12521
Sept. 1993

Correcting for Sampling Bias in the Measurement of Welfare and Poverty in the Côte d'Ivoire Living Standards Survey

Lionel Demery and Christiaan Grootaert

The sampling aspects of a household data set are important to analysts. The early years of the Côte d'Ivoire Living Standards Survey (CILSS) had a sampling bias, which seriously affected estimates of population statistics such as household size. The bias arose from sampling procedures that overrepresented larger dwellings. Assuming that samples drawn in later years were unbiased, a correction procedure is applied that uses weights based on household size. Results from the weighted data are then compared with the unweighted findings to assess the seriousness of the bias. Estimates of household expenditure per capita in the early years of the survey are found to be significantly underestimated, resulting in an overestimation of poverty. The sampling bias also resulted in an underestimation of the upward trend in poverty during 1985-88. The CILSS has been a popular and fruitful data set for policy analysis. These findings, however, cast doubt on the robustness of earlier work. Thus, the effort to trace sampling information is particularly worthwhile for policy-oriented applied research.

Over the years household surveys have proved to be a popular and valuable data source for empirical research in microeconomics. In developing countries household survey data have become more available in the past decade, partly as a result of international programs such as the UN National Household Survey Capability Program, the World Fertility Survey, the World Bank's Living Standards Measurement Study and the Social Dimensions of Adjustment Program, and the USAID-sponsored Demographic and Health Surveys. This availability of survey data has spurred a growing interest in the economics of the household in the context of development economics. The wide array of topics covered include poverty, employment, health, education, and fertility.

Many analysts—economists and other social scientists—devote only a limited amount of attention to the sampling design of the surveys they use; they take the

Lionel Demery is with the Education and Social Policy Department, and Christiaan Grootaert is with the Transportation, Water, and Urban Development Department, both at the World Bank. This article is a result of the research project "Poverty and the Social Dimensions of Structural Adjustment in Côte d'Ivoire, 1985-88: A Policy-Oriented Analysis" (RPO 675-26). It has benefited from comments received from Chris Scott and three anonymous referees. The authors would like to thank Gi-Taik Oh and Meera Venkataraman for excellent computer programming assistance.

© 1993 The International Bank for Reconstruction and Development / THE WORLD BANK

data produced as given, without understanding (or questioning) how the data were obtained. At best, sampling weights are applied to ensure that the results are representative.

This article illustrates the necessity of paying close attention to the sampling aspects of a household survey used in applied microeconomic analysis. This is particularly important when comparisons over time are undertaken. The case study in this article will document that changes in household welfare and in the incidence of poverty observed in Côte d'Ivoire during 1985–88 are modified when corrections are applied for changes in sampling procedures. Likewise, the cross-sectional patterns of welfare and poverty observed in earlier analyses for 1985–86 may be misleading.

The data source for the case study is the Côte d'Ivoire Living Standards Survey (CILSS) conducted between 1985 and 1988. The CILSS was designed to measure living standards and, by repeating the survey each year, to monitor changes in household welfare over time (see Grootaert 1986 for a further discussion of the survey and its content). A disturbing feature of the CILSS results, however, is the recorded decline in mean household size noted by Daho (1992) and Coulombe and Demery (1993). According to these authors this decline was caused by sampling bias in the early years of the survey. If this is indeed the case, there are obvious questions surrounding the robustness of the analyses undertaken with these data.

The CILSS, especially the early years, has been a popular and fruitful data set for policy analysis. The results of studies that estimate population statistics—mean income, mean expenditure, poverty incidence, and so on (see, for example, Glewwe 1987; Kakwani 1990; Kanbur 1990; and Kozel 1990)—are likely to be particularly sensitive to sampling bias. This article assesses the seriousness of the CILSS sampling bias and makes the necessary adjustments to correct it. The findings of other studies that use the data to estimate behavioral relations (such as Deaton 1987, 1989; Glewwe 1991; Grootaert 1987, 1990; and van der Gaag and Vijverberg 1989) may also be affected by the bias, but this issue is not explored here.

Correcting sampling bias involves reweighting the original data, as discussed in section II. Given the important role of the household-size variable in many recent studies of welfare and poverty, section III compares estimates of mean per capita expenditure computed with and without household-size weights. Section IV assesses the implications of these findings for estimates of poverty, and section V examines estimates of basic-needs indicators. But first, section I briefly discusses the results of the CILSS with regard to household size.

I. HOUSEHOLD SIZE IN THE CILSS

Between 1985 and 1988, household size in Côte d'Ivoire declined from 8.31 to 6.33 persons (a fall of 24 percent) according to the “raw” CILSS results (Daho

Table 1. CILSS: Mean Household Size with Three-Month Minimum Residency, by Region, 1985–88
(number of persons)

<i>Year</i>	<i>Côte d'Ivoire</i>	<i>Abidjan</i>	<i>Other cities</i>	<i>Rural</i>
1985	8.31	7.11	8.88	8.54
1986	8.01	6.98	8.86	8.16
1987	7.01	6.18	7.31	7.22
1988	6.33	6.02	6.49	6.36
Percentage decline during 1985–88	24	15	27	26

Source: Coulombe and Demery (1993).

1992; Coulombe and Demery 1993).¹ Table 1 shows that this decline occurred in all regions, with Abidjan reporting the lowest decline (15 percent during the four-year period). These results apply under alternative definitions of “household.” To qualify for inclusion in a household in the CILSS, a person must have resided in the household for 3 months or more during the previous 12 months. This definition underlies the estimates reported in table 1. It is perfectly feasible, however, to use alternative definitions. Table 2 compares estimates of household size using three-month, six-month, and nine-month residency requirements. It is clear that the changes in residency requirements make little difference to the observed decline in household size. If a stricter definition of household membership is applied (for example, requiring that the person has been present at least 6 of the previous 12 months), a decline of 23 percent is still observed. This article

Table 2. Household Size by Residence Criterion, 1985–88
(number of persons)

<i>Year</i>	<i>Minimum number of months present in the household during the preceding 12 months</i>		
	<i>9</i>	<i>6</i>	<i>3</i>
1985	7.77	7.98	8.31
1986	7.50	7.70	8.01
1987	6.69	6.84	7.01
1988	6.05	6.16	6.33
Percentage decline in 1985–88	22	23	24

Source: Authors' calculations.

1. Proper use of the CILSS data requires the application of three sets of weights. The first set is designed to correct for an overrepresentation of better-off households and an underrepresentation of poorer households in the Abidjan sample for 1985–86. The second set is designed to correct for differences in the distribution of primary sampling units across regions in the 1987–88 sampling frame compared with the population distribution derived from the 1988 population census. Unless stated otherwise, all data presented in this article have already been subjected to the application of these two sets of weights. The third set of weights, designed to correct for household-size sampling bias, is the subject of this article.

takes a six-month residency as the qualification for household membership.² This period is the most logical choice for household membership because a shorter qualification period could enable a person to be a member of more than one household during the year and thus be counted more than once in the survey.

The CILSS data collection followed a rotating panel design; that is, each year after 1985, 50 percent of the households were replaced with new ones and the other 50 percent were retained with the intention of revisiting them. The data structure over the four years thus contains three overlapping panels (see Grootaert and Kanbur 1993 for further details). This feature of the data has an important bearing on the interpretation of the observed decline in household size. Table 3 explains how the data in the four years of the survey can be divided into 14 data sets, depending on whether the household forms one of the three panels. Set 1 data refer to households interviewed only once in 1985 and not revisited thereafter. Households in set 2 were interviewed in 1985 with the intention of including them in the first panel but were not found again in 1986, mainly because they had changed residence by the second year. The data in set 5 are the replacements for these households in 1986. The panel households themselves are in sets 3 and 4, for the interviews in 1985 and 1986, respectively. The data sets for the other years are defined in a similar way.

Table 4 reports the sample size and mean household size of each data set.³ The key observation from part B of the table is that the decline in household size is

Table 3. *Data Sets in the CILSS, 1985–88*

<i>Year</i>	<i>Set</i>	<i>Description</i>
1985	1	Nonpanel households
	2	Unmatched households from first panel
	3	First-panel households
1986	4	First-panel households
	5	Unmatched households from first panel
	6	Unmatched households from second panel
	7	Second-panel households
1987	8	Second-panel households
	9	Unmatched households from second panel
	10	Unmatched households from third panel
	11	Third-panel households
1988	12	Third-panel households
	13	Unmatched households from third panel
	14	Nonpanel households

Source: Authors' calculations.

2. Some household members, such as newborns and the household head, are considered members even when they do not fulfill the six-month requirement. Others, such as servants and boarders, are always excluded because they are considered members of separate households in the survey.

3. The difference in the sample sizes of sets 2 and 5 (the unmatched households in the first panel and the 1986 replacement households) needs to be explained. Although 79 households failed to qualify for the first panel, 86 households were selected as replacements because the sample in 1985 was only 1,593 (7 short of a full sample of 1,600 households). Thus, 7 households were selected in addition to the 79 replacement households to restore the sample to 1,600 households in 1986.

Table 4. *Sample Size, Mean Household Size, and Weighted Household Size, by Data Set, CILSS, 1985–88*

Year	Sample or household size							
	A. Sample size (number of households)							
	Nonpanel	Unmatched	First panel		Second panel			
1985	(1) 795	(2) 79	(3) 714	Unmatched	(7) 693	Unmatched	Third panel	
1986		(5) 86	(4) 714	(6) 107		(10) 99	(11) 701	Nonpanel
1987				(9) 107	(8) 693	(13) 99	(12) 701	(14) 800
1988								
	B. Mean household size (number of persons)							
	Nonpanel	Unmatched	First panel		Second panel			
1985	(1) 8.25	(2) 5.25	(3) 8.03	Unmatched	(7) 7.85	Unmatched	Third panel	
1986		(5) 5.14	(4) 7.96	(6) 7.20	(8) 7.36	(10) 5.92	(11) 6.54	Nonpanel
1987				(9) 6.28		(13) 5.31	(12) 6.53	(14) 5.95
1988								
	C. Weighted household size (number of persons)							
	Nonpanel	Unmatched	First panel		Second panel			
1985	(1) 6.77	(2) 4.46	(3) 6.52	Unmatched	(7) 6.36	Unmatched	Third panel	
1986		(5) 4.40	(4) 6.49	(6) 5.88	(8) 6.10	(10) 5.93	(11) 6.53	Nonpanel
1987				(9) 4.92		(13) 5.31	(12) 6.53	(14) 5.95
1988								

Note: Data sets are identified by the numbers in parentheses. See table 3 for a full description of each data set.

Source: Authors' calculations.

almost entirely absent within the panels except for the second panel, which shows a decline from 7.85 in 1986 to 7.36 in 1987. The order of magnitude of the decline in mean household size recorded through nonpanel comparisons across the years cannot be considered as simply reflecting a real trend in household formation and composition in Côte d'Ivoire. For example, mean household size in set 4 (1986) is 7.96 and in set 11 (1987) is only 6.54—a decline of almost 20 percent in just one year is extremely unlikely.

There are three alternative explanations of these observed declines in mean household size:

- Nonsampling errors (measurement errors) arising from either overenumeration in the early years of the survey or underenumeration in later years (the result, for example, of deteriorating field supervision).
- Nonrandom sampling, with oversampling of either large households in the early years or small households in the later years of the survey.
- An underlying reality of demographic change.

There is little evidence that nonsampling errors are responsible. Various tests applied to the data suggest that data quality was maintained during the four years of the CILSS. Field supervision generally improved over time (Daho 1992). Moreover, the fact that the decline in household size arises from recorded declines in nuclear family members does not support the hypothesis that there was an increasing tendency to underenumerate when completing the household roster. If enumerators were tempted to underrecord when completing the household roster, they would most probably have left out more distant relatives and marginal household members (Coulombe and Demery 1993).

The case for sampling bias as the main explanation is based on the observation that, although household size declined throughout the period for all regions, a discontinuity is observed in the data. In particular, the drop in recorded household size is most severe between 1986 and 1987, especially when nonpanel households are considered. This discontinuity coincides with a change in the sampling procedures used in the CILSS. For 1985 and 1986 the sample was drawn from a sampling frame derived from demographic projections based on the 1975 population census. In 1987 the sampling frame was changed: it was derived from an electoral census. In 1988 the frame was again updated, this time derived from preliminary results of the 1988 population census.

The procedures used to draw the sample were also not consistent over time. In mid-1984 a "presurvey" was carried out to list 64 households in each of 100 primary sampling units. (The sampling units had been selected with probabilities proportional to their population and were stratified by region.) For 1985 a sample of 16 households was randomly drawn from the 64-household listing in each primary sampling unit so that a sample of about 1,600 households was achieved for the 100 primary sampling units. In 1986 half the sample (from 50 clusters) was surveyed a second time to form a panel of households, and 800 new households were selected from the other 50 clusters. In 1987 this selection

procedure was changed. A complete listing was made of all households in each primary sampling unit, and 16 households were randomly selected from the list. This change in procedure (and the revised sampling frame mentioned earlier) affected only half the sample of 1987, because the other half of the sample (forming the 1986–87 panel) had been selected in 1986 under the previous sampling regime. Thus in 1987 households in panel set 8 were selected on the original sampling frame and procedure, whereas those in panel set 11 were sampled under a new frame and revised listing procedure (see Daho 1992 for further discussion).

The household-size estimates closely reflect this change in the sampling frame and listing procedures. For 1987 the mean size of households enumerated under the original procedures (that is, data set 8) is 7.36, but it is only 6.44 for households from the revised listing (that is, data sets 9 to 11). As a rough indicator, this difference suggests that 51 percent of the 1985–88 decline in household size (the observed decline being from 7.98 to 6.16—see table 2) is “explained” in 1987 alone by the change in sampling procedures. However, the initial sampling frame and listing procedures applied to the 1985 and 1986 samples and to half the 1987 sample. When these observations (that is, data sets 1 to 8) are pooled, mean household size is 7.77 for the country as a whole. If the other half of the 1987 sample is pooled with the 1988 sample (that is, taking all observations derived under the revised sampling frame and listing procedures, which entails combining data sets 9 to 14), the mean household size is 6.26. In this way, 83 percent of the 1985–88 decline in household size is explained by the change in the sampling procedures.

The main reason for this phenomenon appears to be that the procedures used in listing the 64 households for 1985 and 1986 were biased toward a selection of larger households in the primary sampling units. In particular, the way the second stage of the sampling of the CILSS was implemented in the earlier years could easily have led to sampling bias.⁴ In the selected primary sampling units, where the listing of households was to occur, enumerators were instructed to start the listing process at a random location in the primary sampling unit and from this point to select every *n*th household, that is, with a given fixed “step,” until 64 households were listed. There are two sources of potential bias in this listing procedure. First, the selection of the starting point might not have been random, but subject to motivated bias on the part of the enumerator (such as the selection of a point where there are numerous dwellings or that is easily accessible). Second, in practice, enumerators counted doors to achieve the “step,” rather than counting actual households. This method leads to sample selection bias if the number of doors varies across households. Households with two doors will have twice the probability of selection as those with one door. Given that larger dwellings are more likely to have more doors than small dwellings, counting households on the basis of doors may have caused a bias in the sample,

4. The discussion here draws in part on personal communication with Chris Scott.

leading to overenumeration of large dwellings and, thereby, large households. In fact, mean dwelling size recorded in the CILSS was significantly higher in 1985 and 1986 than in 1987 and 1988, supporting this interpretation.

Coulombe and Demery (1993) have applied tests for the statistical significance of differences in mean household size. The tests indicate that the revision in sampling procedures was responsible for most of the decline. One of the most telling comparisons is between mean household size in the two panel halves of the 1987 sample. This was 7.36 for set 8 (see table 4), which formed the second panel and which is based on the original sampling procedures, and only 6.54 for set 11, which was enumerated under the revised sampling procedures.⁵ This difference is statistically significant at the 1 percent level. (The difference between data sets 8 and 11 was also statistically significant for each region, except Abidjan.)

Although the evidence strongly points to sampling bias in the early years of the survey as the explanation of much of the observed decline in household size, sampling bias does not explain all of the decline. Some household-size declines were observed when the sampling procedures were consistent. The explanation for the remaining portion of the decline can probably be found partly in measurement error and partly in a real phenomenon. The economic hardship during the period, particularly in 1987 and 1988, may have led to breakups in traditional, extended households and their implied mutual support systems. The data do confirm an increase in one- and two-person households over time. Increased migration might also have occurred as a response. If new migrants moved into dwellings not included in the survey's sampling frame, this may have contributed to an apparent decline in household size among the households surveyed. Clearly, these issues require further research.

II. CORRECTING THE SAMPLING BIAS

In many respects, the finding that sampling bias is largely responsible for the astonishing decline in household size in the CILSS is good news. If the decline were due to nonsampling errors, it would be extremely difficult to correct. Moreover, estimates of the size of each household would be suspect, and analysis at the household level would have been seriously flawed. As it is, the analysis suggests that nonsampling errors are not the major problem, so the size of each household can be considered fairly accurate. The inaccuracy lies in the sampling across households: too many large households were enumerated in the early years, creating a bias in the estimate of mean household size. To correct for this, an appropriate reweighting of the sample is called for.

Such a reweighting scheme should correct

- For differences observed within each region (that is, stratum) in estimated household size under the original and revised sampling procedures

5. There is no difference between mean household size in the two panel halves of the 1985 and 1986 samples.

- Only those differences caused by sampling factors; in other words, the part of the observed decline that reflects an underlying reality should be retained in the data.

There are several reweighting schemes that would fulfill these requirements and several variables that could be used as a basis for correcting sample bias. Given that the source of the problem is likely to have originated in the use of doors to count households in the listing, an obvious reweighting scheme could be based on the number of doors in the household. This would not, however, address the problem of the nonrandom starting point. An alternative would be to base the scheme on the size of the dwelling, which would be closely related to the number of doors. But this variable in the CILSS is subject to significant measurement error and inconsistency (in some cases, the outside of the dwelling wall was measured; in others, the inside). A third variable is household size, because this is also likely to be correlated with dwelling size and number of doors. Given the importance of household size in welfare analysis, we chose this variable as the basis of applying correcting weights to the sample.

We considered two methods for weighting the sample to correct for the sampling bias. Both start from the assumption that the distribution of households obtained under the revised sampling frame and listing procedures (that is, for 1987–88) is the true distribution and that the distribution obtained under the original sampling procedures needs to be corrected. This assumption is justified in view of the fact that the use of doors as the selection device is certain to lead to a nonrandom sample, whereas the complete listing that was undertaken in 1987–88 is obviously preferred. It is also consistent with other sources of information on household size. For example, the Côte d'Ivoire 1988 population census reported a mean household size that corresponds closely to the CILSS 1988 estimate.

A first method for computing corrective weights is to use only information for 1987, when the change in sampling procedures occurred. A comparison can be made between the two panels in that year (data sets 8 and 11) or between data set 8 and all other 1987 observations (data sets 9 to 11), which all come from the revised sampling. In both cases, the difference in mean household size would primarily be the result of differences in sampling procedures, and the ratios between the means could be used to correct the 1985–87 observations collected under the original, flawed sampling procedures.

A second method is to pool all available data according to the sampling regime and to generate two distributions of household size with the four years of data: pool I to consist of all observations obtained under the original procedures (that is, data sets 1 to 8 [see tables 3 and 4]) and pool II to comprise the remaining sets (that is 9 to 14).

The main advantage of the first reweighting method is that, by relying only on 1987 data, one can be more confident that the weights correct only the differences in sampling procedure and that they do not include any changes in household size

Table 5. *Calculation of Household-Size Weights for the CILSS, by Region*

Size class (number of persons in house- hold)	Abidjan			Other cities			East Forest			West Forest			Savannah		
	Percentage of households in size class		House- hold-size weights (2)/(1)	Percentage of households in size class		House- hold-size weights (2)/(1)	Percentage of households in size class		House- hold-size weights (2)/(1)	Percentage of households in size class		House- hold-size weights (2)/(1)	Percentage of households in size class		House- hold-size weights (2)/(1)
	Pool I (1)	Pool II (2)		Pool I (1)	Pool II (2)		Pool I (1)	Pool II (2)		Pool I (1)	Pool II (2)		Pool I (1)	Pool II (2)	
1	11.416	11.114	0.974	5.405	11.029	2.040	3.859	6.611	1.713	2.500	5.300	2.120	5.485	8.973	1.636
2	7.897	6.908	0.875	5.856	8.362	1.428	4.851	7.850	1.618	6.071	9.596	1.581	5.612	10.971	1.955
3	8.046	8.513	1.058	7.658	7.875	1.028	7.387	9.662	1.308	6.071	12.747	2.100	9.184	11.139	1.213
4	6.569	11.404	1.736	6.982	10.851	1.554	9.482	10.041	1.059	11.071	14.245	1.287	8.674	11.651	1.343
5	10.407	13.188	1.267	6.644	10.388	1.563	8.490	11.825	1.393	9.107	9.349	1.027	11.480	10.157	0.885
6	11.984	10.247	0.855	8.784	11.979	1.364	9.923	11.936	1.203	10.536	16.055	1.524	10.459	8.080	0.773
7	8.338	10.327	1.239	9.572	6.607	0.690	8.490	7.951	0.937	8.393	8.503	1.013	7.398	6.860	0.927
8	6.798	6.719	0.988	8.896	8.314	0.935	8.269	8.876	1.073	8.929	5.052	0.566	6.633	6.312	0.952
9	8.171	5.622	0.688	7.545	5.560	0.737	7.056	6.278	0.890	10.000	5.547	0.555	7.143	4.517	0.632
10	4.872	5.782	1.187	5.180	3.747	0.723	5.292	4.725	0.893	7.321	3.802	0.519	6.378	4.862	0.762
11	2.352	4.336	1.843	4.054	3.771	0.930	4.741	3.763	0.794	5.714	3.047	0.533	4.464	3.669	0.822
12	2.929	1.037	0.354	3.604	2.325	0.645	3.638	2.274	0.625	2.679	1.406	0.525	3.189	2.104	0.660
13	3.089	1.256	0.407	3.604	2.633	0.731	4.410	2.487	0.564	2.321	2.800	1.206	3.444	1.326	0.385
14	2.144	1.196	0.558	2.928	2.088	0.713	2.977	0.638	0.214	1.071	1.003	0.936	2.934	1.971	0.672
15	1.405	0.409	0.291	3.153	1.268	0.402	1.764	1.424	0.807	1.786	0.651	0.365	2.168	1.936	0.893
16	0.310	1.097	3.541	3.041	0.748	0.246	2.646	1.137	0.430	1.429	0.300	0.210	1.020	1.087	1.066
17	1.136	0.219	0.193	0.788	0.309	0.392	2.095	0.638	0.305	1.071	0.300	0.280	1.148	1.087	0.947
18 or more	2.138	0.628	0.294	6.306	2.146	0.340	4.631	1.886	0.407	3.929	0.300	0.076	3.189	3.297	1.034

Note: Pool I contains households in data sets 1 to 8; pool II contains all other observations (that is, data sets 9 to 14). See table 3 for definitions of data sets.

Source: Authors' calculations.

that might have occurred between 1985 and 1987 and between 1987 and 1988 (that is, changes within the pools). If there was a real decline in household size, the second method would overcorrect for the bias, removing not only the effects of the bias but also some of the real decline. However, because demographic parameters tend to change only very slowly over time, any real changes in a given two-year period are likely to be very small. It would require war or a severe drought, neither of which occurred in Côte d'Ivoire during the period, for mortality to affect household size significantly. Likewise, we are not aware of any major emigration during the period, which might also cause such a decline in household size. This weakens the case for using only the 1987 data set.

There are also disadvantages with the first reweighting method. Data set 8, which provides the households enumerated under the original sampling regime, is the second part of the 1986–87 panel. Taken by themselves, these households do not constitute an unbiased representation of the country, because of panel attrition, which is not a random process. Moreover, we are disturbed by the drop (from 7.85 to 7.36) in mean household size within the second panel. Such a decline is not recorded in the other two panels, and measurement error could be responsible, at least in part. Taking all this together, it does not appear that by itself, data set 8 is a solid basis for developing the corrective weights.

The second method, which relies on pools of all available data, addresses this concern. By incorporating data set 8 into a much larger pool, any measurement error is more likely to cancel out and the attrition effect will also be removed. Moreover, the pooling method has the key advantage that the corrective weights can be based on much larger samples of households. Pools I and II contain 3,881 and 2,507 households, respectively, whereas the 1987 panel data sets, 8 and 11, contain only about 700 households each. This difference is critical if the weights are to be calculated for each of the five strata in the survey (Abidjan, other cities, East Forest, West Forest, and Savannah) and if they are to incorporate differences in the sampling bias over the entire household-size distribution. Neither of these is possible with the first method, because cell sizes would be prohibitively small. More than a third of the cells would contain only 5 observations or less, and only about 10 percent of the cells would have more than 15 observations. We deem this to be an inadequate empirical basis for constructing corrective weights.

For these reasons the reweighting presented here relies on the second method. The data were pooled (as described above), and distributions according to household size were generated for each of the five strata. Household-size weights were then computed as the ratio of the frequencies in pool I and pool II for each size class (households were distributed across 18 size classes). The size distributions and ratios, reported in table 5, show that it was indeed essential to distinguish between strata and household-size classes in calculating the weights. To illustrate how these weights are applied to correct the observations in data sets 1 to 8, table 6 reports the pre- and postweighted distributions of households in East Forest in 1985.

Table 6. *Calculation and Application of Household-Size Weights for East Forest*

Size class (number of persons in house- hold)	Percentage of house- holds in size class		Household- size weights ^a (2)/(1)	Percentage of house- holds in size class, 1985	
	Pool I (1)	Pool II (2)		Un- weighted frequency	Weighted frequency
1	3.859	6.611	1.713	3.3	5.6
2	4.851	7.850	1.618	4.7	7.4
3	7.387	9.662	1.308	7.7	9.9
4	9.482	10.041	1.059	9.3	9.8
5	8.490	11.825	1.393	10.4	14.5
6	9.923	11.936	1.203	10.4	12.5
7	8.490	7.951	0.937	7.4	7.0
8	8.269	8.876	1.073	9.9	10.5
9	7.056	6.278	0.890	6.6	5.8
10	5.292	4.725	0.893	4.9	4.4
11	4.741	3.763	0.794	3.6	2.8
12	3.638	2.274	0.625	2.7	1.8
13	4.410	2.487	0.564	3.0	1.8
14	2.977	0.638	0.214	4.4	1.1
15	1.764	1.424	0.807	1.1	0.9
16	2.646	1.137	0.430	3.3	1.4
17	2.095	0.638	0.305	1.9	0.6
18 or more	4.631	1.886	0.407	5.2	2.1
Mean household size				8.4	6.6
Standard deviation				7.0	5.2

Note: Pool I contains households in data sets 1 to 8; pool II, data sets 9 to 14. See table 3 for definitions of data sets.

a. Household-size weights are calculated by taking the ratio of pool II over pool I frequencies.

Source: Authors' calculations.

Before assessing the implications of these weights for the survey findings, it should be emphasized that the method we retained implies a larger adjustment than the use of 1987 data alone. The difference between the mean household size in data sets 8 and 11 (for Côte d'Ivoire as a whole: $7.36 - 6.54 = 0.82$) or between data sets 8 and 9 to 11 (for Côte d'Ivoire: $7.36 - 6.44 = 0.92$) is significantly less than the difference between pools I and II (for Côte d'Ivoire: $7.77 - 6.26 = 1.51$). Thus reweighting on the basis of the 1987 data alone would have a smaller effect on survey estimates. As we have already suggested, the distinction between pools I and II accounts for 83 percent of the observed decline in household size in the CILSS, whereas using the 1987 data alone would account for only 45 to 51 percent, depending on whether data sets 9 and 10 are excluded.

For the country as a whole, the effect of applying the household-size weights (based on pools I and II) is to reduce mean household size to 6.51, 6.25, and 6.20 persons in 1985, 1986, and 1987, respectively (table 7). For 1985 and 1986 (for which the whole sample was subjected to household-size weights), this reduction amounts to more than 18 percent of the mean computed without the

Table 7. *Mean Household Size with and without Household-Size Weights, by Region, 1985–88*
(number of persons unless otherwise specified)

<i>Area and item</i>	<i>1985</i>	<i>1986</i>	<i>1987</i>	<i>1988</i>
<i>Côte d'Ivoire</i>				
Mean size without weights	7.98	7.70	6.84	6.16
Mean size with weights	6.51	6.25	6.20	n.a.
Difference (percent)	-18.4	-18.8	-9.4	n.a.
<i>Abidjan</i>				
Mean size without weights	6.87	6.57	6.06	5.88
Mean size with weights	6.18	5.90	5.79	n.a.
Difference (percent)	-10.0	-10.2	-4.5	n.a.
<i>Other cities</i>				
Mean size without weights	8.52	8.53	7.20	6.33
Mean size with weights	6.41	6.54	6.21	n.a.
Difference (percent)	-24.8	-23.3	-13.7	n.a.
<i>East Forest</i>				
Mean size without weights	8.44	8.22	7.49	6.46
Mean size with weights	6.60	6.56	6.74	n.a.
Difference (percent)	-21.8	-20.2	-10.0	n.a.
<i>West Forest</i>				
Mean size without weights	8.00	7.63	6.59	5.62
Mean size with weights	6.03	5.76	5.84	n.a.
Difference (percent)	-24.6	-24.5	-11.4	n.a.
<i>Savannah</i>				
Mean size without weights	8.04	7.39	6.70	6.35
Mean size with weights	7.24	6.32	6.21	n.a.
Difference (percent)	-10.0	-14.5	-7.3	n.a.

n.a. Not applicable.

Note: No household-size weights are applied to 1988 data, which were obtained entirely under the revised sampling frame and sampling procedures.

Source: Authors' calculations.

household-size weights. Because only half of the sample in 1987 is subjected to household-size weights, the reduction in the mean is only 9 percent.

The effect of household-size weights is not, however, evenly spread throughout the country (table 7). The upward bias in mean household size appears to have been more of a problem in East Forest and West Forest and in cities other than Abidjan. The order of magnitude of the correction implied by the weights is certainly not trivial—mean household size is reduced by as much as a quarter in these cities and in West Forest. Savannah and Abidjan are least affected by the application of household-size weights, although in both cases mean household size is reduced by more than 10 percent.⁶

Similarly, these adjustments in the data have varying effects on different groupings of households. Table 8 reports mean household size with and without the application of household-size weights for eight socioeconomic groups.⁷ The

6. It may be that applying the Abidjan weights to correct for overrepresentation of better-off households partially corrected the household-size bias.

7. The categorization of households into each of these groups is based on the main source of income for the household and the employment status of the household head (it follows Grootaert 1993).

Table 8. *Mean Household Size with and without Household-Size Weights, by Socioeconomic Group, 1985–88*
(number of persons unless otherwise specified)

<i>Socioeconomic group and item</i>	<i>1985</i>	<i>1986</i>	<i>1987</i>	<i>1988</i>
<i>Côte d'Ivoire</i>				
Mean size without weights	7.99	7.72	6.85	6.17
Mean size with weights	6.51	6.27	6.20	n.a.
Difference (percent)	-18.5	-18.8	-9.5	n.a.
<i>Export crop farmers</i>				
Mean size without weights	9.28	9.16	7.75	6.75
Mean size with weights	6.92	6.97	7.02	n.a.
Difference (percent)	-25.4	-23.9	-9.4	n.a.
<i>Food crop farmers</i>				
Mean size without weights	8.37	7.70	7.14	6.38
Mean size with weights	7.06	6.48	6.54	n.a.
Difference (percent)	-15.7	-15.8	-8.4	n.a.
<i>Public sector employees</i>				
Mean size without weights	7.42	7.68	7.03	7.33
Mean size with weights	6.35	6.52	6.43	n.a.
Difference (percent)	-14.4	-15.1	-8.5	n.a.
<i>Formal private sector employees</i>				
Mean size without weights	7.14	6.40	6.09	5.60
Mean size with weights	6.33	5.75	5.87	n.a.
Difference (percent)	-11.3	-10.2	-3.6	n.a.
<i>Informal private sector employees</i>				
Mean size without weights	5.00	4.54	4.60	4.34
Mean size with weights	4.27	3.62	4.33	n.a.
Difference (percent)	-14.6	-20.3	-5.8	n.a.
<i>Self-employed</i>				
Mean size without weights	7.41	7.51	5.96	5.15
Mean size with weights	5.74	5.78	5.09	n.a.
Difference (percent)	-22.5	-23.0	-14.6	n.a.
<i>Inactive</i>				
Mean size without weights	8.09	7.58	6.71	5.82
Mean size with weights	5.78	6.02	5.86	n.a.
Difference (percent)	-28.6	-20.6	-12.7	n.a.
<i>Unemployed</i>				
Mean size without weights	6.99	4.59	5.74	5.71
Mean size with weights	6.40	4.75	5.38	n.a.
Difference (percent)	-8.4	3.5	-6.3	n.a.

n.a. Not applicable.

Note: No household-size weights are applied to 1988 data, which were obtained entirely under the revised sampling frame and sampling procedures.

Source: Authors' calculations.

mean size of households that farm export crops is reduced by about a quarter, with similar orders of magnitude applying to the self-employed, to workers in the informal sector (at least in 1986), and to the inactive. The overrepresentation of larger households in the sample affects some socioeconomic groups more than others for two reasons. First, the main geographical location is not the same for each group. For example, export crop farmers are located mainly in East Forest and West Forest, where overrepresentation of larger households was

greater, rather than in Savannah, where the sampling bias was not as significant. Similarly, the relatively small effect that household-size weighting has on formal sector workers (both private and public) arises from their being located predominantly in Abidjan. The second reason is that household size is related systematically to socioeconomic category. Very large households—where the reweighting has the most pronounced effect—are more prevalent among the urban self-employed and among households that farm export crops.

III. IMPLICATIONS FOR ESTIMATES OF MEAN HOUSEHOLD EXPENDITURE

The fact that applying household-size weights to the CILSS data set reduces mean household size significantly, with the effects being concentrated among certain areas and socioeconomic groups, suggests that this procedure may have an important bearing on other findings of the survey. We wish to assess the effects of correcting for household-size-related sampling bias on three types of variables: total household expenditure per capita; estimates of the incidence of poverty, based on per capita expenditure as an indicator of welfare; and basic-needs indicators. We intend to answer two important questions. First, does the sampling bias for the early years of the survey (and its correction through the application of household-size weights) significantly affect the means of the variables? Second, are the patterns of expenditure per capita, poverty, and basic-needs fulfillment changed, both regionally and across socioeconomic groups, as a result of the weighting?

There are both theoretical and practical advantages to using total household expenditure as a measure of household welfare, and this is reflected in the popularity of this variable (expressed in per capita terms) in many of the poverty studies conducted with CILSS data (Kakwani 1990; Kanbur 1990; Glewwe 1991). It seems sensible, therefore, to begin by assessing how the household-size weights affect mean total household expenditure per capita.

Table 9 reports two estimates of mean total household expenditure per capita for each year. The first of the two estimates has not been adjusted for household size; in the second estimate household-size weights have been applied. For Côte d'Ivoire as a whole, the effect is to increase mean expenditure by 7.5 percent in 1985, 9.7 percent in 1986, and 3.2 percent in 1987. The lower figure for 1987 is to be expected, given that only half of the sample for that year was subjected to the household-size weights. Clearly, per capita expenditure of larger households is lower than average, so reducing their weight increases the overall mean. The order of magnitude of the error introduced by the sample bias is not insignificant: in 1986, for example, the original estimates of mean per capita expenditure for Côte d'Ivoire as a whole were about 10 percent too low.

Of equal significance is the uneven effect that the household-size weights have across the regions. Predictably, given the fact that mean household-size corrections were least in Abidjan and Savannah, per capita expenditures in these regions are also the least affected by the household-size weights. And again, the

Table 9. *Mean Total Household Expenditure per Capita with and without Household-Size Weights, by Region, 1985–88*
(CFA francs per year unless otherwise specified)

<i>Area and item</i>	<i>1985</i>	<i>1986</i>	<i>1987</i>	<i>1988</i>
<i>Côte d'Ivoire</i>				
Mean expenditure without weights	221,313	204,183	210,286	173,072
Mean expenditure with weights	237,853	223,905	216,965	n.a.
Difference (percent)	7.47	9.66	3.18	n.a.
<i>Abidjan</i>				
Mean expenditure without weights	357,410	291,316	361,544	288,708
Mean expenditure with weights	376,108	312,859	372,361	n.a.
Difference (percent)	5.23	7.40	2.99	n.a.
<i>Other cities</i>				
Mean expenditure without weights	240,638	235,810	237,570	178,165
Mean expenditure with weights	271,864	270,540	250,010	n.a.
Difference (percent)	12.98	14.73	5.24	n.a.
<i>East Forest</i>				
Mean expenditure without weights	156,342	162,980	165,271	152,501
Mean expenditure with weights	164,472	172,341	167,974	n.a.
Difference (percent)	5.20	5.74	1.64	n.a.
<i>West Forest</i>				
Mean expenditure without weights	225,277	184,427	170,298	143,947
Mean expenditure with weights	239,134	204,457	169,776	n.a.
Difference (percent)	6.15	10.86	−0.31	n.a.
<i>Savannah</i>				
Mean expenditure without weights	148,156	147,466	133,947	120,684
Mean expenditure with weights	152,573	154,676	136,061	n.a.
Difference (percent)	2.98	4.89	1.58	n.a.

n.a. Not applicable.

Note: No household-size weights are applied to 1988 data, which were obtained entirely under the revised sampling frame and sampling procedures.

Source: Authors' calculations.

most serious bias appears in other cities and West Forest, at least in 1985 and 1986. Although the correction does not change the rank ordering of the regions, it does noticeably change some differentials. For example, mean per capita expenditure in Abidjan was originally estimated to be 32.7 percent higher than in other cities in 1985 (19.1 percent in 1986). After correcting for the overrepresentation of larger households in the sample in 1985, the differential is reduced to 27.7 percent (13.5 percent in 1986).

Somewhat puzzling is the influence of the household-size weights on West Forest in 1987, where mean per capita expenditure is marginally reduced. The reduction occurs despite a decrease of 11 percent in mean household size (table 7). It seems that, for 1987 at least, the larger households in West Forest enjoyed above-average expenditure levels compared with smaller households, so that reducing their weight reduces the overall mean, albeit marginally. In all other cases, mean expenditures are adjusted upward.

How do these corrections influence the data for socioeconomic groups? The evidence of table 10 is that the sampling bias also has uneven effects on socioeconomic groups. Not surprisingly, mean expenditures of socioeconomic groups

Table 10. *Mean Total Household Expenditure per Capita with and without Household-Size Weights, by Socioeconomic Group, 1985–88*
(CFA francs per year unless otherwise specified)

<i>Socioeconomic group and item</i>	<i>1985</i>	<i>1986</i>	<i>1987</i>	<i>1988</i>
<i>Côte d'Ivoire</i>				
Mean size without weights	221,313	204,183	210,286	173,072
Mean size with weights	237,853	223,905	216,965	n.a.
Difference (percent)	7.47	9.66	3.18	n.a.
<i>Export crop farmers</i>				
Mean size without weights	175,840	166,601	158,791	140,906
Mean size with weights	194,418	182,715	158,434	n.a.
Difference (percent)	10.57	9.67	-0.22	n.a.
<i>Food crop farmers</i>				
Mean size without weights	167,235	155,292	151,795	132,809
Mean size with weights	171,802	165,034	153,771	n.a.
Difference (percent)	2.73	6.27	1.31	n.a.
<i>Public sector employees</i>				
Mean size without weights	394,987	356,881	349,514	254,402
Mean size with weights	431,171	401,988	364,820	n.a.
Difference (percent)	9.16	12.64	4.38	n.a.
<i>Formal private sector employees</i>				
Mean size without weights	346,100	314,902	367,494	278,708
Mean size with weights	364,835	328,060	385,322	n.a.
Difference (percent)	5.41	4.18	4.85	n.a.
<i>Informal private sector employees</i>				
Mean size without weights	223,698	169,723	189,422	144,153
Mean size with weights	230,593	191,132	186,294	n.a.
Difference (percent)	3.08	12.61	-1.65	n.a.
<i>Self-employed</i>				
Mean size without weights	217,439	200,874	206,577	174,333
Mean size with weights	234,910	217,634	212,771	n.a.
Difference (percent)	8.03	8.34	3.00	n.a.
<i>Inactive</i>				
Mean size without weights	233,905	214,911	201,489	191,040
Mean size with weights	253,049	233,127	201,330	n.a.
Difference (percent)	8.18	8.48	-0.08	n.a.
<i>Unemployed</i>				
Mean size without weights	320,954	247,561	193,179	190,503
Mean size with weights	306,624	228,037	203,729	n.a.
Difference (percent)	-4.46	-7.89	5.46	n.a.

n.a. Not applicable.

Note: No household-size weights are applied to 1988 data, which were obtained entirely under the revised sampling frame and sampling procedures.

Source: Authors' calculations.

based largely in Abidjan (mainly formal private sector employees) are affected little by the household-size weights. Similarly, those based predominantly in Savannah (food crop farmers) are also affected less than other groups. However, mean expenditure of export crop farmers (especially those located in West Forest) increases noticeably.

There are some minor changes in the rank ordering of the groups (for example, informal sector employees and the self-employed switch rank order in

1985), but the main effect of the household-size weights is to change differentials in the welfare measure. For example, per capita expenditure in 1985 is recorded as 4.9 percent higher for export crop farmers than for food crop farmers, based on the original biased data. This differential increases to 11.6 percent as a result of the weighting. Similarly, public sector employees appear to have enjoyed expenditure levels that were 11.8 percent higher than those of their counterparts in the formal private sector in 1986, whereas in reality, that is, when the sampling bias was corrected, the difference was 18.4 percent. Such adjustments in measured differentials between the groups are obviously important and underline the significance of correcting for the overestimation of household size in the early years of the CILSS. Indeed, the differences are large enough to be of more than mere academic interest and to have a potential effect on policy interventions. Although the corrections may not imply changes of target groups as such (at least not at a relatively aggregated level), they do affect significantly the amount of resources needed for interventions. Even at the qualitative level, the changes in interregional and intergroup differentials following the corrections may make the difference between triggering interventions or not.

IV. ESTIMATES OF POVERTY

The fact that mean per capita expenditure is generally underestimated as a result of the sampling bias suggests that poverty is likely to have been overestimated in the early years of the survey. In assessing how the household-size weights affect measures of poverty, we compare poverty measures computed with and without the weights.

Estimates of poverty have been computed by Grootaert (1993) for all four years of the CILSS. These estimates utilize a regional price index estimated by Grootaert and Kanbur (1993) and the household-size weights derived here. The poverty line (CFAF128,600) was selected to classify 30 percent of the population as poor in 1985. An alternative line (CFAF75,000) identifies 10 percent of the population as poor—termed the very poor. Grootaert (1993) concludes that the incidence of overall poverty was unchanged between 1985 and 1986 but that the incidence of extreme poverty fell. After 1986 the incidence of both poverty and extreme poverty rose. This was especially true of 1987–88, when the poverty head count ratio—the fraction of total population living in poverty—increased from 34.8 to 45.9 percent (and from 9.1 to 14.1 percent for extreme poverty).

What would these results have been without the application of household-size weights? Tables 11 and 12 estimate the head count ratio with and without the application of the household-size weights (the former as reported by Grootaert 1993) for the whole country and by region for the poor and the very poor, respectively. As expected, the weights noticeably reduce the incidence of poverty. This is especially true for 1986, with estimates of the head count ratio for the

Table 11. *Estimates of Poverty: Head Count Ratio with and without Household-Size Weights, by Region, 1985–88*

Area and item	1985	1986	1987	1988
<i>Côte d'Ivoire</i>				
Head count ratio without weights	0.324	0.346	0.354	0.459
Head count ratio with weights	0.300	0.299	0.348	n.a
Difference (percent)	-7.41	-13.58	-1.69	n.a
<i>Abidjan</i>				
Head count ratio without weights	0.041	0.206	0.078	0.139
Head count ratio with weights	0.034	0.166	0.074	n.a
Difference (percent)	-17.07	-19.42	-5.13	n.a
<i>Other cities</i>				
Head count ratio without weights	0.272	0.277	0.225	0.409
Head count ratio with weights	0.236	0.223	0.244	n.a
Difference (percent)	-13.24	-19.49	8.44	n.a
<i>East Forest</i>				
Head count ratio without weights	0.501	0.425	0.445	0.494
Head count ratio with weights	0.479	0.395	0.435	n.a
Difference (percent)	-4.39	-7.06	-2.25	n.a
<i>West Forest</i>				
Head count ratio without weights	0.209	0.283	0.375	0.553
Head count ratio with weights	0.178	0.200	0.376	n.a
Difference (percent)	-14.83	-29.33	0.27	n.a
<i>Savannah</i>				
Head count ratio without weights	0.523	0.522	0.589	0.652
Head count ratio with weights	0.502	0.481	0.578	n.a
Difference (percent)	-4.02	-7.85	-1.87	n.a

n.a. Not applicable.

Note: The head count ratio is the fraction of the total population living below the poverty line, defined here as per capita annual household expenditure of less than CFAF128,600 in constant 1985 CFA francs. No household-size weights are applied to 1988 data, which were obtained entirely under the revised sampling frame and sampling procedures.

Source: Grootaert (1993) for head count ratio with weights; otherwise authors' calculations.

country as a whole being adjusted downward by 13.6 percent for the poor and 12.3 percent for the very poor.

The original, biased results showed a slow but steady increase in poverty between 1985 and 1987. The corrected results indicate not only that the level of poverty was overestimated, but that the trend was different as well. With the corrected data, poverty is shown to remain steady between 1985 and 1986 and then to rise rapidly in 1987. Clearly, any attempt to link changes in poverty causally to macroeconomic events or policy would produce very different answers, depending on whether the original or corrected results were used. The corrected results imply a much greater success in protecting the poor during the adjustment program and also suggest that the destabilization that started in 1987 following the government's abandonment of the adjustment effort led to much more rapid and severe rises in poverty than the uncorrected results had indicated (see Grootaert 1993 for a detailed analysis of the links between macroeconomic policy and poverty).

Table 12. *Estimates of Extreme Poverty: Head Count Ratio with and without Household-Size Weights, by Region, 1985–88*

<i>Area and item</i>	<i>1985</i>	<i>1986</i>	<i>1987</i>	<i>1988</i>
<i>Côte d'Ivoire</i>				
Head count ratio without weights	0.105	0.073	0.092	0.141
Head count ratio with weights	0.100	0.064	0.091	n.a.
Difference (percent)	-4.76	-12.33	-1.09	n.a.
<i>Abidjan</i>				
Head count ratio without weights	0.006	0.012	0.011	0.000
Head count ratio with weights	0.007	0.014	0.017	n.a.
Difference (percent)	16.67	16.67	54.55	n.a.
<i>Other cities</i>				
Head count ratio without weights	0.092	0.071	0.052	0.073
Head count ratio with weights	0.080	0.055	0.047	n.a.
Difference (percent)	-13.04	-22.54	-9.61	n.a.
<i>East Forest</i>				
Head count ratio without weights	0.145	0.116	0.088	0.139
Head count ratio with weights	0.132	0.095	0.086	n.a.
Difference (percent)	-8.97	-18.10	-2.27	n.a.
<i>West Forest</i>				
Head count ratio without weights	0.023	0.016	0.091	0.161
Head count ratio with weights	0.016	0.019	0.093	n.a.
Difference (percent)	-30.43	18.75	2.20	n.a.
<i>Savannah</i>				
Head count ratio without weights	0.227	0.123	0.195	0.305
Head count ratio with weights	0.226	0.121	0.194	n.a.
Difference (percent)	-0.44	-1.63	-0.51	n.a.

n.a. Not applicable.

Note: The head count ratio is the fraction of the total population living below the extreme poverty line, defined here as per capita annual household expenditure of less than CFAF75,000 in constant 1985 CFA francs. No household-size weights are applied to 1988 data, which were obtained entirely under the revised sampling frame and sampling procedures.

Source: Grootaert (1993) for head count ratio with weights; otherwise authors' calculations.

Just as total household expenditure per capita shows changes following the application of the weights (table 9), so do head count estimates of poverty and extreme poverty (tables 11 and 12). The main differences occur in 1986, with downward adjustments in the poverty estimates for all regions, except for the estimates of extreme poverty in Abidjan and West Forest, which are revised upward. The upward revision suggests that very poor households are smaller than average in these areas. (The household-size weights increase the representation of small households and decrease that of larger households.) Although large households generally have lower levels of expenditure per capita, clearly this is not always the case for the very poor. Why should this be true for 1986 and not for 1985 (when the estimate of extreme poverty is seen to be reduced by the household-size weights) is a matter for conjecture at this stage. The measured incidence of poverty in Abidjan is reduced significantly (by 17.1 percent in 1985 and 19.4 percent in 1986) as a result of the household-size weights, even though mean expenditure in Abidjan is increased only slightly by the application of the

weights. This suggests that expenditures of many of the poor in Abidjan are only marginally below the poverty line.

Tables 13 and 14 report estimates of the incidence of poverty and extreme poverty (respectively) by socioeconomic group. The reduction in the head count estimates of poverty (table 13) resulting from household-size weights is not evenly spread across the groups. Downward adjustments for both 1985 and

Table 13. *Estimates of Poverty: Head Count Ratio with and without Household-Size Weights, by Socioeconomic Group, 1985–88*

<i>Socioeconomic group and item</i>	1985	1986	1987	1988
<i>Côte d'Ivoire</i>				
Head count ratio without weights	0.324	0.346	0.354	0.459
Head count ratio with weights	0.300	0.299	0.348	n.a.
Difference (percent)	-7.41	-13.58	-1.69	n.a.
<i>Export crop farmers</i>				
Head count ratio without weights	0.428	0.414	0.480	0.548
Head count ratio with weights	0.366	0.354	0.477	n.a.
Difference (percent)	-14.48	-14.49	-0.63	n.a.
<i>Food crop farmers</i>				
Head count ratio without weights	0.445	0.450	0.478	0.590
Head count ratio with weights	0.434	0.411	0.473	n.a.
Difference (percent)	-2.47	-8.67	-1.05	n.a.
<i>Public sector employees</i>				
Head count ratio without weights	0.062	0.071	0.075	0.213
Head count ratio with weights	0.049	0.056	0.072	n.a.
Difference (percent)	-20.97	-21.13	-4.00	n.a.
<i>Formal private sector employees</i>				
Head count ratio without weights	0.069	0.131	0.062	0.151
Head count ratio with weights	0.071	0.096	0.061	n.a.
Difference (percent)	2.90	-26.72	-1.61	n.a.
<i>Informal private sector employees</i>				
Head count ratio without weights	0.312	0.507	0.363	0.542
Head count ratio with weights	0.262	0.401	0.364	n.a.
Difference (percent)	-16.02	-20.91	0.28	n.a.
<i>Self-employed</i>				
Head count ratio without weights	0.282	0.332	0.330	0.462
Head count ratio with weights	0.262	0.287	0.333	n.a.
Difference (percent)	-7.09	-13.55	0.91	n.a.
<i>Inactive</i>				
Head count ratio without weights	0.200	0.273	0.321	0.319
Head count ratio with weights	0.183	0.211	0.327	n.a.
Difference (percent)	-8.50	-22.71	1.87	n.a.
<i>Unemployed</i>				
Head count ratio without weights	0.024	0.348	0.393	0.383
Head count ratio with weights	0.041	0.346	0.312	n.a.
Difference (percent)	70.83	-0.57	-20.61	n.a.

n.a. Not applicable.

Note: The head count ratio is the fraction of the total population living below the poverty line, defined here as per capita annual household expenditure of less than CFAF128,600 in constant 1985 CFA francs. No household-size weights are applied to 1988 data, which were obtained entirely under the revised sampling frame and sampling procedures.

Source: Grootaert (1993) for head count ratio with weights; otherwise authors' calculations.

Table 14. *Estimates of Extreme Poverty: Head Count Ratio with and without Household-Size Weights, by Socioeconomic Group, 1985–88*

<i>Socioeconomic group and item</i>	1985	1986	1987	1988
<i>Côte d'Ivoire</i>				
Head count ratio without weights	0.105	0.073	0.092	0.141
Head count ratio with weights	0.100	0.064	0.091	n.a.
Difference (percent)	-4.76	-12.33	-1.09	n.a.
<i>Export crop farmers</i>				
Head count ratio without weights	0.116	0.100	0.144	0.210
Head count ratio with weights	0.086	0.081	0.148	n.a.
Difference (percent)	-25.86	-19.00	2.78	n.a.
<i>Food crop farmers</i>				
Head count ratio without weights	0.147	0.106	0.114	0.197
Head count ratio with weights	0.150	0.101	0.111	n.a.
Difference (percent)	2.04	-4.72	-2.63	n.a.
<i>Public sector employees</i>				
Head count ratio without weights	0.002	0.000	0.009	0.039
Head count ratio with weights	0.003	0.000	—	n.a.
Difference (percent)	50.00	0.00	—	n.a.
<i>Formal private sector employees</i>				
Head count ratio without weights	0.005	0.000	0.013	0.007
Head count ratio with weights	0.008	0.000	0.013	n.a.
Difference (percent)	60.00	0.00	0.00	n.a.
<i>Informal private sector employees</i>				
Head count ratio without weights	0.113	0.000	0.067	0.220
Head count ratio with weights	0.099	0.000	0.067	n.a.
Difference (percent)	-12.39	0.00	0.00	n.a.
<i>Self-employed</i>				
Head count ratio without weights	0.122	0.067	0.067	0.107
Head count ratio with weights	0.115	0.052	0.070	n.a.
Difference (percent)	-5.74	-22.39	4.48	n.a.
<i>Inactive</i>				
Head count ratio without weights	0.071	0.031	0.159	0.072
Head count ratio with weights	0.063	0.028	0.176	n.a.
Difference (percent)	-11.27	-9.68	10.69	n.a.
<i>Unemployed</i>				
Head count ratio without weights	0.000	0.087	0.000	0.088
Head count ratio with weights	0.000	0.137	0.000	n.a.
Difference (percent)	0.00	57.47	0.00	n.a.

n.a. Not applicable; — not available.

Note: The head count ratio is the fraction of the total population living below the poverty line. Here the extreme poverty line is defined as per capita annual household expenditure of less than CFAF75,000 in constant 1985 CFA francs. No household-size weights are applied to 1988 data, which were obtained entirely under the revised sampling frame and sampling procedures.

Source: Grootaert (1993) for head count ratio with weights; otherwise authors' calculations.

1986 are greatest for export crop farmers, public sector employees, and informal private sector employees, and least for food crop farmers. For other groups the effect differs by year. For example, head count poverty estimates for formal private sector employees are not adjusted greatly in 1985 but are revised downward by 26.7 percent in 1986. There are no changes in the rank ordering for 1985 as a result of the household-size weights, but there are some changes induced in 1986. Whereas without the household-size weights the incidence of

poverty in 1986 appears to be greatest among informal private sector employees, this group moves behind food crop farmers when the weights are used.

The main implications for the estimates of extreme poverty (table 14) are that the household-size weights reduce extreme poverty significantly among export crop farmers and informal sector employees in 1985 and among export crop farmers and the self-employed in 1986. The incidence of extreme poverty among food crop farmers is not changed greatly.

Do these findings critically depend on the choice of poverty index and poverty line selected? Dominance tests were applied to the data, comparing the distribution before and after the application of the household-size weights (see Atkinson 1987 and Ravallion 1992 for further details on these tests). Figures 1 and 2 compare the cumulative distribution of household expenditure per capita with and without household-size weights for 1985 and 1986, respectively. For both years, the unweighted distribution lies entirely to the left and above the original one. This first-order dominance indicates that whichever poverty index or poverty line is selected, measured poverty will be greater for the unweighted distribution. This result further supports our conclusion that the original, uncorrected data yield overestimates of poverty in Côte d'Ivoire for 1985 and 1986.

As a methodological point, the revisions in poverty estimates following the application of the corrective household-size weights are the net results of two effects, operating, respectively, on the mean and the distribution. Because house-

Figure 1. *Distribution of Household Expenditure per Capita in Côte d'Ivoire, 1985*

Figure 2. *Distribution of Household Expenditure per Capita in Côte d'Ivoire, 1986*

hold size is related to the level of expenditure, the application of weights changes the mean level of expenditure (see previous section). However, because large households are located with different frequency along the distribution, the application of weights also alters the distribution (even if there is no change in the mean). It is possible to compute the contribution of each effect on the change in the poverty index. Such computations show that the mean effect dominates and is negative (that is, it reduces the poverty index). The signs and magnitudes of the distribution effects vary across regions. As we have seen, the net effect in most cases is to reduce the estimate of poverty incidence.

In summary, more than was the case for the estimate of household expenditure, the measured incidence of poverty is significantly affected by sampling bias in the CILSS. For several socioeconomic groups and regions, uncorrected figures overestimate poverty and extreme poverty by 20 to 30 percent—biases that could well be the difference between whether or not such groups become targets for government programs to alleviate poverty.

The changes we have observed in group differentials and (to a lesser extent) in rank ordering have significance for past studies reporting poverty patterns in Côte d'Ivoire based on CILSS data. For example, Kanbur (1990) decomposes the poverty index for 1985 by socioeconomic group in assessing the effects of struc-

tural adjustment.⁸ The evidence we have presented on estimates of group poverty suggests that Kanbur's findings may not be robust to the corrections for sampling bias we have discussed. For example, Kanbur places export crop farmers second in a rank ordering of extreme poverty (behind food crop farmers). However, the household-size weights significantly affect the estimate of extreme poverty for this group, reducing the head count from 11.6 percent to 8.6 percent (table 14). It is clear that Kanbur's results, which were based on data subject to an uncorrected sampling bias, must be interpreted with care. Similarly, Kakwani (1990) analyzes the statistical significance of differences between group poverty levels. His groupings are based on region and on various characteristics of the household head. The adjustments made by the household-size weights to group poverty differentials suggest that Kakwani's results may well have been different had he accounted for the underlying sampling bias. Similar comments may apply to other studies that rely on money-metric measures of welfare and poverty, such as Glewwe (1987).

V. BASIC-NEEDS INDICATORS

Taking per capita expenditure as a measure of economic welfare, the CILSS data for 1985 and 1986 have been found to underestimate living standards and to overestimate poverty. To what extent does this result also apply to other welfare indicators? This section assesses how selected basic-needs indicators are changed through the application of the household-size weights.

Table 15 reports four basic-needs indicators derived from the CILSS, disaggregated by region.⁹ The household-size weights make very little difference to the indicators. For the country as a whole, the indicators are marginally reduced by the weights. That there is little adjustment arising from household-size weights suggests that the basic-needs indicators are not closely related to household size. An examination of the data confirms this.¹⁰

An interesting feature of the results is the difference in the direction of change between Abidjan and the other regions. For Abidjan, most indicators are adjusted upward by the household-size weights, whereas for the other regions the adjustment tends to be downward. Although most of these adjustments are small, they suggest that larger households in Abidjan have fewer basic needs fulfilled than do smaller households, so that reducing the weight of larger households in the sample increases the indicator. The opposite seems to apply to the

8. Our estimates of household expenditure per capita differ from those reported in Kanbur (1990). This is because of variations in data cleaning and editing procedures.

9. For basic-needs indicators, the disaggregation by region is particularly important, given the role played by the supply of basic-needs services (such as schooling and health clinics), which varies tremendously among regions in Côte d'Ivoire.

10. For example, literacy rates in other cities in 1986 varied from 44.2 percent for households with 1 or 2 members, to 48.1 percent for households with 10 to 17 members, and to 37.3 percent for households with 18 or more members. As an example of a rural area, comparable figures for East Forest are 18.9, 28.0, and 28.5, respectively.

Table 15. *Selected Basic-Needs Indicators with and without Household-Size Weights, by Region, 1985–86*
(percent)

Area and item	1985				1986			
	Adult literacy rate	Net primary enrollment rate ^a	Households with access to tap water	Ill people who consult doctor or nurse	Adult literacy rate	Net primary enrollment rate ^a	Households with access to tap water	Ill people who consult a doctor or nurse
<i>Côte d'Ivoire</i>								
Without weights	31.8	46.4	22.0	47.2	31.8	46.6	22.9	44.8
With weights	30.9	46.2	21.8	45.8	31.9	46.3	22.7	43.9
Difference	-2.83	-0.43	-0.91	-2.97	0.31	-0.64	-0.87	-2.01
<i>Abidjan</i>								
Without weights	54.3	59.8	45.3	63.0	51.9	63.6	49.4	63.6
With weights	55.0	63.3	47.3	62.5	53.1	63.9	48.5	63.6
Difference	1.29	5.85	4.42	-0.79	2.31	0.47	-1.82	0.00
<i>Other cities</i>								
Without weights	46.3	60.7	47.5	55.3	44.5	55.4	46.3	56.0
With weights	45.7	60.6	46.5	53.9	45.5	55.2	46.8	55.8
Difference	-1.30	-0.16	-2.11	-2.53	2.25	-0.36	1.08	-0.36
<i>East Forest</i>								
Without weights	23.6	45.5	10.0	46.7	25.1	49.3	8.8	49.7
With weights	21.3	45.0	8.5	43.8	24.2	48.5	7.2	47.2
Difference	-9.75	-1.10	-15.00	-6.21	-3.59	-1.62	-18.18	-5.03
<i>West Forest</i>								
Without weights	21.3	43.9	—	41.1	21.4	37.6	—	41.9
With weights	20.1	43.2	—	39.6	30.6	36.7	—	39.0
Difference	-5.63	-1.59	—	-3.65	42.99	-2.39	—	-6.92
<i>Savannah</i>								
Without weights	10.0	21.1	1.8	30.4	11.3	21.0	3.0	21.6
With weights	9.9	21.3	1.8	30.3	10.7	20.6	2.9	22.1
Difference	-1.00	0.95	0.00	-0.33	-5.31	-1.90	-3.33	2.31

—Not available.

a. Percentage of primary school-age children in primary school.

Source: Authors' calculations.

other regions—larger households on balance have slightly more basic needs fulfilled.

Although the household-size weights make very little difference to the mean value of these indicators, the correction of the 1985 and 1986 data sets is nevertheless important for basic-needs analysis. Table 16 disaggregates the indicators by poverty status, showing how they are affected by household-size weights for the very poor, the midpoor, and the nonpoor. Although the weights make little difference to the measures for the nonpoor, they are seen to have a far more significant effect on the basic-needs indicators of the poor, especially the very poor. For example, literacy rates among the very poor are adjusted downward by 12 percent in 1985 and 10 percent in 1986. Similarly, primary school enrollment rates for the very poor are adjusted downward by 10 percent in 1985 and 5 percent in 1986.

Table 16. *Selected Basic-Needs Indicators, by Poverty Status, 1985–86*
(percent)

	1985			1986		
	Without house- hold-size weight	With house- hold-size weight	Differ- ence	Without house- hold-size weight	With house- hold-size weight	Differ- ence
<i>Adult literacy rate</i>						
Very poor	17.1	15.0	-12.3	14.9	13.4	-10.1
Midpoor	21.8	19.3	-11.5	23.9	23.1	-3.3
Nonpoor	37.2	36.3	-2.4	36.7	36.2	-1.4
All	31.8	30.9	-2.8	31.8	31.9	0.3
<i>Net primary enrollment rate^a</i>						
Very poor	27.3	24.6	-9.9	24.5	23.2	-5.3
Midpoor	39.8	39.4	-1.0	37.8	36.8	-2.6
Nonpoor	51.8	51.9	0.2	53.2	52.5	-1.3
All	46.4	46.2	-0.4	46.6	46.3	-0.6
<i>Households with access to tap water</i>						
Very poor	1.3	1.5	15.4	2.3	4.4	91.3
Midpoor	9.9	7.9	-20.2	8.6	7.7	-10.5
Nonpoor	29.2	28.7	-1.7	31.2	29.4	-5.8
All	22.0	21.8	-0.9	22.9	22.7	-0.9
<i>Ill people who consult a doctor or nurse</i>						
Very poor	32.8	30.7	-6.4	30.9	29.2	-5.5
Midpoor	38.2	36.5	-4.5	32.4	30.3	-6.5
Nonpoor	51.5	49.7	-3.5	51.5	49.7	-3.5
All	47.2	45.8	-3.0	44.8	43.9	-2.0

Note: The very poor are those below the extreme poverty line, CFAF75,000; the midpoor are those above the extreme poverty line but below the poverty line, CFAF128,600. The poverty lines are in 1985 constant CFA francs.

a. Percentage of primary school-age children in primary school.

Source: Authors' calculations.

The adjustments underscore again the need to ensure that any sampling bias is identified and corrected. In the case of the CILSS, the corrections for basic-needs indicators are only quantitatively important for the poor, but this result must be considered unusual. In most cases, basic-needs fulfillment, like welfare in general, is negatively correlated with household size, and thus sampling bias affecting household selection would be expected to give misleading estimates of the indicators at the national level as well. At any rate, the magnitudes of the corrections noted for the poor in Côte d'Ivoire are sufficiently important to affect the targeting of programs of basic-needs provision to the poor, in the sense that the needs of the poor, and especially the very poor, are higher than previously estimated.

VI. CONCLUDING OBSERVATIONS

This article illustrates the need for analysts to investigate the properties of the data in use, including its sampling design, although this does not appear to be a frequent practice among users of household survey data. Our study of the Côte d'Ivoire Living Standards Survey, 1985–88, reveals flawed sampling procedures, which have not been corrected by previous users of the data. The flawed procedures have resulted in biases in estimates of household size, which in turn have yielded biased estimates of household expenditure and of poverty. Basic-needs indicators reported for the country as a whole have been less affected, but the bias has proved to be more pronounced at the lower end of the distribution.

The correction procedure described here applied suitably constructed household-size weights to the data. The quantitative adjustments to estimated variables proved to be nontrivial and underlined the seriousness of problems arising from sampling biases. For example, the head count estimate of poverty in Côte d'Ivoire was found to have been overestimated by 14 percent in 1986. The bias proved to differ widely across regions and socioeconomic groups and was frequently in the order of magnitude of 20 to 30 percent. Such differences are not merely of academic interest; they can affect policy interventions that are guided by survey results.

The correction of sampling bias also affects time-series analysis of CILSS data. Original results displayed a gradually rising trend in poverty in Côte d'Ivoire from 1985 to 1987. In fact, poverty did not change (or even fell marginally) from 1985 to 1986, although it did increase rapidly in 1987. These results clearly affect the assessment of macroeconomic policies conducted during the period, especially because in 1987 the government of Côte d'Ivoire abandoned a previously sustained adjustment program. The overestimation of poverty in 1985 obviously also means that the total increase in poverty from 1985 to 1988 was underestimated. Taking the head count ratio as an example, poverty was recorded as increasing from 32.4 percent in 1985 to 45.9 percent in 1988 (an increase of 41.7 percent). However, the weighted head count for 1985 is only 30.0 percent, so poverty in fact increased by 53 percent during the period. The

underestimation of the trend in poverty is even clearer for regional estimates. In West Forest, for example, the unweighted data indicate an increase of 164.6 percent, whereas the weighted data indicate an increase of 210.7 percent. Without applying household-size weights, time-series analyses of CILSS data are certain to be subject to errors of this order.

The CILSS data, especially for 1985–86, have been widely used to research welfare and poverty issues. To our knowledge, none of this research has applied any corrections for the sampling bias noted here. The failure to do so casts serious doubts on the robustness of many of the findings reported earlier, particularly because our results have shown that the order of magnitude of the bias is not trivial and will affect welfare and poverty measures regardless of the poverty index and poverty line chosen. More important, policy recommendations made earlier may need to be revised. Our analysis has concentrated on measures of income and expenditure and on the poverty indexes that are computed thereby. We have not assessed how the weighting procedure we selected would affect results based on multivariate analyses, which many of the studies using the CILSS have used.

As a practical matter, what can users of household survey data do to detect possible sampling problems? Our experience suggests that the first line of defense is a thorough knowledge of the country being researched. It was this knowledge that alerted us to the surprisingly large household-size estimates in 1985 and 1986, estimates that were out of line with all other information, such as the 1988 population census.¹¹ Second, if multiple years of data are available, “jumps” in a given variable are a clear and easily noted signal. Third, and obviously, a critical review of sampling procedures and any changes therein is essential—even if data have been used before. We recognize that such a review is difficult in many cases because survey documentation tends to be poor. We found that direct contact with survey staff was extremely helpful in clarifying matters. Notwithstanding the difficulties, tracing sampling information can be worth the effort when research aims to be of more than mere methodological interest.

REFERENCES

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

Atkinson, Anthony B. 1987. “On the Measurement of Poverty.” *Econometrica*, 55:749–64.

Coulombe, Harold, and Lionel Demery. 1993. *Household Size in the Côte d'Ivoire: Sampling Bias in the Living Standards Survey*. LSMS 97. Washington, D.C.: World Bank.

11. Lachaud (1988) also expressed concern about the high CILSS estimate of household size compared with his survey findings for Abidjan.

- Daho, Bakary. 1992. "La Qualité des Données de l'Enquête Permanente Auprès des Ménages de Côte d'Ivoire." World Bank, Africa Technical Department, Washington, D.C. Processed.
- Deaton, Angus. 1987. "The Estimation of Own and Cross Price Elasticities from Household Survey Data." *Journal of Econometrics* 36(1/2):7-30.
- . 1989. "Looking for Boy-Girl Discrimination in Household Expenditure Survey Data." *The World Bank Economic Review* 3(1):1-16.
- Glewwe, Paul. 1987. *The Distribution of Welfare in the Republic of Côte d'Ivoire in 1985*. LSMS 29. Washington, D.C.: World Bank.
- . 1991. "Investigating the Determinants of Household Welfare in the Côte d'Ivoire." *Journal of Development Economics* 35: 307-37.
- Grootaert, Christiaan. 1986. *Measuring and Analyzing Levels of Living in Developing Countries: An Annotated Questionnaire*. LSMS 24. Washington, D.C.: World Bank.
- . 1987. "Côte d'Ivoire's Vocational and Technical Education." wps 19. World Bank, Population and Human Resources Department, Washington, D.C. Processed.
- . 1990. "Returns to Formal and Informal Vocational Education in Côte d'Ivoire: The Role of the Structure of the Labor Market." *Economics of Education Review* 9(4):309-19.
- . 1993. "The Evolution of Welfare and Poverty under Structural Change and Economic Recession in Côte d'Ivoire, 1985-88." wps 1078. World Bank, Africa Technical Department, Washington, D.C. Processed.
- Grootaert, Christiaan, and Ravi Kanbur. 1993. "A New Regional Price Index for Côte d'Ivoire Using Data from the International Comparisons Project." wps 1080. World Bank, Africa Technical Department, Washington, D.C. Processed.
- Kakwani, Nanak. 1990. *Testing for Significance of Poverty Differences with Application to Côte d'Ivoire*. LSMS 62. Washington, D.C.: World Bank.
- Kanbur, Ravi. 1990. *Poverty and the Social Dimensions of Structural Adjustment in Côte d'Ivoire*. SDA 2. Washington, D.C.: World Bank.
- Kozel, Valerie. 1990. *The Composition and Distribution of Income in Côte d'Ivoire*. LSMS 68. Washington, D.C.: World Bank.
- Lachaud, J. P. 1988. "Pauvreté et marché du travail urbain: le cas d'Abidjan (Côte d'Ivoire)." Discussion Paper DP/9/1988. International Institute of Social Studies, Geneva. Processed.
- Ravallion, Martin. 1992. *Poverty Comparisons: A Guide to Concepts and Methods*. LSMS 88. Washington, D.C.: World Bank.
- van der Gaag, Jacques, and Wim Vijverberg. 1989. "Wage Determinants in Côte d'Ivoire: Experience, Credentials, and Human Capital." *Economic Development and Cultural Change* 37(2).