

M1/137

E.2

CONFERENCIA DE ESTADISTICOS
INTERNACIONALES DE LAS AMERICAS

PERIODO DE SESIONES 3/8 OCTUBRE, 1983
BUENOS AIRES - ARGENTINA



045-18/VIII/83-140
Original: Español



SISTEMA DE ESTADISTICAS SOCIALES Y DEMOGRAFICAS

SESD:

La generación de información de
corrientes a partir de Encuestas
a los Hogares

Documento preparado en el Instituto Nacional de Estadística y
Censos (INDEC) por Alejandro RUPNIK

M1/137
782



INDICE

INTRODUCCION

A - Antecedentes de la investigación.

1. Los primeros trabajos en torno al SESD
2. Avances a nivel de América Latina.
3. Antecedentes en la Argentina.

B - Las Encuestas a los Hogares como fuente de información para el SESD

1. Encuesta Permanente de Hogares (E.P.H.)
2. Implementación de la Planilla de Identificación y Seguimiento.
3. Tratamiento de información básica, ciclo de seguimiento.

C - Método de obtención de información dinámica.

1. Construcción de la Estructura de información básica.
2. Generación de archivos unificados.
3. Esquema matricial.

CONCLUSIONES



INTRODUCCIÓN

En el Instituto Nacional de Estadística y Censos (INDEC) se vienen desarrollando desde algunos años tareas tendientes a la implementación del Sistema de Estadísticas Sociales y Demográficas (SESD) propuesto por Naciones Unidas.

La fuente de información básica utilizada es la Encuesta Permanente de Hogares (EPH). Dicha encuesta, por su temática cubre un amplio espectro de indicadores sociodemográficos. Por otra parte, el sistema de rotación permite contar con el 75% de viviendas comunes entre dos relevamientos sucesivos, en esto estriba la posibilidad de obtención de información dinámica.

El nivel alcanzado actualmente en lo atinente a la mecanización del método de obtención de flujos, ha permitido lograr un archivo unificado de todas las variables contenidas en el cuestionario familiar e individual.

A través del CENT'S se están generando una serie de tablas que cubren todo el espectro al que lo habilita la información básica, derivándose de ello una visión dinámica que resulta ser sumamente útil para el conocimiento de los cambios que se producen a nivel social, brindando a su vez un enfoque integrado de dichos cambios a través del tiempo.

A. Antecedentes de la investigación

A.1. Los primeros trabajos en torno al Sistema de Estadísticas Sociales y Demográficas (SESD).

La preocupación en años recientes, de integrar los aspectos sociales, políticos y económicos en el análisis del comportamiento de grupos humanos llevó a dedicar especiales esfuerzos a precisar un enfoque que permitiese una descripción sistemática e integral, con un criterio cuantitativo, de los aspectos más importantes del comportamiento de los individuos, aislada o colectivamente, en un determinado marco social.

Con el objeto de satisfacer dicha preocupación es que en su XV período de sesiones, la Comisión de Estadística de las Naciones Unidas se propuso comenzar a trabajar en la elaboración de un proyecto sobre un sistema integrado de estadísticas demográficas, sociales y de fuerza de trabajo que mostrara sus vinculaciones con el Sistema de Cuentas Nacionales.

A consecuencia de esta iniciativa, en mayo de 1970 se publica el primer documento sobre el tema, preparado por R. STONE, en su carácter de consultor de las Naciones Unidas, titulado "Un sistema integrado de Estadísticas de la Población, de los Recursos Humanos y otras materias sociales, y sus vinculaciones con el Sistema de Cuentas Económicas Nacionales."^{1/}

1/ Naciones Unidas, Consejo Económico y Social, E/CN.3/394

Este primer esfuerzo sería completado luego con una serie de trabajos, alguno de los cuales se mencionarán a continuación. A principios de 1971 se publica el trabajo Sistema de estadísticas de la población, de los recursos humanos y otras materias sociales: Series, clasificaciones e indicadores sociales ^{2/}, que trata principalmente la identificación del tipo de información que podría compilarse. ^{3/}

Al mismo tiempo aparecen otras investigaciones, entre las cuales cabe mencionar las efectuadas en los países escandinavos sobre sistemas de datos individualizados, como así también el trabajo de R. STONE Demographic accounting and model building. ^{4/}

A mediados de julio de 1973, se publica Towards a System of social and demographic statistics que resume los trabajos anteriores encarados por la Oficina de Estadísticas de las Naciones Unidas tomando como base el mismo, en el XVIII período de sesiones de la Comisión de Estadística se aprueba la versión definitiva del Sistema de Estadísticas Sociales y Demográficas. ^{5/}

A.2. Avances a nivel de América Latina

En cuanto a los esfuerzos realizados en la región, debe mencionarse en primer término la reunión efectuada en Santiago de Chile en 1972, auspiciada conjuntamente por la

^{2/} Naciones Unidas, Consejo Económico y Social ST/STAT/49, 1971.

^{3/} Una revisión de los trabajos anteriores apareció como Naciones Unidas Consejo Económico y Social Sistema de Estadísticas demográficas y Sociales y sus vinculaciones con el sistema de Cuentas Nacionales E/CN.3/432, 1972.

^{4/} OECD, París, 1971.

^{5/} Hacia un sistema de Estadísticas sociales y Demográficas ST/ESA/ATAT/SER F/18, Nueva York, 1975

CEPAL y la Oficina de Estadísticas de las Naciones Unidas. Se presentó allí el documento elaborado por J. SORROUILLE Sistema de Estadística Sociodemográficas. Un ejemplo ilustrativo.^{6/}

A esta le siguió la efectuada en México en octubre de 1974, y que contó con el auspicio de la Escuela Latinoamericana de Sociología (ELAS) y el Centro Latinoamericano de Demografía (CELADE). Su objetivo fue iniciar la discusión sobre los sistemas de organización de la información, analizando las posibilidades que podrían tener para la región enfoques ya propuestos en otros ámbitos. Se constituyó allí el grupo de trabajo sobre información sociodemográfica. Un compendio de los trabajos presentados fue publicado en 1977.^{7/} El grupo tuvo su segunda reunión en Asunción en 1976.

La tercer reunión se realizó en Lima del 21 al 25 de mayo de 1979; estuvo principalmente consagrada al tema "Los Censos de Población y Vivienda en la década de 1980 en América Latina". Un resumen de las ponencias y comentarios fue publicado en 1981.^{8/}

A.3. Antecedentes en la Argentina

En el Instituto Nacional de Estadística y Censos durante 1971 se comenzó a realizar un estudio acerca de los objetivos, contenido y posibilidades de implementación del SESD tomando como referencia los primeros documentos sobre

^{6/}Naciones Unidas, CEPAL ST/ECLA/ Conf. 44/L5, 1972

^{7/}Consejo Latinoamericano de Ciencias Sociales: Hacia un Sistema integrado de Estadísticas en América Latina, México, 1977.

^{8/}Consejo Latinoamericano de Ciencias Sociales: Los Censos de Población y vivienda en la década de 1980 en América Latina, Buenos Aires 1981.

el tema.

En una primera etapa los trabajos abordados estuvieron vinculados a la realización de experiencias referentes a los distintos aspectos relativos al enfoque y funcionamientos propuestos para el SESD, con el propósito de lograr una mayor comprensión y evaluación del mismo.

Estas experiencias iniciales se abordaron a partir de datos sociodemográficos básicos trabajándose con cruces múltiples y obteniéndose la información de una única fuente estadística. (ver punto B.1)

Estos primeros pasos se concretan parcialmente y dan origen al documento Sistema Integrado de Estadísticas Socio demográficas y Sociales: Algunas experiencias preliminares en la aplicación del Sistema.^{9/}

En una segunda etapa se abordaron distintos temas con vistas a su implementación concreta.

A través de los trabajos encarados se intentó profundizar en un triple objetivo:

- 1) El funcionamiento de las secuencias
- 2) La modalidad de trabajo y relaciones entre distintas unidades de cuenta
- 3) La experiencia práctica que permitiese trabajar en la complementación entre diversas fuentes estadísticas

en éste sentido se realizaron dos informes en los cuales se plasmaron dichas inquietudes.

Uno de los trabajos estuvo orientado hacia el área educativa, enfocándose desde la perspectiva explicitada en punto 3), el otro informe al cual se hace referencia se circunscribe a la secuencia de hogares respondiendo a los puntos 1) y 2) mencionados precedentemente; ambos trabajos se presentaron en SESD: Aportes para el avance en su aplicación.^{10/}

A partir de la información contenida en la Planilla de Identificación y Seguimiento (para un mayor detalle ver punto B.2) se implementó la investigación sobre Secuencia de Hogares, SESD: Secuencia de Hogares ^{11/}, que centra su análisis en las características demográficas de los mismos y de sus miembros.

En una tercera etapa, en la cual se trabajó con información proveniente de la "Planilla de Vuelco Preliminar", conocida internamente con el nombre de (R₂) se produjo un documento que analiza algunos aspectos del subsistema de ocupación. En el mismo se estudia la población en cuanto a su intervención en el mercado de trabajo. Aquí se analiza la movilidad entre categorías ocupacionales entre ramas de actividad y entre las distintas condiciones de actividad, poniendo énfasis en las posibilidades analíticas que brinda la información dinámica. SESD: Secuencia Activa, Posibilidades Analíticas.^{12/} Cabe destacar que el sistema mecánico de obtención de las matrices de movilidad utilizado no resultaba lo suficientemente flexible cuanto era deseable; en el punto C se explicita la metodología implementada y que satisface ampliamente los objetivos perseguidos.

^{10/}INDEC, Buenos Aires 1975

^{11/}INDEC, Buenos Aires 1976

^{12/}INDEC, Buenos Aires 1979

B. Las Encuestas a los Hogares como fuente de información para el SESD

Las encuestas a los hogares pueden ser una fuente de información de gran importancia para el SESD por dos aspectos fundamentales. Uno de ellos es el que da la posibilidad de obtener información integrada dado que dichas encuestas indagan sobre características básicas y que hacen a distintas áreas temáticas (Hogar, Trabajo, Educación, etc.). El otro es el que brinda la posibilidad de obtener información dinámica, ya que en general en este tipo de relevamientos se suele utilizar un diseño muestral de tipo rotativo, donde los grupos de rotación permanecen en el panel por un tiempo (t) o reentran al mismo más adelante (t+n); $n = 1, 2, \dots, r$. Esto permite, en el caso de tener identificados a los hogares e individuos, obtener información sobre cambios a través de la comparación entre dos momentos distintos en el tiempo.

B.1. Encuesta Permanente de Hogares (E.P.H.) ^{13/}

El Instituto Nacional de Estadísticas y Censos (INDEC) realiza a través de la E.P.H. investigaciones continuas en base a una muestra de hogares particulares en 28 centros urbanos de todo el país, seleccionados en forma aleatoria a través de la cual se investigan diversos aspectos económico-sociales prioritarios, lo que hace de ella una encuesta de objetivos múltiples.

Esta encuesta fue concebida como instrumento de concreción de un sistema integrado de indicadores sociales. Desde su temática se rescatan aquellos atributos básicos que indican las condiciones socio-económicas de la población más significativas.

^{13/} Encuesta Permanente de Hogares: Marco teórico y metodológico de la investigación temática, INDEC, Buenos Aires 1980.

Los campos que abarca son:

- a) Características demográficas
- b) Características ocupacionales
- c) Características migratorias
- d) Situación habitacional
- e) Características educacionales
- f) Distribución de ingresos

Los atributos que se quieren obtener de la población son inherentes a los individuos, por lo tanto el individuo constituye la unidad elemental de análisis, siendo la vivienda la unidad de listado.

La posibilidad concreta de generar un mecanismo para obtener la información dinámica, se deriva del sistema de rotación que tiene el panel de la muestra. El mismo se efectiviza de forma tal que de una encuesta a otra, un 25% de las viviendas son reemplazadas. Para describir el mecanismo de rotación resulta ventajoso ejemplificarlo. Supóngase que la muestra inicial consta de 100 unidades de viviendas divididas en 4 grupos de 25 viviendas cada uno, que se identifican con las letra A, B, C, y D. El siguiente diagrama muestra entonces como de una onda a la otra hay un grupo que abandona el marco muestral y otro que se incorpora.

Cuadro N° 1

Ondas	Grupos de unidades de viviendas intervini- entes en la encuesta
PRIMERA	A B C D
SEGUNDA	B C D E
TERCERA	C D E F
	D E F G
n	E F G H

Es decir, que existe un traslape equivalente al 75% de las unidades entre dos ondas consecutivas.

B.2. Implementación de la Planilla de Identificación y seguimiento

Con la finalidad de explotar las posibilidades enunciadas precedentemente fue necesario realizar algunas adecuaciones al diseño del cuestionario familiar de la E.P.H., así como a las instrucciones para su llenado, las que fueron pensadas de forma tal que cada vivienda, cada hogar y, dentro de ellos, los individuos tengan un número identificatorio fijo a lo largo de todo el período en que participan en la encuesta.

Para ello se creó, como parte del mecanismo funcional de la encuesta, un elemento que cumple un papel fundamental: La planilla de Identificación y Seguimiento.

La misma contiene los datos referentes a Identificación y Ubicación de las viviendas, los hogares, como así también aquellos correspondientes a los individuos que lo componen. Es actualizada onda a onda anotándose en ella todas las modificaciones que a lo largo del tiempo se presentan, de forma tal que se visualiza con más acento la última, ya que es a partir de ésta que se prepara el material que saldrá a campo.

Esta "Planilla de Identificación y Seguimiento" resulta ser un instrumento de suma utilidad para el control interno de la Encuesta.

B.3. Tratamiento de información básica, ciclo de seguimiento

El conjunto de las "Planillas de Identificación y Seguimiento" constituyen un archivo que permite contar con la información histórica necesaria a través de contener el hogar y su composición; esta planilla de Identificación y Seguimiento

to se genera simultáneamente con la incorporación de las viviendas que participan por primera vez (las que pertenecen a los grupos designados con las letras E, F, G y H, ver cuadro N° 1).

Previamente a la salida a campo se preparan en forma si multánea los cuestionarios familiares y la Planilla de Identificación y Seguimiento correspondientes.

Una vez realizada la tarea de relevamiento los cuestionarios familiares ingresan a gabinete (las tareas de gabinete pueden verse en el esquema N° 1).

Luego de las etapas de recepción y análisis, se encuentran en condiciones de ser utilizadas para cumplimentar en la Planilla de Identificación y Seguimiento, las variables que hacen a la identificación de las viviendas, hogares y componentes de las mismas.

Para la preparación de los Cuestionarios Familiares en una próxima salida a campo, habrá un grupo de viviendas nuevas, para las cuales se realiza el proceso anteriormente des cripto y un grupo de viviendas que ya han participado para las que son extraídas del archivo mencionado precedentemente las Planillas de Identificación y Seguimiento correspondientes y es en base a ellas que se preparan los mismos, cumplimentando las variables referentes a Identificación, Ubicación y Com ponentes del hogar, obteniéndose un cuestionario familiar por hogar.

Concluida la etapa de relevamiento, los cuestionarios fa miliares ingresan a Gabinete, efectuada la recepción y el análisis de los mismos, éstos se encuentran en condiciones de ser apareados con su correspondiente Planilla de Identificación y Seguimiento preexistente, es a partir de ese momento que con la utilización de instrucciones específicas se realiza la tarea de Actualización.

Una vez concluidas las etapas descriptas puede decirse que se cerró un "Ciclo de Seguimientos". estando en condiciones de iniciar el siguiente. (Esquema N° 1).

X. Método de obtención de información dinámica

C.1. Construcción de la estructura de información básica.

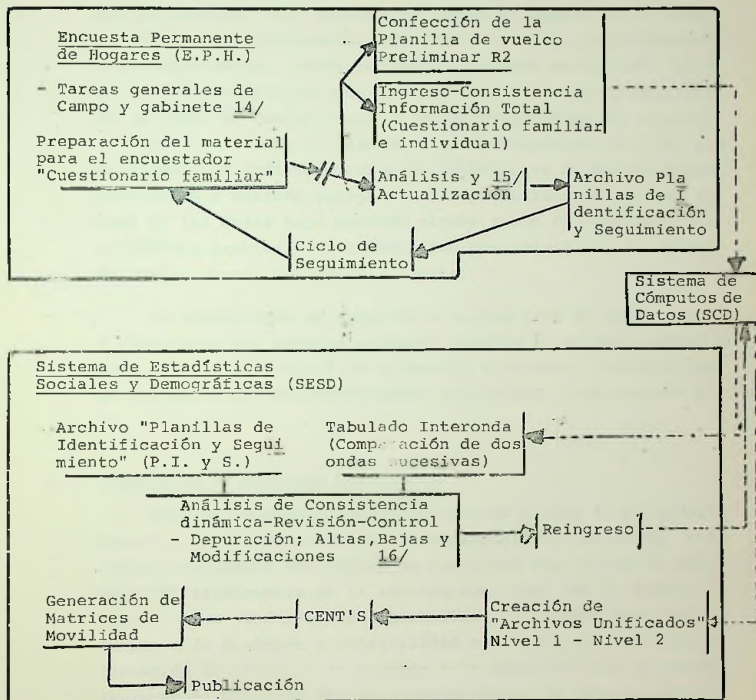
En continuidad con los trabajos ya realizados en el sentido de la obtención de información dinámica (flujos o corrientes) a partir del seguimiento de las viviendas en la E.P.H., es que fueron proyectadas una serie de etapas que permitieran finalmente la exploración exhaustiva de información sobre flujos en las distintas temáticas que abarca la encuesta a través del apareamiento automático de la información.

De acuerdo a lo antedicho se tuvo como objetivo prioritario la implementación de un conjunto de programas que conformasen un sistema lo suficientemente flexible a los efectos de generar matrices de movilidad.

Se comenzó con la tarea que hace a la creación del "Archivo Unificado de Nivel 1". Es aquel que surge a partir de la información contenida en la "Planilla de vuelco preliminar" (R2), ya que estas planillas tienen, por un lado la ventaja de ser cumplimentadas y perforadas en forma inmediata a la realización de la encuesta, y por otro lado, contienen los datos de identificación necesarios para su acoplamiento. Es así que previamente a la realización del mismo se genera un "Tabulado Interonda" que cuenta con la información correspondiente a dos ondas sucesivas a nivel de individuos y de hogares.

Este tabulado es la resultante de comparar entre dos ondas sucesivas la información contenida en cada una de ellas que no posee como contrapartida datos correspondientes a individuos u hogares en la otra, es decir está integrado

TRATAMIENTO DE INFORMACION BASICA Y OBTENCION DE FLUJOS



14/ E.P.H. Manual del Encuestador - Cuestionario Familiar o Individual, Códigos e instrucciones para la codificación de la información. INDEC, Buenos Aires 1974

15/ E.P.H. Planilla de Identificación y Seguimiento - Normas para su llenado actualización y utilización. INDEC, Buenos Aires 1974.

16/ SESD: Manual para la consistencia dinámica. INDEC, Buenos Aires 1975 (inédito)

en algunos casos por individuos aislados de los hogares a los que pertenecen y por hogares completos en otros. La no existencia de esta información como contrapartida, puede deberse a diversas causas, entre las cuales podemos mencionar: para el caso de individuos, que sean salidos; entrados u omitidos en la etapa de vuelco y por lo tanto no aparecen en alguna de las dos ondas. Para el caso de hogares completos, al igual que en el caso de individuos cabe la alternativa de que el hogar sea salido o entrado agregándose la posibilidad de que en alguna de las ondas haya mediado alguna razón de no entrevista; es posible también que la causa sea por omisión en la etapa de vuelco aunque ésta es muy remota.

La información es sometida a un análisis de consistencia y coherencia que permite corregir errores en algunas variables como sexo, edad y nivel de educación alcanzado, efectuándose el reingreso de las correcciones realizadas. (Ver esquema N° 1).

C.2. Generación de archivos unificados

Como se ha mencionado anteriormente y dado el nivel de desarrollo alcanzado en cuanto a la metodología utilizada para contar finalmente con tablas de movilidad fue necesario reformular el tratamiento de la información, dado que se estaba frente a dos limitaciones importantes, la primera de ellas referida a la dinámica y versatilidad necesaria para generar información de flujos y la segunda a la imposibilidad de profundizar en el análisis de las diversas áreas de interés social.

Por lo antes expuesto se generaron archivos de microdatos con las siguientes características: el primero de ellos llamado "Archivo Unificado de Nivel 1". Este reconoce como información básica la proveniente de la mencionada "Planilla de Vuelco Preliminar" (R₂), que cubre un número limitado de variables; las ventajas del mismo se basan en la inmediatez con que se cuenta con dicha información.

ESQUEMA N° 2

Esquema Matricial

Principio del año Θ Principio del año $\Theta + t$	Exterior o resto del mundo	Interior o país o zona ESTADOS INICIALES	Stocks o existencias finales
Exterior o resto del mundo	α	d'	
Interior o país o zona ESTADOS FINALES	b	S	Λ_n^t
Stocks o existencias iniciales		n'	

Los símbolos utilizados tienen el siguiente significado:

$\alpha_{1 \times 1}$ (escalar): total de individuos que han entrado y salido a la zona durante el período t ; por lo tanto no estaban al comienzo de Θ ni al comienzo de $\Theta + t$.

$b_{r \times 1}$ (vector columna): personas que entraron durante el período y permanecen al comienzo de $\Theta + t$.

$d'_{1 \times r}$ (vector fila): personas que estaban en la zona al comienzo del año Θ y salieron durante el período t .

t : indica el período utilizado, puede ser $t = 1$ si se tr
baja anualmente.

r : indica la cantidad de estados en consideración.

$S_{r \times r}$ (matriz cuadrada) $= (S_{ij})_r^{i,j=1}$: contiene a las personas sobrevivientes durante el período t , es decir aquellas que estaban al comienzo de θ y de $\theta + t$.

s_{ij} (elementos de S): cantidad de personas que al comienzo del período estaban en el estado j y al final se encuentran en el i .

$n'_{1 \times r}$ (vector fila): existencias iniciales en los distintos esta
dos, o sea distribución de la población al comienzo del pé
ríodo.

$A^{t n_{r \times 1}}$ (vector columna): existencias finales en los distintos es
tados, o sea distribución de la población al final del pe
ríodo.

En el segundo, denominado "Archivo Unificado de Nivel 2", la información contenida en el mismo resulta ser la suma de to das las variables que releva la E.P.II.

La configuración de los registros que constituyen ambos ar chivos consta de información atinente a identificación y datos referentes al momento (t) y (t+1), todo ello a nivel de indivii duo.

Precisamente esta configuración permite, a través del uso del CENT'S 17/, sortear la primera limitación a la que se ha h echo referencia precedentemente.

En cuanto a la segunda es obvio que al contar con la configuración de los registros a la que se alude anteriormente y que resulta imprescindible, puede avanzarse en el análisis de las diversas áreas temáticas que cubre la E.P.H. Por otra par te permitiría además, entre otras, la sistematización del estud dio de la no respuesta, el cálculo del sesgo posible por efect to de los casos caídos y la evaluación del comportamiento del instrumento de captación de información básica.

Para concluir y haciendo referencia al sistema de programas aludido en el punto C.1., cabe destacar que además de los que generan los archivos unificados lo conforman otros, a sa ber:

- Programa de consistencia y coherencia del proceso definit ivo.
- Programa para el cálculo de variancias (estimador para la variancia del estimador a nivel de totales) para las ciu dades con diseño probabilístico en dos etapas de selecc ción.
- Programa para la incorporación de variables al "Archivo Unificado de Nivel 1".
- Programa para ajuste de matrices por el método biproporc cional o Ras.

17/ Sistema de Tabulación de Censos. International Stadistical Programs Center. Bureau of the Census. Washington, Noviembre 1979.

C.3. Esquema Matricial

La forma de presentación de los datos a través de las matrices de movilidad permite visualizar de una manera simple la relación entre existencias iniciales y finales, a través de las corrientes que a lo largo del tiempo han tenido lugar, facilitando así la interpretación dinámica de la información (ver esquema N° 2).

CONCLUSIONES

- Las encuestas dirigidas a los hogares resultan una fuente de información importante en el área social si se toma en cuenta la diversidad de temáticas que abarca, así como el hecho que recaba información sobre los hogares e individuos en forma permanente e integrada. La importancia de esta fuente es aún mayor si se toma en consideración que no existen generalmente otras fuentes de información social con dichas características.
- La obtención de datos sobre flujos sociales a partir de las encuestas de hogares se presenta aquí como una posibilidad de gran interés, más aún por el hecho que esto es factible de ser realizado con costos mínimos.
- Sería aconsejable, por lo tanto, que los diferentes países volcaran esfuerzos en el estudio de esta posibilidad de obtener información dinámica, así como sobre los temas para los cuales los resultados a obtenerse serán relevantes.
- Sin embargo es conveniente tomar en cuenta la necesidad de adecuar dicha información a los fines propuestos. Al tratar de utilizar las encuestas de hogares como fuentes de información dinámica surgen una serie de inconvenientes que normalmente no son registrados cuando se trata a esta fuente como aislada en el tiempo.
- En definitiva, evaluando los distintos aspectos a que se ha hecho referencia, y dada la creciente preocupación de los países por disponer de una visión lo más ajustada posible de su realidad

social, el SESD adquiere las características de un proyecto de relevancia dentro de su área de aplicación.

- Por otra parte, la implementación del SESD apunta a proporcionar información dinámica e integrada y a facilitar la elaboración de indicadores sociales necesarios a los fines de la planificación económica y social, posibilitando, no sólo elaborar políticas sociales con mayores elementos de juicio, sino estudiar su impacto y evolución.