The World Bank

PREMnotes

May
1998

NUMBER 5

Poverty

Using disaggregated poverty maps to plan sectoral investments

By combining survey and census data to create poverty maps that show where needs are greatest, policymakers can focus scarce development rescources.

A poverty map is a geographical profile that shows where poverty is concentrated in a country, and thus where policies might have the greatest effect on poverty. Many developing country policymakers use poverty maps when planning public investments in education, health, sanitation, water, transport, and other sectors. Social funds or education and health ministries often use poverty maps because geographically targeted investments are thought to reach many poor citizens and to have wide-ranging spillover effects in depressed areas.

A poverty map is most useful if it can be constructed at a fine level of geographical disaggregation. But achieving such fine levels of disaggregation requires very large datasets. It is rare to find survey data that both cover a large sample and provide detailed information on household welfare. In general, there is a tradeoff between size and quality because both goals are financially and administratively costly.

How can this obstacle be overcome? One way is to combine sample survey data with census data to predict consumption and poverty for all households covered by the census. Constructing a poverty map using census data but basing it on an ad hoc welfare indicator—for example, one that measures access to public services—can be risky. Thus this note describes an alternative approach that could help the design of poverty maps and improve the allocation of sectoral investment.

Basic needs poverty maps

Census data generally do not include income or expenditure information because collecting such data for an entire country would be too expensive. Thus policymakers have sought alternative welfare indicators on which to base their poverty maps. In Latin America, as well as in Africa and Asia, poverty maps used to rank regions have been based on indexes of welfare constructed by combining different information from the census, such as access to water, electricity, or sanitation, or education of the household head. This type of welfare indicator, often called a "basic needs" indicator, is generally constructed in an ad hoc manner and is limited by the qualitative nature of a lot of census data.

Individuals or households selected using a basic needs indicator can be quite different from those selected using a comprehensive consumption indicator. Ecuador's National Statistical Institute developed a basic needs indicator in 1994 in response to a government request to develop a directory of poor households to which to target compensatory transfers for an increase in the price of gas. The indicator was constructed at the household level and consisted of an index aggregating five variables—access to water, access to sanitation and waste disposal services, education of the household head, and a crowding index (the number of people per bedroom).

A poverty map is most useful if it can be constructed at a fine level of geographical disaggregation

A Living Standards Measurement Survey for Ecuador allows us to compare the basic needs indicator used by the institute with a consumption welfare indicator (table 1). For this experiment we selected the bottom 20 percent of the population according to both the basic needs indicator and the consumption indicator. An entry of 100 percent in the first quintile would indicate that the two indicators are perfectly matched. Yet only 41 percent of households identified by the basic needs indicator as making up the poorest 20 percent of households are among the bottom 20 percent using the consumption indicator. Thus, if the objective is to reach the poor as defined by consumption levels, the leakage of resources from an allocation based on this criterion would be quite high—nearly 60 percent of resources would go to unintended beneficiaries, with almost 10 percent going to the two richest quintiles.

The new approach can be used only if certain requirements are met

Combining survey and census data

One alternative to the arbitrary and imprecise basic needs indicator is to combine survey and census data. The basic idea is to develop structural models that characterize consumption as a function of housing, employment, household characteristics (such as size and composition), education variables, and the like based on the survey. The models—that is, the estimated parameters—are then applied to the census data to predict household consumption. Consumption is not in the original census dataset, but is artificially added using structural relationships derived from the survey. This approach can only be used if certain data requirements are met. First, a household survey that includes consumption and other household characteristics, such as Ecuador's Living Standards Measurement Survey, must be available, and should correspond to roughly the same period covered by the census. Second, unit record–level census data must be available. Third, a sufficient number of variables that are used to predict consumption must be available in both the survey and the census. Otherwise it would not be possible to use the models based on the surveys to predict census consumption data.

To impute expenditures using the census, the first step is to estimate a regression model of consumption using household survey data. As noted, the only variables that can be used to predict consumption are variables that are also available in the census. In Ecuador these potential predictors consisted of demographic variables such as household size and its age and sex composition, education and occupation information for each family member, housing quality data (materials, size), access to public services such as electricity and water, primary language spoken in the household, and household location. Models were derived for each subregion at which the Living Standards Measurement Survey achieved representativity.

The second step consists of applying the parameter estimates from the regressions to the census data. It is then possible to derive consumption estimates based on household characteristics from the census. These consumption data make it possible to estimate the probability of each household in the census dataset being poor. Since these estimates are derived from an (imperfect) model, an estimation insecurity must be taken into account. The incidence of poverty can be calculated as the mean, over all the households in a given region of the census, of the household-specific estimates.

How can the accuracy of the prediction model be assessed? Apart from conducting tests based on the survey itself—that is, judg-

Table 1 Distribution of the bottom 20% of households in Ecuador using a basic needs indicator, by consumption expenditure quintile

Actual per capita consumption, by quintile	*Beneficiary households*	
	Percentage	*Cumulative*
Bottom 20%	41.4	41.4
20–40%	29.5	70.9
40–60%	19.5	91.4
60–80%	8.0	98.4
Top 20%	1.6	100.0

Source: Hentschel and others forthcoming.

ing the fit of the model—one way is to compare the estimated incidence of poverty using census data with the incidence of poverty for the survey. Table 2 reports such comparisons for Ecuador for the regions at which the survey was representative.

Except for rural Oriente, which appears to be much poorer in the Living Standards Measurement Survey than in the census, the poverty rates from the census (with the predicted consumption variable) are quite similar to those from the survey. (Rural Oriente is a sparsely populated jungle area, so it does not substantially influence overall results.) Rankings of the eight regions are not identical across the two datasets, but both identify rural areas as being poorer than urban areas.

Disaggregated consumption-based poverty maps

The methodology outlined here allows the construction of a poverty map, based on consumption expenditures, at a level of disaggregation below the broad regions for which standard household surveys are suitable. For example, Ecuador has nearly 400 cantons, each with some local autonomy and administration, and these cantons can be divided into well over 1,000 *parroquias* (parishes). Working with the census data, it is possible to calculate canton or parroquia poverty rates to determine where poverty is concentrated.

Taken to an extreme, census data could, in principle, be used to identify individual poor households and to target transfers directly to these households. But this methodology does not allow for such microtargeting of beneficiaries. Because household consumption is only predicted, a fair amount of insecurity remains with each individual estimate. At a more aggregated level such errors can be compensated for by combining geographical targeting with other ways of identifying beneficiaries. But at the household level such errors are much more important. Moreover, these objections are in addition to the well-known arguments against microtargeting, which focus on the effect that such policies could have on the behavior and self-esteem of potential beneficiaries.

Despite the caution against microtargeting, it is undoubtedly useful to develop a poverty map that is disaggregated below broad regions. The optimal level of disaggregation depends on several factors. One is the purpose the poverty map is expected to fulfill. Is it, for example, intended to identify government administrative areas so that the desired level of disaggregation is some level of local government? Or is it intended to identify poor villages or neighborhoods so that community-level projects (such as public infrastructure) can be better targeted?

The optimal level of disaggregation depends on several factors

A second consideration is whether the parameter estimates from a regression model estimated at, say, the regional level, can be assumed to apply to subregions. Throughout this exercise it is implicitly assumed that, within a region, the model of consumption is the same for all households regardless of their province, county, or community. This assumption cannot be tested, and at fine levels of disaggregation it might be inaccurate. The desired degree of disaggregation will also depend on the

Table 2 Incidence of poverty in Ecuador's Living Standards Measurement Survey (LSMS) and census
(standard errors in parentheses)

LSMS rank	*LSMS poverty incidence*	*Census poverty incidence*	*Census rank*
1. Rural Oriente	0.67	0.53 (0.04)	1
2. Rural Costa	0.50	0.42 (0.04)	3
3. Rural Sierra	0.43	0.47 (0.03)	2
4. Guayaquil	0.29	0.28 (0.07)	4
5. Quito	0.25	0.28 (0.08)	5
6. Urban Costa	0.25	0.24 (0.09)	6
7. Urban Oriente	0.20	0.21 (0.11)	8
8. Urban Sierra	0.19	0.23 (0.09)	7

Source: Hentschel and others forthcoming.

availability of other information on poverty that might be available locally. Finally, other methods of local targeting, such as self-targeting, are more important and effective at certain levels of disaggregation.

Disaggregation is likely to lead to more sharply defined poverty profiles. In Ecuador a province-level poverty map showed that calculated poverty rates for the provinces within a region varied substantially. Thus disaggregation reveals poverty patterns not observed at a broader level.

Disaggregation is likely to lead to more sharply defined poverty profiles

Applications

The most useful application of this methodology is to plan sectoral investment by combining consumption-based poverty maps with other indicators of well-being, opportunity, and access. For example, a map documenting regional patterns of access to primary health care centers could be combined with a consumption-based poverty map. Such a map might help policymakers decide how to prioritize efforts to expand access to primary health centers—that is, health investment planning can give priority to the poorest areas with the lowest health care coverage. Furthermore, a close correlation between, say, regional patterns of rural poverty and road access might offer clues on possible causes of poverty. This type of exercise could be undertaken for a wide range of indicators, including health and education levels, ethnicity and indigeneity, access to infrastructure and other public services, and land quality and ecology.

This note was written by Jesko Hentschel (Economist, Poverty Group, PREM Network) and Peter Lanjouw (Economist, Development Research Group, Development Economics). It is based on a forthcoming Policy Research Working Paper by Jesko Hentschel, Jean Olson Lanjouw, Peter Lanjouw, and Javier Poggi called "Combining Census and Survey Data to Study Spatial Dimensions of Poverty."

PREM network
Poverty Reduction and Economic Management

This note series is intended to summarize good practice and key policy findings on PREM-related topics. PREMnotes are distributed widely to Bank staff and are also available on the PREM website (http://prem). If you are interested in writing a PREMnote, email your idea to Kim Murrell. For additional copies of this PREMnote please contact the PREM Advisory Service at extension 87736.

Prepared for World Bank staff