PUBLIC POLICY FOR THE
Privatesector

Note No. 212 June 2000

Infrastructure Reform, Better Subsidies, and the Information Deficit

Andrés Gómez-Lobo, Vivien Foster, and Jonathan Halpern

In developing countries the provision of water and sanitation services is often subsidized. These subsidies take the form of a general underpricing of water, numerous cross-subsidies, and inefficient billing and collection. An essential part of infrastructure reform is the redesign of subsidies. In the design of an optimal subsidy scheme the key decisions are the choice of eligibility criteria, the level of the subsidy, and the budgetary requirements. However, the lack of consistent and reliable data sets which combine socioeconomic and water consumption information may be an important obstacle to making good decisions, undermining efforts to provide affordable water services for the poor. This Note discusses the type of information required, where it can be found, and ways to deal with shortcomings in the data. To illustrate, the Note draws on data from World Bank work in Panama.

There are three types of information required to make informed choices on subsidy and other policies in the water sector. These are:

- *Willingness-to-pay data.* Willingness to pay is the maximum amount that a household would be prepared to spend to secure access to a given quantity of the service. Thus, in economic terms, it represents the limit of affordability of the service. A reasonable rule would be to set subsidies to cover the shortfall between a vulnerable household's willingness to pay for a basic level of consumption and the associated bill.
- *Water consumption data.* These describe the pattern of demand for different types of households. They can be used to establish the basic consumption level that will be subsidized. Water consumption data are also essential for establishing the distributional incidence of cross-subsidy schemes such as rising block tariffs.
- *Socioeconomic data.* These should ideally include household income or expenditure levels, poverty lines, and receipt of other welfare benefits, as well as general indicators of basic needs, such as the quality of housing and its associated facilities, and family wealth, such as durable goods ownership and the property value of the dwelling. This information is useful for determining the target population group for the subsidy, studying the targeting properties of different eligibility criteria (such as geographic location and wealth indicators) and determining the proportion of household income that is spent on water and sanitation.

The ideal data source would combine all three types of information for the same households. However, this ideal data set is rarely available, so the required information often has to be collated from different sources.

What sources?

One source of information is a water company's client database. This will typically contain monthly information on the number of customers, the value of their bills, the level of consumption for customers with water meters, the location of customers, and the tariff structure applicable to each customer. Client databases are generally the best source of information on

consumption levels, which can be used to create a frequency distribution of consumption patterns for metered customers, and to determine seasonal trends in consumption.

However, client databases have several drawbacks as a source of information. They do not provide information on the consumption patterns of informal connections. Neither are client databases very informative as to the consumption level of clients without meters. However, their most important drawback is the lack of data on the socioeconomic characteristics of each household. This limits their usefulness in studying the targeting properties of different subsidy designs.

Another very useful source of information is the data from household expenditure or socioeconomic surveys. These surveys are routinely undertaken by the statistical offices of most countries in order to update consumer price indices and to measure poverty levels or obtain other socioeconomic information. Household surveys are probably the best source of information on poverty levels, demographic characteristics, and general socioeconomic information. Many countries use some form of the Living Standards Measurement Study (LSMS) survey, which was developed by the World Bank in 1980 and has subsequently been adopted (with minor variations) in more than 20 developing countries, including Panama. But household surveys also have drawbacks: they are often outdated and the geographic divisions used may not match the population for whom the subsidy is being designed.

A third source is national censuses. These usually include a wide range of demographic, durable goods ownership and other socioeconomic information, and have the advantage over household surveys that they are representative at all levels of disaggregation, including very fine geographic divisions. However, census data have drawbacks too. They can be quite out of date (censuses are done every ten years or so). They usually measure income and expenditure poorly if at all. Finally, they provide no better information on water consumption and bills than household surveys.

Of the three types of data required for subsidy design, willingness-to-pay information is probably the most difficult to obtain. One possibility is to use a surrogate market approach. In this case, willingness to pay is inferred from the behavior of households in a market for a good that is a complement or substitute for public water supply. The obvious example is demand for non-piped sources of water such as bottled water or truck-based vendors. This method was used in the Dominican Republic to provide a crude first-cut estimate of the demand function for piped water services. Households without a connection to the public network were typically buying water for US$6.33 per cubic meter from water tankers, and were consuming approximately 7 cubic meters per household per month, an expenditure of US$44.3 per month. Households with a connection to the public network paid around US$0.13 per cubic meter for, on average, 37 cubic meters per household per month, an expenditure of US$4.93 per month. These two situations essentially describe two points on the demand curve for water, which can be used to make inferences about the overall demand curve and the corresponding willingness to pay.

Another alternative is to undertake a specific willingness-to-pay survey. These surveys use a "contingent valuation" approach to deduce users' back-of-the-envelope preferences for the service. Many such surveys have been done in Latin America.

If a willingness-to-pay survey is not undertaken, the final option is to use parameters taken from other regional or international studies of willingness to pay, or to use a benchmark parameter. It is striking that the results from a range of studies in different countries are comparatively consistent, with the willingness of poorer households to pay for piped water supplies invariably falling between 3 and 5 percent of household income. However, there is no guarantee that such rules of thumb will be applicable to the area under study.

The characteristics of these data sources are summarized in table 1. Most often, the ideal data-

TABLE 1 DATA SOURCES AND THEIR RELATIVE ADVANTAGES

Data source	Information provided	Advantages	Disadvantages
Water company client database	Consumption	Detailed and usually reliable	No information on informal connections No data on socioeconomic characteristics Absence of information on unmetered customers
Household expenditure surveys	Socioeconomic variables Access to water supply Expenditure on water	Very detailed data to make judgment of household's economic well-being	Information may be outdated May not be representative at the required geographical segmentation No data on physical water consumption Expenditure data may be unreliable
Census	Socioeconomic variables Access to water supply	Representative of all zones and geographical levels	May be outdated No income or expenditure data No data on physical water consumption
Willingness-to-pay surveys	Willingness to pay Consumption data Some socioeconomic data	Detailed control of information gathered Very flexible	Expensive to conduct Methodological problems
International parameters	Willingness to pay	Quick and inexpensive	Parameters may not be correct or relevant
Surrogate market studies	Willingness to pay	Quick and inexpensive	May not be very reliable Required primary data may not be available

base will not be available, so the required information will have to be collated from different sources. The strategy adopted in Panama, for example, was to use multiple sources of information and integrate them to the greatest degree possible. The 1997 LSMS survey was used as the basis of the analysis because it was the only source that combined data on socioeconomic conditions with data on water use. The client records of the water utility, Instituto de Acueductos y Alcantarillados Nacional (IDAAN), which contained information on the location of clients, the corresponding tariff structure, and the consumption each month (whether measured, estimated or imputed) was used to check the accuracy of the consumption levels that were inferred from the expenditure data in the LSMS survey. Finally, since neither of these sources had information on willingness to pay, a special contingent valuation survey was undertaken.

Dealing with incomplete data sets

Data sets are often incomplete. Creative data analysis can overcome this problem.

Inferring unmeasured consumption. The consumption of households without meters cannot be directly observed, yet it is often relevant for subsidy design, especially when metering coverage is expected to increase in the near future. It cannot be assumed that consumption in households without meters is the same as in households with meters. One approach is to estimate unmeasured consumption by examining the behavior of a client in the months after a meter is installed. It is reasonable to conjecture that the household will change its consumption gradually after the installation of the meter. The more randomly new meters are installed, the more representative of the market as a whole such figures will be. Data from Panama showed that the average monthly consumption for households facing the residential tariff declined by about 25 percent over the first four months of metering.

Inferring consumption from expenditure data. The LSMS survey used in the Panama case study provided direct information on how much households spent on water, not how much water they used. Moreover, the information on expenditure

was deficient in a number of respects. It did not say whether customers had a meter or not, nor detail the tariff structure under which they were paying. As a result, the estimate of household water consumption inferred from the LSMS survey diverged substantially from the consumption recorded by the water utility. Notwithstanding these deficiencies, an attempt was made to retrieve consumption information from the expenditure data contained in the LSMS survey by making some reasonable assumptions about the specific tariff that each household paid from among the many different tariff structures charged by the company. On this basis, water consumption could be estimated for each household (albeit imperfectly) by applying the corresponding tariff structure to the water expenditure figure.

Inferring consumption when data were missing. Socioeconomic data and the water consumption data often come from different sources and a statistically valid method for combining them is required. A useful method has been developed by Rajah and Smith (1993) in their work on the distributional impacts of introducing water metering in the United Kingdom. These authors had access to two data sets:

- A large-scale nationally representative household survey with a very rich set of information on income and other socioeconomic indicators, but scant information about water consumption.
- A small-scale detailed local study of water consumption with a very limited range of socioeconomic variables.

With the aim of incorporating a consumption variable into the household survey, Rajah and Smith estimate a water consumption function from the second data set based on socioeconomic variables that are common to both data sets. On the basis of this equation, it was then possible to predict consumption levels in the large-scale household survey. This method was used in Panama. About 17 percent of households in the LSMS survey stated that they did not pay their water bills and, hence, their reported expenditure was equal to zero. Taking the inferred consumption levels from that part of the sample which reported water expenditure, a consumption function was estimated on the basis of the numerous socioeconomic variables contained in the survey. The estimated equation was then used to predict the expenditure level of households with missing values for this variable.

Conclusion

This Note has summarized some methods of acquiring, or producing, the information needed to design a water subsidy scheme. Without such information, and a simulation of the probable effect of different subsidies, there is little guarantee that a subsidy system, however well intentioned, will have the desired effect. Once a suitable data set has been constructed, a simulation model can be created using simple spreadsheet software, providing figures on which a design can be based. Armed with this information, the policymaker should be in a position to design a subsidy program that reaches the intended beneficiaries, provides them with the level of financial support which is strictly necessary, meets the overall budgetary restrictions, and does not have excessive administrative costs.

References

Foster, Vivien, Andrés Gómez-Lobo, and Jonathan Halpern. Forthcoming. "Designing Direct Subsidies for Water and Sanitation Services: Lessons from Panama." Policy Research Working Paper. World Bank, Washington, D.C.

Gómez-Lobo, Andrés, Vivien Foster, and Jonathan Halpern. Forthcoming. "Information and Modeling Issues in Designing Water and Sanitation Subsidy Schemes." Policy Research Working Paper. World Bank, Washington, D.C.

Rajah, N., and S. Smith. 1993. "Distributional Aspects of Household Water Charges." *Fiscal Studies* 14(3):86–108.

Vivien Foster (vfoster@worldbank.org) and Jonathan Halpern (jhalpern@worldbank.org), Latin America and the Caribbean Region, and Andrés Gómez-Lobo (agomezlo@decon.facea.uchile.cl), Department of Economics, University of Chile

To order additional copies please call 202 458 1111 or contact Suzanne Smith, editor, Room F11K 208, The World Bank, 1818 H Street, NW, Washington, D.C. 20433, or Internet address ssmith7@worldbank.org. The series is also available on line (www.worldbank.org/html/fpd/notes/).

Printed on recycled paper.