

## APÉNDICE

### A. Bases de datos

A continuación se presenta una breve descripción de los distintas bases de datos de que disponemos y de la información que puede obtenerse de ellas.

#### 1. Barro-Lee

Base de datos sobre la educación alcanzada en distintos niveles en la población masculina y femenina. Contiene estimaciones del nivel educativo de la población por edad (mayor de 15 años y mayor de 25 años) en 126 países del mundo. (Ver Robert Barro y J. W. Lee, "International Measures of Schooling Years y Schooling Quality", AER, Papers y Proceedings, 86(2), pp. 218-223, y también "International Data on Education", manuscrito.) Los datos se presentan para periodos quinquenales en el lapso 1960-1990.

#### 2. *Conteo de Población y Vivienda 1995*

Los datos proporcionados por el Censo provienen de dos fuentes: la Enumeración, donde se consideran todos los hogares del país, y la Encuesta, para la que se tomó una muestra de 80,000 hogares en el territorio nacional (2,500 por estado). La Enumeración proporciona datos sobre la distribución geográfica de la población y las características de los hogares en el nivel local, mientras la información de la Encuesta se refiere a familias, servicios sociales, servicios de salud, migración y muchos otros temas de nivel estatal.<sup>1</sup>

#### 3. ENCASEH

La Encuesta de Características Socioeconómicas de los Hogares, ENCASEH, fue el principal instrumento de obtención de datos de **PROGRESA** con el fin de identificar a las familias beneficiarias.

Diseño de la muestra: el universo de ese estudio se construyó mediante hogares/personas pertenecientes a las ciudades que de acuerdo con **PROGRESA** cumplían con las condiciones necesarias por presentar altos niveles de pobreza (según el "Censo" y el censo de la población de 1995), situadas en los estados de Guerrero, Hidalgo, Michoacán, Puebla, Querétaro, San Luis Potosí y Veracruz. Se usó un conjunto de muestras estratificadas y conglomeradas. Se reunió información de 506 localidades y 24,700 hogares. En tanto el principal objetivo del estudio fue identificar los resultados e impactos de **PROGRESA**, la muestra se dividió al azar en dos grupos: uno que fue

---

<sup>1</sup> En el caso de Chiapas hay algunos problemas; el Censo consigna un número de población menor al real debido a la imposibilidad de aplicar la Enumeración en ciertas áreas.

beneficiario del programa y sumó 320 localidades (tratadas) y otro que no, de 186 localidades (grupo de control). Los beneficiarios del programa fueron los hogares más pobres de las localidades tratadas.

Esta encuesta se llevó a cabo antes de que los hogares se incorporaran al programa. Se hizo una entrevista en todos los hogares de la localidad. La información que se obtuvo en la ENCASEH sirvió de punto de partida para la evaluación de las condiciones de pobreza de los hogares, y permitió la identificación de los hogares beneficiarios del programa.

La encuesta presenta información de diversa naturaleza sobre las características de los hogares, referida a los siguientes temas: identificación del jefe del hogar y de los responsables de los gastos y el cuidado de los niños; características demográficas del hogar y capital humano de sus miembros; actividades comerciales e ingresos de los miembros; condiciones habitacionales y disponibilidad de servicios; posesión de bienes. Hay también un cuestionario de localidad.

#### **4. ENCEL**

La Encuesta de Evaluación de los Hogares (ENCCEL) se elaboró como instrumento para evaluar **PROGRESA**. La evaluación consiste en hacer cada seis meses una encuesta de seguimiento en los mismos hogares donde se aplicó la ENCASEH. Y en tanto cada seis meses se reunió información sobre los hogares, se creó un panel de hogares.

#### **5. Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 1992, 1994, 1996, 1998, 2000, y 2002**

La ENIGH reúne información socioeconómica sobre hogares de manera periódica y sistemática. Esta información es representativa (si bien únicamente para el año 2002, de acuerdo con la clasificación de la CONAPO) de los niveles nacional, rural-urbano y de estratos marginales. La encuesta tiene como objetivo principal generar información sobre el estado actual de la estructura de ingresos y gastos; la estructura de ingresos y gastos financieros; el valor de bienes y servicios para autoconsumo; las características socioeconómicas de los miembros de los hogares; sus condiciones de trabajo, y las características del hogar. Para cada año, el proceso de muestreo se articuló en estratos, multietápico y por conglomerados. La unidad de muestreo final es el hogar junto con todos sus miembros. En cada etapa, la probabilidad de selección es proporcional al tamaño de la unidad de muestreo, por lo que es necesario usar factores de ponderación para obtener las estimaciones adecuadas.

#### **6. Muestra del XII Censo 2000**

Para este censo se usaron dos cuestionarios; el básico, que se aplicó a toda la población mexicana, y el aumentado (encuesta), que se aplicó sólo a una muestra. El diseño de muestra para la encuesta permite generar información de nivel municipal para la mayoría de los indicadores que se obtuvieron con el cuestionario aumentado. En cada municipio, el muestreo se hizo por conglomerados de etapa única, es decir, se seleccionaron áreas geográficas completas (unidades de muestreo primarias). Estas unidades varían según los tipos de áreas (rurales o urbanas) y en el diseño de muestra que se aplica a cada una. De este modo, las unidades de muestreo podrán ser AGEBs, bloques o localidades rurales.

Debido a los objetivos principales de la muestra, se calculó un tamaño de muestra de aproximadamente 700 hogares por municipio. Sin embargo, en los municipios con número igual o menor de hogares, se hizo un ajuste para “población finita”. De este modo, para Chiapas, Hidalgo, Jalisco y Veracruz el mínimo tamaño de muestra por municipio se fijó en 500 hogares; mientras que en Oaxaca, Yucatán, Puebla y Tlaxcala el mínimo se ubicó entre 150 y 250 hogares. En estos estados, la encuesta se aplicó a todos los hogares en los municipios con menos de 1,000 habitantes. Lo anterior dio como resultado que en todo el país se entrevistaran 2.2 millones de hogares.

Por último, dado el diseño del muestreo, hay que usar factores de ponderación para obtener cálculos confiables.

## **7. Anexo Estadístico del Informe del Presidente de los años 2001, 2002 y 2003**

En el Anexo Estadístico del Informe del Presidente se presenta una compilación de datos del Gobierno Federal. La información proviene de los registros administrativos de diversos gobiernos y secretarías. También se incluyen distintos indicadores de nivel estatal.

## **8. Encuestas de empleo: ENE, ENET y ENEU**

Las encuestas de empleo tuvieron dos etapas de diseño de muestreo. En la primera se reunió la información año por año y estuvieron representadas las áreas rurales y urbanas. Desde 1996, la Encuesta Nacional de Empleo, o ENE, y desde 1987 la Encuesta Nacional de Empleo Urbano, o ENEU, reúnen información sobre las características de la fuerza de trabajo en México en los niveles nacional y urbano, respectivamente. Estas encuestas son la base para el cálculo de las estadísticas de empleo oficiales. La ENE y la ENEU comparten las mismas variables y estructura. Ambas incluyen variables contextuales de nivel individual como educación, participación en el trabajo, horas trabajadas, tipo de contrato y beneficio recibido, algunas preguntas sobre cobertura de seguridad social o fuerza de trabajo y otras variables básicas como género, estado civil y edad.

La ENEU reúne los mismos conjuntos de información que la ENE, pero apunta únicamente a la población urbana; esto significa que la encuesta se restringe a las áreas urbanas con más de 100,000 habitantes. La cobertura geográfica de la ENEU fue cambiando con la inclusión de nuevas áreas urbanas, y actualmente cubre las 45 mayores áreas urbanas de México. La mayor ventaja de la ENEU es que la encuesta se basa en un panel de individuos a los que puede hacerseles un seguimiento por cinco trimestres. La ENEU está estructurada como un panel rotativo. En cada muestra trimestral hay cinco cohortes, cada una en una etapa diferente de completación del ciclo de entrevistas: un quinto de la muestra en su primera entrevista, un quinto en su última (quinta) entrevista, y tres quintos de la muestra en etapas intermedias. La ENEU condujo entrevistas cuatrimestrales muy extensas en las 16 mayores áreas metropolitanas entre 1992 y 2000. La muestra se selecciona de manera que resulte representativa geográfica y socioeconómicamente. El cuestionario cubre ampliamente temas como participación en el mercado de trabajo, salarios y horas trabajadas, que tradicionalmente se encuentran en este tipo de encuestas de empleo. Al mismo tiempo, una variable de identificación de hogar permite la construcción de ingresos de hogares.

En la segunda etapa, a partir de 2000, la ENEU desaparece y la ENE incorpora las bases de datos de empleo tanto de la ENE como de la ENEU, se transforma en la ENET y tiene una estructura de panel que permite hacer el seguimiento de cada individuo durante cinco trimestres. Específicamente, desde el segundo trimestre de 2001, la ENE consiste en una base de datos transversal y es representativa de los siguientes niveles agregados, mutuamente excluyentes:

- a) Nivel nacional.
- b) Cuatro tamaños de localidades (estos niveles permiten dividir la muestra en áreas rurales y urbanas).
- c) 48 grandes ciudades.
- d) 32 entidades federativas mexicanas.

ENET es la sigla correspondiente a la nueva versión de la ENE, que a principios de 2000 se transformó en una encuesta trimestral (Encuesta Nacional de Empleo Trimestral).

## **9. Polity IV**

El proyecto Polity evolucionó desde la década de los setenta y se volvió una fuente muy usada de datos longitudinales y transversales sobre las características de la autoridad de las formas de gobierno modernas. Se le usa mucho por sus evaluaciones del grado de democracia y autocracia de las estructuras políticas de los Estados-nación. Reúne información anual codificada sobre régimen y características de autoridad de todos los países independientes (con una población mayor de 500,000 habitantes) en el sistema global de naciones y cubre el lapso 1800-2002. El sitio web de Polity IV tiene sede en el

Centro para el Desarrollo Internacional y Manejo de Conflictos (*Center for International Development and Conflict Management*) de la University of Maryland, College Park.

#### **10. Conjunto de datos de investigación gubernamental Instituto del Banco Mundial (IBM)**

Este conjunto de datos presenta cálculos de seis dimensiones de gobierno, y cubre 199 países y territorios durante cuatro periodos: 1996, 1998, 2000 y 2002. Estos indicadores se basan en cientos de variables individuales que miden las percepciones sobre el gobierno, tomadas de 25 fuentes de datos distintas elaboradas por 18 organizaciones diferentes. Estas mediciones particulares de gobierno se asignan a categorías que designan dimensiones fundamentales de gobierno, y el modelo de componentes no observados se usa para construir seis indicadores agregados de gobierno en cada uno de los cuatro periodos. Las estimaciones puntuales de las dimensiones de gobierno se presentan junto con los márgenes de error para cada país en cada uno de los cuatro periodos. Los indicadores de gobierno presentados son una actualización y una expansión del trabajo previo del IBM parte de un proyecto de investigación sobre indicadores que se empezó en 1998 (Kaufmann, Kraay y Zoido-Lobatón 1999a,b y 2002).

#### **11. Indicadores de desarrollo mundial 2002**

El Banco Mundial recurre a una gran variedad de fuentes para producir las estadísticas publicadas en los Indicadores de Desarrollo Mundial. Los datos sobre deuda externa llegan al Banco Mundial directamente de los países miembro en vías de desarrollo a través del Sistema de Información de Deudores (*Debtor Reporting System*). Los otros datos se obtienen mayormente de la Organización de las Naciones Unidas y sus organismos especializados, del Fondo Monetario Internacional (FMI) y de informes nacionales para el Banco Mundial. Los cálculos del Banco también se usan para mejorar aspectos como la actualidad o la cohesión de los datos. Para la mayoría de los países, los cálculos de cuentas nacionales se obtienen de los gobiernos miembro mediante misiones económicas del Banco Mundial. La mayoría de los datos de fuentes nacionales se obtienen de archivos administrativos comunes, encuestas especiales o censos periódicos.

## B. Cálculo de Subsidio de Electricidad (RES)

**Cobertura.** El informe de la ENIGH de 1992 a 2002 se refiere al acceso a la electricidad (en 2002 también identifica consumidores que tienen un contrato formal), así como gasto de hogares en electricidad. Al contrario de lo que ocurre con los ingresos y los gastos totales, cuyos niveles la ENIGH sub-informa de manera significativa en comparación con las Cuentas Nacionales, la ENIGH *sobre-informa* mensualmente los niveles de gasto eléctrico en relación con los registros administrativos por una magnitud cercana a 2. Esto puede explicarse porque en la mayoría de los hogares se paga una cuenta de electricidad bimestral pero tal vez se olvida dividirla cuando se informa mensualmente sobre los gastos a la ENIGH.

**Gasto público.** Aplicamos el esquema de tarifa residencial que cubre la mayoría de los hogares y cerca de la mitad del gasto de electricidad residencial total (tarifa 1) para asignar un consumo de electricidad (kWh) a los hogares en base a su informe de gasto (divido por dos, por el motivo mencionado en el párrafo anterior). La gran mayoría del gasto y los hogares restantes corresponde a las tarifas 1A-1C, que son bastante similares a la última para que este cálculo también resulte válido en su caso. Para determinar los costos de abastecimiento consideramos que la tasa impositiva más alta para consumo residencial refleja costos marginales verdaderos, lo que implica un subsidio total para consumidores residenciales cercano a las estimaciones del Gobierno Federal. Para el periodo 1992-2000 esto implica un único esquema impositivo con tres tasas marginales progresivas, pero la reforma fiscal de 2002 implica tres esquemas diferentes, haciendo más difícil la determinación directa de los niveles de consumo de electricidad para este año. Por esta razón, para determinar la distribución del subsidio en el último año aplicamos las tarifas de 2002 a la estructura de consumo estimada para 2000 y consideramos que se había mantenido básicamente estable entre los dos años.