HOW WELL CAN METHOD SUBSTITUTE FOR DATA?
FIVE EXPERIMENTS IN POVERTY ANALYSIS

Martin Ravallion

No one doubts that good data are essential to sound policymaking. Alas, data are invariably faulty. Methodological solutions to data inadequacies have often been proposed and implemented, but they have been tested only rarely. Yet the methods that are used may well determine the direction of policy. For example, the particular survey method used—and the way nonsurvey data are interpreted—may be critical in assessing whether a country's strategy for reducing poverty is working. This article shows how counterfactual experiments can help test the reliability of various methods of dealing with common data problems. Well-designed methods—and they need not be very complicated—can help get around the problem, although it appears that substituting method for data is a long way from being perfect.

Objective data obtained from representative surveys of living conditions are widely used to stimulate public awareness of poverty and motivate government actions to benefit the poor. Yet analysts and policymakers routinely find that these data are deficient in one or more important respects and must find credible methods for dealing with those deficiencies. Various solutions have been proposed that rely on certain regularities in living conditions and use a relatively small number of more easily measured variables—such as membership in certain predefined socioeconomic groups—to infer the missing data.

Hidden differences in living standards may well confound such efforts. Even though the partitions commonly used in assessing poverty—such as land ownership or region of residence—reveal large disparities, these disparities may be weak indicators for targeting the poor. In fact, some recent research indicates that variations between socioeconomic groups are often dwarfed by differences within such groups (see, for example, Datt and Ravallion 1993; Ravallion and
Surprisingly little effort has gone into assessing the performance of routine tools of poverty analysis. Yet the methods that are chosen to deal with inadequacies in the data may well influence the conclusions analysts draw, including the implications for policy. For example, in research on Indonesia, it has been found that the method used to establish the poverty lines by region can dramatically alter the structure of the resulting regional poverty profile and hence the regional priorities for alleviating poverty; indeed, regional rankings produced by two common methods were virtually uncorrelated (Ravallion and Bidani 1994). Here the method chosen to make up for missing data would markedly influence the precise policy prescription.

This article evaluates some of these techniques by applying them in situations where the data are available and asking how well they work. Fortunately, not all countries are missing the same data, so it is possible to infer how a particular remedial method would work in a “data-poor” setting by comparing how well that same method performs in a “data-rich” setting. Knowing about the strengths and weaknesses of the many methods that have been used to generate information in data-poor settings can also help researchers make choices in collecting new data. Nationally representative socioeconomic surveys can be expensive; good substitutes, if they can be found, could produce savings.

Evaluating Methodology

There are three generic problems with existing poverty data. First, the specific surveys used are sometimes flawed; there may be problems with the questions asked or the sample selected. Second, it is often the case that too few surveys were conducted or they are incomparable over time. Third, researchers often lack complementary data that are needed to interpret the results; for example, making consistent welfare comparisons from survey data can be difficult if data on prices are not also available.

Ingenious—and sometimes simple—methods of dealing with these data problems can often be devised, although how well they work is unknown. One way to test how well the analyst can circumvent these data problems is to devise a counterfactual experiment. The objective is to try the method in a setting where it is not in fact needed—because the required data are already available—and see how well it performs. The following section outlines five such experiments that test how well methodology can substitute for data.

Several observations about these experiments are in order. First, their aim is not to develop “quick and dirty” methods; in some cases the econometric analysis that is used is quite sophisticated. Second, these experiments share the generic problems of any empirical research. That is, the results are to some extent
data-specific, and their robustness in other settings is not known. And measurement errors may be a problem, as always. The following experiments try to assess how well the selected methods perform against data that meet the prevailing standards of best practice but are unlikely to be perfect.

Third, even without measurement errors, reasonable people may disagree about what should be measured. Here the measure of poverty is defined as a lack of command over market goods. Although this measurement is undeniably important, it is clearly not the only dimension of well-being; command over nonmarket goods, such as some publicly provided services, may be an important omission in conventional measures (for further discussion, see Ravallion 1993). Nor do the experiments reported here constitute a comprehensive list of data problems. For example, Chaudhuri and Ravallion (1994) show the difficulties of inferring standards of living through the use of single cross-sectional surveys; and Haddad and Kanbur (1990) study the inequality within households that is often hidden in conventional surveys.

Experiment 1: A Poverty Profile with Missing Prices

Policy discussions about poverty are often informed by a poverty profile that shows how poverty measures vary by location, sector of employment, or some other household characteristic. Unfortunately, deficient data and methods often misinform such discussions. One of the most common—and potentially serious—problems is how to make consistent comparisons of what a given unit of money is worth to poor people in one place versus another.

Ideally, studies that compare incomes or expenditures over time or in different regions should adjust for differences in prices. Almost all statistical agencies now monitor prices over time in constructing the Consumer Price Index, but monitoring differences in prices by localities at a given time is far less common. This data gap can be a serious concern in developing countries where the prices of the same goods and services differ widely within the country. Without local price data, one might be tempted to ignore these differences in calculating whether people are poor. How would this omission affect the results? Is an alternative method available that would provide the missing information?

Indonesia is an interesting case study for examining these questions, given its geography and likely locational price differentials. For this experiment, researchers set out to measure the extent of poverty by region and sector of employment, with and without adjusting for differences in the cost of basic needs. Indonesia’s Central Bureau of Statistics provided the raw survey and price data. Using the method described in Ravallion and Bidani (1994), the nominal poverty line was adjusted for regional differences in the cost of a bundle of basic foods considered sufficient to meet minimum food-energy (caloric) requirements for good health and normal activities. An allowance for differences in nonfood spending was added as well, consistent with the spending behavior of the households involved. The results were then compared with the costs of the same items based
on average national prices without adjusting for regional differences in the cost of living.

The results are given in Figure 1, which ranks all provinces (each one split into urban and rural areas) by the head-count index (that is, the percentage of people who are living below the poverty line) based on nominal expenditures, and then plots the results against corresponding estimates with an adjustment for differences in the cost of basic consumption needs. So, if the ranking of regions does not change when adjusting for the cost of living, the lines are everywhere positively sloped from left to right; if there is also agreement on the cardinal values of the poverty measures, the lines increase linearly on a 45-degree line (indicated by the straight line in Figure 1).

It can be seen from Figure 1 that both the poorest urban area (point A) and the poorest rural area (point B) are correctly identified. But in several cases, adjustments for differences in the cost of living lead to a change in ranking. For example, the second-poorest rural area after adjusting for cost-of-living differences (point C) is only the eleventh-poorest without adjustment. The rank correlation between the head-count index using local poverty lines (with an adjust-

---

**Figure 1. Effect of Cost-of-Living Differences on Indonesia’s Poverty Profile by Province**

<table>
<thead>
<tr>
<th>Adjusted percentage of population below poverty line</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
</tr>
<tr>
<td>50</td>
</tr>
<tr>
<td>40</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>10</td>
</tr>
</tbody>
</table>

Note: Each point is one province (urban or rural). The rank correlation between the adjusted and unadjusted figures is 0.88. Point A is the poorest urban area; point B is the poorest rural area; point C is the second-poorest rural area after adjustment for cost-of-living differences.

Source: Author’s calculations using Indonesia’s National Socioeconomic Survey for 1990.
ment for prices) and that using the national mean poverty line (that is, based on national mean prices) is 0.88.

If the experiment is repeated using sector of employment instead of location (urban or rural), price adjustment has much less effect because prices do not vary by employment (figure 2). Also note that for the sectoral profile, if differences in urban and rural costs of living can be incorporated correctly, the results with and without the adjustment for other regional differences show considerable agreement.

It is interesting to compare these results with another method that is commonly used to compensate for missing price data when setting poverty lines, the "food-energy-intake method." Indonesia and many other countries use this method, which defines the poverty line in each region as the level of expenditure at which a predetermined level of food-energy intake is typically met in that region. This technique has the advantage that it does not require any data on prices. But because people in richer regions or sectors tend to have more expensive tastes, and so spend more to reach any given caloric intake, the method can yield differences in poverty lines far in excess of the cost-of-living differences of the poor (Ravallion 1993). Ravallion and Bidani (1994) compared the food-energy-intake method with a price-adjusted poverty profile for Indonesia and

---

**Figure 2. Effect of Cost-of-Living Differences on Indonesia's Poverty Profile by Employment**

![Graph showing the effect of cost-of-living differences on poverty profile by employment.](image)

*Note: Each point is a subgroup of households defined by principal sector of employment.*

*Source: Author's calculations using Indonesia's National Socioeconomic Survey for 1990.*

Martin Ravallion 203
found that the differences in poverty lines between urban and rural sectors implied by the food-energy-intake method were large enough to cause a reversal in all poverty measures between urban and rural areas and numerous reversals in the ranking of provinces. After adjusting for differences in the cost of basic consumption needs, Ravallion and Bidani found a greater incidence, depth, and severity of poverty in rural areas than in urban areas. The food-energy-intake method indicated the reverse, however, with more poverty in urban areas. The ranking of regions (with each province divided into urban and rural areas) by the two methods was very different.

These inconsistencies arise because food-energy intake is determined by many other factors besides real incomes, such as relative prices, tastes, and activity levels, that can influence the poverty lines obtained by the food-energy-intake method in ways that have nothing to do with differences in purchasing power over basic consumption needs. Group A may be able to afford more of all basic needs than Group B, and yet, because of the differences in the way the two groups allocate their money (say, group A buys more expensive calories because they taste better or are more convenient), the food-energy-intake method can find that group A is poorer (Ravallion 1993). The method solves the problem of missing price data, but at a high cost in terms of the usefulness of the resulting poverty profile for informing policy decisions.

Some methods of dealing with missing price data work far better than others. And in some cases the best practice turns out to be the simplest. In the experiment above, not adjusting at all for regional differences in the cost of living gave better predictions of which regions or sectors were poorest than did the more sophisticated food-energy-intake method. Even without data on prices by location, one can often make a reasonable guess as to how much more it would cost the rural poor to achieve the same standard of living in urban areas and to build this calculation into the poverty profile. This assumption can be based on a quick price survey that compares a few typical urban and rural areas.

Experiment 2: Monitoring Poverty Despite Gaps and Lags in Survey Data

There are often long lags between surveys, making it difficult to track a country's performance in fighting poverty. So developing a set of easily monitored, aggregate indicators of poverty that could serve as an alternative would be useful. An experiment in India compared the information provided by a set of aggregate indicators with survey data to see whether the findings from the two techniques were consistent.

This experiment used data supplied by India's National Sample Survey Organization, which has provided tabulations of the distribution of consumption per person since the mid-1950s. The surveys have been carried out at varying frequencies ranging from less than a year to more than five years. (For example, data were not collected between mid-1978 and 1983.) The consumption mea-
sures are comprehensive, and the surveys appear to follow sound and consistent practices. The National Sample Survey (NSS) has been the main instrument for monitoring poverty in India, and a large body of literature uses these data (reviewed by Ravallion and Datt 1995). Analysts have had to rely almost exclusively on the survey agency's own tabulations of the raw data, which typically are not published until two to three years after the survey. Only rarely have the raw data been released, and in those cases the lags between collection and release appear to have been even longer.

Given these lags and the gaps between survey rounds, the question is whether poverty in India can be monitored using more readily available leading indicators from other sources. The answer would also help determine how frequently survey data should be collected.

Some simple forecasting experiments based on NSS data can try to answer the question. The analysis focuses solely on the rural poor, whose income is derived from their own farms (which often do not produce enough for their own food needs) and from wages received from working on other farms. Thus two potential leading indicators for monitoring poverty are the agricultural wage rate and agricultural output per acre. Both are typically available well in advance of tabulations from the NSS data.

Poverty measures, such as the head-count index, show that poverty increased in rural India during two periods—the early 1970s and the late 1980s (Ravallion and Datt 1995). How well could data on average wages and farm output have predicted these setbacks for the poor? The forecasting model used here to answer that question is based on a model of rural poverty outlined in detail in Ravallion and Datt (1995). The salient features of the model for the current discussion are summarized in box 1.

The model tracks the data well, but good within-sample performance need not mean good post-sample forecasts. To estimate the accuracy of this methodology as a predictor, two types of forecasts were considered.

- **Type 1**: This experiment entailed making forecasts one year ahead, assuming that the lag in the availability of NSS tabulations on the distribution of per capita consumption is one year longer than the lag in the availability of data on the real agricultural wage rate and agricultural output per acre. The extra lead time for these two variables would make it possible to predict poverty measures before the NSS data are released.

- **Type 2**: This experiment comprises a sequence of dynamic forecasts, again assuming that real wage and agricultural output data are available but that the NSS data are only available at longer intervals; one must thus rely partly on past forecasts of the poverty rate. (The previous year’s poverty measure is one of the variables used to forecast the next year’s poverty measure; see box 1.)

Table 1 gives the results of three forecasting exercises for the head-count index. (The results of forecasting two other poverty indexes—the poverty gap and
A separate model was estimated for each of three poverty measures; the head-count index, showing the percentage of the population below the poverty line; the poverty gap index, reflecting the average distance below the poverty line; and the squared poverty-gap index, in which those distances are squared to make the measure sensitive to changes in inequality among the poor (Foster, Greer, and Thorbecke 1984).

The model was dynamic, in that the most recent past poverty measure was used as one of the predictors. Consumption-based poverty measures may not adjust instantaneously to current variables, so this feature is important. A nonlinear regression method was used to deal with the uneven spacing of the surveys over time (Ravallion and Datt 1995). The lagged poverty measure was a persistently significant predictor of the current measure.

The other two predictors were the real agricultural wage rate (deflated by the Consumer Price Index for Agricultural Laborers), and average agricultural output per acre in both the current and the last survey year. A time trend was also included to pick up the effects of any time-trended omitted variables. Although other variables were tried but had no additional explanatory power, it is clear that there are omitted variables. For example, from the data available it was not possible to construct a series for employment on rural public works schemes (which have historically been important in famine relief). Furthermore, lags and other limitations of existing public finance data appear to entail at least as large a lag as India’s National Sample Survey (NSS) data, thus diminishing their usefulness as leading indicators. These forecasts could probably be improved, but it would entail considerable extra effort.

The model was first fitted to data for a core sample comprising all previous NSS rounds from 1958-59 until the date beyond which the forecasts were to be made. This entailed re-estimating the model for each experiment. Within-sample performance of the models was good; $R^2$ was 0.98 or higher when using the full sample of twenty surveys, and $R^2$ for the subsamples used for forecasting was in the range 0.90-0.98.

The results for type 1, in panel A, show forecasts of the poverty measures for 1987-88 (using a model based on the data up to and including 1986-87); for 1988-89 (using the data up to 1987-88); and for 1989-90 (using the data up to 1988-89). The percentage forecasting error ranges from a 7 percent underestimation to a 2.5 percent overestimation. The model failed to forecast the (small) increase in the head-count index between 1986-87 and 1987-88, predicting a drop instead.

Panel B gives type 2 forecasts for the same three years based on the same model as the 1987-88 forecast in panel A (using the seventeen survey rounds up to 1986-87). From 1987-88, new forecasts are drawn from the previously forecast poverty measure rather than from the actual measure. (This is the correct way to assess forecasting performance when there are gaps of more than a year between survey rounds.) Naturally, in this case, past forecasting errors are built into future forecasts, so a sizable drift of the forecasts from the actual values is possible relative to year-ahead forecasts. The first year in panel B is, of course,
### Table 1. Forecasting Poverty in Rural India

<table>
<thead>
<tr>
<th>Year (Panels A and B)</th>
<th>One year ahead (Type 1, Panel A)</th>
<th>Three years ahead (Type 2, Panel B)</th>
<th>The mid-1970s reversal (Type 2, Panel C)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Forecast</td>
<td>Percent error</td>
</tr>
<tr>
<td>1986-87</td>
<td>34.19</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>1987-88</td>
<td>34.99</td>
<td>33.50</td>
<td>-4.3</td>
</tr>
<tr>
<td>1988-89</td>
<td>34.63</td>
<td>32.07</td>
<td>-7.4</td>
</tr>
<tr>
<td>1989-90</td>
<td>29.65</td>
<td>30.38</td>
<td>2.5</td>
</tr>
</tbody>
</table>

n.a. Not applicable.

**Note:** The table gives both actual (survey-based) and forecasted (model-based) measures of the head-count index, the percentage of people living below the poverty line.

**Source:** Author's calculations.
the same as in panel A. The second year is similar. But the third year drifts considerably farther; the underestimation of poverty in 1988–89 magnifies the downward bias in the 1989–90 forecast.

We know that the living standards of the poor deteriorated in the early 1970s as a result of a drought in western India. How well does the model forecast this event? Real agricultural wages and average farm yields fell in the early to mid-1970s. (See panel C for type 2 forecasts from the model using data up to 1970–71.) The model predicts the increase in poverty to 1973–74, but it overestimates the effect by a wide margin. Poverty did not increase nearly as much as the aggregate data on wage rates and yields suggest. The model correctly predicts the end of the crisis, showing a decline in poverty by the late 1970s. By the next survey round, however, it has drifted off track by a wide margin.

In summary, even highly aggregated data on agricultural wages and yields contain information that is relevant to predicting poverty outcomes before new household survey data are available or in lieu of them. But these are by no means perfect indicators. Although poverty measures can be forecast with tolerable precision one year ahead of the release of the survey data, quite sizable drift can arise after just a year or two. In one case a model using these indicators fails to predict an increase in poverty, while in the other it does predict it, but suggests far worse outcomes than the eventual survey data. Possibly the forecast could be improved by adding more variables, although the number of surveys available and the lags and gaps in the availability of other data will no doubt constrain such efforts in practice. It is likely that methods such as these are cheaper than conducting new surveys, but this saving must be weighed against the loss of accuracy in monitoring poverty more than a year ahead.

Experiment 3: Rapid Appraisal Methods

In a comprehensive socioeconomic survey, the bulk of the interview time is devoted to various consumption and income components. For example, a typical survey for the World Bank's Living Standards Measurement Study contains many pages of questions on consumption and income sources (Ainsworth and van der Gaag 1988; also see Demery and others 1992 on a survey instrument developed for African countries). Furthermore, calculating consumption and (particularly) income from the raw data is not easy. Some analysts argue that there are far better ways to assess living standards—and a wide range of options, including those described by Grootaert and Marchant (1991), Kumar (1993), Chambers (1994), and Narayan and Srinivasan (1994, ch. 20).

During the last ten years or so, a technique known as "rapid-appraisal" has emerged as a popular alternative to more comprehensive surveys based on statistical sampling. The idea is to use very short interviews and direct observation to obtain a set of objective and subjective indicators of welfare. Some of the rapid appraisal methods include focus group and community interviews as well as less formal individual or household-level interviews (see Kumar 1993). Here
I focus solely on rapid household-level surveys, including subjective questions. Some exponents are convinced that rapid appraisal can be just as accurate as a full-blown survey (Chambers 1994)—an assertion that is difficult to verify unless both methods are applied to the same sample. Such an approach is rarely taken, however.

For this experiment, some typical rapid-appraisal questions were added to a high-quality conventional survey: the 1993 Jamaica Survey of Living Conditions, collected by the Statistical Institute of Jamaica. The survey already included many objective questions commonly covered in rapid-appraisal interviews; to these were added subjective questions on whether respondents considered the family's consumption of various categories of goods (food, housing, clothing, transport, health care, and schooling) to be "less than adequate," "more than adequate," or "just adequate." These questions, which required less than one page of the questionnaire and roughly five minutes of extra time, produced a response rate of virtually 100 percent. Forty-four percent of sampled households responded that their food consumption was "less than adequate for their family's needs," although the proportion rises to about two-thirds for the poorest quintile. A similar pattern in the differences between responses by "poor" and "rich" is evident from the answers to most of the other subjective questions (table 2).

How well can these subjective indicators predict a standard objective measure of consumption expenditure derived from the same survey? A regression model can be used to predict consumption per person, based on answers to subjective questions—several of which turn out to be highly significant predic-

<table>
<thead>
<tr>
<th>Answer</th>
<th>All respondents</th>
<th>Poorest quintile</th>
<th>Richest quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food is inadequate</td>
<td>44.2</td>
<td>67.1</td>
<td>19.1</td>
</tr>
<tr>
<td>Housing is inadequate</td>
<td>48.7</td>
<td>71.7</td>
<td>26.2</td>
</tr>
<tr>
<td>Clothing is inadequate</td>
<td>40.6</td>
<td>64.1</td>
<td>16.3</td>
</tr>
<tr>
<td>Transportation is inadequate</td>
<td>51.5</td>
<td>65.4</td>
<td>34.0</td>
</tr>
<tr>
<td>Health care is inadequate</td>
<td>45.0</td>
<td>62.3</td>
<td>23.0</td>
</tr>
<tr>
<td>Schooling is inadequate</td>
<td>51.8</td>
<td>58.2</td>
<td>52.1</td>
</tr>
<tr>
<td>Food and housing are inadequate</td>
<td>32.5</td>
<td>57.2</td>
<td>10.4</td>
</tr>
<tr>
<td>Food, housing, and clothing are inadequate</td>
<td>25.7</td>
<td>48.9</td>
<td>6.8</td>
</tr>
<tr>
<td>Food, housing, clothing, and transport are inadequate</td>
<td>20.0</td>
<td>40.8</td>
<td>4.9</td>
</tr>
<tr>
<td>Food, housing, clothing, transport, and health care are inadequate</td>
<td>18.4</td>
<td>40.2</td>
<td>3.6</td>
</tr>
<tr>
<td>Food, housing, clothing, transport, health care, and schooling are inadequate</td>
<td>13.7</td>
<td>32.4</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Note: Survey respondents were ranked according to household expenditure per person.
Source: Author's calculations based on the primary survey data from the 1993 Jamaica Survey of Living Conditions, collected by the Statistical Institute of Jamaica.
tors of consumption. The combined explanatory power of these subjective indicators is still rather modest, however; they could only explain 20 percent of the variance in log consumption per person.

Can the model's explanatory power be increased by adding some easily monitored objective indicators? For instance, the interviewer can readily observe certain characteristics of the dwelling and can ask simple "yes or no" questions about it and about the respondent's possessions. From the rest of the Jamaican survey, I picked twenty-five short questions that could probably be answered in ten to fifteen minutes, including:

- Information supplied from the interviewer's own observations, such as location (Kingston, other urban, rural); type of dwelling (detached, semi-detached, apartment); and material of outer wall (wood, stone, brick)
- Questions, probably answerable by any adult household member, including number of people in the household; level of schooling completed by the head of the household (primary, secondary, tertiary); number of rooms in the dwelling; whether it is owned or rented; whether the household owns or shares a toilet; whether it owns or shares a kitchen; source of drinking water (inside tap, outside tap, public standpipe, well); and ownership of various consumer durables (such as telephone, television, washing machine, refrigerator).

Adding this set of easily measured objective indicators to the regression model for predicting consumption considerably improves the total explanatory power; the additional indicators explain 57 percent of the variance in consumption.

To test the performance of these rapid-appraisal indicators further, I reestimated the model on a smaller sample and used it to predict the consumption of the rest of the original sample. Such an experiment is closer to a real-world setting in which relationships revealed by a small rapid-appraisal survey of a few easily monitored indicators are used to make predictions about a far larger population. (Some actual or proposed methods of targeting transfers aimed at reducing poverty have worked this way.) For this experiment, the augmented model described above (combining subjective and short-answer objective questions) was reestimated on a random subsample comprised of every fourth household in the full sample. This model was then used to predict consumption per person in each household in the reserve sample, which included the remaining three-quarters of the original sample (1,270 households).

Table 3 compares the actual and predicted values for households grouped by quintiles of consumption per person. (So, for example, 69 of the 254 households in the poorest quintile in terms of actual consumption were predicted by the model to be in the second poorest quintile.) The model predicted 46 percent of the households in the poorest quintile to be in other quintiles, although it predicted 81 percent to be in the poorest two quintiles (54 percent correctly in the poorest quintile plus 27 percent incorrectly in the second poorest). If one splits all the quintiles into deciles (not shown here), one finds that in 60 percent of
### Table 3. Joint Distribution of Actual and Predicted Values

<table>
<thead>
<tr>
<th>Actual quintile</th>
<th>Poorest quintile</th>
<th>Second quintile</th>
<th>Third quintile</th>
<th>Fourth quintile</th>
<th>Richest quintile</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poorest</td>
<td>137 (54)</td>
<td>69 (27)</td>
<td>36 (14)</td>
<td>11 (4)</td>
<td>1 (0)</td>
<td>254</td>
</tr>
<tr>
<td>Second</td>
<td>66 (26)</td>
<td>83 (33)</td>
<td>60 (24)</td>
<td>35 (14)</td>
<td>10 (4)</td>
<td>254</td>
</tr>
<tr>
<td>Third</td>
<td>32 (13)</td>
<td>61 (24)</td>
<td>70 (28)</td>
<td>66 (26)</td>
<td>25 (10)</td>
<td>254</td>
</tr>
<tr>
<td>Fourth</td>
<td>8 (3)</td>
<td>31 (12)</td>
<td>60 (24)</td>
<td>86 (34)</td>
<td>69 (27)</td>
<td>254</td>
</tr>
<tr>
<td>Richest</td>
<td>1 (0)</td>
<td>10 (4)</td>
<td>28 (11)</td>
<td>61 (24)</td>
<td>154 (61)</td>
<td>254</td>
</tr>
</tbody>
</table>

Note: There are 1,270 households in the sample; 254 in each quintile. Figures in parentheses are percentages of the total for that quintile.

Source: Author's calculations based on the primary survey data from the 1993 Jamaica Survey of Living Conditions, collected by the Statistical Institute of Jamaica.

The model correctly predicts the decile of those actually in the poorest decile, although only 25 percent are predicted to be in the fourth and fifth deciles. Among the middle quintiles (see, for example, the third quintile in table 3), the extent of misclassification is sizable, as indicated by the relatively large entries above and below the diagonal. The model does best for the richest quintile, predicting 61 percent of those households accurately.

When the experiment was repeated without the subjective welfare questions, the method's performance deteriorated, but not by much. Most of the work is being done by the rapid-appraisal objective indicators.

There are undoubtedly other subjective or objective proxies that are easily surveyed and that could add to the explanatory power of this method. Against this possibility are two reasons to believe that these experiments have probably overstated the performance of rapid-appraisal methods. First, these experiments show what the best result would be in predicting consumption from such indicators; the weights on the various rapid-appraisal indicators have been chosen optimally (in the sense of minimum sample error variance in predicting consumption). In practice, analysts typically do not have a complete integrated survey on which to estimate optimal parameters. In such cases they would have to select a set of weights by some subjective means. With little else to go on, the set of weights chosen for this purpose may be far from ideal.

Second, rapid-appraisal strategies often do not use rigorous sampling methods; indeed, many practitioners explicitly reject such methods (see Casley's 1993 comments, however). This means that sampling biases will add to the imprecision of the rapid-appraisal method in practice.
Overall, these experiments suggest that a rapid-appraisal survey lasting, say, twenty minutes on a relatively small sample can yield some significant predictors of the more expensive objective surveys. A significant share of the variance of consumption will almost certainly be left unexplained, however, and the predictions from such methods are unlikely to be highly correlated with actual standards of living. As with other short-cut methods described above, these limitations to the accuracy of rapid-appraisal methods will have to be weighed against the potential cost savings.

**Experiment 4: Decomposing Aggregate Indicators without an Integrated Survey**

Many socioeconomic indicators are available only in a highly aggregated form, that is, by provinces or countries. Yet one would like to know how they vary between socioeconomic groups, such as the poor and the nonpoor, or between urban and rural residents. This calculation would be easy if one had a survey that collected all the relevant information for the same households. Is it possible to estimate the decomposition reasonably well without such a survey?

Bidani and Ravallion (forthcoming) propose a method that can be used when the distribution of the population across the relevant subgroups (such as urban and rural) is known. An econometric model is estimated in which the observed aggregate indicator (such as the infant mortality rate) across countries or regions is modeled as a function of the population distribution by subgroups (such as the proportion of poor). Using this approach, the infant mortality rate of poor and nonpoor households can be estimated when all that is known is the overall average infant mortality rate for each country and the proportion of people who are poor. Bidani and Ravallion apply the method to find out how health indicators differ between the poor and the nonpoor. They find that the average life expectancy of the poorest two-thirds of the population in their sample of thirty-five countries is nine years less than that of the nonpoor and that their children are 50 percent more likely to die before their first birthday. Prescott and Jamison (1985), who use a similar technique to examine mortality rates and health-service availability in China, found that rural residents had worse health and used health care less often than did urban residents. (The Prescott-Jamison method is not used here; for further discussion, see Bidani and Ravallion forthcoming.)

As with the earlier experiments, the most convincing way to test for bias in this method is through a counterfactual experiment. For this experiment, the accuracy of the urban-rural breakdown of provincial socioeconomic aggregates was tested. Many statistical agencies publish socioeconomic data at the level of the local administrative jurisdictions, but they rarely break down these figures into the categories one would like, such as urban and rural. Indeed, it may be administratively difficult to do so. Yet based on data in those countries that provide such information, large differences in socioeconomic outcomes, as well as in access to and use of public services, are known to exist between urban and rural areas.
How accurately can the urban-rural decomposition of socioeconomic aggregates be estimated? Five experiments were carried out for this article; four used data from provincial Indonesia, and one used data from the state level in India. In all cases, the aggregate provincial indicator, the actual urban and rural values of the indicator, and the shares of the urban and rural populations in the province or state were known.

The aggregate indicators were first related by a simple regression to the share of the population in urban and rural areas (using the “random coefficients” estimators described in Bidani and Ravallion forthcoming). The estimated coefficients represent the average urban and rural indicators across all the provinces. Table 4 gives the results for the actual and estimated decompositions of the means in each of the five cases. In only one case—the percentage of well-nourished children in Indonesia—are the estimates very close to the actual means. In the remaining cases, the rural mean is underestimated and the urban mean is overestimated.

Clearly, one of the problems with this test is that relevant factors that influence the socioeconomic indicator may have been omitted, thereby altering the model’s estimates of the effect of the included variables. I therefore revised the model to account for some of these factors, as shown in table 4, with significant improvements in the estimated relationships. For food expenditure in Indonesia, the augmented regression used total consumption expenditure and average household size. For mean expenditure in Indonesia, the augmented regression used mean income. For income in Indonesia, the augmented regression used mean total expenditure. For the percent below a nutrition-based poverty line in India, the augmented regression used mean expenditure and the Gini index of expenditure. In the augmented models, I have deliberately aimed by trial and error to obtain the best estimates of the subgroup means using these explanatory variables from available data sets. In practice—when the true means are, of course, unknown—one would be unlikely to do as well as these estimates.

This experiment is an example in which the simplest methodology is likely to produce the least reliable results. The estimated urban and rural means are substantially closer to the actual means under the augmented regressions than under the simpler unaugmented regressions (table 4). Clearly, the urban-rural decompositions can be heavily biased unless the model is expanded to include a potentially wide range of other relevant variables. If augmentation can be done, the decomposition method is capable of yielding reasonable substitutes for the missing data.

**Experiment 5: Assessing the Prevalence of Poverty without a Survey**

This experiment was designed to show how accurately aggregate economic and social indicators estimate the prevalence of poverty. For this study, the proportion of people below the poverty line in each of twenty-two countries (estimated from household survey data) was compared with forecasts of the same
### Table 4. Actual and Estimated Average Indicators for Rural and Urban Areas

<table>
<thead>
<tr>
<th>Item</th>
<th>Actual Rural</th>
<th>Actual Urban</th>
<th>Regressing on population shares alone Rural</th>
<th>Urban</th>
<th>R^2</th>
<th>Regressions augmented to include other variables Rural</th>
<th>Urban</th>
<th>Adjusted R^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Percentage of Indonesian children under five deemed to be well-nourished in 1987</td>
<td>46.16</td>
<td>59.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Per capita food expenditure in Indonesia (in rupiah/person/month)</td>
<td>17,690</td>
<td>22,735</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Mean expenditure in Indonesia (in rupiah/person/month by province)</td>
<td>25,379</td>
<td>41,167</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Income in Indonesia (in rupiah/person/month)</td>
<td>18,922</td>
<td>31,453</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Percentage of population below nutrition-based poverty line in India (by state, 1983)</td>
<td>28.42</td>
<td>24.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

n.a. Not applicable.

**Note:** Random coefficient regression methods were used; see Bidani and Ravallion (forthcoming) for details on the estimation method.

**Source:** For item 1, BPS (1989); for items 2 and 3, authors' calculations from the National Socioeconomic Survey of Indonesia, 1990; for item 4, authors' calculations from the National Socioeconomic Survey of Indonesia, 1984, with data on population proportions based on BPS (1989); for item 5, Datt and Ravallion (1993).
poverty measure in each country derived only from aggregate data. The poverty line was set at $1 a person a day (converted into local currencies to assure purchasing power parity in 1985). Using the methodology documented in Ravallion, Datt, and van de Walle (1991), regressions were estimated on each of the twenty-two possible samples, obtained by leaving out one country at a time. These are counterfactual experiments, meaning that the poverty measure for the particular country is assumed not to be known, although it is known for each of the other twenty-one countries. The predictor variables were private consumption per capita from the national accounts evaluated at purchasing power parity and at official exchange rates, the level of urbanization, the infant mortality rate, life expectancy at birth, and the proportion of women in the labor force. Some of these variables are of interest in their own right; infant mortality, for example, can be an important indicator of well-being (such as access to public health care), but here I am concerned only with the variable’s usefulness in predicting consumption poverty.

Figure 3 shows that the predictions from the aggregate data are highly positively correlated with the survey-based estimates; the simple correlation coefficient is 0.87. Nonetheless, the absolute errors in predicting the prevalence of

![Figure 3. Actual and Predicted Poverty Measures across Countries](image)

Percentage of population consuming less than US$1 per day

- Survey estimate
- Predicted

Country (ranked by survey-based estimate)

Note: The predicted measure for each observation is based on a regression that excludes that observation.

Source: Author’s calculations from primary data for twenty-two countries (Ravallion, Datt, and van de Walle 1991).
poverty are often quite large. Indeed, the average absolute error as a percentage of the original survey estimate was 49 percent. The poorest few countries were correctly identified, although there was considerable reranking. For example, the aggregate indicators suggested that the eleventh poorest country (based on the survey estimates) was almost the least poor. For the country with the largest absolute error, the survey estimate of the head-count index was 57 percent, while the aggregate economic and social indicators predicted a figure of 39 percent.

These discrepancies arise from two factors. First, the extent of inequality varies from country to country, and this variation is hard to pick up without distributional data from a survey. Second, the relation between social indicators and consumption-based poverty measures differs from country to country; some countries where a large share of the population is poor provide effective public health care and thus have good social indicators, such as low infant mortality, while others do not. These differences can make it hard to assess the extent of poverty without a household survey. It appears that the readily available economic and social aggregates can give, at best, a rough idea of the prevalence of poverty.

Conclusions

Much effort goes into making up for inadequate data on poverty; the tools used range from various rapid-appraisal methods to sophisticated econometric models. But in using such methods, how do we know that the cure is any better than the disease? Or how do we decide which cure works best? Surprisingly little effort has gone into rigorous testing of these methods. New approaches—aiming to implement some new methodological paradigm or just to improve an existing survey—are routinely introduced without prior testing. Little more than blind faith guides the policymaker’s interpretation of results.

In many ways the experiments reported here represent a limited investigation. It is not feasible to span all the methodological choices that matter, although these experiments may be suggestive in other applications. The results may not be robust to measurement error or applicable in other settings; there are still precious few opportunities for tests of this sort. And there are gains and losses that I have not been able to quantify. The imprecision introduced by imperfect data and methods may appreciably lower the expected benefits from efforts at targeting the poor. Nor have I addressed the costs of data collection, which must be weighed against the benefits of greater accuracy in measurement.

Nonetheless, the results do suggest that there are some relatively easily monitored proxies for poverty that can help identify and monitor poverty when survey data are unavailable or inadequate. These proxies rely on certain plausible theoretical or common sense relationships that are expected to hold between
conventional survey-based poverty data and more aggregated data from other sources (or from simpler rapid-appraisal methods). The experiments reported here suggest that data inadequacies can be at least partially surmounted by careful use of such methods (table 5). Even with rather poor data, the situation facing a creative analyst is rarely hopeless.

The results also point to the limitations of these methods. Omitted variables may well account for such a large share of the variation in standards of living that the method tends to be a poor substitute for better data. Even a rather narrowly defined objective indicator of welfare, such as consumption or income, is determined by a large number of variables, only a small subset of which are amenable to monitoring by these methods. In addition, the variation in welfare at any given income level is so broad that one must remain somewhat pessimistic about the scope for really credible monitoring of poverty in many data-poor settings.
<table>
<thead>
<tr>
<th>Data problems</th>
<th>Proposed solutions</th>
<th>The tests</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. You have a good consumption survey, but no locational cost-of-living index.</td>
<td>Either (1) ignore the problem and use one nominal poverty line in all regions, or (2) anchor the lines to food-energy requirements so that the local line is the expenditure level at which requirements are met on average in each region.</td>
<td>First estimate the regional poverty profile for Indonesia using locational data on the costs of basic consumption needs, then see how well the profile can be predicted without these data.</td>
<td>Method (1) is better than (2). The errors in ranking using (1) need not be large, particularly if one can at least make a reasonable assumption for the urban-rural cost-of-living difference. But method (2) can be way off the mark when aiming to construct a profile of absolute poverty.</td>
</tr>
<tr>
<td>2. There are gaps and lags in survey-data availability.</td>
<td>Identify proximate causes of poverty that can be monitored using more readily available aggregate data.</td>
<td>Using time series data for rural India, see how well the model predicts beyond the period over which it was estimated. Use real agricultural wage rate and average farm yield as the predictors.</td>
<td>The wage rate and farm yield are good predictors of rural poverty measures for India. One can forecast poverty fairly well one year ahead of the survey. But sizable drift can arise after one year.</td>
</tr>
<tr>
<td>3. A full-size objective socioeconomic survey is not feasible, but a rapid appraisal is.</td>
<td>Use both subjective and objective rapid-appraisal methods, from short questionnaires, and small samples.</td>
<td>A full-sized conventional survey for Jamaica also included a short module of subjective and simple objective questions, so the results can be compared. Compare predictions for consumption.</td>
<td>Roughly half of the variance in consumption can be explained this way. But the method still entails large errors; for example, in 60 percent of cases the rapid-appraisal method incorrectly predicts the decile of those actually in the poorest decile. The method was better at identifying rich households than poor ones.</td>
</tr>
</tbody>
</table>
4. You would like to know how various indicators differ between subgroups but your survey did not ask for those indicators.

5. You do not have any kind of survey, but you want to at least assess how prevalent poverty is in a country.

- Use econometric analysis to decompose the average indicator by subgroups; the population mean indicator is regressed on the proportions of people who are poor and nonpoor (say), allowing for the structure of the regression errors.

- Data on the urban-rural decomposition of socioeconomic aggregates by region for Indonesia and India can be used to test the method. Actual values are compared with the model's predicted decomposition.

- For countries in which you have the required data, construct a regression model to predict the poverty rate as a function of the readily observed socioeconomic aggregates.

- For each of 22 countries with surveys, see how well poverty can be predicted using aggregate data. For each country, calibrate the model on the other 21 countries.

- Simple models perform badly. But provided the model is augmented to allow for other determinants, these data suggest that reasonable predictions are possible.

- The experiment suggests that the ranking will rarely be far off, but that the absolute forecasting errors will be large in some cases.
Notes

Martin Ravallion is with the Policy Research Department of the World Bank. He would like to thank Benu Bidani, Shaohua Chen, Gaurav Datt, Christiaan Grootaert, Margaret Grosh, Stephen Howes, Emmanuel Jimenez, Dominique van de Walle, Jacques van der Gaag, and Qing-hua Zhao. Statistical agencies in twenty-three countries have helped by collecting data used here, but special mention should be made of the Statistical Institute of Jamaica for agreeing to add to its survey the questions used in Experiment 3.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.


