

Beyond the direct use of survey data for policy, household surveys provide much of the raw material for modeling and trying to understand household behavior. I shall discuss a number of such studies and their results as I illustrate the various techniques.

### *1.1.3. Survey design and its implications for analysis*

I shall use the "typical" household income and expenditure survey as my example, but many of the arguments can be applied to other types of surveys too. Such surveys typically collect data on a household basis — a household usually being defined as a group of people who share the same "cooking-pot" —and ask how much was spent over some reference period on a lengthy list of consumption items; the reference period can be anything from a day to a year, and sometimes varies by category of expenditure, with shorter recall periods for high frequency items like food, and longer periods for unusual purchases, like clothes or durable goods. In countries with near universal literacy, households can be asked to keep diaries; otherwise enumerators verbally ask respondents to recall individual purchases. Data are also collected on the respondents, at the very least covering the numbers, sexes, and ages, or people in the household. This can be extended to a range of household characteristics, such as education, occupation, and race. Data are frequently also collected on quantities consumed as well as expenditures, at least for readily measurable goods such as food. There will also be data on location, and perhaps more if the enumerators collect and retain data on the environment, for example on the size of the village, whether it has a school, and so on. Such surveys are sometimes carried out on an annual basis [Taiwan, Korea, India until 1973/1974 and since 1991], *but* more usually are done at intervals, often quinquennially, on the grounds that consumption patterns and levels of living and poverty do not change very quickly. The surveys are typically nationally representative, with each remaining in the field for a year, although there are also many special purpose surveys that are restricted in geographical coverage, and which last for a period shorter than a year.

Households are chosen at random, but there is a wide range of designs. The simplest is where each household in the country has an equal chance of being selected, but such simplicity is uncommon, if only because there are very different costs of obtaining *data* from different types of households. Rural households are more widely scattered than urban households, and in many LDCs, there are some households that live in inaccessible (sometimes even dangerous) areas. Any sample design that minimizes cost for a given degree of precision (or equivalently maximizes precision at given cost) will therefore lead to oversampling of urban and under sampling of rural households. Beyond this,