APÉNDICES

APÉNDICE I:
ANÁLISIS METODOLÓGICO APLICADO A LA ENCUESTA DE LA DEUDA SOCIAL ARGENTINA

La Encuesta de la Deuda Social Argentina (EDSA) contiene un cuestionario multipropósito que se aplica sobre una muestra probabilística de población adulta de 18 años y más, estratificada de manera no proporcional por nivel socioeducativo de las unidades residenciales de grandes áreas metropolitanas del país. En cuanto al universo geográfico del estudio, la muestra es representativa de dos tipos fundamentales de centros urbanos: 1) el Área Metropolitana de Buenos Aires (AMBA), y 2) las Áreas Metropolitanas del Interior del país con más de 200 mil habitantes (Ciudades del Interior). En cuanto al universo social objeto de estudio, la estratificación socioeconómica se efectuó a partir de la clasificación de conglomerados residenciales de hogares según el perfil educativo predominante de los jefes de hogar en las unidades censales. De esta manera, quedaron clasificados cinco espacios residenciales socioeducativos (ERS): ERS Muy Bajo, ERS Bajo, ERS Medio Bajo, ERS Medio y ERS Medio Alto.

Respondiendo a una estrategia de investigación comparada de tipo longitudinal (seguimiento de casos panel), hasta el momento la EDSA fue aplicada en cuatro oportunidades. Las mediciones se realizaron con una periodicidad anual durante el mes de junio de 2004 (línea de base), 2005 y 2006. A seis meses de 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.

La cantidad total de casos fue de 1.100 en las tres primeras mediciones, y de 1.500 en junio de 2006. A partir de estos casos, se obtuvo información acerca de algo más de 4.400 miembros de los hogares relevados en las primeras mediciones. En la última medición, la ampliación de la muestra permitió alcanzar a un total de 5.676 personas. En todos los casos, el marco muestral utilizado para el diseño de la muestra y la selección de los casos 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.

El procesamiento de los datos y su análisis estadístico se realizó tomando en cuenta la naturaleza compleja de la muestra, utilizando el módulo de muestras complejas del SPSS en su versión 13.

AI.1 Diseño muestral estratificado por criterios socio-residenciales y regionales

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, para 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 radios censales).

- *Regiones Metropolitanas:* se consideró a partir de la diferencia de dos grandes sistemas socio-económicos urbanos. Por una parte, el Área Metropolitana del Buenos Aires o AMBA (incluyendo en la misma a la Ciudad de Buenos Aires y a 24 partidos del conurbano bonaerense) y, por otra, las principales ciudades del interior del país, aquellas con más de 200 mil habitantes, tomadas en esta etapa de investigación de manera global (Gran Córdoba, Gran Salta, Gran Resistencia, Gran Mendoza, Bahía Blanca y Neuquén-Plottier.)

- *Estratos Socioeconómicos Residenciales:* se definió según las características socioeducativas de las unidades censales como medida de proximidad a una estratificación socioresidencial 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 y un estrato de comparación de clase media-alta.

En función de la estratificación residencial se utilizó como variable de estratificación muestral principal el porcentaje de jefes de hogar con educación secundaria completa por radio censal (datos del Censo Nacional de Población, Hogares y Viviendas 2001 provistos por el INDEC). A partir de estudios previos y aplicación de análisis de correlación con variables como el NBI (necesidades básicas insatisfechas), porcentaje 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 radios de siete 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. Habiéndose aplicado este criterio, la estructura socioeconómica pudo quedar 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 ponerle un criterio de diferenciación socio-cultural a este segmento frente a los sectores de clase media más tradicionales.

Figura AI.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 radio censal		**Segunda condición:** % de hogares con horno a microondas
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. 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 radio censal y región metropolitana (Área Metropolitana del Gran Buenos Aires y Ciudades del Interior). Los recursos disponibles determinaron, para las tres primeras mediciones, un tamaño de muestra de 1.100 casos. A partir de esta cantidad de observaciones, se procedió a la selección de 184 puntos residenciales de muestra. 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.

La estratificación según criterio de aglomeración y estrato socioeducativo de los puntos de muestra fue uniforme: se asignaron la mitad de los casos al AMBA y la otra mitad a las Ciudades del Interior, distribuyéndose 300 casos para cada estrato muestral. Los casos fueron distribuidos proporcionalmente según la participación de cada área metropolitana en cada estrato muestral (1).

A continuación se presentan tres cuadros que 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 junio de 2006.

Figura AI.2a: Distribución de la población de 18 años o más representada por la EDSA, según espacio residencial socioeducativo y conglomerado urbano.

Junio de 2006

	ERS Muy bajo	ERS Bajo	ERS Medio bajo	ERS Medio	ERS Medio alto	TOTAL
AMBA	**1 793 920**	**2 166 391**	**1 217 719**	**731 790**	**2 308 012**	**8 217 832**
Ciudades del Interior	**300 964**	**696 017**	**681 602**	**200 387**	**494 993**	**2 373 963**
Gran Córdoba	102 082	247 493	296 698	88 946	231 095	966 314
Gran Mendoza	89 906	159 785	145 088	48 815	122 704	566 298
Gran Salta	21 291	100 986	77 534	31 744	53 802	285 357
Gran Resistencia	49 561	72 478	61 572	11 296	24 097	219 004
Bahía Blanca	20 964	67 849	62 565	5 039	42 089	198 506
Neuquén	17 160	47 426	38 145	14 547	21 206	138 484
TOTAL	**2 094 884**	**2 862 408**	**1 899 321**	**932 177**	**2 803 005**	**10 591 795**

Figura AI.2b: Distribución de los puntos muestra de la EDSA, según espacio residencial socioeducativo y conglomerado urbano.

Junio de 2006

	ERS Muy bajo	ERS Bajo	ERS Medio bajo	ERS Medio	ERS Medio alto	TOTAL
AMBA	**25**	**25**	**25**	**25**	**25**	**125**
Ciudades del Interior	**25**	**25**	**25**	**25**	**25**	**125**
Gran Córdoba	8	9	11	11	11	49
Gran Mendoza	8	6	5	6	6	31
Gran Salta	2	4	3	4	3	16
Gran Resistencia	4	3	2	1	2	12
Bahía Blanca	2	2	2	1	1	9
Neuquén	2	2	1	2	2	9
TOTAL	**50**	**50**	**50**	**50**	**50**	**250**

Figura AI.2c: Distribución de los casos relevados por la EDSA, según espacio residencial socioeducativo y conglomerado urbano.

Junio de 2006

	ERS Muy bajo	ERS Bajo	ERS Medio bajo	ERS Medio	ERS Medio alto	TOTAL
AMBA	**150**	**150**	**150**	**150**	**150**	**750**
Ciudades del Interior	**150**	**150**	**150**	**150**	**150**	**750**
Gran Córdoba	47	52	64	66	66	295
Gran Mendoza	47	35	32	36	37	187
Gran Salta	11	22	18	24	18	93
Gran Resistencia	26	16	14	6	10	72
Bahía Blanca	10	14	14	6	7	51
Neuquén	9	11	8	12	12	52
TOTAL	**300**	**300**	**300**	**300**	**300**	**1 500**

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

AI.2 Procedimiento de selección de los casos de la muestra

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 seleccionados mediante un sistema de cuotas de sexo y edad. En general, para cada punto de muestra quedaron asignadas 6 viviendas (una unidad de observación por vivienda). Seleccionado el individuo, se le aplicó un formulario personal, relevándose además características de la vivienda y de los demás miembros del hogar, identificando a cada uno de los miembros 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, la segunda a la probabilidad de que el radio seleccionado sea o no suplente y la última 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 (2) (Muiños, 2005).

AI.3 Estudio de datos de panel, supervivencia y rotación de casos

Dado los objetivos del Barómetro, entre los cuales se incluye el estudio dinámico de trayectorias, en cada onda de aplicación de la EDSA se buscó mantener en la muestra a un panel formado por los mismos respondentes y hogares relevados en la línea de base. Pero como es común en este tipo de estudios, entre cada medición tuvo lugar una pérdida de casos con una rotación aproximada del 25% de los mismos. Los principales motivos registrados fueron por rechazo, fallecimiento o ausencia prolongada del informante o mudanza de hogar. Los casos perdidos fueron reemplazados a través de una selección por muestreo sistemático de otra vivienda en el mismo punto muestral, seleccionándose al interior de la misma, como respondente, al individuo perteneciente al mismo grupo de edad y sexo del caso perdido.

La EDSA de control logró la reentrevista del 75% de los casos relevados en la línea de base. La encuesta aplicada en junio de 2005 tuvo un saldo de sobrevivientes de 65% de casos, respecto de la primera onda. Si se consideran los casos reentrevistados en 2005 que habían participado en alguna de las dos ondas de 2004, la supervivencia fue del 74%. En la última medición, el 52% de los casos comparables (1.100) fueron encuestados también en 2004 y 2005.

La siguiente figura da cuenta del porcentaje de casos por espacio residencial socioeducativo que se han mantenido en la muestra a lo largo de los cuatro relevamientos.

Figura AI.3: Panel superviviente de los casos relevados por la EDSA a lo largo de los cuatro relevamientos, según espacio residencial socioeducativo.

Junio 2004 a Junio de 2006

	Junio 2004	Diciembre 2004	Junio 2005	Junio 2006
ERS Muy Bajo	300	238	198	185
% sobre linea de base	100.0	79.3	66.0	61.7
ERS Bajo	300	241	206	164
% sobre linea de base	100.0	80.3	68.7	54.7
ERS Medio Bajo	300	228	207	152
% sobre linea de base	100.0	76.0	69.0	50.7
ERS Medio Alto	200	116	102	68
% sobre linea de base	100.0	58.0	51.0	34.0
Total	**1100**	**823**	**713**	**569**
% sobre linea de base	**100.0**	**74.8**	**64.8**	**51.7**

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

AI.4 Consistencias internas e imputación de ingresos de no respondentes

Una etapa previa a la construcción de indicadores y su análisis, involucró la consistencia de los datos obtenidos por el relevamiento de la EDSA. En primer lugar, se llevaron a cabo controles dentro de cada cuestionario con el propósito de detectar inconsistencias en las repuestas de los individuos, especialmente en lo referido a bloques temáticos. 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 el 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. En cambio, como se observa en la siguiente figura, no es significativa la variación en el tiempo del porcentaje de no respuestas.

Figura AI.4: Ingresos familiares de los hogares e ingresos laborales de los ocupados.

Junio de 2004 a junio de 2006

	Junio de 2004	Junio de 2005	Junio de 2006
Ingresos familiares			
Porcentaje de no respuestas[1]	12.5	16.6	13.6
Promedio de ingresos relevados	725	888	1 256
Promedio de ingresos estimados	731	908	1 331
Ingresos laborales			
Porcentaje de no respuestas[2]	17.7	19.3	23.5
Promedio de ingresos relevados	604	695	1 095
Promedio de ingresos estimados	597	699	1 105

[1] Porcentaje calculado sobre el total de hogares.
[2] Porcentaje calculado sobre el total de ocupados.

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

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 respondentes 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 y otra para el ingreso 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 de considerar la recepción de asistencia en forma monetaria o no monetaria de parte de organismos públicos y privados.

AI.5 Clasificación de los casos para la presentación de los resultados

En el presente informe, se ha optado por presentar una clasificación de los datos en términos de un criterio de clasificación doble. (3)

Por un lado, se aplicó un criterio regional, clasificándose los casos según el conglomerado urbano donde estos fueron relevados. En particular, se subdividieron los resultados entre la región Metropolitana de Buenos Aires (AMBA) y las grandes Ciudades del Interior del país (4). El segundo criterio de clasificación buscó medir el Estrato Socioeconómico del Hogar encuestado, a partir de considerar los activos principales del mismo en dos niveles: atributos propios del hogar y atributos del entorno residencial donde éste se encuentra localizado. Para el hogar se utilizó la información de la EDSA, empleándose los datos brindados por el encuestado acerca de sus miembros y de la vivienda en la que habitan. Para el entorno residencial, se optó por las características del vecindario en donde se ubica el hogar mediante la identificación de las características del radio censal al que éste pertenece (la fuente de datos utilizada fue el Censo de Población de 2001).

La construcción del estrato socioeconómico implicó una serie de decisiones metodológicas. En particular, el esquema clasificatorio propuesto procuró integrar los atributos del hogar y del vecindario en materia de dotación de capital físico y educativo. Para ello, las variables utilizadas para la definición de estos aspectos fueron:

- **Capital educativo del hogar:** fue definido operativamente como el promedio de años de educación del hogar, medido a partir del nivel educativo del jefe, cónyuge y demás integrantes con más de 24 años de edad.

- **Capital físico del hogar:** fue definido operativamente como un indicador compuesto de las características de la vivienda. Los aspectos considerados son cuatro: la propiedad, la calidad de los materiales, el equipamiento doméstico y la cantidad de miembros del hogar en relación a los

ambientes de la vivienda. Estos cuatro aspectos se combinaron para formar un índice de capital físico del hogar, cuyos valores más altos corresponden a las viviendas en mejor situación relativa y los menores a las viviendas precarias, con tenencia irregular, mal equipadas y con espacio habitacional insuficiente.

- **Capital educativo del vecindario:** fue definido operativamente como el porcentaje de jefes de hogar con educaciónsecundaria completa o más en el radio censal de localización del hogar.

- **Capital físico del vecindario:** fue definido operativamente como el porcentaje de hogares en viviendas clasificadas como casas tipo A (5) o departamentos, en ambos casos con equipamiento de horno de microondas en el radio censal de localización del hogar.

La alternativa de clasificación seleccionada fue combinar estas variables por el método de análisis factorial. Mediante la aplicación de la técnica de componentes principales se extrajo el primer componente de las variables consideradas. A partir de los valores de este primer factor se generaron cuartiles. Estos cuatro grupos ordenados se caracterizan por representar niveles crecientes de capital educativo y de capital físico, incluyendo tanto el capital individual como el capital del entorno residencial donde habitan los individuos y hogares que forman parte del estudio (6) (Maletta, 2006). Los cuatro grupos resultantes quedaron clasificados en: 1) Estrato Muy Bajo, 2) Estrato Bajo, 3) Estrato Medio Bajo y 4) Estrato Medio Alto.

Figura AI.5: Variables utilizadas para la clasificación según estrato socioeconómico.

Junio de 2006

	Muy bajo	Bajo	Medio bajo	Medio alto	**Total**
Capital educativo del vecindario: Porcentaje de hogares cuyo jefe tiene secundario completo o más en el radio censal	12.8	24.3	46.5	78.8	**40.7**
Capital físico del vecindario: Porcentaje de hogares en casas tipo A o departamentos con microondas en el radio censal	13.9	20.0	34.7	59.7	**32.1**
Capital educativo del hogar: Años de educación promedio de los miembros del hogar mayores de 24 años	6.6	8.6	10.6	13.3	**9.8**
Capital físico del hogar: Índice de capital físico del hogar	0.8	1.2	1.5	1.7	**1.3**

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

Este procedimiento supone una mayor facilidad a la hora de replicar la clasificación obtenida en los relevamientos previos. Para esto se aplicaron para las ondas anteriores los mismos puntos de corte

de la estructura cuartílica obtenida en la EDSA de junio de 2006. Por lo tanto, el procedimiento utilizado implicó un agrupamiento no fijo de los hogares del panel, ya que se trasladaron hacia atrás los puntos de corte estimados para el primer componente y no la pertenencia de clase observada en junio de 2006.

Al trasladar a las mediciones anteriores los puntos de corte que definen los estratos socioeconómicos, surgió el problema de que el factor para el cual se deben obtener los cuatro grupos no es estrictamente comparable (7). En particular, esto se debe a que los casos agregados con la ampliación de la muestra llevada a cabo en junio de 2006 presentan –por definición- capitales educativos y físicos (a nivel del hogar y del entorno residencial) significativamente 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 400 casos adicionados en junio de 2006, manteniendo constantes los valores observados en las variables de clasificación. Es decir, los valores exactos de junio de 2006 de capital educativo y físico del vecindario y del hogar fueron replicados para las mediciones anteriores. Este procedimiento permitió disponer de 1500 casos comparables en todas las ondas. Aplicándose los puntos de corte que definen los cuartiles de junio de 2006, se obtuvieron cuatro grupos comparables en cada medición. Luego de realizar la reclasificación de los casos de todas las ondas, definiendo cuatro grupos en cada medición a partir de la replicación de los puntos de corte, se eliminaron del análisis las 400 observaciones agregadas.

Este tipo de procedimiento supone que los grupos construidos constituyen cuartiles, en tanto que dividen a la muestra (ponderada) en cuatro partes iguales, únicamente para el relevamiento de junio de 2006. En las mediciones iniciales existe una menor representación, en términos de casos relevados, de los estratos socioeconómicos más altos (véase la siguiente figura).

Figura AI.6: Distribución de los casos de la EDSA según estrato socioeconómico.

	Junio de 2004	Junio de 2005	Junio de 2006
Total	**1 100**	**1 100**	**1 500**
Muy Bajo	339	318	348
Bajo	325	345	349
Medio Bajo	276	283	489
Medio Alto	160	154	314

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

Pero la aplicación de estos procedimientos impacta en la comparabilidad de los resultados a nivel agregado entre ondas. La estimación de valores comparables para la incidencia general de un determinado atributo, así como para los subtotales por conglomerado urbano, se ve afectada por las diferencias muestrales entre ondas. Para resolver este problema se aplicó a las ondas anteriores -siempre para las estimaciones agregadas- la estructura de ponderaciones correspondiente a junio de 2006. Esto implica que, en todas las ondas, los totales y subtotales se obtuvieron mediante el promedio ponderado de los valores de cada estrato socioeconómico.

AI.6 Análisis de incidencia, desigualdad y variaciones netas y brutas

El diseño de la EDSA permite la disponibilidad de un conjunto de datos de panel. Esta característica tiene importantes consecuencias, ya que admite realizar tres tipos de análisis que fueron aplicados en este informe con objetivos complementarios.

En primer lugar, se analizó la incidencia de los indicadores, tanto en forma general como de manera diferencial para cada criterio de clasificación (estratos socioeconómicos y conglomerados urbanos). En función de esto se calcularon tasas de recuento, que se computan de manera muy simple como el cociente entre la cantidad de casos que poseen determinado atributo y el total de casos observados. En concreto, se realizó la estimación de las proporciones de incidencia de cada indicador para el total de los casos, además de consignarse los valores correspondientes a cada conglomerado urbano y cada estrato socioeconómico.

Por otra parte, dado que uno de los objetivos principales de este estudio consiste en evaluar la magnitud de las desigualdades existentes entre estratos socioeconómicos y entre conglomerados urbanos, se procuró generar una medida comparable de la magnitud de tales diferencias. En esta oportunidad, se utilizaron dos formas paralelas de medir desigualdades:

(a) Una primera medida de las brechas existentes es el riesgo relativo, un ratio (razón o cociente) entre las incidencias de cada indicador para cada grupo. El riesgo relativo como fue definido en esta oportunidad compara la incidencia de determinado indicador en uno de los grupos con la incidencia del mismo indicador en otro grupo. En particular, se compara el estrato socioeconómico Muy bajo con el Medio alto, por un lado, y AMBA con las Ciudades del Interior, por el otro. El riesgo relativo 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. 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 estructura socioeconómica (o el conglomerado urbano) y el indicador analizado. Valores mayores que 1 (RR>1) suponen una asociación directa

(positiva) entre la estructura socioeconómica (conglomeración urbana) y el indicador analizado, mientras que valores menores que 1 (RR<1) sugieren una relación inversa (negativa) entre la estratificación socioeconómica (conglomeración urbana) y el indicador medido. (8)

(b) Una segunda forma de evaluar las desigualdades en las incidencias de los indicadores de la Deuda Social es a partir de las diferencias relativas. Esta medición de la desigualdad consiste en la simple sustracción del porcentaje de incidencia del indicador en el estrato Muy bajo (AMBA) menos el del Medio alto (Ciudades del Interior). De esta manera, se obtiene una medida absoluta de la diferencia entre un grupo y el otro. Por ser absoluta, está medida en las mismas unidades de las tasas que compara (en este caso, puntos porcentuales). Esta medida será igual a cero (DR=0) cuando la incidencia del indicador en cada estrato sea la misma, lo cual implica una ausencia de asociación entre la estructura socioeconómica (la conglomeración urbana) y el indicador analizado. Valores mayores que cero (DR>0) sugieren una asociación directa entre el estrato socioeconómico (la conglomeración urbana) y el indicador pertinente, mientras que valores menores que cero (DR<0) suponen una relación inversa entre el estrato socioeconómico (la conglomeración urbana) y el indicador.

Al mismo tiempo, con el objetivo de determinar la importancia de las diferencias entre las incidencias según estratos socioeconómicos o conglomerado metropolitanos, se aplicaron pruebas de significatividad. La prueba elegida para la evaluación de la significatividad de las diferencias es la prueba de diferencia de proporciones para muestras independientes. 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. Además, las categorías de los grupos utilizados para la clasificación, estratos socioeconómicos y conglomerados metropolitanos, son independientes entre sí. 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 distintos grupos. El software proveyó la estimación de las proporciones y sus correspondientes desvíos estándar, utilizados para el contraste de las hipótesis. (9)

Con el fin de estudiar la evolución temporal de los indicadores, se analizaron los cambios netos a lo largo del tiempo. Para la mayor parte de los indicadores se presentó la evolución temporal de los valores a partir de las mediciones de junio de 2004, junio de 2005 y junio de 2006. Sin embargo, en algunos casos las preguntas utilizadas para la construcción del indicador no figuraban en alguna medición, lo cual obligó a presentar la evolución incompleta (10). Para determinar la importancia de los cambios temporales en la incidencia de los indicadores se calculó la diferencia absoluta entre las proporciones observadas en cada momento. En particular, se analizaron tres diferencias:

- **Variación 2004-2006:** muestra la variación (aumento o disminución) absoluta total en el porcentaje del indicador entre los dos extremos del período analizado (variación punta a punta).

- **Variación 2004-2005:** muestra la variación (aumento o disminución) absoluta de la incidencia del indicador durante el primer año del análisis.

- **Variación 2005-2006:** muestra la variación (aumento o disminución) absoluta de la incidencia del indicador durante el segundo año del análisis.

Con el propósito de determinar la importancia de las variaciones temporales en los niveles de incidencia se recurrió a una prueba de significatividad. Si bien el proceso coherente con el resto del desarrollo del análisis hubiese sido la realización de una prueba de diferencia de proporciones para muestras relacionadas (11), 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 utilizó la prueba para muestras independientes descripta anteriormente.

Para completar el análisis fue necesario estudiar los flujos de cambio en la incidencia de los indicadores. Este tipo de exploración implicó un análisis de los casos panel, conformado por aquellos individuos que continuaron en la muestra en más de una EDSA. Los indicadores que se elaboraron en esta instancia tienen la categoría de transiciones temporales. Básicamente, se trabajó con transiciones que muestran el cambio entre dos momentos del tiempo. Siempre que los datos estuvieran disponibles, éstas se definieron para representar el flujo interanual de los individuos entre situaciones en las cuales se presentaba el atributo medido por el indicador y situaciones en las cuales no se presentaba tal atributo. La construcción anual permitió, por un lado, evitar posibles problemas de estacionalidad en los datos, además de admitir el transcurso de un período suficiente para observar cambios significativos. De acuerdo a la combinación de las mediciones quedaron definidos tres tipos de transiciones:

- **Transición 2004-2006:** momento de inicio junio o diciembre de 2004, momento de llegada junio de 2006.

- **Transición 2004-2005:** momento de inicio junio o diciembre de 2004, momento de llegada junio de 2005.

- **Transición 2005-2006:** momento de inicio junio de 2005, momento de llegada junio de 2006.

En todos los casos, las unidades se clasificaron conforme a la situación observada en el momento de inicio y en el momento de llegada:

(a) **Se mantuvo sin presentar el atributo:** el individuo no presentó el atributo medido por el indicador ni en el momento de inicio ni en el de llegada.

(b) **Dejó de presentar el atributo:** el individuo presentaba el atributo medido por el indicador en el momento de inicio, pero dejó de presentarlo en el momento de llegada,

(c) **Empezó a presentar el atributo:** el individuo no presentaba el atributo inicialmente, pero empezó a presentarlo en el momento de llegada,

(d) **Se mantuvo presentando el atributo:** el individuo presentó el atributo medido por el indicador en ambos momentos.

Para determinar si cualquiera de estas situaciones ocurría de manera diferencial en alguno de los estratos socioeconómicos se aplicaron pruebas de diferencia de proporciones para muestras independientes, de la misma manera que la descripta anteriormente. Para la determinación de las probabilidades de transición entre las distintas situaciones se calcularon tasas específicas de cambio, relacionadas a los procesos de entrada y salida. En todos los casos las diferencias entre el estratos socioeconómico muy bajo (por ser el de mayor vulnerabilidad) y el resto de los estratos se midieron utilizando las mismas medidas antes mencionadas, cuya significación se estimó con pruebas de diferencias de proporciones.

Notas

(1) Durante las tres primeras mediciones los casos correspondientes al ERS Medio Alto fueron menores (200 en lugar de 300) y 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, a la vez que equiparó la cantidad de casos relevados en cada estrato muestral. 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 en Junio de 2006.

(2) 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 versión 13, utilizados para el cálculo de los estimadores y las pruebas de significancia.

(3) La clasificación de los casos utilizada para los análisis realizados en los informes anteriores tuvo base en la estratificación muestral. Es decir, los casos fueron presentados en los tabulados y demás procedimientos desagregados según su espacio residencial socioeducativo de pertenencia.

(4) La cantidad de casos relevados, determinada a partir de restricciones presupuestarias, imposibilita un análisis representativo y significativo de los datos con una desagregación geográfica menor.

(5) El Censo de Población de 2001 define a una Casa tipo A como aquella vivienda con salida directa al exterior (sus habitantes no pasan por pasillos o corredores de uso común) construida originalmente para que habiten personas. Este tipo de vivienda no tiene condiciones deficitarias (se consideran condiciones deficitarias: tener piso de tierra o ladrillo suelto u otro material, no tener provisión de agua por cañería dentro de la vivienda o no disponer de inodoro con descarga de agua).

(6) 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.

(7) El problema de la comparabilidad no reside en el diferente número de casos de las diversas mediciones (1100 para las dos primeras, 1500 para 2006), sino que surge porque la muestra original no cubría todo el espectro socioeconómico, pues dejaba fuera un estrato ("clase media"). Hay además cambios en el número de casos del espacio residencial Medio alto.

(8) 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 la interpretación del mismo, 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.

(9) 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_x 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_o, 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 (5%).

(10) En los casos en los que no existía la medición de la Línea de Base, se utilizaron los valores de la medición de control (diciembre 2004) en su lugar. En otros casos, únicamente figuraban los valores correspondientes a la primera y a la última medición.

(11) Las muestras son relacionadas porque se evalúan los mismos sujetos en dos momentos del tiempo.