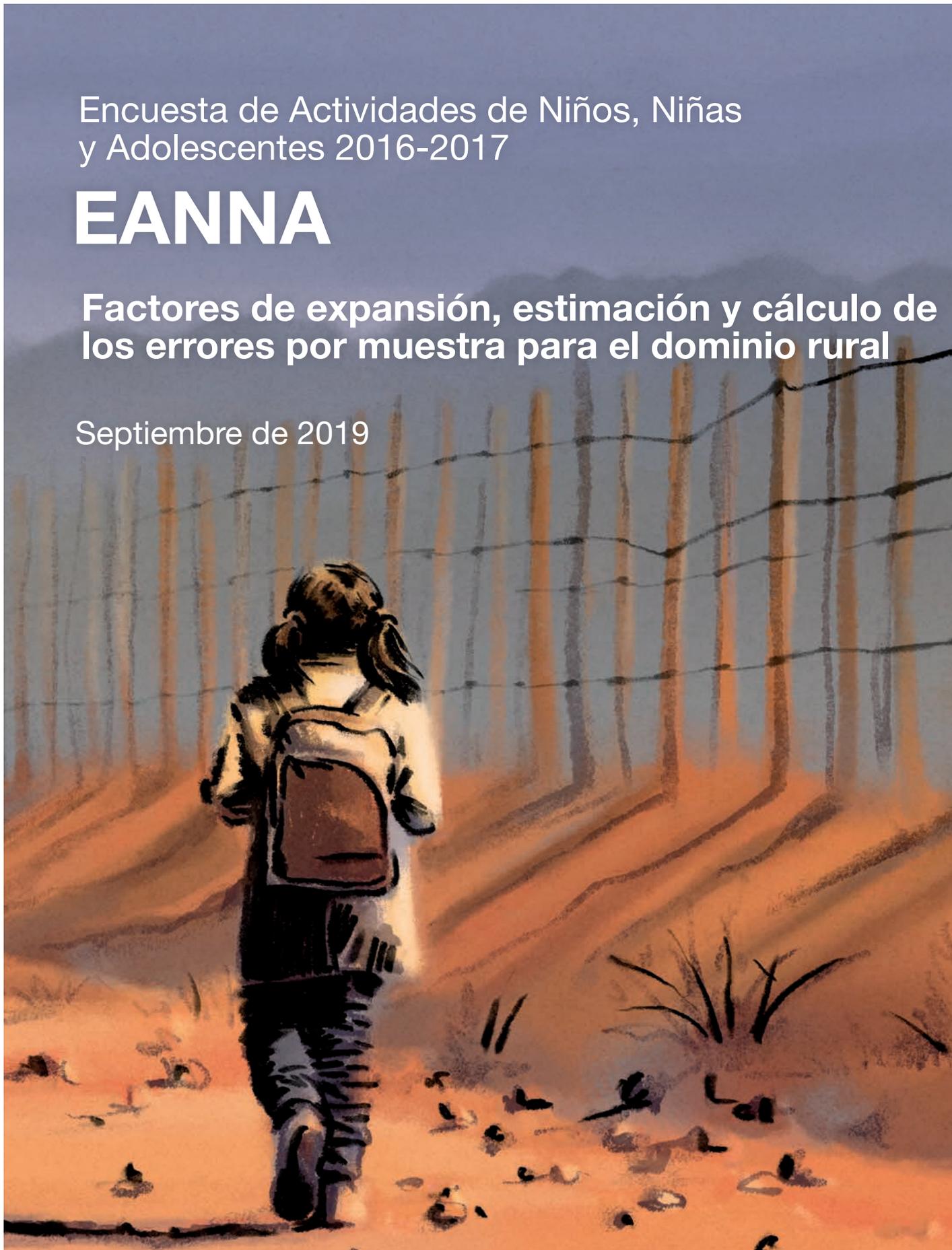


Encuesta de Actividades de Niños, Niñas
y Adolescentes 2016-2017

EANNA

Factores de expansión, estimación y cálculo de
los errores por muestra para el dominio rural

Septiembre de 2019



Encuesta de Actividad de Niñas, Niños y Adolescentes 2016-2017

Factores de expansión, estimación y cálculo de los errores por muestra para el dominio rural

Septiembre de 2019

Instituto Nacional de Estadística y Censos (INDEC)

Esta publicación fue realizada por Hugo Delfino, integrante del equipo de la Dirección General de Estudios Macroeconómicos y Estadísticas Laborales, dependiente de la Secretaría de Gobierno de Trabajo y Empleo del Ministerio de Producción y Trabajo.

ISBN en trámite

Libro de edición argentina



Esta publicación utiliza una licencia Creative Commons. Se permite su reproducción con atribución de la fuente.

Responsable de la edición: Lic. Jorge Todesca

Director técnico: Mag. Pedro Lines

Directora de la publicación: Mag. Silvina Viazzi

Coordinación de producción editorial: Lic. Marcelo Costanzo

Buenos Aires, septiembre de 2019

Publicaciones del INDEC

Las publicaciones editadas por el Instituto Nacional de Estadística y Censos pueden ser consultadas en www.indec.gov.ar y en el Centro Estadístico de Servicios, ubicado en Av. Presidente Julio A. Roca 609 C1067ABB, Ciudad Autónoma de Buenos Aires, Argentina. El horario de atención al público es de 9:30 a 16:00.

También pueden solicitarse al teléfono (54-11) 5031-4632

Correo electrónico: ces@indec.gov.ar

Calendario anual anticipado de informes: <https://www.indec.gov.ar/indec/web/Calendario-Fecha-0>



Índice

1. Introducción	4
2. Diseño muestral en el dominio rural	4
3. Dominios geográficos de estimación y tamaño de la muestra	7
4. Determinación y ajuste de los factores de expansión	8
4.1 Ajuste por elegibilidad dudosa	9
4.2 Ajuste por no respuesta	9
4.3 Ajuste por calibración	10
Anexo I. Total de UPM, USM y UTM	11
Anexo II. Distribución de la muestra de viviendas seleccionadas por jurisdicción	12
Anexo III. Tasa de respuesta de los hogares	13

1. Introducción

El Instituto Nacional de Estadística y Censos (INDEC) conjuntamente con el Ministerio de Trabajo, Empleo y Seguridad Social (MTEySS) realizaron la Encuesta de Actividades de Niños, Niñas y Adolescentes (EANNA) 2016-2017, con el objetivo de medir y diagnosticar la situación del trabajo infantil en el país¹.

Se trata de la segunda encuesta de este tipo aplicada en el país y la primera de carácter nacional, ya que cubre a toda la población de la Argentina, tanto la residente en zonas urbanas como la que integran las áreas rurales.

Esta publicación es una guía de referencia básica de la metodología adoptada para determinar los factores de expansión que se emplean en las estimaciones oficiales, y que permite estimar sus errores de muestreo en el dominio de la encuesta.

La EANNA Rural es un estudio basado en una muestra probabilística en la que se seleccionan viviendas de áreas rurales. En la Argentina, se clasifica como población rural a la que se encuentra agrupada en localidades de menos de 2.000 habitantes o dispersa en campo abierto.

En esta sección se describe el diseño muestral, el cálculo de los factores de expansión con sus ajustes por elegibilidad, rechazo y calibración de la EANNA Rural.

El presente documento se centra en las características y aspectos metodológicos que son específicos al ámbito rural. En aquellos tópicos que corresponda, se hace referencia a la metodología especificada en la nota técnica N° 1 *Encuesta de Actividad de Niños, Niñas y Adolescentes 2016-2017. Factores de expansión, estimación y cálculo de los errores por muestra para el dominio urbano*.

2. Diseño muestral en el dominio rural

El diseño de la muestra fue probabilístico, a cada unidad de la población se le asoció una probabilidad distinta de cero de ser seleccionada para integrar la muestra.

El Censo 2010 fue el marco general para la elaboración de la muestra de la EANNA Rural. Brinda información de la cantidad de viviendas y población para cada provincia a nivel de radio censal, pero, debido al tiempo transcurrido y a que no se dispone de datos sobre la localización de cada una de las viviendas, se requirió obtener información actualizada del total de viviendas particulares presuntamente ocupadas en los radios censales seleccionados y elaborar un listado de viviendas a partir del cual seleccionar las viviendas a encuestar.

En consecuencia, un objetivo técnico fue realizar un diseño muestral que posibilite lograr un marco apropiado a efectos de reducir al máximo el sesgo por el tiempo transcurrido. Se consideró que el diseño probabilístico más eficiente es un muestreo estratificado de conglomerados a tres etapas.

¹ Ver https://www.indec.gob.ar/ftp/cuadros/sociedad/eanna_2018.pdf.

Estrategia muestral

Como la incidencia de la poblacional rural dispersa es muy diferente según la región del país de que se trate, se definieron dos estratos: radios con población agrupada y radios con población dispersa. Al interior de cada estrato/región se eligieron radios con probabilidad proporcional a la población objeto de estudio.

Definición de radios censales considerados

- Radio urbano: con población agrupada únicamente, y conformado por manzanas o sectores pertenecientes a una localidad (incluye radios de localidades de menos de 2.000 habitantes que son consideradas localidades con población rural).
- Radio rural: con población dispersa únicamente, y donde las viviendas se distribuyen en campo abierto en forma diseminada.
- Radio mixto: con población rural dispersa en campo abierto, y con población agrupada en pequeños poblados o en bordes amanzanados de localidades².

En la primera etapa de la investigación, en cada dominio se eligió una muestra de radios dentro de cada estrato, con probabilidad proporcional al tamaño. En ellos, se construyó un marco de muestreo de viviendas particulares presuntamente ocupadas (habitadas).

A los radios con población dispersa, debido a la elevada diseminación geográfica que suelen tener las viviendas, se los subdividió en áreas de menor extensión denominadas “segmentos”. A los radios con población agrupada en localidades si bien no se los subdividió, por razones de homogeneidad en la nomenclatura, también se los denominó “segmentos”, con probabilidad 1 de entrar en la muestra.

En la segunda etapa, se seleccionó la submuestra de segmentos dentro de cada radio. En los radios de población agrupada se incluyó toda el área y en los de población dispersa se seleccionaron hasta dos segmentos.

En la tercera etapa, se eligió la submuestra de viviendas presuntamente habitadas.

En resumen, en el presente diseño se definieron los radios como unidades primarias de muestreo (UPM); los segmentos, como unidades secundarias de muestreo (USM) y las viviendas particulares, como las unidades terciarias de muestreo (UTM).

En total fueron elegidos 359 radios distribuidos en las 23 provincias del país incluidas en el estudio³.

² Ver <https://geoservicios.indec.gob.ar/codgeo/index.php?pagina=definiciones>.

³ Para ver un detalle por región de las UPM, USM y UTM que componen la muestra maestra e involucradas, ver Anexo I.

Cuadro 1. Cantidad de radios seleccionados por tipo

Radios	Total	Porcentaje
Total	359	100
Solo con población agrupada	76	21
Solo con población dispersa	160	45
Mixtos (población agrupada y dispersa)	123	34

Fuente: INDEC, *Encuesta de Actividades de Niños, Niñas y Adolescentes 2016-2017*.

Elaboración del marco de muestreo

En las localidades que integran la muestra (población agrupada), se realizó el conteo e identificación de las viviendas particulares ocupadas siguiendo los criterios habituales en estos casos para el correcto recorrido del área a relevar. Posteriormente, se seleccionaron de manera sistemática las viviendas a encuestar.

La elaboración del marco de viviendas en radios con población rural dispersa en el territorio planteó particulares desafíos: desde la complejidad de abordar áreas extensas⁴, con terrenos irregulares y gran dispersión de viviendas, hasta problemas de accesibilidad (acceso por caminos de tierra, huellas y senderos de difícil o nulo tránsito vehicular, ausencia de caminos internos que interconecten las viviendas o caseríos, localización de viviendas en áreas remotas, etc.). No se cuenta con antecedentes de elaboración de un marco de viviendas rurales dispersas de esta magnitud en nuestro país.

Para resolver este desafío se diseñó una estrategia metodológica innovadora que consistió en identificar, mediante el uso de sistemas de información geográfica (SIG) e imágenes satelitales de Google Earth y Bing, techos de estructuras edilicias (posibles viviendas) y de esa manera poder construir el marco de muestreo.

Con base en las pruebas realizadas, se sabe que los techos son un buen indicador de las viviendas particulares ocupadas, aunque pueden encontrarse discrepancias, que al llegar a campo, permiten actualizar dicho marco. Para ello se realizan las siguientes actividades:

- ✓ Identificación en terreno de las viviendas particulares ocupadas y registro de la información en mapas y planillas operativas. Se recorren los segmentos de manera completa y se verifica, entre otros datos, si los techos marcados corresponden o no a viviendas particulares ocupadas.
- ✓ Incorporación de las viviendas habitadas que no habían sido previamente detectadas a través de las imágenes satelitales y eliminación de las marcas que no correspondan.
- ✓ Volcado de la información en la cartografía o en croquis realizados a mano alzada y en planillas diseñadas a tal fin.

⁴ Los radios rurales con población dispersa seleccionados en la muestra de la EANNA Rural tienen en promedio una superficie de 289 km², el radio de menor superficie ocupa 2 km² en la provincia de Buenos Aires y el valor máximo corresponde a un radio de Neuquén, con 5.500 km².

Una vez obtenido el marco actualizado, se procede a sortear las viviendas a encuestar por el método de muestreo sistemático.

3. Dominios geográficos de estimación y tamaño de la muestra

La población objetivo de la encuesta está compuesta por los hogares⁵ y las personas que en 2017 residían en viviendas particulares⁶ ubicadas en localidades que, de acuerdo con datos del Censo Nacional de Población, Hogares y Viviendas de 2010, tenían menos de 2.000 habitantes o estaban dispersas en el territorio. Se excluyó de la población en estudio a las personas que vivían en viviendas colectivas o en viviendas particulares móviles.

La encuesta tiene como dominios de estimación al total país y las 5 regiones estadísticas:

- Cuyo: Mendoza, San Juan y San Luis.
- Noreste: Chaco, Corrientes, Formosa y Misiones.
- Noroeste: Catamarca, Jujuy, Salta, Tucumán, La Rioja y Santiago del Estero.
- Pampeana: Córdoba, Santa Fe, Entre Ríos, La Pampa y resto de partidos de Buenos Aires.
- Patagonia: Chubut, Neuquén, Río Negro, Santa Cruz y Tierra del Fuego.

La muestra seleccionada para la EANNA Rural 2016/2017 está constituida por 11.833 viviendas, distribuidas por dominio de estimación según el siguiente cuadro⁷:

⁵ Hogar es la persona o grupo de personas, parientes o no, que habitan bajo el mismo techo y comparten los gastos de alimentación u otros esenciales para vivir (por ejemplo: alquiler, expensas comunes, impuestos, servicios de luz, gas, agua, teléfono).

⁶ Vivienda es el recinto fijo o móvil destinado a alojar personas. Las viviendas pueden haber sido construidas o adaptadas para ser habitadas o bien ser utilizadas con ese fin. Según el régimen de convivencia de las personas que las ocupan, las viviendas se clasifican en:

- Viviendas particulares: destinadas a alojar personas que viven en un régimen de tipo familiar (incluye hoteles familiares, pensiones o inquilinatos).
- Viviendas colectivas: vivienda destinada a alojar personas que viven en un régimen institucional (no familiar), regulado por normas de convivencia de carácter administrativo, militar, religioso, de salud, de reclusión, de trabajo, de educación, etc.

⁷ Para ver la distribución de la muestra de viviendas por jurisdicción ver Anexo II.

Cuadro 2. Distribución de la muestra de viviendas por región

Región	Cantidad de viviendas
Cuyo	2.240
Noreste	2.100
Noroeste	2.120
Pampeana	2.923
Patagonia	2.450
Total del país	11.833

Fuente: INDEC, *Encuesta de Actividades de Niños, Niñas y Adolescentes 2016-2017*.

4. Determinación y ajuste de los factores de expansión

Distintos factores contribuyen a que las ponderaciones o factores de expansión para todas las unidades no sean uniformes. Entre ellos, los señalados en relación con la complejidad del diseño muestral y la asignación territorial de la muestra, o los ajustes por cobertura y no respuesta que se expondrán.

El diseño muestral empleado en la EANNA Rural determina el factor de expansión inicial de cada vivienda, y surge de la multiplicación de las inversas de las probabilidades de inclusión de cada una de las etapas de selección definidas en el apartado 2. Por lo tanto, el factor de expansión de la k -ésima vivienda ubicada en la j -ésima USM dentro de la i -ésima UPM se lo define como:

$$w^{(0)}_{ijk} = f_{1i} f_{2ij} f_{3ijk},$$

donde,

f_{1i} es la inversa de la probabilidad de inclusión de la i -ésima UPM;

f_{2ij} es la inversa de la probabilidad de inclusión en la segunda etapa de muestreo de la j -ésima USM dentro de la i -ésima UPM seleccionada;

f_{3ijk} es la inversa de la probabilidad de inclusión en la última etapa de muestreo de la k -ésima vivienda dentro de la j -ésima USM de la i -ésima UPM seleccionada.

Tal como se menciona en la sección 4 del documento *Encuesta de Actividad de Niños, Niñas y Adolescentes 2016-2017. Factores de expansión, estimación y cálculo de los errores por muestra para el dominio urbano.*, las principales fuentes de error de una encuesta son los de cobertura, a causa de la deficiencia o desactualización del marco de muestreo, la no respuesta, los errores de medición presentes cuando se indagan las características de interés a la población, y los que surgen en el procesamiento y edición de los datos.

Aun tomando estos recaudos, estos errores requieren ajustes previos a la etapa de estimación, con el objeto de aumentar la calidad.

4.1 Ajuste por elegibilidad dudosa⁸

En el cuadro 3, se presentan los resultados de la encuesta en el ámbito rural en relación a la cantidad de viviendas elegibles, no elegibles y de elegibilidad dudosa, por dominio de estimación y a nivel nacional, que intervienen con sus factores de expansión iniciales, $w^{(0)}_{ijk}$, en los cálculos del factor a_1 .

Cuadro 3. Cantidad de viviendas elegibles, no elegibles y de elegibilidad dudosa por región

Región	Viviendas en la muestra	Viviendas elegibles	Viviendas no elegibles	Viviendas de elegibilidad dudosa
Cuyo	2.240	2.206	16	18
Noreste	2.100	2.072	12	16
Noroeste	2.120	2.079	16	25
Pampeana	2.923	2.564	54	305
Patagonia	2.450	2.270	57	123
Total del país	11.833	11.191	155	487

Fuente: INDEC, *Encuesta de Actividades de Niños, Niñas y Adolescentes 2016-2017*.

Como por diseño la probabilidad de selección de un hogar dentro de una vivienda es la probabilidad de selección de dicha vivienda, en adelante, los factores de expansión corresponden a cada hogar identificado dentro de las viviendas elegibles. Es decir que el factor de expansión de cada vivienda es afectado a cada uno de los hogares que la componen, por lo tanto, el peso del l -ésimo hogar de la k -ésima vivienda en la j -ésima USM dentro de la i -ésima UPM es $W^{(1)}_{ijkl} = W^{(1)}_{ijk}$.

4.2 Ajuste por no respuesta⁹

En la EANNA Rural solo se consideraron 23 jurisdicciones¹⁰ como PROVINCIA y ESTRATO de UPM (2 categorías).

Para ilustrar la cantidad de unidades de la muestra que con sus factores de expansión

⁽¹⁾ $_{kl}$ se involucran en los cálculos de los factores de ajuste a_2 , el siguiente cuadro totaliza los hogares con y sin respuesta registrados en la encuesta a nivel nacional¹¹ y por dominio de estimación.

⁸ Ver este mismo apartado en el documento para el dominio urbano.

⁹ Ver este mismo apartado en el documento para el dominio urbano.

¹⁰ Se excluyen la CABA y los 31 partidos de la provincia de Buenos Aires que componen el conurbano.

¹¹ En el Anexo III se presentan las tasas de respuesta a nivel nacional y por dominio calculadas bajo el estándar habitual de la AAPOR.

Cuadro 4. Total de hogares en viviendas clasificadas como elegibles, hogares con y sin respuesta, por región

Región	Hogares elegibles	Hogares con respuesta	Hogares sin respuesta
Cuyo	2.219	2.031	188
Noreste	2.121	1.966	155
Noroeste	2.110	2.016	94
Pampeana	2.696	2.483	213
Patagonia	2.317	2.076	241
Total del país	11.463	10.572	891

Fuente: INDEC, *Encuesta de Actividades de Niños, Niñas y Adolescentes 2016-2017*.

4.3 Ajuste por calibración

Los factores $w^{(2)}_{ijkl}$ son calibrados de forma tal que satisfacen totales de población surgidos de proyecciones poblacionales para el conjunto de variables sociodemográficas utilizadas, deduciendo la población urbana estimada por la EANNA Urbana¹².

¹² Los totales poblacionales proyectados fueron calculados a partir de datos censales de población según CNPyV 2010 al 15 de noviembre de 2016 y determinados por la Dirección Nacional de Estadísticas Sociales y Poblacionales del INDEC.

Anexo I. Total de UPM, USM y UTM

Cuadro 5. Total de UPM, USM y UTM, por jurisdicción

Región	Cantidad de radios seleccionados (UPM)	Cantidad de segmentos seleccionados (USM)	Cantidad de viviendas seleccionadas (UTM)
Cuyo	70	120	2.240
Noreste	70	114	2.100
Noroeste	70	104	2.120
Pampeana	79	112	2.923
Patagonia	70	103	2.450
Total del país	359	553	11.833

Fuente: INDEC, *Encuesta de Actividades de Niños, Niñas y Adolescentes 2016-2017*.

Anexo II. Distribución de la muestra de viviendas seleccionadas por jurisdicción

Cuadro 6. Cantidad de viviendas seleccionadas por jurisdicción

Región	Provincia	Cantidad de viviendas seleccionadas
Cuyo	Mendoza	1.568
	San Juan	448
	San Luis	224
Total de Cuyo		2.240
NEA	Corrientes	450
	Chaco	480
	Formosa	300
	Misiones	870
Total del NEA		2.100
NOA	Catamarca	300
	Jujuy	270
	La Rioja	120
	Salta	390
	Santiago del Estero	530
	Tucumán	510
Total del NOA		2.120
Pampeana	Buenos Aires	851
	Córdoba	851
	Entre Ríos	407
	La Pampa	111
	Santa Fe	703
Total Pampeana		2.923
Patagonia	Chubut	560
	Neuquén	665
	Río Negro	1085
	Santa Cruz	105
	Tierra del Fuego	35
Total de la Patagonia		2.450
Total de viviendas seleccionadas		11.833

Fuente: INDEC, *Encuesta de Actividades de Niños, Niñas y Adolescentes 2016-2017*.

Anexo III. Tasa de respuesta de los hogares¹³

Cuadro 7. Tasa de respuesta por regiones y total del país

Región	RR5
	%
Cuyo	91,5
Noreste	92,6
Noroeste	95,5
Pampeana	91,8
Patagonia	89,6
Total del país	92,1

Fuente: INDEC, *Encuesta de Actividades de Niños, Niñas y Adolescentes 2016-2017*.

¹³Ver este mismo apartado en el documento para el dominio urbano.