

# Diseño muestral, Construcción de indicadores y estructura metodológica de la EDSA

## CAPITULO METODOLÓGICO LA ENCUESTA EDSA DEL BICENTENARIO

De Grande Pablo

Quartulli Diego

Tinoboras Cecilia

La Encuesta de la Deuda Social Argentina (EDSA) aborda el estudio de las privaciones y realizaciones en materia de desarrollo humano de la sociedad argentina a través de una muestra probabilística estratificada no proporcional de población adulta de 18 años y más. Hasta el momento, la EDSA fue aplicada en siete oportunidades. Las mediciones se realizaron con una periodicidad anual durante el mes de junio de 2004 (línea de base),<sup>1</sup> 2005, 2006, 2007, 2008 y 2009.<sup>2</sup> Para el adecuado examen de los problemas interdisciplinarios que aborda el estudio de la Deuda Social desde la perspectiva del Desarrollo Humano y Social, se aplica un cuestionario multipropósito con amplia cobertura temática, incluyendo test psicosociales que permiten el diagnóstico, el monitoreo y la evaluación del impacto de las políticas públicas y las estrategias familiares.

### AM.1 Diseño muestral <tit. niv. 1>

La EDSA cuenta con un diseño muestral estratificado según criterios socio-residenciales y regionales. La estratificación es de tipo no proporcional y adoptó dos criterios de clasificación. En primer lugar, uno de *conglomeración urbana* y, en segundo lugar, otro de *estratificación por nivel socioeconómico de las unidades residenciales* (unidades censales) correspondientes a los conglomerados estudiados. En cuanto al universo geográfico del estudio, la muestra es representativa de dos tipos de conglomerados:

<viñeta>

- (1) Área metropolitana del Gran Buenos Aires (GBA);
- (2) Áreas metropolitanas del interior del país con más de 200 mil habitantes (ciudades del interior).

<fin de viñeta>

---

<sup>1</sup> A seis meses de relevada la primer onda existió una medición intermedia (diciembre de 2004), cuya finalidad fue la de operar como control a los datos relevados en la línea de base.

<sup>2</sup> La medición 2009 fue realizada en el mes de octubre. Si bien se modificó en el mes de relevamiento, en el primer semestre fue aplicada una sub-muestra (en el mes de mayo), que contó con la particularidad de que se realizó telefónicamente y únicamente a casos entrevistados en las ondas anteriores.

La estratificación socioeconómica se realizó a partir de una clasificación de conglomerados residenciales (unidades censales) según el perfil educativo predominante de los jefes de hogar. De esta manera, se clasificaron cinco espacios residenciales socioeducativos (ERS): *ERS Muy Bajo*, *ERS Bajo*, *ERS Medio Bajo*, *ERS Medio* y *ERS Medio Alto*. El diseño de la EDSA también responde a una estrategia de investigación comparada de tipo *longitudinal no rotatoria* (seguimiento de casos panel).

La cantidad total de casos relevados fue de 1.100 en las tres primeras mediciones, de 1.500 en junio de 2006 y de 2.520 en 2007 y 2008, y 2130 en 2009. A partir de estos casos, se obtuvo información acerca de algo más de 4400 miembros de los hogares relevados, en las primeras mediciones. En la medición de 2006, la ampliación de la muestra permitió alcanzar a un total de 5.676 personas, mientras que la totalidad de personas relevadas en 2007 fue de 9.071, en 2008 de 8940 y en 2009 de 7513. En todas las mediciones, el marco muestral utilizado fue la información censal a nivel de radio, correspondiente al Censo Nacional de Población, Hogares y Vivienda de 2001, realizado por el Instituto Nacional de Estadísticas y Censos de la República Argentina (CNVHyP, 2001). El procesamiento de los datos y su análisis estadístico se realizó tomando en cuenta la naturaleza compleja de la muestra –utilizando para ello el módulo de Muestras Complejas del SPSS–.

En función de la particular importancia que reviste para este programa de investigación relevar las heterogeneidades y desigualdades existentes en la estructura socio-residencial urbana, la EDSA se planteó un diseño capaz de aproximarse a los diferentes grupos vulnerables a partir de dos factores estructurantes de condiciones y oportunidades de inclusión social: a) *uno de tipo regional / metropolitano* (medido por el grado de concentración urbana de las ciudades); y b) *otro de tipo socioeconómico* (medido según las características socioeducativas de los puntos muestra).

<viñeta>

(a) *Regiones Metropolitanas*: se consideró a partir de la diferencia de dos grandes sistemas socioeconómicos urbanos. Por una parte, el Gran Buenos Aires o GBA (incluyendo a la Ciudad de Buenos Aires y a 24 partidos del conurbano bonaerense), y, por otra, algunas de las principales ciudades del interior del país, aquellas con más de 200 mil habitantes (Gran Córdoba, Gran Salta, Gran Resistencia, Gran Mendoza, Bahía Blanca y Neuquén-Plottier). En la medición del año 2007 se incorporaron al grupo del Interior dos aglomerados adicionales: Gran Rosario y Paraná, que se mantuvieron vigentes en 2007-2008. En 2009 fue incorporado sólo Gran Rosario como aglomerado adicional.

(b) *Estratos Socioeconómicos Residenciales*: se definió según las características socioeducativas de las unidades censales como medida de proximidad a una estratificación socioeconómica residencial de la población y sus hogares. Siguiendo este criterio, se diferenciaron tres estratos socioeducativos vulnerables (muy bajo, bajo y medio-bajo), un estrato de vulnerabilidad media (medio) y un estrato de comparación no vulnerable (media-alta).

<fin de viñeta>

En función de la estratificación residencial se utilizó como principal variable criterio el porcentaje de jefes de hogar con educación secundaria completa por punto muestra. A partir de estudios previos y de la aplicación de un análisis de correlación (con variables como el NBI –necesidades básicas insatisfechas–, proporción de adolescentes que no asisten a la

escuela secundaria, tasa de desempleo, entre otras dimensiones), dicho indicador mostró ser un criterio altamente confiable para representar diferentes probabilidades de vulnerabilidad socioeconómica. La aplicación de este indicador sobre el marco muestral formado por los puntos muestra de las mencionadas áreas metropolitanas permitió definir –para una distribución observada de tipo multimodal– cinco tipos teóricamente relevantes de espacios residenciales.

Los puntos de corte o rangos de concentración asignados a cada segmento fueron seleccionados buscando representar las diferentes modas presentes en la distribución (figura AM.1). Habiéndose aplicado este criterio, la estructura socioeconómica quedó representada a partir de la definición de cinco estratos residenciales socioeducativos (ERS) medidos al nivel de radio censal: 1) Radios de Clase Muy Baja (muy bajo nivel educativo); 2) Radios de Clase Baja (bajo nivel educativo); 3) Radios de Clase Media Baja (nivel educativo medio bajo, con baja incidencia de horno microondas); 4) Radios de Clase Media (nivel educativo medio o medio alto, con alta incidencia de horno microondas); y 5) Radios de Clase Media Alta (alto nivel educativo). Cabe agregar que en el caso de los radios de “clase media alta” o “nueva clase media” también se utilizó el porcentaje de hogares que poseen horno a microondas como forma de diferenciar socio-culturalmente a este segmento frente a los sectores de clase media más tradicionales.

**FIGURA AM.1**  
**DEFINICIÓN DEL ESTRATO MUESTRAL PARA LA EDSA**

Definición nominal	Definición operacional		
	Primera condición: % de hogares con jefe con educación secundaria completa (o más) en el punto muestra		Segunda condición: % de hogares con horno a microondas en el punto muestra
ERS Muy Bajo	0% a 11%		
ERS Bajo	12% a 27%		
ERS Medio Bajo	28% a 46%	y	0% a 34%
ERS Medio	28% a 46% 47% a 64%	y	35% y más, o
ERS Medio Alto	65% y más		

Fuente: EDSA, Observatorio de la Deuda Social Argentina. UCA.

Con el objetivo de que la población seleccionada resultara representativa de los distintos tipos de espacios socioeconómicos residenciales, el muestreo fue de tipo polietápico, estratificado de acuerdo a los niveles socioeducativos del punto muestra y región metropolitana (Gran Buenos Aires y ciudades del interior). Los recursos disponibles determinaron, para las tres primeras mediciones, un tamaño de muestra de 1.100 casos.<sup>3</sup> A partir de esta cantidad de observaciones, se procedió a la selección aleatoria ponderada de 184 radios censales (puntos muestra). Se tomó como decisión que a cada radio seleccionado

<sup>3</sup> Durante las tres primeras mediciones el ERS Medio no fue relevado. La ampliación de la muestra ocurrida en junio de 2006 permitió completar la cobertura de la estructura socioeconómica mediante el muestreo del estrato medio. Mientras que las tres primeras mediciones resultaban representativas de aproximadamente 8 millones de personas mayores de 18 años, la inclusión de los casos adicionales permite ampliar esta cifra hasta los 10,6 millones. En términos de los hogares alcanzados, las tres primeras mediciones predicaban acerca de 3,2 millones, incrementándose a 4,5 a partir de junio de 2006.

se le asignaran 6 unidades de estudio. La medición de junio de 2006 permitió una ampliación de la muestra a un total de 1.500 casos, alcanzando los puntos muestrales a 250. Durante 2007, la inclusión de los dos aglomerados adicionales y la extensión de la muestra para Gran Córdoba y Gran Mendoza llevaron el total de casos relevados hasta 2.520,<sup>4</sup> distribuidos en 420 radios. Al mismo tiempo, este cambio modificó la estructura de la muestra para facilitar posteriores análisis por aglomerado. En el año 2008 se mantuvo el mismo número de casos y el marco muestral de aglomerados y puntos muestrales de la encuesta. Ahora bien, en 2009 se modificó la inclusión de los dos aglomerados adicionales –excluyendo Paraná–, lo que dejó un total de casos de 2130 y se utilizó en consecuencia el mismo diseño muestral de 2008.

Las figuras AM.2a, AM.2b y AM.2c exponen la distribución de la población representada por la muestra, los casos seleccionados y los puntos muestra, según conglomerado urbano y espacio residencial socioeducativo, para la medición de octubre de 2009.

**FIGURA AM.2a**  
**DISTRIBUCION DE LA POBLACION DE 18 ANOS O MAS REPRESENTADA POR LA EDSA, SEGUN ESPACIO RESIDENCIAL SOCIOEDUCATIVO Y CONGLOMERADO URBANO.**  
**Año 2009**

	ERS Muy Bajo	ERS Bajo	ERS Medio Bajo	ERS Medio	ERS Medio Alto	TOTAL
<b>AMBA</b>	<b>1.913.134</b>	<b>2.239.406</b>	<b>1.190.931</b>	<b>2.290.162</b>	<b>746.972</b>	<b>8.380.605</b>
<b>Ciudades del Interior</b>	<b>433.043</b>	<b>950.335</b>	<b>871.066</b>	<b>668.815</b>	<b>258.965</b>	<b>1.770</b>
Gran Córdoba	91.583	235.621	284.605	229.352	88.946	930.107
Gran Mendoza	89.906	159.785	145.088	122.704	48.815	566.298
Gran Salta	22.175	100.986	78.461	52.390	31.102	285.114
Gran Resistencia	49.561	72.478	61.572	24.097	11.296	219.004
Bahía Blanca	18.695	64.197	62.565	42.089	5.039	192.585
Neuquén-Plottier	23.480	65.918	46.604	30.629	17.407	184.038
Gran Rosario	137.643	251.350	192.171	167.554	56.360	805.078
<b>TOTAL</b>	<b>2.346.177</b>	<b>3.189.741</b>	<b>2.061.997</b>	<b>2.958.977</b>	<b>1.005.937</b>	<b>11.562.829</b>

Fuente: EDSA, Observatorio de la Deuda Social Argentina. UCA.

<sup>4</sup> Debido a un problema con la recolección de datos en el aglomerado Paraná durante la medición de 2007 debieron descartarse tres casos, con lo cual el total llegó a 2517.

**FIGURA AM.2b**  
**DISTRIBUCIÓN DE LOS PUNTOS MUESTRA DE LA EDSA, SEGÚN ESPACIO RESIDENCIAL SOCIOEDUCATIVO Y**  
**CONLIGERADO URBANO**  
**Año 2009**

	ERS Muy Bajo	ERS Bajo	ERS Medio Bajo	ERS Medio	ERS Medio Alto	TOTAL
<b>AMBA</b>	25	25	25	25	25	125
<b>Ciudades del Interior</b>	46	46	46	46	46	230
Gran Córdoba	11	11	11	11	11	55
Gran Mendoza	13	13	13	13	13	65
Gran Salta	3	3	3	3	3	15
Gran Resistencia	2	2	2	2	2	10
Bahía Blanca	2	2	2	2	2	10
Neuquén-Plottier	2	2	2	2	2	10
Gran Rosario	13	13	13	13	13	65
<b>TOTAL</b>	71	71	71	71	71	355

Fuente: EDSA, Observatorio de la Deuda Social Argentina. UCA.

**FIGURA AM.2c**  
**DISTRIBUCIÓN DE LOS CASOS RELEVADOS POR LA EDSA, SEGÚN ESPACIO RESIDENCIAL**  
**Año 2009**

	ERS Muy Bajo	ERS Bajo	ERS Medio Bajo	ERS Medio	ERS Medio Alto	TOTAL
<b>AMBA</b>	150	150	150	150	150	750
<b>Ciudades del Interior</b>	276	276	276	276	276	1.770
Gran Córdoba	66	66	66	66	66	330
Gran Mendoza	78	78	78	78	78	390
Gran Salta	18	18	18	18	18	90
Gran Resistencia	12	12	12	12	12	60
Bahía Blanca	12	12	12	12	12	60
Neuquén-Plottier	12	12	12	12	12	60
Gran Rosario	78	78	78	78	78	390
<b>TOTAL</b>	426	426	426	426	426	2.130

Fuente: EDSA, Observatorio de la Deuda Social Argentina. UCA.

## Margen de error de la muestra <tit. niv. 2>

El error de una muestra aleatoria simple de 1.100 casos, correspondientes a las tres primeras mediciones, es de  $\pm 2.95\%$  (para la estimación de una proporción poblacional del 50% y un nivel de confianza del 95%). El incremento de la muestra en 2007 a un total de 2520 casos y el mantenimiento del mismo número (N) de casos para 2008 lleva este valor a  $\pm 1,95\%$ . Sin embargo, la estratificación no proporcional, que reduce el error de las comparaciones entre estratos, aumenta el error total de la muestra nacional (y también el de los estratos de nivel socioeconómico). Bajo el diseño estratificado original, el margen de error total era en 2004 de  $\pm 3.48\%$  (para la estimación de una proporción poblacional del 50% y un nivel de confianza del 95%), mientras que actualmente se ha visto disminuido hasta  $\pm 3,06\%$  para los aglomerados comparables, y  $\pm 2,96\%$  para el total de aglomerados considerados.

Dados los cambios ocurridos en la muestra utilizada para el relevamiento de la EDSA, es necesario analizar cómo son afectados los niveles de precisión de los resultados presentados. Los márgenes de error se reducen en las sucesivas muestras realizadas, y esta disminución se debe a dos motivos principales. Por un lado, la inclusión en 2006 del estrato Medio supone una disminución del error en el Gran Buenos Aires desde  $\pm 4,34\%$  en 2004 hasta  $\pm 3,81\%$  en 2007 y 2008, mientras que en las ciudades del interior esta disminución es del  $\pm 4,43\%$  en 2004 al  $\pm 3,38\%$  en 2007 y 2008. En este último conglomerado opera también un incremento en la cantidad de casos relevados.

El segundo motivo de importancia es la inclusión en 2007 de dos aglomerados adicionales correspondientes al interior del país. Mientras que el error para el Gran Buenos Aires se mantiene (recibe solamente la influencia de la inclusión del estrato adicional), en ciudades del interior disminuye desde  $\pm 3,38\%$  (margen de error en 2007, 2008 y 2009 de los aglomerados comparables con 2004) hasta  $\pm 2,81\%$ , cerca de dos puntos menor al de 2004 para el mismo conglomerado.

En la figura AM.3 se presentan la cantidad de casos relevados, el error estándar y el margen de error correspondientes a cada estrato y conglomerado en la primera y última mediciones.

**FIGURA AM.3**  
**MÁRGENES DE ERROR PARA LA MUESTRA DE LA EDSA**  
**Evolución 2004-2009**

	Año 2004			Año 2009 <sup>1</sup>			Muestra ampliada <sup>2</sup> Año 2009		
	n	Error estándar	Margen de error	n	Error estándar	Margen de error	n	Error estándar	Margen de error
<b>Gran Buenos Aires</b>									
ERS Medio Alto	100	5,00	9,80	150	4,08	8,00	150	4,08	8,00
ERS Medio	s/d	s/d	s/d	150	4,08	8,00	150	4,08	8,00
ERS Medio Bajo	150	4,08	8,00	150	4,08	8,00	150	4,08	8,00
ERS Bajo	150	4,08	8,00	150	4,08	8,00	150	4,08	8,00
ERS Muy Bajo	150	4,08	8,00	150	4,08	8,00	150	4,08	8,00
<b>Total</b>	<b>550</b>	<b>2,13</b>	<b>4,34</b>	<b>750</b>	<b>1,83</b>	<b>3,81</b>	<b>750</b>	<b>1,83</b>	<b>3,81</b>
<b>Ciudades del Interior</b>									
ERS Medio Alto	100	5,00	9,80	198	3,55	6,96	276	2,66	4,06
ERS Medio	s/d	s/d	s/d	198	3,55	6,96	276	2,66	4,06
ERS Medio Bajo	150	4,08	8,00	198	3,55	6,96	276	2,66	4,06
ERS Bajo	150	4,08	8,00	198	3,55	6,96	276	2,66	4,06
ERS Muy Bajo	150	4,08	8,00	198	3,55	6,96	276	2,66	4,06
<b>Total</b>	<b>550</b>	<b>2,13</b>	<b>4,43</b>	<b>990</b>	<b>1,59</b>	<b>3,38</b>	<b>1.380</b>	<b>1,19</b>	<b>2,81</b>
<b>Total urbano</b>									
ERS Medio Alto	200	3,54	7,99	348	2,68	6,47	426	2,23	6,02
ERS Medio	s/d	s/d	s/d	348	2,68	6,68	426	2,23	6,25
ERS Medio Bajo	300	2,89	5,87	348	2,68	5,69	426	2,23	5,04
ERS Bajo	300	2,89	6,42	348	2,68	6,32	426	2,23	5,77
ERS Muy Bajo	300	2,89	6,95	348	2,68	6,99	426	2,23	6,57
<b>Total</b>	<b>1.100</b>	<b>1,51</b>	<b>3,48</b>	<b>1.740</b>	<b>1,20</b>	<b>3,06</b>	<b>2.130</b>	<b>1,00</b>	<b>2,96</b>

<sup>1</sup> Los resultados no incluye la ciudad de Rosario.

<sup>2</sup> Los resultados incluye la ciudad de Rosario.

s/d: sin dato. El ERS Medio no fue relevado durante 2004.

Fuente: EDSA, Observatorio de la Deuda Social Argentina. UCA.

## AM.2 Procedimiento de selección de los casos de la muestra <tit. niv. 1>

La selección de las unidades censales para cada espacio residencial se realizó mediante un muestreo aleatorio de radios con probabilidad proporcional al tamaño de la población de 18 años y más de cada aglomerado considerado. Las manzanas o puntos de muestra barrial al interior de cada radio y las viviendas de cada manzana o barrio se seleccionaron aleatoriamente a través de un muestreo sistemático, mientras que los individuos dentro de cada vivienda fueron escogidos mediante un sistema de cuotas por sexo y edad. Para cada punto muestra quedaron asignados 6 hogares. Una vez seleccionado el individuo, se le aplicó un formulario personal, relevándose características de la vivienda y de los demás miembros del hogar, identificando a cada uno de los individuos en función de su posición respecto al jefe laboral.

Si se ignora el efecto de la selección de viviendas y de los individuos al interior de las mismas, este diseño muestral polietápico genera diferentes probabilidades de selección, y por ende también, ponderadores constantes dentro de cada espacio. Estas probabilidades de selección tienen la forma:

$$P_{i,j} = \frac{R_s 6}{2P_e}$$

Aproximadamente, donde:

$P_{i,j}$  = Ponderador del espacio  $i$  en la ciudad  $j$ ;

$R_s$  = Cantidad de radios seleccionados;

$P_e$  = Población total del estrato.

Estas probabilidades de selección pueden descomponerse, también aproximadamente, en probabilidades de selección de cada etapa:

$$PR_{i,j} = \frac{R_s P_s}{P_e} \frac{1}{2} \frac{6}{P_i}$$

donde el primer cociente se refiere a la probabilidad de selección del radio, el segundo a la probabilidad de que el radio seleccionado sea o no suplente y el último es la probabilidad de seleccionar 6 individuos en la población del radio. Esta última, supone que la selección de manzanas, viviendas e individuos derivan, en última instancia, en equiprobabilidad para los individuos, dentro de cada radio<sup>5</sup> (Muiños, 2005).

---

<sup>5</sup> La descomposición de las probabilidades de selección por espacio y área urbana permitió determinar con mayor precisión los parámetros a utilizar en el módulo de muestras complejas del SPSS, utilizados para el cálculo de los estimadores y las pruebas de significancia.

### AM.3 Consistencias internas e imputación de ingresos de no respondientes<tit. niv. 1>

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. 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. En segundo lugar, se realizó una consistencia temporal para las reentrevistas. Este procedimiento implicó el control de las respuestas de los mismos sujetos a lo largo del tiempo, supervisando errores de carga, valores faltantes y respuestas incoherentes temporalmente.

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 (Gasparini y Sosa Escudero, 2001). En el caso de la EDSA, se confirmó una relación directa y significativa entre el estrato residencial socioeducativo (ERS) y la no respuesta. Asimismo, (figura AM.4) es considerable la variación en el tiempo del porcentaje de no respuestas, particularmente en la última medición.

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).<sup>6</sup>

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 total de los hogares. Particularmente, el análisis de regresión para el ingreso laboral tomó en cuenta variables de índole demográfica (*sexo, grupos de edad, situación conyugal*) y socioeconómica (*nivel educativo, situación ocupacional, ocupación principal, jefatura de hogar, etc.*), además de tenerse en cuenta los espacios residenciales socioeducativos (ERS) del diseño muestral. El análisis para el ingreso del hogar incluyó otras variables como *cantidad de componentes, población económicamente activa dentro del hogar, tipo de familia, ciclo vital del hogar, clima educativo, características del jefe de hogar*, además considerar la *recepción de asistencia en forma monetaria o no monetaria* por parte de organismos públicos y privados. Las estimaciones fueron realizadas según 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.

---

<sup>6</sup> El método de estimación utilizado fue el mismo para cada uno de los años de la serie 2004-2009. Para ellos se ajustó un modelo regresión lineal múltiple, a partir de la aplicación del comando MVA (Missing Value Analys) disponible en el módulo de Valores Perdidos del paquete estadístico SPSS. La estimación se hizo utilizando el método de estimación EM.



**FIGURA AM.4**  
**INGRESOS DE LOS HOGARES E INGRESO DE INGRESOS LABORALES DE LOS OCUPADOS**  
 Evolución 2004-2009

	Aglomerados Comparables (*)						Muestra ampliada(**)		
	Año 2004	Año 2005	Año 2006	Año 2007	Año 2008	Año 2009	Año 2007	Año 2008	Año 2009
<b>Ingresos del Hogar</b>									
Porcentaje de no respuestas <sup>1</sup>	14,1	18,4	18,5	28,8	49,2	50,3	29,1	49,5	50,6
Promedio de ingresos laborales	741	823	1003	1255	1455	1823	1261	1450	1825
Promedio de ingresos relevados	694	777	965	1180	1373	1759	1183	1375	1750
Promedio de ingresos estimados	908	954	1135	1395	1560	1927	1409	1541	1940
<b>Ingresos laborales</b>									
Porcentaje de no respuestas	21,3	22,8	21,3	29,7	44,2	45,1	30,0	44,5	45,4
Promedio de ingresos del hogar	863	1080	1326	1715	2044	2365	1725	2051	2343
Promedio de ingresos relevados	795	1033	1254	1568	1831	2297	1582	1842	2276
Promedio de ingresos estimados	1237	1270	1627	2007	2259	2451	2014	2255	2422

(\*) Los resultados no incluye la ciudad de Rosario.

(\*\*) Los resultados incluye la ciudad de Rosario.

<sup>1</sup>Porcentaje calculado sobre el total de hogares.

Fuente: EDSA, Observatorio de la Deuda Social Argentina. UCA.

## AM.4 Comparabilidad intertemporal de los resultados <tit. niv. 1>

El diseño muestral de la EDSA ha sufrido en el transcurso del tiempo algunas modificaciones que pueden afectar la comparabilidad intertemporal de los resultados obtenidos. Si bien la cantidad de casos se ha incrementado, la ampliación de la muestra no tiene efectos nocivos cuando se trata de una mayor cantidad de casos representativos del mismo universo. Sin embargo, dado que el diseño muestral de la encuesta se basa en dos criterios principales, su comparabilidad intertemporal debe evaluarse en estos dos niveles:

<viñeta>

- Comparabilidad de aglomerados;
- Comparabilidad de estratos muestrales.

<fin de viñeta>

El primer punto es de simple consideración: durante la medición de 2007 se agregaron al estudio dos aglomerados adicionales correspondientes al interior del país: Gran Rosario y Paraná. La inclusión de estos dos aglomerados en 2007 y 2008 (y sólo la del primero de ellos para 2009) permite una mejor representación del interior del país, mejorando las estimaciones. Sin embargo, para realizar la comparación de stocks en relación con mediciones anteriores, no pueden tenerse en cuenta estas incorporaciones. Por esta razón, las comparaciones temporales excluyen de 2007, 2008 y 2009 estos aglomerados, presentándose sus resultados por separado, el primero utilizando los aglomerados comparables, el segundo, la inclusión del aglomerados Gran Rosario para los años 2007, 2008 y 2009.

La comparabilidad de los estratos muestrales es un tema más complejo que ya se ha tenido en cuenta en publicaciones anteriores (ODSA, 2007). Como fue aclarado anteriormente, en sus inicios la EDSA cubría una parte de la estructura socioeconómica, dejando afuera el Estrato Medio de la estratificación muestral. Este estrato, caracterizado por hogares que

estuvieran ubicados en radios con una incidencia de entre 28% y 46% de jefes de hogar con educación secundaria completa y un 35% o más de hogares con horno a microondas, o bien entre 47% y 64% de jefes con secundaria completa, fue agregado al análisis en una medición posterior, durante 2006. La inclusión de este estrato muestral permitió el análisis de la estructura socioeconómica completa, pero agregó complejidad a la comparación intertemporal de los resultados. En particular, los valores totales obtenidos a partir de estas muestras constituyen resúmenes de poblaciones objetivo diferentes.

Sin embargo, el procedimiento de selección de los casos basado en radios censales representativos de distintos espacios residenciales socioeducativos permite el relevamiento de hogares que, perteneciendo a un radio censal/punto muestra/barrio con altos niveles educativos, posean niveles educativos bajos o muy bajos, y viceversa. Tomando en cuenta este hecho, se decidió caracterizar a los radios muestrales según el porcentaje de jefes de hogar con secundaria completa efectivamente observados en ese radio, para luego redistribuirlos en cuatro categorías según se describe a continuación:

<viñeta>

- Menor al 12% de jefes con secundaria completa en el radio;
- Entre 12% y 28% de jefes con secundaria completa en el radio;
- Entre 28% y 55% de jefes con secundaria completa en el radio;
- Más de 55% de jefes con secundaria completa en el radio.

<fin de viñeta>

Originalmente, la distribución observada de los casos se compara con una distribución teórica, elaborada a partir de los valores muestrales de las dos últimas mediciones (ambas incluyen a la estructura socioeconómica completa) y que acepta un grado razonable de dispersión de los casos. Ahora bien, el problema que surge es conocido: producto de fallas de medición o de sesgos educativos en la selección de los hogares y entrevistados, el ponderador ajustado NES, si bien corrige buena parte de dicho sesgo tanto a nivel educativo de jefes de hogar como de los entrevistados, no lo hace de manera completa.

Dada la estrecha relación entre la educación y las condiciones sociales, aunque leve, el mencionado error de ajuste afecta las comparaciones interanuales introduciendo un sesgo no aleatorio en las estimaciones a lo largo del tiempo. Asimismo, este error impacta también de manera no aleatoria en las comparaciones entre estratos, afectando nuestras estimaciones sobre la desigualdad social, así como sobre la evolución de la misma. Una adecuada corrección a este problema implicaría volver a construir los ponderadores considerando una serie de controles sobre los niveles de educación de la población con mayor desglose por nivel educativo y aglomerado. Sin embargo, seguir dicha estrategia hubiera implicado impactar en los criterios de clasificación (NES), alterando toda la estructura de análisis de la que actualmente disponemos. Debido a que los tiempos de producción son exigüos y el ponderador NES resuelve la mayor parte del problema, tal revisión no fue posible ni recomendable.

Ante este escenario, se buscó ensayar un procedimiento alternativo de bajo costo tiempo/recurso humano capaz de permitir comparaciones generales e inter clases/NES sin cambiar el modelo de análisis. El procedimiento seguido fue más sencillo en términos de

operaciones lógicas: ajuste de las distribuciones de cada año a un año parámetro, a partir de la aplicación de un coeficiente de corrección que no alterara la participación de cada NES de clasificación (cuartiles) en el total de la muestra ponderada, buscando al mismo tiempo controlar niveles de instrucción más específicos a los seguidos por el procedimiento original.

Es importante señalar que este ajuste no se hizo a partir de parámetros censales ni se basó en las unidades de diseño muestral (educación de los jefes de hogar). Esta decisión se fundamentó en la imperiosa necesidad de no alterar el clasificador (NES), dado que ello habría también implicado mayor uso de recursos/tiempo de trabajo. Por tal motivo se eligió lograr un ajuste parcial tomando como parámetro la medición del año 2007 (el clasificador NES se basó en dicho año para identificar los puntos de corte cuartílicos). Como resultado de este ajuste, se produce un incremento general de los niveles educativos promedio de la muestra en sus primeras mediciones (2004 y 2005), compensándose la ausencia del estrato medio. Este incremento se logra otorgándole un mayor peso a los casos observados dentro de las categorías 3) y 4) anteriores. Así, se logra que los niveles de educación promedio de la muestra sean similares en el tiempo, permitiendo una mayor homogeneidad temporal.

## AM.5 Metodología de elaboración de los Índices de Desarrollo Humano y Social<sup>7</sup> <tit. niv. 1>

La construcción de una medida de desarrollo humano y social exige reducir en un indicador de resumen las diferentes dimensiones que se investigan en este estudio, mediante el reconocimiento de las interrelaciones y correlaciones entre los varios componentes y la asignación de ponderaciones apropiadas. La noción fundamental detrás de esta construcción es que los distintos aspectos observables a través de la encuesta de la Deuda Social Argentina son diferentes dimensiones de un mismo concepto subyacente que no es directamente mensurable, y que denominamos “Desarrollo Humano y Social”.

Este complejo concepto, tal como ha sido estudiado en este informe y en los anteriores, reconoce dos niveles principales con dos dimensiones cada uno. Cada una de las dimensiones es un concepto o construcción intelectual abstracta, que de por sí es inobservable, razón por la cual su medición se efectúa a través de indicadores elaborados a partir de las variables de la encuesta. Dado que estos indicadores son siempre parciales e imperfectos, se utilizan varios de ellos con el objeto de reflejar de manera más completa la variabilidad que presenta el concepto subyacente, el cual puede ser entonces inferido indirectamente a partir de los indicadores observables y sus mutuas correlaciones.

La metodología de elaboración de los Índices de Desarrollo Humano y Social abarca una serie de procedimientos y técnicas estadísticas complejas que han sido expuestas en las publicaciones anteriores. En consecuencia, en el presente anexo metodológico, no serán expuestas.<sup>8</sup> Para la elaboración de los índices de 2009 se utilizó un procedimiento distinto

---

<sup>7</sup> Este punto es un resumen de la descripción pormenorizada de la técnica, realizada en el anexo metodológico 3 del Barómetro de la Deuda Social 2007 (DII-ODSA, 2008).

<sup>8</sup> Para un desarrollo detallado de la metodología de análisis factorial seguida para la elaboración de los índices y de los métodos de agregación utilizados en el período 2004-2008, véase el anexo metodológico 3 de *La Deuda Social Argentina: 2004-2008* (ODSA, 2009).

al de los años anteriores. En la presente medición se tomó la decisión de plantear un modelo de estimación que permitió –en base a una batería de variables predictoras– realizar una proyección de los valores de los índices en 2009. En esta dirección se procedió a realizar dos estimaciones: una primera del *índice de condiciones materiales de vida*, y una segunda, del *índice de integración humano y social*. Una vez estimados ambos, el *índice de desarrollo humano y social* se calculó como un promedio ponderado de ambos índices.

El método de estimación utilizado para hacer estas proyecciones fue un análisis de regresión lineal múltiple. Este procedimiento se realizó a partir de la aplicación del comando MVA (Missing Value Analysis) disponible en el módulo de Valores Perdidos del paquete estadístico SPSS de la Versión 17.0. Este comando permite introducir variables categóricas y cuantitativas en un modelo predictivo. En el presente análisis se utilizó el método de estimación EM, el cual estima las medias, la matriz de covarianza y la correlación de las variables cuantitativas con los valores perdidos a partir de un proceso iterativo. El procedimiento tomó en cuenta un serie de variables *dummy* de tipo demográfico (sexo, grupos de edad, situación conyugal), socioeconómico (nivel educativo, situación ocupacional, ocupación principal, jefatura de hogar, etc.), y otra serie de variables -también bajo un formato *dummy*- vinculadas estrictamente a las dimensiones de análisis propias de cada una de las condiciones que se sintetizan en cada índice: a) en el índice de condiciones materiales de vida se aplicó un set de indicadores de déficit y logros de las dimensiones de hábitat, salud, situación económica de los hogares, trabajo y seguridad social, y b) en el índice de integración humano y social, indicadores de las dimensiones de condiciones psicológicas, vida social y confianza política.

## AM.6 Análisis de las medidas de desarrollo humano y social <tit. niv. 1>

Los Índices de Desarrollo Humano y Social se expresan en una escala de calificación de 0 a 10 puntos, donde cero (0) representa la máxima distancia observada a los umbrales normativos mínimos y diez (10) expresa el acceso a las condiciones establecidas por estos umbrales. Los índices se exponen mediante el cálculo de promedios de estas calificaciones, para los totales y las categorías de las variables de corte seleccionadas.

Se calcularon diferencias de medias como medida de las discrepancias en las calificaciones obtenidas para los diversos grupos definidos por cada variable de corte. La diferencia absoluta se computó mediante la sustracción de los valores promedio del índice para cada categoría respecto de la categoría seleccionada como comparación (categoría de referencia). Esta diferencia absoluta está expresada en la misma unidad de medida que los índices y, por lo tanto, puede interpretarse como puntos de diferencia entre calificaciones.

<viñeta>

- La diferencia de medias es mayor que cero ( $DM > 0$ ) cuando existe una asociación directa entre la variable de corte y el índice respectivo.
- La diferencia de medias es igual a cero ( $DM = 0$ ) cuando el promedio en la categoría de la variable de corte y su categoría de referencia es el mismo, lo cual implica una ausencia de asociación entre la variable de corte y el índice analizado.

- La diferencia de medias es menor que cero ( $DM < 0$ ) en el caso en el que se da una relación inversa entre la variable de corte y el índice estudiado.

<fin de viñeta>

La importancia de estas diferencias se evalúa a partir de la aplicación de una prueba de hipótesis, específicamente la *prueba de diferencia de medias para muestras independientes*.<sup>9</sup> Si disponemos de dos muestras aleatorias independientes, la primera consta de  $n_x$  observaciones de una población con una media de  $\mu_x$  y la media muestral resultante es  $\bar{X}$ . La segunda tiene  $n_y$  observaciones de una población cuya media es  $\mu_y$  y la media muestral resultante es  $\bar{Y}$ . El objetivo es contrastar la hipótesis de que las medias poblacionales  $\mu_x$  y  $\mu_y$  son iguales. Es decir, las hipótesis nula y alternativa son:

$$H_0) \mu_x = \mu_y$$

$$H_1) \mu_x \neq \mu_y.$$

De acuerdo a  $H_0$ , el estadístico de prueba:

$$z = \frac{(\bar{X} - \bar{Y})}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$$

tiene distribución normal estándar. Sin embargo, las varianzas poblacionales son desconocidas. Si se utilizan las varianzas muestrales, esta variable aleatoria sigue una distribución t de Student con  $(n_x + n_y - 2)$  grados de libertad. La regla de decisión es rechazar  $H_0$  a favor de  $H_1$  toda vez que la probabilidad asociada a este estadístico sea menor que el nivel de significación elegido. Cuando los tamaños muestrales son grandes, la aproximación normal es buena incluso al sustituir las varianzas poblacionales por las muestrales (Newbold, 1997).

Habiendo seleccionado un nivel de significación del 5% se optó por utilizar la corrección de Bonferroni para comparaciones múltiples. Consecuentemente, el verdadero nivel de significación de acuerdo a la corrección de Bonferroni es  $\frac{\alpha}{k} = \frac{0,05}{k}$  donde  $k$  es la cantidad de comparaciones realizadas.

## AM.7 Análisis de los indicadores de logros y déficits <tit. niv. 1>

---

<sup>9</sup> Se considera que las muestras son independientes puesto que las categorías de las distintas variables de corte son independientes entre sí. Esto es así con la excepción del estrato socioeconómico que, como se describe en el punto 6 de este mismo anexo, se construye a partir del cálculo de cuartiles de un mismo índice socioeconómico. Si bien no puede afirmarse que estos estratos sean independientes, existen limitaciones en el software utilizado que obligan al empleo de esta versión de la prueba estadística.

El estudio de cada dimensión del Desarrollo Humano y Social se encaró a través de una serie de indicadores elaborados a partir de las preguntas del cuestionario de la EDSA. Estos indicadores se expresaron de manera cualitativa como variables dicotómicas 0-1, donde la unidad representa la presencia de privaciones en el contenido al que refiere el indicador.<sup>10</sup> A partir de estos indicadores se calcularon tasas de recuento, que se computan como el cociente entre la cantidad de casos que poseen determinado atributo y el total de casos observados. Se analizó la prevalencia de los indicadores, tanto en forma general como de manera diferencial para cada variable de corte.

Como una medida de las diferencias existentes entre las categorías de las variables de corte, se calculó el Riesgo Relativo de que la persona estuviera en situación de privación en relación con la categoría de referencia elegida. El Riesgo Relativo es un ratio (razón o cociente) entre las incidencias de cada indicador para cada grupo. Es una medida relativa de interpretación directa, que indica cuántas veces mayor es la incidencia de determinado indicador en un grupo respecto del otro. Su magnitud da idea de la desigualdad existente respecto del indicador entre los grupos comparados.

<viñeta>

- Si el riesgo relativo es mayor que 1 ( $RR > 1$ ), existe una asociación directa (positiva) entre la variable de corte y el indicador analizado.
- Cuando el riesgo relativo es igual a uno ( $RR = 1$ ) la incidencia de un indicador en ambos grupos es la misma, lo cual sería signo de la falta de asociación entre la variable de corte y el indicador analizado.
- Si el riesgo relativo es menor que 1 ( $RR < 1$ ), existe una relación inversa (negativa) entre la variable de corte y el indicador medido.<sup>11</sup>

<fin de viñeta>

Con el objetivo de determinar la importancia de las diferencias entre las incidencias según las categorías de cada variable de corte, se aplicaron pruebas de significación, en particular la *prueba de diferencia de proporciones para muestras independientes*.<sup>12</sup> La selección de este método se basó en que los indicadores de la EDSA son elaborados como tasas de recuento y, por lo tanto, constituyen proporciones. En consecuencia, una prueba de diferencia de proporciones de este tipo permite determinar si las incidencias de cada indicador resultan o no significativamente diferentes para los diferentes grupos.

---

<sup>10</sup> En ciertas dimensiones se optó por diferenciar dentro del déficit dos niveles de gravedad a partir de un umbral adicional, con el objetivo de incorporar la mayor cantidad de información posible en la construcción del índice correspondiente. Estos indicadores ordinales se emplearon únicamente para la elaboración de los índices, mientras que la prevalencia del déficit en cada dimensión se midió a través del empleo de indicadores dicotómicos.

<sup>11</sup> Por tratarse de una medida relativa, el riesgo relativo no está expresado en las unidades de medida de las variables que compara. Esto constituye una ventaja para su interpretación, dado que lo hace comparable entre variables o indicadores. Es decir, si un indicador cualquiera muestra un riesgo relativo alto entre los grupos mencionados esto es signo de desigualdad entre ellos. Otro indicador que tenga un riesgo relativo menor a aquel en el mismo momento, mostrará menor desigualdad, no importa lo diferentes que sean las dimensiones que estos indicadores midan.

<sup>12</sup> Véase la explicación brindada en la nota 2 del punto AM.2 de este mismo anexo.

A continuación se brinda una breve descripción de la *prueba de diferencia de proporciones para muestras independientes*. Si disponemos de dos muestras aleatorias independientes, la primera consta de  $n_x$  observaciones de una población con una proporción de  $p_x$  y la proporción muestral resultante es  $\hat{p}_x$ . La segunda tiene  $n_y$  observaciones de una población cuya proporción es  $p_y$  y la proporción muestral resultante es  $\hat{p}_y$ . El objetivo es contrastar la hipótesis de que las proporciones poblacionales  $p_x$  y  $p_y$  son iguales. Es decir, las hipótesis nula y alternativa son:

$$H_0) p_x - p_y = 0;$$

$$H_1) p_x - p_y \neq 0.$$

De acuerdo a  $H_0$  y si el valor común de las proporciones es  $p_0$ , el estadístico de prueba es:

$$z = \frac{(\hat{p}_x - \hat{p}_y)}{\sqrt{\frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y}}}$$

La regla de decisión es rechazar  $H_0$  a favor de  $H_1$  toda vez que la probabilidad asociada a este estadístico con distribución normal estándar aproximada, a dos colas, sea menor que el nivel de significación elegido. Nuevamente se aplicó la corrección de Bonferroni, seleccionando un nivel de significación del 5% (Newbold, 1997).

## AM.8 Variables de corte seleccionadas <tit. niv. 1>

Tanto las medidas de desarrollo humano y social como los indicadores de déficit 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 de los resultados.

Las variables de corte seleccionadas fueron de dos tipos en relación a la unidad de análisis a la cual refieren. Por tanto, se tomaron variables que refieren a la unidad de análisis de *personas*, por un lado, y *hogares*, por el otro. Las variables del primer grupo que se utilizaron fueron: *sexo*, *grupos de edad*, *niveles de instrucción formal* (dicotómico), *capital de agencia* y *redes sociales*. Las del segundo grupo fueron: *estrato socioeconómico* (categorizado en cuartiles), *estrato socioeconómico* (dicotómico categorizado en los deciles 1 y 10) y *conglomerado urbano*.

## AM.9 Clasificación de los casos para la presentación de los resultados <tit. niv. 1>

El análisis de los datos se realizó a partir de una clasificación en términos del conjunto de variables de corte descritas en el apartado anterior. Sin embargo, dos de los criterios de clasificación se consideran principales. Por un lado, el criterio regional, que clasifica los

casos según el conglomerado urbano donde estos fueron relevados. En particular, se subdividen los resultados entre el Gran Buenos Aires (GBA) y las grandes ciudades del interior del país.

El segundo criterio de clasificación de importancia es el estrato socioeconómico del hogar encuestado. El esquema de clasificación socioeconómica presentado en esta oportunidad sigue lineamientos similares al empleado en el Barómetro de 2007 (ODSA, 2008), que incorporaba los principales activos del hogar en dos niveles, aquellos propios del hogar y los del entorno residencial donde éste se localiza.<sup>13</sup> En este caso se efectuaron dos mejoras sobre la metodología aplicada en aquella publicación.

Por un lado, se otorgó una mayor importancia a un tercer nivel de análisis con el propósito de precisar la situación del entorno inmediato del hogar, asumiendo la heterogeneidad socioeconómica existente en los radios censales. En consecuencia, la posición socioeconómica de un hogar resulta ahora función de tres niveles, en lugar de dos: a) hogar; b) entorno de la vivienda; y c) radio censal. Por otro lado, se dio mayor transparencia al proceso de clasificación socioeconómica de los hogares sustituyendo criterios *ad hoc* por criterios técnicos, al momento de decidir las ponderaciones con las que influyen los factores tomados en cuenta en el análisis. Concretamente, el juego de ponderadores es reemplazado por los coeficientes derivados de la correlación estadística de los indicadores originales, provenientes de un análisis de tipo factorial similar al empleado para la generación de los índices de desarrollo.

El esquema clasificatorio integró los atributos del hogar y del vecindario en materia de dotación de capital educativo y de acceso a las tecnologías de información y comunicación (TIC). Las variables utilizadas para la definición de estos aspectos fueron:

<viñeta>

- *Capital educativo del vecindario*: definido operativamente como el porcentaje de jefes de hogar con secundaria completa o más en el radio censal de localización del hogar.
- *Acceso a tecnologías en el vecindario*: definido operativamente como el promedio de los porcentajes de hogares en el barrio con acceso a medios de comunicación e información. Los bienes considerados son tres: televisión por cable, teléfono celular e Internet.
- *Capital educativo del hogar*: definido operativamente como el total de años de educación del jefe de hogar, como miembro más representativo del mismo.
- *Acceso a tecnologías en el hogar*: definido operativamente como un indicador compuesto del acceso que logra el hogar a cada uno de los bienes tecnológicos mencionados. Estos tres aspectos se combinaron para formar un índice, logrado como la suma simple o inventario de bienes que este posee. De esta manera, los hogares que posean sendos bienes tendrán un puntaje de 3, mientras que aquellos que no posean ninguno tendrán un puntaje de cero.

---

<sup>13</sup> Con este propósito se combinaban dos fuentes de datos, agregando a la información de la EDSA que se emplea a nivel hogar los datos del Censo de Población 2001 para informar acerca del nivel del entorno residencial, a través de las características del radio censal al que el hogar pertenece.



<fin de viñeta>

Posteriormente se combinaron ambos atributos del hogar y ambos atributos del radio censal en dos índices, cada uno logrado a partir de un promedio ponderado que otorga mayor peso al capital educativo (75%) por sobre el acceso a tecnologías (25%).

Estos dos índices correspondientes al radio censal y al hogar fueron luego combinados mediante el procedimiento de análisis factorial similar al empleado para la generación de los Índices de Desarrollo Humano y Social de ODSA con el indicador de vivienda en villa de emergencia o asentamiento precario, representativo de un nivel intermedio referente al entorno próximo al hogar. De este procedimiento se obtuvieron los coeficientes que sirvieron para establecer el juego de ponderaciones para la construcción del índice socioeconómico. La inclusión de este último indicador al análisis permitió la concentración de los hogares en villa de emergencia o asentamiento precario en el cuartil socioeconómico más bajo, otorgando mayor precisión al criterio de clasificación. Una vez obtenido el índice socioeconómico se obtuvieron sus cuartiles, cuatro grupos ordenados que representan niveles crecientes de capital educativo y de acceso a tecnologías de información y comunicación, incluyendo tanto el capital individual como el capital del entorno residencial donde habitan los individuos y hogares que forman parte del estudio<sup>14</sup> (ODSA, 2008). Los cuatro grupos resultantes quedaron clasificados en: 1) Estrato Muy Bajo; 2) Estrato Bajo; 3) Estrato Medio Bajo; y 4) Estrato Medio Alto.

---

<sup>14</sup> La decisión de definir los grupos para la presentación tabular utilizando cuartiles supone una clasificación relativa, razón por la cual la ubicación de cada individuo en la escala depende de la posición de los demás individuos encuestados.

**FIGURA AM.5**  
**VARIABLES UTILIZADAS PARA LA CLASIFICACIÓN SEGÚN ESTRATO SOCIOECONÓMICO.**

	Atributos del hogar		Atributos del vecindario				Hogares ubicados en villa de emergencia (%)	
	Capital educativo	Acceso a TIC	Capital educativo	Acceso a TIC				
	Años de educación del jefe (promedio)	Inventario de bienes tecnológicos (promedio)	Jefes con secundaria completa en el vecindario (%)	Hogares con internet en el vecindario (%)	Hogares con tv con cable en el vecindario (%)	Hogares con tel. celular en el vecindario (%)		Promedio (%)
<b>Estrato socioeconómico 2004</b>								
Muy Bajo	5,9	0,3	10,9	1,9	32,6	13,5	16,0	34,6
Bajo	7,6	0,5	22,7	6,2	43,2	22,2	23,8	0,0
Medio Bajo	10,2	0,9	33,8	11,0	49,6	28,9	29,8	0,0
Medio Alto	13,2	1,7	64,2	34,0	72,7	48,6	51,7	0,0
<b>Estrato socioeconómico 2005</b>								
Muy Bajo	5,9	0,4	11,1	2,1	32,6	13,7	16,1	38,6
Bajo	7,8	0,6	22,2	6,1	43,7	22,0	23,9	0,0
Medio Bajo	10,4	1,1	32,8	10,4	48,7	28,4	29,1	0,0
Medio Alto	13,3	1,8	64,3	34,0	72,1	48,9	51,6	0,0
<b>Estrato socioeconómico 2006</b>								
Muy Bajo	6,5	0,5	12,0	2,5	33,0	14,3	16,6	28,8
Bajo	8,0	1,0	23,2	6,3	44,9	22,9	24,7	0,0
Medio Bajo	9,9	1,4	37,9	13,6	52,8	31,4	32,6	0,0
Medio Alto	13,5	2,2	59,6	29,3	68,0	44,7	47,3	0,0
<b>Estrato socioeconómico 2007</b>								
Muy Bajo	6,0	0,6	12,5	2,5	34,3	14,6	17,1	27,9
Bajo	8,3	1,2	22,8	6,3	44,6	22,8	24,5	0,0
Medio Bajo	9,6	1,7	40,4	15,1	54,7	32,7	34,2	0,0
Medio Alto	13,6	2,3	60,8	30,1	68,2	45,2	47,8	0,0
<b>Estrato socioeconómico 2008</b>								
Muy Bajo	6,4	0,7	12,0	2,4	34,1	14,5	17,0	28,1
Bajo	8,1	1,3	24,9	7,1	45,7	24,5	25,7	0,0
Medio Bajo	9,9	1,8	41,1	15,5	55,1	32,9	34,5	0,0
Medio Alto	13,5	2,4	60,6	29,8	68,2	45,1	47,7	0,0
<b>Estrato socioeconómico 2009</b>								
Muy Bajo	6,9	0,9	12,8	3,2	34,9	15,3	17,8	29,1
Bajo	8,9	1,5	25,1	7,3	45,9	24,7	25,9	0,0
Medio Bajo	10,5	2,8	41,4	15,8	55,4	33,2	34,8	0,0
Medio Alto	14,2	3,5	61,0	30,2	68,6	45,5	48,1	0,0

Fuente: EDSA, Observatorio de la Deuda Social Argentina. UCA.

Este mismo procedimiento se repitió para cada medición. Sin embargo, el factor para el cual se deben obtener los cuatro grupos no es estrictamente comparable.<sup>15</sup> Esto se debe a que los casos agregados con la ampliación de la muestra llevada a cabo a partir de junio de 2006 presentan capitales educativos y accesos tecnológicos (a nivel del hogar y del entorno residencial) superiores a las medias observadas de las mediciones anteriores.

Para superar esta dificultad, se optó por calcular el factor de ordenamiento socioeconómico de los hogares para las mediciones anteriores, agregando los casos adicionales en junio de 2007,<sup>16</sup> manteniendo constantes los valores observados en las variables de clasificación. Es decir, los valores exactos de junio de 2007 de capital educativo y de acceso a tecnologías del vecindario y del hogar fueron replicados para las mediciones anteriores. Este procedimiento permitió disponer de 1740 casos (considerando únicamente los aglomerados comparables) en todas las ondas. Luego de realizar la reclasificación de los casos de todas las mediciones, definiendo cuatro grupos a partir del cálculo de cuartiles, se eliminaron del análisis las observaciones agregadas.

<sup>15</sup> El problema de la comparabilidad no reside en el diferente número de casos de las diversas mediciones, sino que surge porque la muestra original no cubre todo el espectro socioeconómico, pues deja fuera un estrato (“clase media”).

<sup>16</sup> Se seleccionó 2007 como año base para este procedimiento debido a que se trata del primer año en el cual se incorporan las ciudades de Paraná y Rosario, además de agrandarse la muestra en otros aglomerados ya existentes.

Los puntos de corte que definen los cuartiles del índice socioeconómico de 2007 (percentiles 25, 50 y 75) fueron luego aplicados a los índices correspondientes al resto de los años. Mediante un ajuste del ponderador poblacional se devolvió la estructura cuartílica a las clasificaciones resultantes para los años con tamaño de muestra distinto a 2007.

**FIGURA AM.6**  
**DISTRIBUCIÓN DE LOS CASOS DE LA EDSA SEGÚN ESTRATO SOCIOECONÓMICO.**  
**Evolución 2004-2009**

	Serie histórica <sup>1</sup>						Muestra ampliada <sup>2</sup>			
	Año 2004	Año 2005	Año 2006	Año 2007	Año 2008	Año 2009	Año 2007	Año 2008	Año 2009	
Cantidad de casos										
<b>Total</b>	<b>1.100</b>	<b>1.100</b>	<b>1.500</b>	<b>1.740</b>	<b>1.740</b>	<b>1.740</b>	<b>0</b>	<b>2.130</b>	<b>2.130</b>	<b>2.130</b>
Muy Bajo	301	301	327	371	386	388	0	466	478	458
Bajo	308	302	337	390	378	362	0	459	459	455
Medio Bajo	225	231	326	398	379	353	0	466	445	428
Medio Alto	266	266	510	581	597	637	0	739	748	789
Distribución ponderada de los casos (sin ajustar - en porcentaje)										
<b>Total</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>0</b>	<b>100</b>	<b>100</b>	<b>100</b>
Muy Bajo	33	32	26	25	26	23	0	22	22	22
Bajo	30	32	24	24	23	21	0	22	22	21
Medio Bajo	19	20	21	25	23	20	0	22	21	20
Medio Alto	17	17	28	26	28	36	0	35	35	37

<sup>1</sup> Los resultados no incluye a la ciudad de Rosario.

<sup>2</sup> Los resultados incluye a la ciudad de Rosario.

Fuente: EDSA, Observatorio de la Deuda Social Argentina. UCA.

## AM.10 Variaciones temporales <tit. niv. 1>

Con el fin de estudiar la evolución temporal de los Índices de Desarrollo Humano y Social y de los indicadores de déficit, se analizaron los cambios netos a lo largo del tiempo. El cálculo de estos cambios se realizó para la serie histórica de valores comparables,<sup>17</sup> correspondientes a los valores de 2004 a 2007, sin incluir los aglomerados Gran Rosario y Paraná, que fueran incorporados durante la última medición.

Para determinar la importancia de los cambios temporales en cada serie de resultados se calcularon dos tipos de diferencias. En el caso de los Índices de Desarrollo Humano y Social, se calcularon diferencias relativas.<sup>18</sup> Para los indicadores utilizados en la construcción de estos índices, se aplicaron diferencias absolutas<sup>19</sup> entre las proporciones observadas en cada momento. En cada caso, se analizaron dos tipos de variaciones en el tiempo:

<sup>17</sup> Ver punto AM.4 de este mismo anexo.

<sup>18</sup> Las diferencias relativas en el tiempo se calculan como  $\left( \frac{\text{Valor}_t - \text{Valor}_{t-k}}{\text{Valor}_{t-k}} \right) \cdot 100$ , donde  $t$  representa el período actual y  $k$  la cantidad de períodos transcurridos entre los dos momentos del tiempo que se quieren evaluar.

<sup>19</sup> Las diferencias absolutas en el tiempo se calculan como  $\text{Valor}_t - \text{Valor}_{t-k}$ .

<viñeta>

- *Variaciones interanuales*: muestra la variación (aumento o disminución) en el valor entre un año y el siguiente, a lo largo de todo el período analizado.
- *Variaciones respecto del año base*: junio de 2004 es la medición utilizada como Línea de Base para la EDSA. Por este motivo, se calculan las variaciones ocurridas entre este año y el último año de la serie.

<fin de viñeta>

Con el propósito de determinar la importancia de las variaciones temporales en los niveles de prevalencia se recurrió a una prueba de significación. En el caso de los Índices de Desarrollo Humano y Social se aplicaron diferencias de medias. En el caso de los indicadores de privación, se utilizaron diferencias de proporciones. Si bien el proceso coherente con el resto del desarrollo del análisis hubiera sido la realización de pruebas para muestras relacionadas,<sup>20</sup> las limitaciones del software utilizado en términos del cálculo de la covarianza entre las dos series de datos a partir del módulo de muestras complejas, impidió que se llevara a cabo esta prueba. En su reemplazo se utilizaron las pruebas para muestras independientes descriptas anteriormente.

#### AM.11 Curvas de incidencia <tit. niv. 1>

Las curvas de incidencia se utilizaron en los análisis de los *índices* y en el de los *ingresos de los hogares*. En cada uno de los gráficos presentados en cada caso, el eje vertical representa la medida de incidencia correspondiente a cada puntuación o valor. De manera más concreta, dado un puntaje cualquiera  $I_X$  (dentro del rango posible) cada curva representa el porcentaje de personas u hogares que posee una calificación igual o menor a  $I_X$ . De esta manera, se puede observar de forma inmediata la incidencia relativa a todos y cada uno de los puntajes, sin la necesidad de establecer un umbral determinado. Asimismo, dado un límite o puntuación definido por algún criterio (teórico o empírico), este tipo de gráficos permiten realizar la evaluación crítica de la efectividad del valor seleccionado.

Las curvas de incidencia permiten, a su vez, la comparación de los niveles de incidencia en dominios establecidos; en el caso de este informe: *períodos y estratos socioeconómicos*. La condición de dominancia de primer orden indica, para un rango determinado de puntuaciones o valores, que la mayor ocurrencia relativa de un dominio es más elevada (o más baja) si su curva de incidencia se halla siempre por arriba (o por debajo) de la curva de otro dominio.

---

<sup>20</sup> Las muestras son relacionadas porque se evalúan, en cierta proporción, los mismos sujetos u hogares en dos momentos del tiempo.