3,829%	655	3,829%		
Concepción	620	3,936%	181	7,284%	439	4,677%
San pedro	1.403	2,616%	240	6,326%	1.163	2,874%
Cordillera	823	3,416%	251	6,186%	572	4,098%
Guaira	624	3,923%	200	6,930%	424	4,759%
Caaguazú	1.416	2,604%	466	4,540%	950	3,180%
Caazapá	596	4,014%	105	9,564%	491	4,423%
Itapúa	1.758	2,337%	625	3,920%	1.133	2,911%
Misiones	421	4,776%	195	7,018%	226	6,519%
Paraguarí	805	3,454%	205	6,845%	600	4,001%
Alto Paraná	1.708	2,371%	1.067	3,000%	641	3,871%
Central	2.720	1,879%	2.124	2,126%	596	4,014%
Ñeembucú	330	5,395%	168	7,561%	162	7,700%
Amambay	355	5,201%	213	6,715%	142	8,224%
Canindeyú	605	3,984%	201	6,912%	404	4,876%
Pte. Hayes	334	5,362%	77	11,168%	257	6,113%

Fuente: elaboración propia

2.5 Probabilidades de selección de las unidades muestrales

Una vez aplicada la encuesta la muestra está conformada por 15.173 entrevistados - que entregaron información personal y de sus respectivos hogares¹⁰ -, para los cuales el factor de expansión corresponde al inverso de la probabilidad de selección del individuo.

El cálculo de la probabilidad de selección de un individuo en la muestra es el siguiente:

- a) *Probabilidad de selección de la UPM.* La selección se realizará con probabilidad proporcional al tamaño, de acuerdo al Pre CENSO -2012,

$$f_1 = \frac{N_{hu}}{N_h}$$

Donde:

N_{hu} : n° de UPM seleccionadas del estrato h.

N_h : n° de total de UPM del estrato h.

¹⁰ Cabe señalar que en la tercera etapa de selección se logró tener información de 18.333 viviendas que accedieron a contestar la composición del hogar para la selección con la tabla de Kish.

b) *Probabilidad de selección de una vivienda en cada UPM u-ésimo del estrato h.*

$$f_2 = \frac{m_{hu}}{M'_{hu}}$$

Donde:

m_{hu} : n° de viviendas seleccionadas en la UPM u-ésimo del estrato h.

M'_{hu} : n° de viviendas empadronadas en la fecha de actualización de la UPM u-ésimo del estrato h.

c) *Probabilidad de selección de la persona mayor de 15 años en la vivienda i de la UPM u-ésimo del estrato h.* Se procedió a seleccionar a un individuo al interior de la vivienda. De esta forma, para cada persona se determinó la cantidad de sujetos elegibles (sin considerar personas que, pese a tener más de 15 años no deben participar del estudio por otras limitaciones), y el factor de expansión de selección del individuo al interior del hogar es igual a:

$$f_3 = \frac{q}{P_{hui}}$$

Donde:

q : toma valor 1 para personas mayores de 15 años.

P_{hui} : n° ponderado (por q) de personas mayores de 15 años seleccionables en la vivienda i de la UPM u-ésimo del estrato h.

d) *Probabilidad general de selección de la persona k en la vivienda i en la UPM u-ésima del estrato h.*

$$f_k = f_1 * f_2 * f_3$$

2.6 Construcción de pesos (factores de expansión)

Los pesos o factores de expansión, en general, se utilizan para compensar las distintas probabilidades de selección de las unidades muestrales o por cierto tipo de datos faltantes. De este modo, con el uso de estos pesos, la muestra se mantiene representativa de la población objetivo.

2.6.1 Pesos transversales de la Ronda 1 ELPS - Paraguay

En una encuesta longitudinal se pueden construir factores de expansión longitudinales y de corte transversal. En el caso del factor de corte transversal, se construye un factor específico en cada etapa, y se asume que la no respuesta es aleatoria, y por tanto se distribuye la probabilidad de selección de la etapa anterior, entre las unidades muestrales efectivamente encuestadas en la etapa correspondiente.

Con esto se genera un factor de expansión— $w=1 /f_i$ para cada subgrupo.

2.6.2 Factor de ajuste de la Ronda 1 ELPS – Paraguay información EPH

Una vez calculado el factor de expansión, se aplicó un segundo ajuste al factor para garantizar que la distribución de la población estimada a través de la encuesta coincida con la distribución de la población por quinquenios de edad de la EPH 2015. Esto es la proporción entre la distribución de cada edad de la muestra EPH sobre la distribución de las edades de la población estimada con el factor original.

$$f_{edad} = \frac{d_{eph}}{d_{elps}}$$

Donde,

f_{edad}= factor de ajuste a las edades EPH

d_{eph}=distribución de las edades de la muestra EPH

d_{edad}=distribución de las edades de la población estimada con el factor original

Tabla 19: Distribución muestra con factor de selección ELPS según quinquenios

Quinquenios	N	%
15 a 19 años	416.835	15%
20 a 24 años	316.951	11%
25 a 29 años	280.583	10%
30 a 34 años	280.023	10%
35 a 39 años	227.224	8%
40 a 44 años	222.133	8%
45 a 49 años	204.367	7%
50 a 54 años	207.940	7%
55 a 59 años	176.419	6%
60 a 64 años	146.364	5%
65 a 69 años	123.591	4%
70 años o más	206.882	7%
Total general	2.809.312	100%

Fuente: elaboración propia

Tabla 20: Quinquenios de Población de 15 años y más (EPH)

	Frecuencia	Porcentaje
15 a 19 años	747.642	15,2
20 a 24 años	652.123	13,2
25 a 29 años	543.271	11,0
30 a 34 años	501.495	10,2
35 a 39 años	400.451	8,1
40 a 44 años	352.891	7,2
45 a 49 años	344.627	7,0
50 a 54 años	333.076	6,8
55 a 59 años	306.387	6,2
60 a 64 años	239.961	4,9
65 a 69 años	178.964	3,6
70 años o más	325.068	6,6
Total	4.925.956	100,0

Fuente: elaboración propia

2.6.3 Factor de ajuste de la Ronda 1 ELPS – Paraguay información DGEEC

Una vez calculado el factor de expansión de selección con el respectivo ajuste por edad EPH, se aplicó un tercer ajuste al factor para garantizar que la población estimada a través de la encuesta coincida con la población según las Proyecciones de la DGEEC a Octubre del 2015. Dado que no es posible lograr que este ajuste para cada grupo particular de la población se tiene que definir variables claves para las cuales realizar el ajuste por post-estratificación. Se utilizaron las siguientes:

Sexo:

1. Hombre
6. Mujer

Departamento:

0. Asunción
1. Concepción
2. San pedro
3. Cordillera
4. Guaira
5. Caaguazú
6. Caazapá
7. Itapuá
8. Misiones
9. Paraguari
10. Alto Paraná
11. Central
12. Ñeembucú

13. Amambay
14. Canindeyú
15. Pte. Hayes

La razón entre la población efectiva y la población estimada con la encuesta constituye el factor de ajuste delta de la celda correspondiente.

$$f_{\text{delta}} = \frac{t_{\text{DGEEC}}}{t_{\text{ELPS}}}$$

Donde,

f_{delta} = factor de ajuste

t_{DGEEC} = población efectiva

T_{ELPS} = población estimada

Tabla 21: Factor de ajuste por sexo/departamento a Proyección DGEEC OCT 2015 (delta)

<u>Departamento</u>	<u>Hombres</u>	<u>Mujeres</u>	<u>Total</u>
Asunción	2,84774539	1,99319601	0,43227742
Concepción	1,58565891	1,23139229	0,71813157
San pedro	1,53201801	1,05555892	0,7905731
Cordillera	1,77435671	1,24364956	0,67761594
Guaira	1,62972992	1,37455723	0,66822354
Caaguazú	1,52319653	1,25730021	0,72302247
Caazapá	1,50129029	1,26874317	0,72470851
Itapúa	1,44118863	1,29138653	0,73349623
Misiones	1,50000145	1,12741819	0,77559335
Paraguarí	1,57800485	1,26860969	0,70738821
Alto Paraná	1,72078752	1,38760148	0,6498775
Central	2,47465431	1,74372895	0,49064901
Ñeembucú	1,25912439	1,31330344	0,7778726
Amambay	1,80920389	1,24293145	0,67943409
Canindeyú	1,61881412	1,41281739	0,65956872
Pte. Hayes	1,79346101	1,32296673	0,65249667
Total general	1,84480136	1,4413977	0,61735179

Fuente: elaboración propia

Tabla 22: Distribución población de 15 años o más según quinquenios (Proyecciones DGEEC)

<u>Quinquenios</u>	<u>N</u>	<u>%</u>
15 a 19 años	664.861	15%
20 a 24 años	630.139	14%
25 a 29 años	584.932	13%
30 a 34 años	513.265	11%

35 a 39 años	417.825	9%
40 a 44 años	369.765	8%
45 a 49 años	328.546	7%
50 a 54 años	285.185	6%
55 a 59 años	245.305	5%
60 a 64 años	196.531	4%
65 a 69 años	140.165	3%
70 años o más	170.615	4%
Total general	4.547.133	100%

Fuente: elaboración propia

Tabla 23: Distribución muestra con factor de ajuste a al total población DGEEC ELPS según quinquenios - Base de datos marzo 2016

Quinquenios	N	%
15 a 19 años	703.133	15%
20 a 24 años	607.191	13%
25 a 29 años	489.604	11%
30 a 34 años	460.067	10%
35 a 39 años	361.795	8%
40 a 44 años	326.689	7%
45 a 49 años	323.315	7%
50 a 54 años	312.993	7%
55 a 59 años	292.950	6%
60 a 64 años	222.895	5%
65 a 69 años	158.368	3%
70 años o más	288.134	6%
Total general	4.547.134	100%

Fuente: elaboración propia

2.6.4 Factor final de expansión de entrevistado de la Ronda 1 ELPS – Paraguay

Así, el factor de expansión final de cada entrevistado en la primera ronda de la ELPS de Paraguay se obtiene de la siguiente forma:

$$ffinal = w * f_edad * delta$$

Donde,

ffinal = factor final

w = factor orifinal

f_edad = factor de ajuste a las edad EPH

delta = factor de ajuste a las proyecciones de la DGGEC a Octubre del 2015

El valor promedio del factor de expansión es 300, la desviación estándar es 245, mientras que el valor mínimo es 21 y el máximo 3584.

3. Referencias

- DGEEC (2015). Paraguay. Proyección de la Población por Sexo y Edad, según departamento, 2000 – 2025, DGEEC, Paraguay.
- Division, U. N. S. (2008). Designing household survey samples: practical guidelines, United Nations Publications.
- Lynn, P, N. Buck, N., Burton, A. C le and Laurie (2005), A review of methodological Research Pertinent to Longitudinal Survey Design and Data Collection, ISER Working Papers, No. 2005-29.
- Lynn, P., Buck, N., Burton, J., Heather, L., & Uhrig, S. (2006). Quality profile: British household panel survey-version 2.0: Waves 1 to 13: 1991-2003. ISER report, Essex.

4. Anexos

4.1 Actualización del número de viviendas Pre CENSO 2012

El recorrido de la actualización del número de viviendas del Pre CENSO 2012 consistió en la visita previa de cada UPM por parte del supervisor, quien realizaba los siguientes pasos:

1. Con el mapa que tenía la información cartográfica del Pre CENSO el Supervisor llegaba con su equipo para iniciar la visita de la UPM.
2. Desde la esquina noroeste iniciaba el recorrido siguiendo las manecillas del reloj (a la derecha).
3. En dicho recorrido anotaba cada una de las viviendas encontradas enumerando aquellas que eran residenciales ocupadas. Las viviendas señaladas en el mapa pero identificadas como desocupadas se identificaban con una D y aquellas sin moradores con una A. Para chequear ello, el supervisor averiguaba en cada dirección cuál era la situación actual y realizaba la clasificación sobre los mapas originales.
4. Una vez realizado el recorrido completo se procede a identificar la vivienda de partida de la selección.
5. Si el número de viviendas pesquisado es igual al del Pre CENSO 2012 se considera el número de partida propuesto por la hoja de actualización.
6. Si el número de viviendas pesquisado no es igual (sea mayor o menor) se aplica la siguiente forma:

$$\text{Vivienda número 1: TV * aleUPM}$$

Donde,

TV: Total de viviendas de la UPM

aleUPM: número aleatorio de UPM indicado en la hoja de actualización

Ejemplo:

UPM 00XZ

TV: 42

aleUPM: 0,35

$$\text{Vivienda número 1} = 42 * 0,35$$

$$\text{Vivienda número 1} = 14,7$$

$$\text{Vivienda número 1} = 15$$

En este caso las viviendas del primer compacto a encuestar serán las identificadas como n° 15, 16, 17, 18, 19 y 20.

7. Ahora en cada UPM se deben aplicar 12 encuestas divididas en 2 compactos de 6 encuestas.
8. Para definir los compactos se divide el TV por dos y ese resultado es el número de viviendas que debo contar a partir de la vivienda 2 del primer compacto para identificar la vivienda número 1 del segundo compacto.

Ejemplo:

UPM 00XZ

TV: 42

aleUPM: 0,35

Vivienda número 1 de 1er. Compacto = 15

Salto para identificar la vivienda número 1 del 2do. Compacto=42/2

Salto para identificar la vivienda número 1 del 2do. Compacto=21

Entonces, desde la vivienda 16 cuento 21 (16+22=38) y esa vivienda será mi 1era. Vivienda del según compacto.

Por tanto, las viviendas n° 36, 37, 38, 39, 40 y 41 serán las que se deberán encuestar en el 2do. Compacto.

9. Las UPM por definición tienes 50 viviendas máximo en zonas urbanas y 60 viviendas en zonas rurales. En el caso de que existe alguna de ellas con más viviendas se realiza un proceso de partición.
10. Para ello se utiliza la siguiente fórmula:

$$TV/20$$

Donde,

TV: Total de viviendas de la UPM

20: constante

Ejemplo,

UPM: 0WXZ

TV: 100

$$\text{Partición} = 100/20$$

$$\text{Partición} = 5$$

Es decir, la UPM se divide en 5, conformando 5 nuevas UPM de 20 viviendas cada una. Así la UPMn1 estaría de la 1 a la 20, la UPMn2 de la 21 a la 40, la UPMn3 de la 41 a la 60, la UPMn4 de la 61 a la 80 y la UPMn5 de la 81 a la 100.

4.2 Análisis de la actualización de viviendas

Para la construcción de los factores de expansión se utilizó la información de la actualización de viviendas del recorrido ELPS, algunas de las observaciones al respecto con las siguientes.

1. En 84 UPM el número de viviendas del Pre - CENSO 2012 coincide con el número de viviendas actualizadas ELPS - Paraguay.
2. En 1300 UPM el número de viviendas del Pre – CENSO 2012 es mayor que el número de viviendas actualizadas en el recorrido ELPS – Paraguay
3. En 393 UPM el número de viviendas del Pre – CENSO 2012 es menor que el número de viviendas actualizadas en el recorrido ELPS – Paraguay
4. En 15 UPM no se realizaron encuestas por problemas de acceso a las UPM

Tabla 24: UPM sin encuestas logradas

Departamento	UPM sin encuestas logradas	Situación del trabajo de campo
Asunción	0006	Edificio que no permitió acceso
Asunción	0009	Edificio que no permitió acceso
Asunción	0032	Edificio que no permitió acceso
Asunción	0047	Muestra no elegible viviendas desocupadas
Asunción	0050	Edificio que no permitió acceso
Asunción	0055	Muestra no elegible viviendas de uso no habitacional
Asunción	0095	Edificio que no permitió acceso
Alto Paraná	0473	Rechazo
Alto Paraná	0513	Rechazo
Central	0585	Edificio que no permitió acceso
Central	0655	No contacto a las viviendas
Central	0749	Condominio cerrado no permitió acceso
Alto Paraná	1571	Ya no existe la UPM
Canindeyú	1751	Ya no existe la UPM
Pte. Hayes	1800	No se accede por condiciones naturales

Fuente: elaboración propia

4.3 Tabla de kish

ENCUESTA LONGITUDINAL DE PROTECCIÓN SOCIAL DE PARAGUA Y - 2015
Composición de personas en la vivienda

Folio_Encuesta:		Dirección:						Letra:		
INSTRUCCIONES: Anote en orden decreciente según edad a todos los integrantes de la vivienda de 15 o más años y seleccione utilizando la Tabla de Kish uno de ellos para ser encuestado.										
Ord.	1a. N° de hogares en la vivienda.	1b. Nombre de pila de cada integrante seleccionable del hogar.	1c. Parentesco con el Jefe de Hogar	1d. Sexo	1e. Edad (años cumplidos)	1.f Persona seleccionable o excluida	1g. Numere de mayor a menor edad a todos los integrantes seleccionables del hogar de 15 años o más.		1i. Último grado o curso aprobado	1j. Situación ocupacional
	1. Hogar 1 2. Hogar 2 3. Hogar 3	Seleccionable: *Tiene 15 años o más cumplidos al momento de la entrevista *No tiene ninguna discapacidad intelectual	1. Jefe(a) de Hogar 2. Cónyuge o conviviente 3. Hijo(a) 4. Padre/Madre 5. Hermano(a) 6. Yerno/Nuera 7. Nieto(a) 8. Suegro(a) 9. Otros parientes 10. Empleada(o) doméstica(o) 11. Otro no pariente	1. Hombre 6. Mujer	*en el cuestionario preguntar fecha de nacimiento de la persona seleccionada.	1. 15 años y más 2. Discapacidad mental 3. Otra exclusión. Especifique.	1h. Señale con una cruz la persona seleccionada en la Tabla de Kish		1. Ninguna 2. Grado Especial 3. Programas de Alfabetización 4. Educación inicial 5. Educación Escolar Básica (1° al 6°) Ex Primaria 6. Educación Escolar Básica (7° al 9°) 7. Ex Secundaria Básica (1° al 3°) 8. Educación Media (1° al 3°) 9. Ex Bachillerato (4° al 6°) 10. Superior no Universitario 11. Superior Universitario 12. Postgrado 77. No sabe 99. No responde	1. Trabaja 2. Desempleado 3. Busca 1era. Vez 4. Jubilado 5. Ama de casa 6. Estudiante 7. Otro inactivo
	1a.	1b.	1c.	1d.	1e.	1f.	1.g	1.h	1i.	1j.
1										
2										
3										

4.4 Glosario

1. UPM: Unidad primaria muestral.
2. Conglomerado: Consiste en la división previa de la población de estudio en grupos o clases que se suponen heterogéneos respecto a característica a estudiar, de los cuales son seleccionados algunos para ser parte de la muestra.
3. Estrato: Consiste en la división previa de la población de estudio en grupos o clases que se suponen homogéneos respecto a característica a estudiar y que no se solapan.
4. Área: corresponde a la clasificación del territorio bajo la distinción entre urbano – rural.
5. Hogar: Es la persona o conjunto de personas, sean o no parientes que residen habitualmente en una misma vivienda particular, ocupándola total o parcialmente y que atienden en común sus necesidades alimenticias (comparten los gastos para la olla) o de otra índole.
6. Compacto: Agrupación de viviendas que han sido seleccionadas para ser encuestadas
7. Efecto diseño: cociente entre la varianza del estimador que se basa en el muestreo a testear respecto a la varianza del estimador obtenido bajo el supuesto de diseño aleatorio simple.
8. Error muestral: es la desviación de la muestra seleccionada de las verdaderas características, rasgos, comportamientos, cualidades o figuras de toda la población.

