NEW EVIDENCE ON THE POVERTY COUNT

Joseph J. Minarik, The Brookings Institution

New data are available which suggest that the size of the poverty population has been overestimated. The published poverty count is based on the March Supplement of the Current Population Survey (CPS), which is conducted by the Bureau of the Census. The new data are the Panel of Study of Income Dynamics (PSID) collected by the University of Michigan's Survey Research Center (SRC), under contract first from the Office of Economic Opportunity and later from the Department of Health, Education and Welfare.

The poverty counts of the two surveys are shown in Table 1. The left column shows the poverty population as a percentage of the total population according to the Current Population Survey, annually from 1960 to 1972. These figures show meaningful progress against poverty, with some cyclical sensitivity. Column 2 shows a poverty series for part of the same period derived from the Panel Study of Income Dynamics; it is consistently lower. The difference between the two surveys is significant at the ninety-nine percent level each year. This raises an interesting question. Is the new poverty count more accurate than the standard Census Bureau estimate?

To answer this question, it is necessary to examine the Current Population Survey and the Panel Study of Income Dynamics in detail, to determine which of the two data sources is a more accurate representation of the low income population.

I. The Data Sources.

Complete documentation of the two surveys is available from DUALabs and the Survey Research Center.[1] The key differences between the two surveys can be summarized quickly here:

(1) The CPS is collected by the federal government, and the PSID by a private research organization.

(2) The CPS is collected monthly; the PSID, annually.

(3) The PSID interviews the same subjects repeatedly (that is, it is "longitudinal"); the CPS makes no attempt to follow its subjects or to associate their responses over time, although it does interview most of its subjects eight times.

(4) The PSID sample size is about one-tenth that of the CPS.

(5) The PSID is a stratified sample, while the CPS is designed to be a nearly uniform random sample within the constraints imposed by an area sampling technique.

(6) Apart from the first contact with its subjects, the CPS interviews by telephone; the PSID did most of its interviewing in person until the 1972 income year survey.

II. Accounting for the CPS-PSID difference in poverty counts.

The first question on the differences in poverty counts is whether both samples are truly random. The Current Population Survey would seem to be virtually immune from sample design problems. It is based on a decennial one hundred percent sample of the population, the Census, and is updated annually by detailed data on population change. Its large sample size permits interviewing in many different geographical areas, allowing a closer compliance of the sample to the entire population.

On the other hand, perfect representation of the total population was not the primary goal of the PSID. It was designed to provide detailed information on household behavior over time, with particular emphasis on low income units. To that end, the sample was stratified with extra observations of low income households and fewer of older units whose expected lifetime was shorter and whose behavior would not be as complex. Sample design was also constrained by cost considerations; only areas where SRC had interviewers were considered for the sample, and the size of the sample was only about one-tenth of the size of the CPS. Noninterviews in the first two years of the panel were frequent. All of these constraints would lead to greater differences between the PSID and the population as a whole.

The question therefore reduces to whether the PSID is significantly different from the CPS in some important respect. In order to answer this question, tests of goodness of fit were performed between the two surveys. The distributions of the age, race, sex, and schooling of the heads of households in the survey were tested because those characteristics explain most of the systematic variation in incomes of households and therefore are of central importance in determining the poverty status of the households. The hypothesis that the distributions are the same must be rejected in every case except the sex of household heads in 1971 and 1972.[2]

The impact of the differences between the two surveys is not clear from these results, however. The PSID undercounts aged households, which should reduce the poverty count, but it also overcounts blacks and racial minorities, which should increase it. Household heads with elementary school educations are undercounted, but high school dropouts are overcounted. Therefore it is not clear whether the PSID poverty count should be too high or too low on the basis of the inaccuracy in the distribution of these variables. In order to answer this question, a set of weights was derived from the two samples in order to force the joint distribution of age, race, sex and education of household heads in the PSID to be identical to that in the CPS; the qualitative change in the poverty count using these weights would give some idea of just what the effect of the PSID inaccuracies is.

Table 2 shows the results of these experiments for the three years for which both samples were available to me. In 1967, the PSID probably had a lower poverty count because of the inaccuracy of the sample relative to the actual distribution of the variables we chose. In 1971 and 1972, however, the poverty counts are not significantly altered by the weights applied in this experiment, and indicate that it is more likely that the PSID would overestimate rather than underestimate poverty.

These results would seem to indicate that the PSID can yield valid estimates of the number of the poor. The early years of the survey were confused by the high noninterview rate and a shortened questionnaire, but in later years the more stable panel and improved techniques yielded acceptable results. While the aggregate poverty counts are not seriously biased by the inaccuracies of the sample, less aggregated results by age and race must be used with caution.

So if sampling inaccuracies do not account for the difference between the CPS and the PSID poverty counts, what does? One possibility is the fact that income reporting in the PSID matches more closely with national accounting aggregates than does that of the CPS in at least one important respect. The aggregates of reported income in Table 5 indicate that public assistance income was better reported in the PSID than in the CPS. Because public assistance is directed at low income households, the greater PSID reporting should have an impact on poverty counts, as should the higher PSID reporting of labor and asset income. Social Security income is less well reported in the PSID than the CPS until 1972, but a rough adjustment for the undercounting of heads of households over 60, mentioned above, accounts for almost all of the margin between the two surveys. The PSID also asks for the amounts of private transfers received - charity, help from relatives, and the like. The CPS asks for such transfers only if they are made on a regular basis. The differences in reporting of such transfers, which are included in the "other" category in Table 3, may be due to either the conceptual differences between the two surveys or to better reporting in general by the PSID.

It would be useful to know whether any of the differences in income reporting between the two surveys would account for the differences in the poverty counts. The easiest way to estimate the effect of the reporting differences would be to assume that all incomes in the CPS on any single type are underreported by the same fixed proportion. Under this simplifying assumption all incomes reported in the CPS could be adjusted by fixed proportions according to their type, to make CPS reporting equal to that of the PSID. One exception was made to this general technique: The Social Security income adjustment was modified to compensate for the undercounting of family heads over sixty in the PSID.

The results of this experiment, shown in Table 4, indicate that such a coarse adjustment of CPS underreporting to PSID levels accounts for almost half of the difference in the poverty counts for 1971, but only about one quarter of the difference in 1972. The additional income imputed to compensate for underreporting lowers the 1971 CPS poverty population by more than two million, but by only one and one-half million in 1972.

The experiment described above has obvious weaknesses. It assumes that all income types are underreported by fixed proportions across all households; in fact, more households zero income of any given type in the CPS than in the PSID. It also imputes income by uniform proportions across all income classes; those proportions may or may not be meaningfully accurate for households just below the poverty line. To get a bit closer to the real problem, a further experiment will try to impute income to marginally poor CPS households on the basis of the experience or marginally poor PSID households, with some disaggregation on the basis of household characteristics.

The reporting data suggest that labor and asset incomes are reported fairly well in both surveys, but that the PSID does substantially better in public assistance and miscellaneous income. This would suggest that households at the same level of total income would report more transfers in the PSID; that is, the "mix" of income by source would be different. The next experiment will take the income "mix" of marginally poor PSID households as given, apply it to the marginally poor CPS households, and examine the impact of this change on the poverty counts.

Table 5 shows income by source for households from both surveys whose income was between 90 and 100 percent of their poverty requirement in 1971. The households are disaggregated into those with head sixty-five years old or older, those with female head and children present, and all others. Not only are the mixes of income quite different between types of households, but also between the two surveys within household types; the differences are clearly great enough to have an impact on the poverty counts.

In order to apply the PSID income mix to the CPS households, the following experiment was performed: Both surveys were divided by type of household, as described above, and households within each type were divided according to the ratio of their income to needs standard, in units of one-tenth; so households whose income was between 0.0. and 0.1 of their needs standard were in one class, those between 0.1 and 0.2 were in a second class, and so forth. In each such class the mean labor and asset income, public transfer income, private transfer income, total income and needs standard were calculated. Those observations with incomes from 0.6 to 1.4 of the poverty standard were arbitrarily selected as "marginally poor" and used for the present analysis; the other observations were discarded.

Next, a simple model of income determination was formulated for each household type. It was assumed that the basic income source for the aged was public transfers (in particular Social Security), and that labor, asset and other income, which were considered endogenous, could be expressed as a function of public transfers and the needs standard, which were considered predetermined. The rationale for this formulation is that most people over sixty-five probably consider Social Security as their basic means of support, and work or seek private help if they feel that their Social Security payment does not meet their needs. For all other families, the model was changed to make labor and asset income predetermined, and public transfers endogenous. The rationale for this change is that most nonaged households support themselves mainly through labor income and seek transfers to fill any needs not met by labor income.

The equations suggested by this model were

fitted for the PSID for income years 1971 and 1972, in each year for the mean incomes of the three household types described above. Independent variables were deleted from the equations where necessary to minimize the standard errors of the predictions. The results were then used to predict the mean incomes by type for the equivalent classes of households in the CPS. Where these predictions would raise the mean income of a particular subcategory of CPS households (for example, aged households whose income is between .90 and 1.0 of their poverty thresholds) above their mean needs standard, then those households were considered to be in fact nonpoor in light of the PSID reporting data, and the CPS aggregates were adjusted accordingly. The result of this process, shown in Table 6, shows that better than eighty percent of the difference in the poverty counts is accounted for by the exercise in both years. While this exercise has obvious limitations, it does demonstrate that the differences between the two surveys in income reporting in general and in reporting among the poor and near poor in particular have much to do with the differences in poverty counts.

III. Conclusions.

The results presented thus far indicate that the Panel Study of Income Dynamics provides an adequate sample for investigations on the number of poor persons in the United States, and that its greater income reporting leads to a lower poverty count than that obtained by the CPS. These results, while of interest in themselves, raise the further question of why the PSID income reporting is greater. Some tentative answers will be raised here which, surprisingly enough, feed back to the original question of the number of the poor in the United States.

There are several reasons why one might suppose that the PSID might obtain better income reporting than the CPS. For one thing, the Census Bureau's attachment to the federal government may inhibit the reporting of income. The reinterviewing in the PSID may lead to greater cooperation by the interviewees; the panel members may well remember their incidental income items better, knowing that they will be interviewed in the coming year; and the smaller sample permits longer interviews and more detailed questioning. These suggestions are beyond the control of either organization and their importance is totally unverifiable.

There are other factors which relate to the design of the two surveys and which raise important issues. For one, the CPS collects data on the income of the "current" population only; it does not inquire about the income of decedents from households in the March survey. This "creates" a certain amount of "widows and orphans" poverty, as follows: Suppose that a husband who provides all support in a husband-wife-children family dies late in one year or even in January or February of the following year, and that the family is surveyed by the CPS in March. The CPS inquires only of the income of the surviving family members in the previous year; if they had none, then the family is recorded as having zero income and in poverty for the previous year, even if the husband's income before death put them substantially above the poverty level. The same would happen with any family separation other than death.

This factor suggest that part of the difference in poverty counts between the two surveys is probably due to a subtle difference in procedures.[3] (The PSID asks retrospective questions about the incomes of former family members and thus does not have such problems.) This is a reminder that our poverty counts through the CPS are constrained by many practical limitations of large-scale data collection. Many of these limitations form a conceptual "wedge" between a true measure of economic well being and the minimum cash income level which our poverty concept represents. This has two implications: First, it is important to keep these limitations in mind and view the poverty concept as a practical and necessarily somewhat arbitrary measure; and second, it must be understood that many subtle procedural differences between the two surveys, even beyond the one cited here, could contribute to the difference in the poverty counts observed.

A further issue which needs to be aired is the consideration of in-kind government transfers in the counting of the poor. At present, benefits from food stamps, medicaid and public housing are not considered in any way in determining who among the respondents to the CPS are poor. These programs are designed to raise the real incomes of persons whose incomes are otherwise unacceptably low. So long as their contributions to the economic well-being of the poor are ignored, they lie open to ill-considered political criticism. These programs should be given credit for their contributions in fighting poverty, particularly the poverty of the intact family with a working head which is otherwise ineligible for public support.

To sum up: The PSID shows a smaller poverty population than the CPS. The difference between the two counts cannot be assigned to differences in demographic representation, but can be explained in large part by differences in income reporting in the two surveys. One procedural difference between the two surveys was cited as accounting for an uncertain share of the difference in measured numbers of poor; other, more subtle differences may or may not account for the rest. Despite the sensitivity of the process of counting the poor to procedural details, the results here do suggest that we may have made more progress than heretofore believed against poverty.

Footnotes

1. See note b, Table 1, and note a, Table 2.
2. Results of these tests are available on request from the author.
3. Frank Burns of the Census Bureau has researched this question and reports that the existing Census data do not provide a sufficient base for estimating the sensitivity of the measured poverty count to this procedure.

TABLE 1. Persons in Poverty as a Percentage of all Persons, 1960-1972.

	Bureau of the Census[a]	Panel Study of Income Dynamics[b]
1960	22.2	
1961	21.9	
1962	21.0	
1963	19.5	
1964	19.0	
1965	17.3	
1966	14.7	
1967	14.2	11.6
	(.22)	(.64)
1968	12.8	9.9
	(.22)	(.62)
1969	12.1	10.3
	(.22)	(.60)
1970	12.6	10.4
	(.22)	(.59)
1971	12.5	10.2
	(.22)	(.59)
1972	11.9	8.6
	(.22)	(.52)

[a] Published estimates, from U.S. Bureau of the Census, Current Population Reports, Series P-60, No.80, Table 11; No.86, Table L, 1,2; No. 83, p. 3; No. 95, Table 1. Estimates are not perfectly comparable over time; see No. 95, page 3 for details. Standard errors in parentheses from Current Population Reports, Series P-60, No. 98, Table A-9.

[b] Estimates derived by the author from the Panel Study of Income Dynamics, six year household sample. All responsibility for interpretation and use of these data is mine. Documentation is available as A Panel Study of Income Dynamics, Volumes 1 and 2, Institute for Social Research, The University of Michigan, Ann Arbor, Michigan. Because the poverty thresholds used in the PSID are different from those used by the Census Bureau, the offical census thresholds were applied to the PSID in this study. Standard errors in parentheses from James N. Morgan, et al., Five Thousand American Families - Patterns of Economic Progress, Volume 1; Ann Arbor Institute for Social Research, The University of Michigan, Table B-1.

TABLE 2. Weighted and Unweighted PSID Poverty Rates and 99% Confidence Interval, 1967,1971 and 1972.[a]

		PSID unweighted	PSID weighted[b]
1967	Upper Bound	13.21	
	Estimate	11.61	12.43
	Lower Bound	10.01	
1971	Upper Bound	11.81	
	Estimate	10.21	9.56
	Lower Bound	8.61	
1972	Upper Bound	10.16	
	Estimate	8.56	7.95
	Lower Bound	6.96	

a. Documentation of the PSID is cited in note 6, Table 1. Estimates from the CPS were derived by the author from the CPS Public Use Sample, March Supplement, for income years 1967, 1971 and 1972. All responsibility for interpretation and use of these data is mine. Documentation is available from DUALabs, Arlington, Virginia.

b. Weights were constructed in order to alter the joint distribution of the age, race, sex and education of household heads in PSID to be identical to that in the CPS. The smaller classificatory intervals of the variables shown in the note to Table 3 were used. The weights were constructed by first applying each surveys own weights, amd then determining the proportion of each sample which falls in each of the cross tabulated cells. The PSID observations were then weighted by the ratio of the CPS frequency to the PSID frequency.

TABLE 3. CPS and PSID Income Reporting as Percent of Control Aggregates, 1967, 1970, 1971 and 1972[a]

	1967		1971		1972	
	CPS	PSID	CPS	PSID	CPS	PSID
Total Income						
Percent Reported	83.7	87.3	87.7	94.7	88.8	95.7
Labor and Asset Income						
Percent Reported	84.9	88.5	89.4	96.3	90.4	97.1
Social Security[b]						
Percent Reported	93.0	c	87.4	80.7	93.3	94.3
Public Assistance[b]						
Percent Reported	75.1	c	71.0	84.6	71.2	77.4
Other						
Percent Reported	56.1	74.8	69.4	82.9	69.0	77.9

a. For data sources, see note b, Table 1, and notes a and e, Table 2. Preliminary income control aggregates were made available by Dan Radner of the Bureau of Economic Analysis, Department of Commerce. All responsibility for the interpretation and use of the aggregates is mine.

b. The PSID presents detail of transfer income received by the head of household and spouse, if any. Transfers of any kind received by other family members are presented as one figure, not on detail. Such transfers were allocated among the "Social Security," "Public Assistance," and "Other" categories in the same proportions as the transfers received by head and spouse in the same year.

c. The PSID survey in 1967 did not include detail of transfer payments. All transfers and transfer control aggregates are allocated to "Other" for the 1967 PSID only.

TABLE 4. Adjusted Poverty Counts, in Thousands, 1971 and 1972[a]

CPS sample with PSID reporting[b]

	1971	1972
CPS actual	25,808	24,556
CPS adjusted	23,451	22,969
PSID actual	20,965	17,672
Percent of CPS-PSID difference explained	48.7	23.1

a. For data sources, see note b, Table 1, and note a, Table 2.
b. See text for details of reporting adjustment.

TABLE 5. Incomes, by Source, of Marginally Poor Households in CPS and PSID, 1971, Percent[a]

Households with Aged Heads, Total Income .90-.99 of Poverty Threshold

	Income Type		
	Earned and Asset	Public Transfer	Other
CPS	14.1	84.4	1.5
PSID	11.6	71.6	16.8

Households with Female Heads and Children Present, Total Income .90-.99 of Poverty Threshold

	Income Type		
	Earned and Asset	Public Transfer	Other
CPS	36.2	53.7	10.1
PSID	28.6	60.1	11.4

All Other Households, Total Income .90-.99 of Poverty Threshold

	Income Type		
	Earned and Asset	Public Transfer	Other
CPS	80.3	18.8	0.9
PSID	80.4	17.0	2.5

a. For data sources, see note b, Table 1, and note a, Table 2.

TABLE 6. CPS Sample With PSID "Mix" of Income Types for Marginally Poor Households.[c]

1971

	CPS Actual	CPS Adjusted	PSID Actual	Percent Difference Explained
Total	25,808	21,931	20,965	80.1
Aged Head	5,262	3,362	4,010	151.8
Female Head With Children	8,541	6,564	6,242	86.0
Other	12,005	12,005	10,713	0.0

1972

	CPS Actual	CPS Adjusted	PSID Actual	Percent Difference Explained
Total	24,556	18,694	17,672	81.6
Aged Head	4,562	2,835	3,782	221.4
Female Head With Children	8,880	6,622	5,314	63.3
Other	11,115	9,237	8,577	74.0

a. See text for details of income mix adjustment.