

# ANEXO METODOLÓGICO

## LA ENCUESTA DEL DEUDA SOCIAL ARGENTINA DEL BICENTENARIO (2010-2016). INFORME 2007/2010-2011.

DIEGO QUARTULLI <sup>78</sup>

La Encuesta de la Deuda Social Argentina - Serie del Bicentenario (2010-2016) (EDSA – Bicentenario) aborda el estudio de las privaciones y realizaciones en materia de desarrollo humano de la sociedad argentina a través de una muestra multipropósito probabilística, conglomerada y estratificada no proporcional de hogares particulares urbanos. Dada el diseño de esta encuesta, se busca que la misma permita comparaciones temporales, tanto con respecto a la información brindada por la Encuesta de la Deuda Social Argentina - Serie EDSA 2004-2009, como con relación a los resultados que arroje la misma a lo largo de la nueva Serie 2010-2016.

Hasta el momento la EDSA en todas sus versiones fue aplicada en nueve oportunidades desde el año 2004 hasta la actualidad, aunque cabe destacar que a partir de 2007, y principalmente desde la edición 2010, se introdujeron una serie de importantes cambios metodológicos, los cuales dieron finalmente lugar a la EDSA Serie Bicentenario. Pero si los contenidos del cuestionario han sufrido modificaciones a lo largo de los años, la mayor parte de los ítems relevados se han mantenido casi sin cambios, en especial desde 2007, justamente con el objeto de permitir comparaciones en el tiempo.<sup>2</sup>

1 Se agradece la colaboración de Agustín Salvia, Isidro Adúriz, Julieta Vera y Pablo de Grande en la elaboración de algunos apartados de este anexo.

2 Los cuestionarios de la EDSA aplicados desde 2004 hasta la

En cuanto al marco muestral, si bien la EDSA Bicentenario cubre actualmente un total de 20 aglomerados urbanos, mantiene en su estructura muestral a siete de los ocho grandes aglomerados urbanos relevados por la serie EDSA 2007-2009.<sup>3</sup>

Es debido a estos resguardos metodológicos, más allá incluso de los cambios que también experimentó la EDSA-Bicentenario entre 2010 y 20011, que el presente informe ofrece información comparativa de manera consistente sobre la evolución de un conjunto amplio de indicadores sociales entre 2007, 2010 y 2011. Sin embargo, cabe señalar que para hacer posible estas comparaciones, los resultados correspondientes a 2007 son estimaciones que contienen adecuaciones estadísticas producto de la aplicación de un método de “empalme” entre series.

En este marco, el presente anexo metodológico habrá de dar cuenta de estas y otras cuestiones metodológicas relevantes con el objetivo de exponer los procedimien-

fecha pueden consultarse en <http://www.uca.edu.ar/index.php/site/index/es/universidad/investigacion/programa-observatorio-de-la-deuda-social-argentina/barometro-deuda-social/edsa/>

3 Para mayor información sobre el diseño muestral de la EDSA Serie 2007-2009, ver el anexo metodológico del informe del Barómetro de la Deuda Social Argentina No. 6 en ODSA (2010). En cuanto a la estrategia seguida para la ampliación de la muestra a partir de la EDSA Bicentenario, remitirse al anexo metodológico del informe del Barómetro de la Deuda Social Argentina, Serie Bicentenario (2010-2016), Año I, en Salvia, Adaszko, Donza, et al. (2011b).

**Figura AM1.1.**

**CANTIDADES DE PERSONAS Y HOGARES SEGÚN AGLOMERADO URBANO PARA EDSA 2007-2009 Y EDSA BICENTENARIO (2010-2016)**

Proyecciones a 2010. Censo de Viviendas, Hogares y Población 2001.

GRUPO	AGLOMERADOS	EDSA 2007-2009			EDSA BICENTENARIO (2010-2016)		
		POBLACIÓN TOTAL	POBLACIÓN HASTA 17 AÑOS	HOGARES	POBLACIÓN TOTAL	POBLACIÓN HASTA 17 AÑOS	HOGARES
CABA	CAPITAL FEDERAL				2.725.094	566.982	1.024.231
	CONURBANO NORTE				2.580.198	828.309	714.962
CONURBANO BONAERENSE (*)	CONURBANO OESTE				3.171.247	1.032.478	864.842
	CONURBANO SUR				3.484.064	1.138.612	956.177
	<b>TOTAL</b>				<b>9.235.509</b>	<b>2.999.399</b>	<b>2.535.981</b>
<b>TOTAL GRAN BUENOS AIRES</b>		11.960.603	3.566.381	3.560.212	11.960.603	3.566.381	3.560.212
OTRAS AEREAS METROPOLITANAS	GRAN ROSARIO	1.151.988	345.639	343.417	1.151.988	345.639	343.417
	GRAN CÓRDOBA	1.365.418	425.073	384.365	1.365.418	425.073	384.365
	GRAN SAN MIGUEL DE TUCUMÁN Y TAFÍ VIEJO				775.181	274.105	185.833
	GRAN MENDOZA	847.838	275.054	225.062	847.838	275.054	225.062
	MAR DEL PLATA				534.308	149.905	170.758
RESTO URBANO	GRAN SALTA	465.762	179.246	108.723	465.762	179.246	108.723
	GRAN PARANÁ				247.203	81.679	69.086
	GRAN RESISTENCIA	357.688	137.719	90.223	357.688	137.719	90.223
	GRAN SAN JUAN				423.645	146.657	105.527
	NEUQUÉN-PLOTIER-CIPOLETTI	293.540	105.172	82.492	293.540	105.172	82.492
	ZÁRATE				86.963	27.759	24.140
	LA RIOJA				144.020	55.151	34.003
	GOYA				67.179	25.773	17.013
	SAN RAFAEL				109.041	33.760	31.045
	COMODORO RIVADAVIA				134.438	46.861	37.306
USHUAIA Y RÍO GRANDE				97.139	39.026	27.016	
	BAHÍA BLANCA	301.572	76.433	105.342			
<b>TOTAL CIUDADES DEL INTERIOR</b>		<b>4.783.806</b>	<b>1.544.336</b>	<b>1.339.624</b>	<b>7.101.354</b>	<b>2.348.579</b>	<b>1.936.009</b>
<b>TOTAL GENERAL</b>		<b>11.960.603</b>	<b>5.110.717</b>	<b>4.899.836</b>	<b>19.061.954</b>	<b>5.914.960</b>	<b>5.496.221</b>

tos que apoyan los resultados de este estudio. Los contenidos centrales del mismo serán el diseño muestral, la selección de los casos, la elaboración de ponderadores, las adecuaciones estadísticas para la comparación temporal, la definición de las variables de clasificación, la estimación de ingresos y horas trabajadas no declaradas y el margen de error muestral de las estimaciones.

**AM1. DISEÑO MUESTRAL**

El objetivo de la EDSA Serie Bicentenario fue poder contar, a partir de 2010, con una muestra representativa de gran parte de los hogares de la Argenti-

na urbana. Por esta razón la identificación de los hogares incluidos en la muestra se apoyó en un diseño muestral polietápico que primero conglo-mera y luego estratifica a la población objetivo.

En la primera etapa se tomaron en cuenta criterios demográficos a partir de la definición en primer lugar de un conjunto de aglomerados a incluir (por región y tamaño), y en una segunda etapa como modo de estratificar al interior de cada conglomerado muestral se consideraron criterios socio-educativos clasificando a los hogares del aglomerado según la tasa de educación de los jefes de hogar del radio censal registrado en el último censo nacional de vivienda, hogares y población disponible con desagregación por radio (INDEC, 2001).

**Figura AM1.2**

**PORCENTAJES DE HOGARES SEGÚN ESTRATIFICACIÓN MUESTRAL PARA ÁREAS METROPOLITANAS Y CENTROS URBANOS DE MÁS DE 200.000 HABITANTES DE LA EDSA 2007-2009 Y EDSA BICENTENARIO (2010-2016).**

PROYECCIONES A 2010. CENSO DE VIVIENDAS, HOGARES Y POBLACIÓN 2001.		
ESTRATOS MUESTRALES	HOGARES DEL MARCO MUESTRAL (POR CADA AGLOMERADO), ORDENADOS POR NIVEL EDUCATIVO DEL JEFE DE HOGAR	
	2007-2009	EDSA BICENTENARIO
1 - MUY BAJO	0% A 11%	12.5% (1/8)
2 - BAJO	2% A 27%	25% (1/4)
3 - MEDIO	28% A 46%	25% (1/4)
4 - MEDIO ALTO	47% A 64%	25% (1/4)
5 - ALTO	65% A 100%	12.5% (1/8)

**Figura AM1.3**

**PORCENTAJE DE HOGARES SEGÚN ESTRATO MUESTRAL PARA AGLOMERADOS DE MENOS DE 200.000 HABITANTES DE LA EDSA BICENTENARIO 2010-2011**

ESTRATOS MUESTRALES	HOGARES DEL MARCO MUESTRAL (POR CADA AGLOMERADO)
1 - BAJO	33.3% (1/3)
2 - MEDIO	33.3% (1/3)
3 - ALTO	33.3% (1/3)

La aplicación de estos procedimientos buscó generar una reducción del margen de error en cada estrato permitiendo comparaciones más precisas entre ellos. Por otra parte, al ser un muestreo estratificado no proporcional, esto exigió que posteriormente fuese necesario introducir correcciones a través de un sistema de ponderación.

Los centros urbanos que fueron considerados en la selección de hogares de la EDSA Bicentenario, siguiendo el criterio demográfico son los que se observan en la figura AM1.1.

Como se observa en la figura AM1 existen algunos cambios entre los diseños muestrales de la EDSA 2007-2009 y la EDA Bicentenario aunque todos, a excepción del caso de Bahía Blanca, permiten mantener la comparabilidad en el tiempo y aumentar el dominio empírico de la muestra.

Con el objeto de hacer más eficiente la selección de los casos, el diseño muestral de la EDSA Bicentenario aplicó, al igual que en la EDSA serie 2007-2009, un criterio de estratificación de los radios censales a partir

del nivel educativo de los jefes de hogar. La información censal permitió ordenar los radios muestrales de cada aglomerado según el porcentaje de hogares con jefe con secundario completo, y posteriormente, agrupar a los mismos según esta variable.

A partir de este ordenamiento, se formaron 5 grupos poblacionales en el caso de los aglomerados con al menos 200 mil habitantes, y 3 grupos poblacionales en el caso de los aglomerados menores a ese tamaño. Para mejorar la captación de los casos extremos (de mayor y menor nivel socioeducativo), en el caso de los grandes aglomerados se generaron 3 grupos centrales con el 1/4 de los casos cada uno y 2 grupos en los extremos con 1/8 cada uno de la población total. De esta forma, los estratos quedaron conformados según el tipo de aglomerado como se detalla en las figuras AM 1.2 y AM 1.3

Para el caso particular de la EDSA Bicentenario, una vez segmentado el marco muestral de cada aglomerado según los criterios de estratificación descritos, se procedió a establecer la cantidad de radios / puntos muestra correspondientes a cada uno. Dado que por factores de representatividad se estableció un tamaño de muestra de 5.712 hogares, con una asignación esperada de 6 hogares por punto muestra, quedaron determinados para la primera etapa del proceso de selección un total de 952 radios, los cuales fueron distribuidos entre los aglomerados siguiendo un criterio de no proporcionalidad. El número de radios asignados a cada aglomerado dependió de la manea en que se determinaron los dominios de representatividad estadística y de la necesidad de poder predicar con sobre dominios dependiendo del número de hogares esperados en cada caso. Las cantidades de radios / puntos muestra y de hogares asignados a cada aglomerado urbano fueron los que se detallan en la figura AM 1.4.

En la figura AM1.4 se observa de una manera más precisa lo comentado en la presentación de la figura AM1.1, acerca de las diferencias entre la EDSA 2007-2009 y la EDSA Bicentenario, al observarse como se distribuyeron los hogares y los puntos muestras dentro de los aglomerados comparables y los no comparables. De esta manera se aprecia cómo en la EDSA Bicentenario no sólo aumentaron la cantidad de aglomerados sino que también en el caso de los aglomerados comparables también se aumentó la cantidad de hogares y puntos muestras permitiendo estimaciones, no sólo con un mayor dominio empírico sino también con un menor margen de error.

**Figura AM1.4**

**CANTIDAD DE CASOS DE HOGARES Y PUNTOS MUESTRALES SEGÚN AGLOMERADO URBANO**

Valores para EDSA 2007-2009 entre paréntesis

GRUPO	AGLOMERADOS	EDSA 2007-2009		EDSA BICENTENARIO (2010-2016)	
		HOGARES	PUNTOS MUESTRA	HOGARES	PUNTOS MUESTRA
CABA	CAPITAL FEDERAL			444	74
	CONURBANO NORTE			432	72
CONURBANO BONAERENSE (*)	CONURBANO OESTE			432	72
	CONURBANO SUR			432	72
	<b>TOTAL</b>			<b>1296</b>	<b>216</b>
<b>TOTAL GRAN BUENOS AIRES</b>		750	125	1740	290
OTRAS AEREAS METROPOLITANAS	GRAN ROSARIO	390	65	624	104
	GRAN CÓRDOBA	330	55	624	104
	GRAN SAN MIGUEL DE TUCUMÁN Y TAFÍ VIEJO			624	104
	GRAN MENDOZA	390	65	624	104
	MAR DEL PLATA			192	32
	GRAN SALTA	90	15	192	32
	GRAN PARANÁ			192	32
	GRAN RESISTENCIA	60	10	192	32
	GRAN SAN JUAN			192	32
	NEUQUÉN-LOTIER-CIPOLETTI	60	10	192	32
RESTO URBANO	ZÁRATE			54	9
	LA RIOJA			54	9
	GOYA			54	9
	SAN RAFAEL			54	9
	COMODORO RIVADAVIA			54	9
	USHUAIA Y RÍO GRANDE			54	9
	BAHÍA BLANCA	60	10		
	<b>TOTAL CIUDADES DEL INTERIOR</b>		<b>1380</b>	<b>230</b>	<b>3972</b>
<b>TOTAL GENERAL</b>		<b>2130</b>	<b>355</b>	<b>5712</b>	<b>952</b>

**AM2 SELECCIÓN DE LOS CASOS DE LA MUESTRA**

Definida la cantidad de puntos muestrales a seleccionar en cada aglomerado-estrato, se tomó la lista de radios correspondientes con esta estratificación y se realizó un sorteo con reposición de puntos muestras. En esta lista los radios fueron ponderados según la cantidad de hogares registrada en el censo, para así mantener la misma probabilidad de selección de cada hogar.

Por último se generó una lista de números aleatorios con un N igual a la cantidad de puntos muestras que se quiera seleccionar y una vez seleccionados estos se ve-

rificó a qué radio muestral pertenecía. De esta forma se procedió con todos los estratos y aglomerados hasta completar la lista de puntos muestrales, uniendo a cada punto muestral requerido con un código de radio censal. Luego, a partir de contar con la cartografía correspondiente se hizo una selección aleatoria de punto muestra dentro de cada radio muestral seleccionado.

Para el caso de la EDSA Bicentenario 2011, a diferencia de la edición de 2007, no se volvió a sortear todos los puntos muestra cómo se describió más arriba, sino que se optó, como primera instancia, conservar los mismos punto muestra de la EDSA 2007-2009 y la EDSA Bicentenario 2010, y re-encuestar a los mismos hogares e individuos entrevistados en 2010

como parte de una estrategia que permitiera, por un lado, captar datos para futuros estudios longitudinales (casos panel), y por otro, reducir considerablemente los gastos de logística<sup>4</sup>. Como consecuencia de lo anterior, aproximadamente el 30% de los casos de 2011 fueron casos panel.

En caso de que la reencuesta no se hubiera podido llevar a cabo, sea por imposibilidad de encontrar al hogar y/o al mismo encuestado (después de un intento y dos revisitas), el caso fue reemplazado por otro hogar del mismo punto muestra y con las mismas características de sexo y edad del entrevistado, mediante un muestreo sistemático.

Por último cuando el punto muestral anterior no hubiera coincidido con los datos de la estratificación muestral, por ejemplo al cambiar de manera rotunda la trazabilidad urbana, el punto muestra fue reemplazado por otro con las mismas características “esperadas” del anterior, mediante un sorteo aleatorio dentro del conglomerado correspondiente.

## SELECCIÓN DE HOGARES

---

### *Ciudad de Buenos Aires*

Para la selección de los hogares de la Ciudad de Buenos Aires, en la EDSA Bicentenario se tuvo en consideración dos aspectos diferenciales. Por un lado la alta tasa de rechazo de las entrevistas por “timbreo” especialmente a medida que se asciende de clase social y especialmente en los hogares cuya vivienda es parte de un edificio. Por otro lado la alta penetración del servicio telefónico “fijo” exceptuando los hogares con residencia en áreas marginales.

Estos dos supuestos permitieron que salvo en el estrato muestral más bajo la selección de los hogares se hiciera sobre una base de datos telefónica con el punto muestra censal geo-referenciado. De esta forma para los puntos muestras seleccionados de esos estratos se pudo obtener un padrón de teléfonos, lo que permitió luego efectuar un “timbreo” telefónico de los hogares de cada punto muestra en forma aleatoria a partir del padrón. Posteriormente a los hogares seleccionados se les aplicó el cuestionario en forma telefónica.

---

4 En la edición EDSA Bicentenario 2010 tampoco se volvieron a sortear todos los puntos muestra del marco muestral sino sólo aquellos *nuevos* puntos muestra provenientes de nuevos aglomerados o de cambios ad-hoc del marco muestral propio de la evolución de la sociedad.

En el caso del estrato más bajo se siguió el mismo procedimiento de visita domiciliaria que se utilizó en el resto de los aglomerados urbanos.

### *Resto de los aglomerados urbanos*

En los demás aglomerados del país para EDSA Bicentenario y en todos los aglomerados de la EDSA 2007-2009, por cada punto muestral seleccionado se elaboró un mapa ubicando geográficamente el radio censal en la cartografía del aglomerado. En el mapa se identificaron las manzanas del radio y se seleccionó una como punto de inicio en forma aleatoria. De la misma, a su vez, se seleccionó una esquina desde la cual comenzar el rastreo de hogares. Cuando el radio no presentaba manzanas identificables, se aplicó el mismo procedimiento segmentando el mismo en espacios regulares equivalentes a una manzana.

Una vez indicada la esquina desde la cual comenzar el trabajo de campo, el encuestador visitó el punto muestral realizando un ‘timbreo’ de los hogares de la manzana y registrando en la hoja de ruta los domicilios visitados. En caso de agotar la manzana sin encontrar los casos requeridos, el encuestador continuó con las 8 manzanas adyacentes a la manzana indicada.

## SELECCIÓN DE RESPONDENTES

---

Tras la selección del hogar, la selección del individuo a encuestar (“el respondiente”) se realizó con un criterio de identificación de cuotas de edad y sexo para controlar estas variables en la captación de la información y obtener en la muestra estimadores cercanos a los parámetros poblacionales<sup>5</sup>. Para ello, cada encuestador dispuso de un set aleatorio de 6 casos a relevar por cada punto muestral asignado, indicando para cada uno de ellos el grupo de edad y sexo esperados.

Los hogares en los que el encuestador encontraba no habitaran adultos que se correspondan con ninguna de las cuotas solicitadas, fueron descartados en el campo, continuándose el timbreo desde el siguiente hogar. Si el hogar por el contrario contaba con casos correspondientes con las cuotas pero di-

---

5 Es conocido que en todos los casos resulta dispar la tasa de no-respuesta (fundamentalmente por ausencia) de algunos subgrupos de población que deviene en su sub o sobre-representación (varones-jóvenes-activos es el caso más paradigmático, por mencionar un ejemplo). Las cuotas y las revisitas se inscriben como procedimientos idóneos para reducir estos sesgos.

chas personas no se encontraban en el hogar en ese momento, se realizaba un máximo de tres visitas (2 revisitas al hogar) buscando encontrar a la persona identificada para realizar la encuesta.

---

### AM3. ELABORACIÓN DE PONDERADORES

---

Como se indicó antes (AM1 y AM2), debido a que se seleccionó un diseño muestral estratificado pero con una conglomeración no proporcional, debe ponderarse y expandirse para devolverle, a cada hogar efectivamente seleccionado, la probabilidad real de haber sido seleccionado y de esa manera corregir sesgos de las estimaciones muestrales frente a los parámetros poblacionales de la población objetivo. Para conseguir este objetivo se procedió a elaborar un coeficiente  $n/N$  a nivel de cada radio seleccionado (en donde “n” son los hogares efectivamente seleccionados en la muestra y “N” son los hogares esperados según los parámetros poblacionales) y luego expandirlo al total de hogares del aglomerado, a los fines de establecer la corrección  $n/N$  para el total de la muestra.

Si una muestra estuviera libre de sesgos de selección, el proceso de ponderación a nivel de hogares habría finalizado en este punto, ya que, a esta altura, tendríamos la certeza de que cada hogar, en el total de la muestra, asume el peso que le corresponde de acuerdo a su probabilidad de selección.

Independientemente de lo anterior también se efectuaron arreglos bajo la suposición de que es imposible contar con una distribución libre de sesgos específicamente en lo referido al sesgo de “no-respuesta”. Si bien los diseños muestrales y los trabajos de campo prevén estrategias para disminuir las tasas de no-respuesta (revisitas, cuotas, concertaciones), los sesgos logran atenuarse pero no corregirse en su totalidad.

Como consecuencia de lo anterior se corrigieron o calibraron los pesos o factores de expansión iniciales (inversa de la probabilidad de selección) con la ayuda de información auxiliar conocida o preestablecida a partir de registros o fuentes externas consideradas como válidas como los datos censales e información *ad-hoc* proveniente del campo de la EDSA Bicentenario 2010.

Esta segunda corrección atiende a considerar las diferencias entre la muestra observada y la esperada de acuerdo con los atributos de los hogares y/o las personas que componen los hogares seleccionados. Para

ello se utilizó el procedimiento de “calibración por marginales fijos” (Neville y Sarndall, 1992) que estima las frecuencias “condicionales” de una tabla de contingencia según los parámetros poblacionales conocidos.

En el caso de la base de individuos componentes del hogar (que ya trae consigo los primeros factores de ponderación) se efectúa una segunda calibración que involucra la distribución por sexo, grupos de edad (0-4, 5-17, 18-29, 30-49, 50 y más) a nivel de cada aglomerado. En el caso de la base de respondentes, la segunda calibración tomó en cuenta, además de la distribución por sexo, grupos de edad, la condición de actividad de los respondentes, descendiendo a nivel de cada estrato-aglomerado que integra la muestra.

Finalmente, la base hogares fue la última en recibir su ponderador final, ya que absorbe las calibraciones previas efectuadas sobre sus componentes. Este último procedimiento, que traslada las correcciones realizadas a nivel de componentes hacia la base de hogares buscó evitar que los hogares compuestos por segmentos sobre o subrepresentados generen distorsiones en los resultados finales a nivel de indicadores relativos a los hogares. Esto es importante, ya que dos hogares que a priori cuentan con la misma probabilidad de selección, solapan un posible sesgo de acuerdo con su composición y, es probable que cuenten con diferencias significativas en una buena parte de sus características y comportamientos; por lo tanto, de no mediar una calibración, se llegaría a resultados posiblemente sesgados.

En forma complementaria, al trabajarse sobre diversas unidades de análisis (hogares, personas, niños, etc.) y el diseño muestral estar hecho sobre hogares es recomendable la construcción de un ponderador o expansor específico para cada unidad de análisis utilizada. En este caso, tanto para 2010 como para 2011 se hicieron tanto expansores (que incluyen una ponderación en su interior) como ponderadores (que sólo “corrigen o ajustan” la cantidad de casos muestrales a la proporción de casos esperados en la población objetivo sin expandir) referidos al total de la población objetivo.

En ambos casos se realizó una expansión simple de los casos, corrigiendo posibles desvíos en las cuotas y asegurando que el estrato muestral, la edad y el sexo de la muestra seleccionada expresara *cantidades* y/o proporciones que se correspondieran con los parámetros conocidos del total de personas y hogares de cada aglomerado.

---

#### **AM4. ADECUACIONES ESTADÍSTICAS PARA LA COMPARABILIDAD TEMPORAL**

---

En pos de conseguir una comparación más genuina entre las diferentes muestras se aplicaron coeficientes que permitieran reducir la discrepancia cuyo origen fuera la falta de comparabilidad de aglomerados y la falta de comparabilidad de estratos muestrales.

Una alternativa para lograr la comparabilidad de los aglomerados hubiera sido reducir el análisis a los aglomerados comparables, esto es, utilizar sólo los aglomerados que se repiten en las ediciones analizadas. Esta estrategia, si bien es correcta desde el punto de vista metodológico, habría tenido como defecto que en casos como el de la EDSA no poder aprovechar la nueva información proveniente de los nuevos aglomerados a partir de 2010.

Por el lado de la comparabilidad de los estratos muestrales el problema a resolver es más complejo. Al momento de revisar el diseño muestral para el campo de la EDSA Bicentenario 2011 todavía se contaba con el marco muestral construido con los parámetros poblacionales del censo de 2001. Luego de más de 10 años de evolución social, estudios específicos demostraron el crecimiento de determinadas franjas poblacionales especialmente de villas y asentamientos urbanos más allá de lo observado en el censo de 2001 (Cravino et al., 2008, Falcón y Raffo, 2011).

Complementariamente a las publicaciones específicas en la actualización del marco muestral también influyó la experiencia adquirida en la salida a campo de la serie 2010, especialmente en lo referente a los puntos muestras excluidos debido a no coincidir con lo esperado según el marco muestral diseñado originalmente con los datos censales de 2001. De esta manera es de suponer que en la próxima edición de la EDSA la modificación del marco muestral sería menor a la ahora realizada debido a que en 2010 el marco muestral se enfrentó a una realidad social con 10 años de evolución y en 2011 se encontrará sólo con la evolución social de un sólo año, gracias a la corrección basada en datos secundarios específicos y en la experiencia de primera mano obtenida en el campo de la serie 2010 de la EDSA Bicentenario.

El problema inmediatamente posterior a construir un ponderador más actualizado para 2011<sup>6</sup> fue replicar el mismo para los datos 2010, bajo el supuesto de que la mayoría de los cambios observados en el marco muestral se efectuaron en el período 2001-2011 y no sólo en 2011, con la diferencia que en 2010 la muestra ya fue realizada y no es posible reasignar casos, con lo que la opción más razonable fue corregir el ponderador, cambiando de este modo los datos publicados anteriormente sobre 2010. Mantener los datos publicados de la serie 2010 hubiera representado una pérdida de oportunidad de corregir algunas estimaciones a luz de nueva información<sup>7</sup>.

El siguiente problema a resolver era la comparabilidad con los datos de 2007 y la estrategia que se siguió fue la siguiente. Si bien se amplió el marco muestral en 2010, los principales cambios de la muestra provinieron de la agregación de nuevos aglomerados, manteniendo constantes los ya utilizados con anterioridad, a excepción del aglomerado Bahía Blanca. Esto último permitió cotejar los datos obtenidos para 2007 (con el ponderador de 2007) con los datos de 2010 (con el nuevo ponderador) siempre para los aglomerados comparables. Luego, mediante un coeficiente de empalme que se calculó realizando una razón entre los valores cotejados para los aglomerados comparables, se realizó una estimación para los valores de 2007, multiplicando dicho coeficiente con los valores obtenidos para 2010 (con el nuevo ponderador) permitiendo interpretar los datos al total muestral de 2010-2011 y no sólo a los aglomerados comparables.

---

#### **AM5. DEFINICIÓN DE LAS VARIABLES DE CLASIFICACIÓN UTILIZADAS**

---

Los indicadores de desarrollo humano y social fueron analizados de acuerdo a su distribución según ciertas variables de corte, seleccionadas debido a su carácter estructural y a su importancia en la capacidad explicativa y/o predictiva de los resultados.

---

<sup>6</sup> Los criterios generales que se tuvieron en cuenta en la construcción del ponderador de 2011 se encuentra en el apartado metodológico N° 3 del presente anexo.

<sup>7</sup> Esta práctica es usual en los organismos de estadísticas de población y en general a la publicación de una nueva información, por ejemplo un censo nacional, le sigue una rectificación tanto de estimadores retrospectivos como prospectivos.

Esquemáticamente se pueden subdividir entre aquellas variables que refieren a propiedades de los aglomerados, de los hogares y aquellos que lo hacen a propiedades de los individuos componentes del hogar. En el análisis de los datos expuestos en el Informe sobresalen los siguientes criterios de clasificación:

- » Para el caso de los individuos se destacan los análisis en base al sexo, la edad agrupada en grupos etáreos (18-29; 30-49; 50 y +) y el nivel educativo dicotomizado en dos niveles educativos. Ver figura AM5.a.
- » Para el caso de las propiedades de los hogares el informe privilegió los análisis en base a la condición socio-residencial y al estrato socio-económico. En el primer caso el criterio socio-residencial da cuenta de formas diferenciales de habitar y de vivir en el área urbana. Se categorizó en tres formas distintas de urbanización con diferente grado de formalidad y de acceso a recursos de infraestructura y de servicios (ver más detalle en el siguiente cuadro. En el segundo caso la construcción es más compleja y se explicita a continuación. Ver figura AM5.b.
- » Por último también se utilizó un criterio geo-demográfico, que clasificó los casos según el tipo de aglomerado urbano donde estos fueron relevados. En particular, se subdividen los resultados entre CABA, Conurbano Bonaerense, Otras áreas metropolitanas y Resto urbano. Ver figura AM5.c

## ESTRATO SOCIO-ECONÓMICO

Como se aclaró anteriormente dentro de las propiedades de los hogares utilizadas en el informe se destaca el concepto de estrato socioeconómico del hogar encuestado. Aquí tendiendo como objetivo construir un *indicador* de corte con un alto poder discriminante con el menor número de variables en su interior se aplicó una técnica de “reducción de dimensiones” para variables categóricas. Específicamente se aplicó en primer lugar un Análisis de Componentes Principales para Variables Categóricas (CATPCA) que posibilitó la optimización de puntajes óptimos que luego sirvió de insumo para realizar un análisis factorial con extracción de un factor. Este indicador se utilizó tanto en la EDSA 2007-2009 como en la EDSA Bicentenario.

La noción fundamental detrás de esta construcción es que los distintos aspectos *observables* a través de las variables de la EDSA son diferentes dimensiones de un mismo concepto subyacente que no es directamente mensurable

**Figura AM 5.a**

### NOMBRE, DESCRIPCIÓN Y CATEGORÍAS DE LAS VARIABLES REFERIDAS A PERSONAS

Variable	Descripción	Categorías
SEXO	SE REFIERE AL SEXO DE LOS MIEMBROS DEL HOGAR	- VARÓN - MUJER
EDAD	SE REFIERE A LOS GRUPOS DE EDAD DE LOS MIEMBROS DE LOS MIEMBROS DEL HOGAR	- 0 A 5 AÑOS, - 6 A 12 AÑOS, - 13 A 17 AÑOS, - 18 A 34 AÑOS, - 35 A 59 AÑOS, - 60 AÑOS Y MÁS.
NIVEL EDUCATIVO	SE REFIERE A LA EDUCACIÓN DE LOS MIEMBROS DEL HOGAR	- SECUNDARIO INCOMPLETO - SECUNDARIO COMPLETO

**Figura AM 5.b**

### NOMBRE, DESCRIPCIÓN Y CATEGORÍAS DE LAS VARIABLES REFERIDAS A CARACTERÍSTICAS DEL HOGAR

Variable	Descripción	Categorías
ESTRATO SOCIO-ECONÓMICO	SE PROCEDE A REALIZAR UNA CLASIFICACIÓN EN CUARTILES TOMANDO EN CONSIDERACIÓN CRITERIOS EL NIVEL EDUCATIVO, LA CONDICIÓN OCUPACIONAL, EL ACCESO A BIENES Y TECNOLOGÍA Y CARACTERÍSTICAS DE LA VIVIENDA DEL HOGAR. LA CONSTRUCCIÓN DE ESTA VARIABLE SE DESCRIBE CON DETALLE EN EL AM5.	- MUY BAJO - 1º CUARTIL - BAJO - 2º CUARTIL - MEDIO BAJO - 3º CUARTIL - MEDIO ALTO - 4º CUARTIL
CONDICIÓN RESIDENCIAL	REPRESENTA TRES MODALIDADES DIFERENTES DE URBANIZACIÓN CON GRADOS DIVERSOS DE PRESENCIA DEL ESTADO EN LO QUE HACE A LA PLANIFICACIÓN, LA REGULACIÓN Y LA INVERSIÓN PÚBLICA EN BIENES URBANOS.	- VILLAS O ASENTAMIENTOS PRECARIOS - TRAZADO URBANO DE NIVEL SOCIOECONÓMICO BAJO - TRAZADO URBANO DE NIVEL SOCIOECONÓMICO MEDIO
SEXO DEL JEFE	SE REFIERE AL SEXO DEL JEFE DEL HOGAR	- VARÓN - MUJER
NIVEL EDUCATIVO	SE REFIERE A LA EDUCACIÓN DEL JEFE DE HOGAR	- SIN SECUNDARIO COMPLETO - CON SECUNDARIO COMPLETO
NIÑOS EN EL HOGAR	SE REFIERE A LA PRESENCIA DE INDIVIDUOS DE MENOS DE 18 AÑOS EN EL HOGAR	- SIN NIÑOS EN EL HOGAR - CON NIÑOS EN EL HOGAR

y que denominamos “Estrato socio-económico” que sería el “factor” común a todas las dimensiones.

Las variables *observables* que se tomaron en consideración, luego de un proceso de selección en donde



**Figura AM 5.c**

**NOMBRE, DESCRIPCIÓN Y CATEGORÍAS DE LAS VARIABLES REFERIDAS A CARACTERÍSTICAS DEL OS AGLOMERADOS**

<b>CONGLOMERADO URBANO</b>	SE CONSIDERAN CUATRO GRUPOS DE AGLOMERADOS PRINCIPALES, DE LOS CUALES LA EDSA ES REPRESENTATIVA.	- CIUDAD AUTÓNOMA DE BUENOS AIRES,
		- CONURBANO BONAERENSE (NORTE, OESTE Y SUR)
		- OTRAS ÁREAS METROPOLITANAS (GRAN CÓRDOBA, GRAN ROSARIO, GRAN MENDOZA, GRAN TUCUMAN)
		- RESTO URBANO (MAR DEL PLATA, GRAN SALTA, GRAN PARANÁ, GRAN RESISTENCIA, GRAN SAN JUAN, NEUQUÉN - PLOTTIER - CIPOLETTI, ZÁRATE, LA RIOJA, GOYA, SAN RAFAEL, COMODORO RIVADAVIA Y USHUAIA).

se analizaron las correlaciones e intercorrelaciones entre un conjunto mayor de posibles candidatas, fueron las siguientes:

- » Para “Acceso a bienes y tecnologías” se tomaron en consideración a distintas variables como tener televisión con cable, computadora con internet, auto, microondas, tarjeta de crédito, etc. Previamente estos bienes se combinaron para formar un “score”, logrado como la suma simple estandarizada (se estandarizó dividiendo el puntaje obtenido por el número de bienes sumados) de bienes que este posee. De esta manera, los hogares que poseen todos los bienes tiene un puntaje de uno (1), mientras que aquellos que no poseen ninguno se les asigna un puntaje de cero (0).
- » Para “Clima educativo del hogar” se tomó en consideración la educación lograda por el jefe de hogar operacionalizada a través de los años de educación.
- » Para “Condición ocupacional del hogar” se tomó en consideración la ocupación del principal sostén del hogar y la tasa de empleo del hogar. Esta última se construyó dividiendo la cantidad de aportantes sobre la cantidad de individuos del hogar.
- » Para “Nivel socio-residencial” se tomó en consideración propiedades más “estructurales” que hacen al bienestar de sus ocupantes y que se consideran menos insensibles a cambio económicos como si el barrio posee trazado urbano o no.

Por lo dicho anteriormente acerca del funcionamiento de la técnica factorial, la posición social de cada hogar resultó una función autoponderada de los valores de las variables observadas seleccionadas.

Una vez que se calculó para cada hogar su respectivo nivel de estratificación socio-económica se procedió a su clasificación en cuartiles ordenando a los hogares según niveles crecientes de capital educativo, acceso a bienes y tecnologías de información y comunicación y capital ocupacional incluyendo tanto el capital del hogar como el capital del entorno residencial donde habitan los individuos y hogares que forman parte del estudio, obteniendo las siguientes categorías: 1) Estrato Muy Bajo; 2) Estrato Bajo; 3) Estrato Medio Bajo; y 4) Estrato Medio Alto. Cabe destacar que todo el procedimiento se realizó con los datos ponderados<sup>8</sup> debido a que una ponderación posterior desequilibraría una estructura de cuartiles previa.

**AM6. IMPUTACIÓN DE INGRESOS Y HORAS TRABAJADAS NO DECLARADAS**

Una etapa previa a la construcción de indicadores de privación y su análisis involucró la consistencia de los datos obtenidos por el relevamiento de la EDSA 2007-2009 y la EDSA Bicentenario. En primer lugar, se llevaron a cabo controles con el propósito de detectar inconsistencias en las respuestas de los individuos, especialmente en lo referido a los bloques temáticos del cuestionario.

Un problema frecuente de las encuestas de hogares que requirió especial consideración fue el tratamiento de las respuestas de ingreso, tanto del ingreso laboral como del ingreso del hogar. Es usual en este tipo de encuestas que no todos los individuos entrevistados respondan a las preguntas de ingresos. Este fenómeno puede sesgar las estimaciones de desigualdad si, por un lado la no respuesta depende del ingreso, y por otro, si el porcentaje de no respuesta varía en el tiempo (Salvia y Donza, 1999) (Gasparini y Sosa Escudero, 2001). En el caso de la EDSA, se confirmó una relación directa y significativa entre el estrato socio-económico y la no respuesta. Ver Figura AM6.a.

<sup>8</sup> Ver apartado metodológico n°3 de este anexo metodológico para detalles específicos sobre la construcción de los ponderadores.

**Figura AM 6.a****NO DECLARANTES DE INGRESOS Y DE HORAS SEMANALES TRABAJADAS SEGÚN ESTRATO SOCIO-ECONÓMICO**

EDSA 2007-2009 y EDSA Bicentenario (2010-2016). En porcentajes.

ESTRATO SOCIO ECONÓMICO	HOGARES QUE NO DECLARARON EL TOTAL DE INGRESOS (EN PORCENTAJES)		
	2007	2010	2011
MUY BAJO	25,50%	17,40%	11,70%
BAJO	33,50%	20,60%	14,50%
MEDIO BAJO	32,60%	24,00%	14,70%
MEDIO ALTO	41,80%	34,00%	23,90%
<b>TOTAL</b>	<b>33,60%</b>	<b>24,00%</b>	<b>16,20%</b>

FUENTE: EDSA 2007-2009 Y EDSA BICENTENARIO (2010-2016). OBSERVATORIO DE LA DEUDA SOCIAL ARGENTINA. UCA. NOTA: LOS INGRESOS DE LOS AÑOS 2007 Y 2010 FUERON DEFLACIONADOS APLICANDO UN ÍNDICE DE PRECIOS ALTERNATIVO AL OFICIAL PROVENIENTE DE CENTROS Y EQUIPOS DE INVESTIGACIÓN (IPC 7 PROVINCIAS CENDA/IPC).

Para poder resolver el problema de no respuestas se realizó la estimación de un modelo de regresión que permitió efectuar la imputación de ingresos a los no respondientes a partir de los ingresos de las personas en condiciones laborales, demográficas y socioeconómicas similares (Salvia y Donza, 1999).

Siguiendo este método, se realizaron dos estimaciones separadas, una para el ingreso laboral del individuo adulto seleccionado –en caso de estar ocupado- y otra para el ingreso laboral total del hogar. En ambos casos las estimaciones fueron realizadas por separado para cada uno de los estratos socioeconómicos subdividiendo así la población. Es decir, en base al modelo de predicción considerado se procedió a efectuar estimaciones para cada uno de los estratos socioeconómicos con el objeto de mejorar la precisión de las estimaciones de ingresos. En cualquier caso, se ajustó el nivel de ingresos del hogar y laboral si el valor estimado quedó fuera del rango (no el monto) de ingresos declarado por el respondiente.

- » El análisis de regresión para el ingreso laboral tomó en cuenta variables demográficas (*sexo, grupos de edad*), socioeconómicas (*nivel educativo, situación ocupacional, ocupación principal, jefatura de hogar, cantidad de horas semanales trabajadas, rango de ingreso laboral, etc.*), psicológica (*capacidad de pensar proyectos, déficit de creencias de control, déficit de conformidad con las propias capacidades, salud mental*).
- » El análisis en el caso de la estimación para el ingreso del hogar incluyó otras variables como *condición residencial, situación de hacinamiento, tipo de*

*hogar, ciclo vital, rango de ingreso del hogar, características del jefe (sexo, edad, nivel educativo, condiciones ocupacionales), acceso a servicios básicos e infraestructura urbana (agua corriente, red de gas, red de cloacas, calles pavimentadas) además de considerar la recepción de asistencia en forma monetaria o no monetaria por parte de organismos públicos y privados. Ver Figura AM6.b*

- » Adicionalmente, se realizó la estimación de un modelo de regresión que permitió efectuar la imputación de la cantidad de horas semanales trabajadas a los ocupados no respondientes de esta información. Este análisis también incluyó variables demográficas (*sexo y edad*), socioeconómicas (*nivel educativo, situación ocupacional, ocupación principal, jefatura de hogar, ingreso laboral, etc.*) y psicológicas (*capacidad de pensar proyectos, déficit de creencias de control, déficit de conformidad con las propias capacidades, salud mental*). Del mismo modo que en la imputación de ingresos laborales e ingresos del hogar no declarados, la estimación de horas trabajadas no declaradas fue realizada para cada uno de los estratos socioeconómicos.

**AM7. ESTIMACIÓN DE ERRORES.**

En la EDSA Bicentenario (2010-2016), al ser una muestra multipropósito no hay una sola variable a utilizar y por lo tanto no hay un sólo margen de error a calcular. Por otro lado al realizarse sobre un diseño de una muestra compleja, cada análisis posee un tipo específico de margen de error. En el caso particular de la EDSA, al ser una muestra polietápica hay “factores de diseño” como la conglomeración y la estratificación que aumentan o reducen el margen de error del esperado bajo un muestreo aleatorio simple.

Teniendo en cuenta lo anterior, por una cuestión de espacio y comunicacional, sólo se seleccionaron tres variables que se consideraron las más relevantes en los análisis del Barómetro como lo son las variables de Estrato socio-económico, Condición socio-residencial y Aglomerado Urbano.

Por la misma razón para los resultados que siguen se ha supuesto un factor de diseño igual a 1, entendiendo por esto la cancelación de los efectos negativos de la conglomeración y los positivos de la estratificación sobre el margen de error muestral total aunque estos se

**Figura AM 6.b**

**MEDIA DE INGRESOS Y DE HORAS SEMANALES TRABAJADAS  
-CON Y SIN ESTIMACIÓN- SEGÚN ESTRATO SOCIO-ECONÓMICO**

EDSA 2007-2009 y EDSA Bicentenario (2010-2016).

ESTRATO SOCIO-ECONÓMICO	MEDIA DE INGRESOS DE LOS HOGARES (EN PESOS CONSTANTE A DICIEMBRE DE 2011)					
	2007		2010		2011	
	SIN ESTIMACIÓN	INCLUYE ESTIMACIÓN	SIN ESTIMACIÓN	INCLUYE ESTIMACIÓN	SIN ESTIMACIÓN	INCLUYE ESTIMACIÓN
MUY BAJO	2426	2408	2117	2543	2464	2735
BAJO	3239	3490	3105	3221	3418	3532
MEDIO BAJO	3902	4246	4255	4353	4733	4798
MEDIO ALTO	5741	6440	6781	6927	8100	8005
<b>TOTAL</b>	<b>3778</b>	<b>4212</b>	<b>3817</b>	<b>4314</b>	<b>4592</b>	<b>4768</b>

FUENTE: EDSA 2007-2009 Y EDSA BICENTENARIO (2010-2016). OBSERVATORIO DE LA DEUDA SOCIAL ARGENTINA. UCA. NOTA: LOS INGRESOS DE LOS AÑOS 2007 Y 2010 FUERON DEFLACIONADOS APLICANDO UN ÍNDICE DE PRECIOS ALTERNATIVO AL OFICIAL PROVENIENTE DE CENTROS Y EQUIPOS DE INVESTIGACIÓN (IPC 7 PROVINCIAS CENDA/IPC).

**Figura AM 7**

**MEDIA DE INGRESOS Y DE HORAS SEMANALES TRABAJADAS  
-CON Y SIN ESTIMACIÓN- SEGÚN ESTRATO SOCIO-ECONÓMICO**

EDSA 2007-2009 y EDSA Bicentenario (2010-2016).

ESTRATO SOCIO-ECONÓMICO	2007		2010		2011	
	HOGARES	ERROR*	HOGARES	ERROR*	HOGARES	ERROR*
MUY BAJO	466	4.49	1318	2.7	1398	2.62
BAJO	459	4.52	1414	2.61	1525	2.51
MEDIO BAJO	466	4.49	1594	2.45	1548	2.49
MEDIO ALTO	739	3.54	1327	2.69	1241	2.78
<b>TOTAL</b>	<b>2130</b>	<b>2.01</b>	<b>5653</b>	<b>1.3</b>	<b>5712</b>	<b>1.3</b>

CONDICIÓN SOCIO RESIDENCIAL	2007		2010		2011	
	HOGARES	ERROR*	HOGARES	ERROR*	HOGARES	ERROR*
VILLA O ASENTAMIENTO PRECARIO	78	11.1	150	8.0	270	5.96
TRAZADO URBANO DE NSE BAJO	954	3.17	2682	1.89	2902	1.82
TRAZADO URBANO DE NSE MEDIO	1096	2.96	2838	1.84	2540	1.94
<b>TOTAL</b>	<b>2130</b>	<b>2.12</b>	<b>5670</b>	<b>1.3</b>	<b>5712</b>	<b>1.3</b>

AGLOMERADO URBANO	2007		2010		2011	
	HOGARES	ERROR*	HOGARES	ERROR*	HOGARES	ERROR*
GBA	750	3.58	1722	2.36	1737	2.35
RESTO DE INTERIOR	1440	2.58	3960	1.56	3975	1.55
<b>TOTAL</b>	<b>2130</b>	<b>2.12</b>	<b>5682</b>	<b>1.3</b>	<b>5712</b>	<b>1.3</b>

MARGEN DE ERROR TOTAL CALCULADO SOBRE UNA MUESTRA ALEATORIA SIMPLE PARA LA ESTIMACIÓN DE UNA PROPORCIÓN POBLACIONAL DEL 50% Y UN NIVEL DE CONFIANZA DEL 95%. FUENTE: EDSA 2007-2009 Y EDSA-BICENTENARIO (2010-2016). OBSERVATORIO DE LA DEUDA SOCIAL ARGENTINA. UCA.

calcularon con la máxima variabilidad posible bajo el principio de aleatoriedad simple. Estos márgenes fueron calculados siempre sobre los casos encontrados y no sobre los casos ponderados y/o expandidos.

A continuación se reproducen los márgenes de error total de las principales variables utilizadas en el Informe calculado en base a una proporción poblacional de un 50 % y un nivel de confianza del 95%.

Cabe destacar que para el total de la muestra de la EDSA Bicentenario 2011 (5712 casos) el margen de error total es de 1,3 puntos porcentuales mientras que para la EDSA 2007-2009 (2130 casos) es de 2,01 puntos porcentuales.

- » Para el caso del Estrato socio-económico (figura AM7.a), al ser una clasificación en cuartiles, todas las categorías poseen una escasa diferencia en su margen de error total que gira en torno a los 2,6 puntos porcentuales para el caso de la EDA Bicentenario y en 4 puntos porcentuales para la EDSA 2007-2009.
- » Para el caso de Condición socio-residencial (figura AM7.b) puede apreciarse claramente como sus desiguales cantidades se traducen en desiguales márgenes de error totales arrojando 5,96 puntos porcentuales para las villas o asentamientos precarios, un 1,82 para la categoría de trazado urbano bajo y un 1.94% para trazado urbano medio para la EDSA Bicentenario 2011. Cabe advertir la notable mejoría en el tiempo en la reducción del margen de error total para el caso de las villas o los asentamientos precarios al comenzar con 11.10 puntos porcentuales en la EDSA 2007-2009, pasar a 8 puntos porcentuales en 2010 y llegar 5.96 puntos porcentuales de margen de error total en la EDSA Bicentenario 2011.
- » Para el caso específico de la variable “Aglomerado urbano” (figura AM7.c) en la EDSA Bicentenario 2011 el margen de error total es de 2,35 puntos porcentuales para GBA y de 1,55 puntos para el resto urbano integrante de la muestra. Para el caso de la EDSA 2007-2009 el margen de error total fue de 3,58 puntos porcentuales para GBA y de 2,58 puntos porcentuales para el resto del interior.