Rationale and Construction

of Poverty Measures

in the

General Social Survey

Ethan Ligon

NORC

University of Chicago

GSS Methodological Report No. 57

September, 1988

> *There are only two families in the world,*
> *the Haves and the Havenots.*

> – Cervantes, *Don Quixote*, II

Cervantes may have been figuratively correct in his claim that the world contains only two families, insofar as we can always distinguish two sorts of people: the Haves and the Havenots, the bourgeoisie and the proletariat, the empowered and the unempowered, the rich and the poor. However, the precise line which separates classes of people may be difficult to pin down; the actual assignment of all families or individuals to one class or another will in practice always involve either arbitrary assignment or arbitrary distinction.

It may be a truism of society that "ye have always the poor with you,"[1] but "the poor" may in fact be more or less impoverished, and are certainly not homogeneous in their poverty. A lack of wealth is the defining characteristic of poverty, and people possess differing amounts of wealth. To arbitrarily claim that those with less than a certain amount of wealth are poor, and all others are non-poor, smacks of oversimplification.

Despite the difficulty, the arbitrariness, and the oversimplification required to define precisely who the poor are, there is considerable demand for just such a definition. This demand has three principle sources: from those who formulate or criticize policies relating to the phenomenon of poverty, from those who implement such policies, and from those who research the phenomenon of poverty.

Those who formulate policy have an interest in precise definition because the relief of poverty is a frequently stated and restated policy objective. The "War on Poverty," "The Great Society," and "two chickens in every pot," are slogans of only a few recent assaults on the phenomenon of poverty; the history of the twentieth century contains many more.

Once policy-makers are finished with their task of formulating some policy directing that such and such be done with, to, or for the poor, it

---

[1]Matthew 26:2

1

becomes necessary for those who administer welfare programs and the like to be able to distinguish -- on a finer level than that required for policy formulation -- who precisely it is that they should administer *to*. This is a rather demanding task, for while generalities of population composition, location, and size will quite probably suffice for the policy-maker, the administrator must decide whether or not a *particular* (rather than general or statistical) entity should have such and such done with, to, or for it. Furthermore, computational cost is very likely a constraint for the administrator; guide-lines must in general be simple, easy to administer, and above all be concrete.

Finally, researchers have two chief needs for a poverty dichotomization. First, the analysis and understanding of the outcome of a particular policy will probably require the researcher to use the poverty dichotomization assumed by the policy-maker. Second, a researcher might be interested in "The Poor" as some sort of sociological or class entity, independent of administrative definitions. The latter researcher must place some faith in the objective, rather than merely administrative or arbitrary, existence of such a definition as well as in its proper delineation.

Possible Poverty Measures

What might a reasonably objective, non-arbitrary, poverty definition look like? It depends on the understanding one has of the nature of poverty. For some purposes, an absolute definition (e.g. those with less than x wealth are poor) will suit best. For other purposes, we may be interested in relative deprivation, and so may define poverty as a lack of wealth relative to that of the general population.

If, indeed, we "have always the poor with us," then a suitable definition might be a relative one; the x percent of the population with the lowest wealth or income. However, for many uses this definition is not suitable; if everyone had the same amount of wealth, then our criterion would be meaningless. Likewise, if variation exists, but even the poorest member of society has an income sufficient to meet his needs, then can such a member truly be called impoverished?

An alternative relative definition without the drawbacks of the fixed percentile variety is this: the poor are those whose wealth or income is less than a certain percent of the median wealth or income of the population. This definition has its attractions: it allows the theoretical possibility of having *no* poor; the number of poor will fall as variation in wealth or income falls; and yet built into the definition is a *relative* concept of poverty; a mere general increase in wealth will not eradicate poverty without a necessary decrease in variation.

Such a relative definition of poverty has its attractions, but is imperfect insofar as the x percent we choose to denote poverty is entirely arbitrary. Should it be 10%, 50%, 120%? This relative definition fails to take into account the absolute nature of *need*, and this is something that most widely used contemporary definitions have taken into account.

What constitutes need? Needs will clearly vary across individuals, but we can say with reasonable certitude that one *needs* adequate shelter, clothing, and nutrition. The need for shelter and clothing will vary with the clime, but the need for nutritional adequacy does *not* vary greatly across time or space; therefore, it is often the case that poverty definitions will use nutritional adequacy as their benchmark. However, needs other than food exist, even if they are less quantifiable and more variable; thus the poverty line is generally drawn at a level which is some multiplier of the amount of money required to maintain some level of nutritional adequacy.

The Federal Government's Measurement of Poverty

This latter approach is precisely that taken in recent years by various arms and agencies of the U.S. government[2]. The nutritional adequacy of a diet is specified by the U.S. Department of Agriculture, which has also calculated the cost of various adequate diets. The cost and adequacy of these diets were based originally on a 1955 USDA food study (updated in 1965) which also yielded a rough multiplier for low-income families of 3; that is, low-income fami-

---

[2]For a more complete history of the formulation and evolution of the poverty index, see Orshansky 1970.

lies with a nutritionally adequate diet typically spent about 1/3 of their income on food.[3]

The first official promulgation of a "poverty-line" in the U.S. was undertaken by the President's Council of Economic Advisors. In their 1964 report which officially initiated Johnson's War on Poverty, the CEA designated as poor any family of two or more persons with an annual combined income of less than $3000, and any individual living alone with an annual income of less than $1500.[4]

The crudity of the CEA measure of poverty was quick to be criticized[5]; the fact that the $3000 figure was applied to families regardless of size was particularly troublesome. The outcome of such criticism was the formulation, by Mollie Orshansky of the Social Security Administration, of a more refined "poverty matrix", which took into account not only family size, but also the number of children, the sex and age of the head of the household, and the farm or non-farm nature of the household.

This more refined poverty measure was soon adopted by the Office of Economic Opportunity, "as a working tool for budget and planning purposes, and in large measure also as a guideline for eligibility for anti-poverty pro-grams."[6] By 1969, the Census Bureau had been assigned the task of publishing the Orshansky poverty measure as an annual statistical series. Since that time, there has only been one major change in the index: for 1981, the matrix eliminated distinctions based on farm or non-farm status, eliminated the sex of head of household distinction, and expanded the matrix to account for families of nine members or more (formerly seven or more).

The poverty index and figures generated by the Census Bureau are some-what limited in their analytical usefulness, however. The Census Bureau, for reasons of confidentiality, releases only aggregated poverty figures and some selected tables. Microdata released by the Census Bureau including poverty status of respondents tend to be heavily censored to protect confidentiality,

---

[3]Orshansky 1963, pp. 8-9

[4]CEA, 1965. The numbers given by the CEA are in turn roughly based on numbers generated by the Social Security Administration (Orshansky 1963).

[5]See Friedman 1965, pp. 3;29-42 for a concise summary of such criticisms.

[6]Orshansky 1977, p. 233

and to have only a small number of variables of limited usefulness for many purposes. While the federal government alone maintains a very large number of databases with poverty information[7], few if any provide the detail and breadth of non-income related variables that the General Social Survey (GSS) does.

Construction of the GSS Poverty Measure

In constructing variables relating to poverty status for the GSS, we have striven to balance two somewhat incompatible goals: compatibility with the census measure, and consistency across time. For a variety of reasons, it was not possible to completely meet either of these goals.

The desire to create a good, consistent time series led us to adopt a single form of the Orshansky matrix, rather than attempting to emulate the Census Bureau by using several different forms of matrices for different years. Because of its relatively simple construction and contemporary usage, we selected the matrix currently in use by the Census Bureau, adopted in 1982[8], and used it to extend our poverty calculations back through 1971 and forward through 1987. For this reason, census figures and GSS figures prior to 1981 are slightly less comparable than they are in later years. The differences, however, should be quite small; the Census Bureau calculated the poverty rate for 1980 using both matrices, and found that the use of the new matrix yielded an overall increase in the poverty rate of only 0.2%[9].

The Census Bureau has made use of some information on households in determining poverty status that the GSS does not collect data on, o for which GSS and census data are too dissimilar to permit precise comparison. An example of such a characteristic is the sex of the head of household, which was used by the Census Bureau as a dimension in its poverty matrix prior to 1982 The GSS did not collect this piece of information from respondents for several early years (1972-74, and half of the 1976 sample), and for no years

---

[7]Citro & the Bureau of the Census, 1976, is a 230 page document devoted to nothing other than providing a list of abstracts of such databases.

[8]See *supra*, p. 4 for a brief explanation of the evolution of the Orshansky matrix, or see Orshansky 1970 for a more complete summary.

[9]Bureau of the Census 1982, pp. 2-5.

is this information included in the final machine-readable file compiled by
the GSS staff. Similarly, the GSS does not collect information on current
farm/non-farm residence, which the Census Bureau also used prior to 1981. It
is largely because of these sorts of shortcomings that the GSS poverty defini-
tion is based on the poverty matrix adopted by the Census Bureau in 1982,
which was more spare in its number of dimensions (eliminating the farm/non-
farm and sex/marital status of the head of household distinctions).

The underlying variables used are the same as those of the Census Bureau
(household income, family size, number of children, and age of head of house-
hold); however, the instruments used to measure these underlying variables may
differ substantially from those used by the Census Bureau. For example, the
age of the head of household is not a component of the GSS cumulative file, so
we have used as a proxy the age of the respondent. The definition of
"household" used by the GSS differs in some regards from that used by the
Census Bureau, and the census definition itself has undergone a number of
changes in the last two decades. The most critical item collected from the
respondent is the household income for the past year: the GSS elicits and
records less precise figures for this than does the CPS, and uses a less
precise definition of household income.

The actual variables used to construct poverty status for the GSS are
these[10]: HOMPOP and UNRELAT (used to derive the number of related household
members); INCOME72, INCOME, INCOME77, INCOME82, and INCOME86 (categorized
household income -- each version, used in different years, differs somewhat in
its categories); BABIES, PRETEEN, and TEENS (summed to yield the number of
children in the household); and AGE (used to determine whether or not the head
of the household is older than 64).

The first step in determining the poverty status of a household is to
determine the critical income below which the household is considered poor.
This calculation for the GSS is done using the census poverty matrix of 1982
as a base, and adjusting the matrix each year for inflation (as measured by
the Consumer Price Index[11]). The variables which are taken into account in

---

[10]See Davis & Smith, 1988 for the precise wording of and other information relating to these mnemonics.

[11]See Appendix A for a table of the relevant price indices, and a list of the calculated poverty matrices.

calculating the poverty index for a given family are the number of family members (HOMPOP minus UNRELAT), the number of children (the sum of BABIES, PRETEEN, and TEENS), whether the respondent's age is greater than 64 years (AGE), the year, and the price index. For example, consider a family of four in 1987, consisting of two adults younger than 65, and two children. By consulting the poverty matrix for 1987 (see Appendix A), we can find precisely what the poverty threshold for such a family is; in this case, it is $11,519.

The second step in determining the poverty status of the family is to compare the relevant poverty index calculated previously to the family income. If family income is less than the calculated family-specific poverty index, then the family is considered poor; otherwise the family is non-poor. The difference between family income and the relevant poverty number can be termed the family's "income deficit[12]." Returning to our example above, if the annual family income is $20,000, then the family is non-poor, and has an income deficit of -$8481.

At this point, a complication arises. The various income variables on the GSS are all categorical variables (i.e. recorded income is only accurate to the nearest $1000, $2000, $2500, $5000, or $25,000, depending on year and income), and the top category in each year is open-ended (e.g. $60,000 or more). The categorical nature of the variable makes it impossible to assign poverty status with certainty if the relevant poverty number happens to lie within the same category as recorded income.

In creating a poverty-status variable, three different methods were considered to deal with the problem of indefinite assignment within a single income category. The first -- and most straight-forward -- method was simply to create a poverty-status variable with three, rather than two, non-missing categories: poor, borderline, and non-poor. This rather conservative approach had the advantage of not foisting questionable assignments onto unsuspecting researchers, and in fact emphasizing the drawbacks of using a binary measure as a proxy for a continuous variable. The principal disadvantage lay in the difficulties that would face researchers who wished to make a concrete poverty

---

[12]Note that if a family's income is greater than their poverty threshold, then by stipulation, they have a negative income deficit. It might make more sense to speak of an income surplus, but income deficit is used to remain consistent with census usage.

assignment; the middle category of "borderline" would be impossible to recode with any accuracy into the poor and non-poor categories.

These difficulties led to the formulation of a second variable, in which all borderline cases would be allocated either to the poor or the non-poor category. The only difficulty was how to accomplish the allocation. Three possibilities presented themselves: we could simplistically assign the income for all families within a single category to the midpoint of that category, we could assign income within a category not to the midpoint, but rather to the mean income within the category (where that mean would have to be estimated from some other data source, such as the CPS data), or we could assume some distribution for income, centered about a calculated mean, and assign income within each category according to this distribution. The latter two solutions seemed preferable on theoretical grounds, but the first had the advantage of simplicity. After calculating the mean income within each category for 1980, it was found that the mean of each category tended to be remarkably similar to the midpoint[13]. This finding justified use of the simple midpoint approach. The third approach we considered involved the construction not of a simple poverty/non-poverty variable, but rather of an categorical "income-deficit" measure. The advantages of this approach were several. The problem of what to do with indefinite cases was obviated; these cases could simply be assigned an income-deficit category of "-k to +k," where k would be equal to one half of the finest income category. Perhaps most importantly, such an approach

---

[13]Means for each category were calculated using the March CPS income data for 1980. The only large differences between mean and midpoint occur in the most extreme categories, where the assignment of poverty status is generally a foregone conclusion. The persistent bias which apparently places the midpoint above the mean may be the result of rounding -- note that if the respondent rounds income to the nearest thousand, this will bias the mean downward within a category for most categories. A comparison of means and midpoints follows:

Category	Midpoint	Mean	%Difference
< $999	NA	-$361.59	NA
1000-2999	$2000	2255.28	12.7%
3000-3999	3500	3499.07	-0.0%
4000-4999	4500	4447.81	-5.2%
5000-5999	5500	5455.81	-4.4%
6000-6999	6500	6447.94	-5.2%
7000-7999	7500	7457.86	-4.2%
8000-9999	9000	8936.33	-6.4%
10000-12499	11250	11188.26	-2.5%
12500-14999	13750	13701.46	-1.9%
15000-17499	16250	16167.12	-3.3%
17500-19999	18750	18687.84	-2.5%
20000-22499	21250	21134.15	-4.6%
22500-24999	23750	23686.04	-2.6%
25000-34999	30000	29207.30	-7.9%
35000-49999	42500	40878.04	-10.8%
> 49999	NA	61504.88	NA

also allows the differentiation of levels of poverty, rather than a simple identification of poverty as a binary state-variable.

Imputation of Poverty Status

A perennial difficulty in measuring income in surveys is nonresponse bias; those with high income are less likely to answer questions regarding their income[14]. Furthermore, questions about income typically have very high item nonresponse rates, aggravating the problem. The cumulative (1974-1988) non-response to the GSS family income item is 8.3% -- this compares to an average item non-response rate of about 1.7% across the same years. Thus, we have imputed poverty status for many cases, in order to produce better estimates of the poverty rate.

There are two stages to the imputation. First, the relevant variable for individual *respondent's* income is compared to the family's poverty index. Since the respondent's income is a component of family income, we can defin-itely assign any respondent's family to non-poverty status if the respondent's income is greater than the family's poverty index. This "safe" imputation allows us to make a poverty status assignment for either of the two methods outlined above which only indicate poverty status; about 14.3% of cases miss-ing information on family income have information on the respondent's income, and of these, the respondent's income is high enough to assign the family to the non-poor class for 24% of the cases.

Second, we have constructed a model with which to estimate income-defic-it, based on the characteristics of respondents who *did* respond to the family

---

[14]For families in the March 1969 CPS, the relationship between income and non-response to income items seems quite strong; the following figures are from Ono 1969, Table B-1.

Total Money Income	Allocated Share of Population	Non-Response Rate
None	0.3%	15.9%
$1 to 999	1.5%	23.4%
1000 to 1999	3.4%	15.7%
2000 to 2999	5.1%	16.0%
3000 to 3999	6.1%	17.1%
4000 to 4999	6.0%	17.8%
5000 to 5999	6.9%	17.2%
6000 to 6999	7.6%	17.1%
7000 to 9999	23.4%	17.0%
10000 to 14999	25.0%	19.3%
15000 to 24999	12.1%	24.1%
25000 and over	2.6%	35.2%

income item. This model allows us to properly impute income-deficit (and consequently poverty status), reducing non-response bias, assuming that the variables incorporated into the model[15] are well correlated with the factors which affect non-response. Using this latter imputation, we are able to assign poverty status to 80.5% of the cases missing one or more of the variables necessary to construct poverty status. The effect of these imputations is to reduce the number of cases for which poverty status cannot be assigned from 8.9% to 1.5% of all cases.

Poverty Variables in the GSS

The poverty variables actually constructed are three, corresponding to the three approaches discussed above. These variables and their categories are shown in Table I. Of these, POVLINE is the simplest, categorizing families simply as poor or non-poor. POVLINE1 has the feature of categorizing borderline cases separately; these are the cases which cannot be definitely assigned to either poor or non-poor status, because of the lack of precision in the various GSS income measures. Note also that POVLINE1 has separate categories for those cases whose poverty status is imputed; when using POVLINE1 these can be either recoded to include imputed cases, excluded for those who prefer not to use imputed data, or included in the analysis but whose imputation is noted. If the researcher is using one of the other two poverty variables, POVLINE or INCDEF, POVLINE1 can still be used as an imputation flag. INCDEF is the income-deficit variable, which describes the difference between a family's income and the relevant poverty number for that family. The categories used for INCDEF are somewhat crude, but this crudity reflects the crudity of comparison between the various family income variables on the GSS and the

Table I: Poverty Variables Constructed for the GSS

```
POVLINE     (1)Poor (inc. imputes)
            (2)Non-Poor     "
            (9)NA


POVLINE1    (1)Poor
            (2)Poor - Imputed
            (3)Non-Poor
            (4)Non-Poor - Imputed
            (5)Borderline
            (6)Borderline - Imputed
            (9)NA


INCDEF      (1) < -10,000
(includes   (2) -10,000 to -5000
all         (3) -4999 to -1000
imputed     (4) -999 to 999
cases)      (5) 1000 to 4999
            (6) 5000 to 10,000
            (8) > 10,000
            (9) NA
```

---

[15]See Appendix B for elucidation of the income-deficit model.

poverty index.

Technical Shortcomings of the GSS Poverty Measures

This paper will not undertake here a general critique of the Orshansky poverty matrix on which the GSS measure is based, nor will it dwell on any short-comings in the measurement or conception of poverty on the part of the Census Bureau. What this paper will do is point out some difficulties in *comparing* GSS poverty figures with similar figures generated by the Census Bureau. While such comparisons may be fruitful (some such comparisons are made in the next section), it would be wise to remember that the two measures are not necessarily strictly commensurable. Measurement differences and conceptual differences exist between the CPS and GSS measures for every variable used to assign poverty status; these differences are not so great as to make com-parisons meaningless, but they certainly do make them less than perfect.

There are five dimensions in the poverty index: family income, family size, number of children, age of householder, and year. Of these five dimensions on the GSS, only year is similar enough to the census use to assure trouble-free comparisons, and even year has a small complication that the re-searcher should always remember. Both the Census Bureau and the GSS collect income information in about March of each year, and both ask about income retrospectively; that is, they ask the respondent what his or her income was *last* year. Thus, when examining GSS poverty data, the researcher should remember that, for example, the 1988 GSS data reflects the respondent's 1987 poverty status and income.

While conceptually the comparison of income to a family-specific poverty number is an extremely simple, straight-forward step, in practice it becomes much more problematic. To begin with, the GSS notion of income is rather ill-defined. Respondents are presented with a list of income categories, identified by letter, and asked[16]:

> In which of these groups did your total *family* income, from *all* sources, fall last year before taxes, that is? Just tell me the letter.

---

[16]The wording of the income question on the 1972 GSS differed slightly. Respondents were asked:

"In which of these groups did your total family income, from all sources, fall last year—1971—before taxes, that is?"

Interviewers were further instructed that:

> Total income includes interest or dividends, rent, Social Security,
> other pensions, alimony or child support, unemployment compensation,
> public aid (welfare), armed forces or veteran's allotment.[17]

This formulation is rather similar to the question asked by the Census Bureau
on the basic Current Population Survey questionnaire:

> Which category on this card represents the total combined income of all
> members of this FAMILY during the past 12 months? This includes money
> from jobs, net income from business, farm or rent, pensions, dividends,
> interest social security payments and any other money income received by
> members of the FAMILY who are 14 years of age or older?[18]

However, the annual poverty figures issued by the Census Bureau are based on
responses to the March CPS, which includes an income supplement to the basic
questionnaire. This supplement breaks down family income into eleven sources
of income[19], which presumably leads to more accurate reporting of income,
and -- by recalling to the respondent almost all possible sources of direct
income -- probably leads to a less biased estimate of income than the item on
the basic questionnaire.

Family size in the GSS is defined as HOMPOP minus UNRELAT, that is, the
number of members of the respondent's household minus the number of those
members who are unrelated to the respondent. The census definition is more
complicated, much of the complexity centering around the definition of a
housing unit (or dwelling unit prior to 1980). However, for the large major-
ity of cases, GSS and census definitions probably correspond well on family
size. The number of cases where differences in definition produce different
counts of family members *and* in which the different count results in a
difference in assignment of poverty status is probably minuscule. Differences
in the count of the number of children within a family should be almost solely
due to differences in measuring family size.

Age of householder is the single most difficult census variable to
replicate for the GSS, however, it is the least critical. The GSS doesn't

---

[17]Ibid.

[18]U.S. Dept. of Commerce, Bureau of the Census, 1978A, p. 231

[19]Bureau of the Census 1981, pp. 237-247

specify a head of household or householder as the Census Bureau does; the best we can do is to produce the age of the respondent. For this reason, the GSS calculation of poverty status takes age into account only if the respondent lives alone, guaranteeing that he or she is the head of household. The Census Bureau takes age into account if the household has one or two members, but differences in the poverty matrix resulting from age of householder in households with two members are very small. The GSS poverty matrices use a weighted average of the poverty threshold for a family of two.

Comparison of Census and GSS Poverty Figures

If one undertakes to compare aggregate estimates of poverty as measured by the GSS and by the Census Bureau, the results are somewhat reassuring. When one compares the number of persons in the U.S. living below the poverty line as estimated by the GSS and by the Census Bureau, only small differences occur (See Table I). Moreover, the GSS seems to follow all the same trends that the Census Bureau does: when the census estimate increases, so does that of the GSS; when
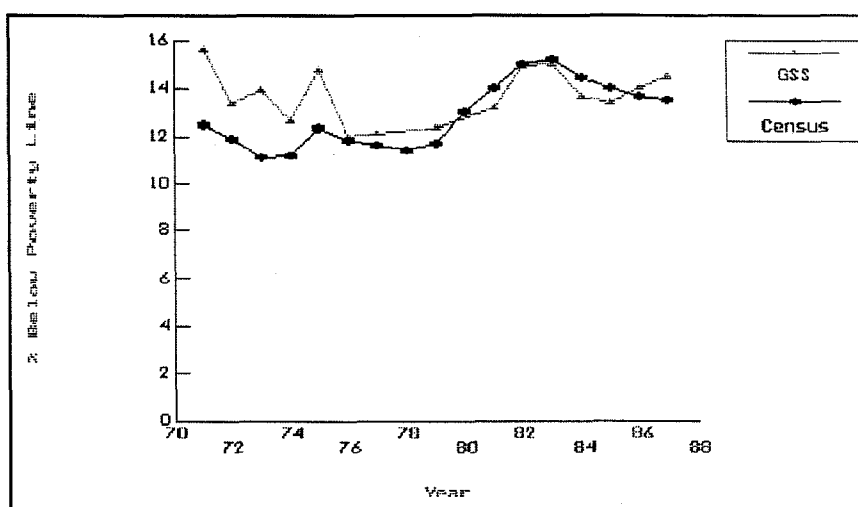
Table II: Comparison of GSS and Census Poverty Figures

	GSS						
	POVLINE			POVLINE1			
	Imputed	Not Imputed	Imputed		Not Imputed		Census
YEAR	POOR	POOR	POOR	BORDERLINE	POOR	BORDERLINE	POOR
1971	15.6	13.6	12.2	8.1	10.5	8.3	12.5
1972	13.3	12.0	10.8	5.4	9.7	5.7	11.9
1973	13.9	13.0	10.2	4.7	9.3	4.8	11.1
1974	12.6	11.7	8.7	6.4	7.8	6.6	11.2
1975	14.8	14.0	11.8	6.4	10.7	6.8	12.3
1976	12.0	11.1	8.5	4.2	7.6	4.2	11.8
1977	12.1	11.6	10.3	5.4	9.9	5.7	11.6
1978							11.4
1979	12.3	12.3	10.1	3.9	10.2	3.9	11.7
1980							13
1981	13.1	13.7	12.2	4.2	12.7	4.4	14
1982	14.9	15.7	12.6	3.3	13.2	3.5	15
1983	15.0	15.6	13.7	4.2	14.2	4.4	15.2
1984	13.6	14.3	11.9	3.5	12.6	3.8	14.4
1985	13.3	14.3	12.9	3.5	13.8	3.5	14.0
1986	14.0	14.8	13.6	1.6	14.4	1.8	13.6
1987	14.5	15.5	12.0	3.6	12.9	3.9	13.5

the census estimate
falls, so does the GSS's.

Interestingly, the GSS figures are consistently higher than the census
figures in early years, though the two series tend to converge in later years.
However, it may be that it is the later years that are in some sense 'ab-
normal'; there are *a priori* reasons, discussed briefly above, for believing
that census measures of income will yield consistently higher estimates of
family and respondent income than will the GSS measure. Furthermore, examina-
tion of median family income as estimated by the GSS and the Census Bureau
seems to support this hypothesis (See Table III). If the GSS does in fact
underestimate family income (relative to the CPS), then we would expect GSS
poverty figures to be consistently higher than census figures. GSS figures
are higher in every year but one, but the difference between census and GSS
figures shifts substantially downward

after 1976.   The reasons for this
shift are unclear. It seems suspicious
that the period of the shift (1976)
corresponds to the first year in which a
full probability sampling method was used
exclusively; such sampling methods have
some characteristics that might explain
the shift in the GSS poverty estimates.
For instance, it is known that the GSS
full probability sample tends to under-
sample central cities; if those people
within a central city are more likely to
be poor than the general population, then
that might serve as an explanation.[20]

The flaw in this argument is that
for two years, the GSS was conducted

Table III: Median Family Income

| | Median Family Income | |
Year	Census	GSS
1971	$10,299	$8000-9999
1972	11,110	10,000-14,999
1973	12,045	10,000-14,999
1974	12,910	10,000-14,999
1975	13,730	10,000-14,999
1976	14,970	10,000-12,499
1977	16,020	12,500-14,999
1978	17,650	NA
1979	19,595	15,000-17,499
1980	21,039	NA
1981	22,391	15,000-17,499
1982	23,442	17,500-19,999
1983	24,679	20,000-22,499
1984	26,649	20,000-22,499
1985	27,734	20,000-22,499
1986	29,470	22,500-24,999
1987	30,853	25,000-29,999

using a split sample. Roughly half of the total sample was sampled using a
full probability design, and the other half a block-quota design. If the

---

[20]See Stephenson, 1978 for an examination of this and other differences between the use of full probability
sampling and quota sampling on the GSS.

shift in GSS poverty figures were due to a design effect, then the difference should show up in these years. In fact, the 1975 sample does support this interpretation -- the full probability half of the sample has a poverty rate of 10.3%, while the block-quota half yields a poverty rate of 13.3%. However, a dramatic switch occurs in 1976, with a full-probability poverty rate of 15.1%, and a block-quota rate of 10.8%. Thus, the source of the downward shift in GSS measures of the poverty rate remains unclear.

Conclusion

Any researcher wishing to know more about the group the *government* defines as poor will find that the GSS data will provide a largely coincident group for which a very rich set of data has been collected. As has been emphasized throughout this paper, one should beware of oversimplistic analyses based on the GSS poverty measures; often direct use of income variables is more appropriate. However, the poverty measures presented here, particularly the income-deficit measure, do have some advantages over simple income measures. If one is interested in quality of life sorts of questions, then the poverty measures -- income-deficit particularly -- may serve as a good proxy for disposable income. Almost any analysis of the causes and concomitants of poverty could profitably use the GSS poverty measures, particularly if the researcher has no quarrel with the government's definition of poverty.

Appendix A: Poverty Matrices by Year

1987

Fam Size	Number of Children								
	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	5909								
1 (> 65)	5447								
2	7606	7829							
3	8884	9142	9151						
4	11714	11906	11519	11558					
5	14127	14333	13894	13554	13347				
6	16249	16313	15978	15655	15176	14892			
7	18697	18813	18411	18130	17608	16999	16330		
8	20911	21096	20715	20372	19911	19312	18688	18529	
9	25155	25276	24940	24658	24195	23557	22980	22837	21958

1986

Fam Size	Number of Children								
	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	5701								
1 (> 65)	5255								
2	7338	7553							
3	8571	8820	8829						
4	11302	11487	11113	11151					
5	13630	13828	13405	13077	12877				
6	15677	15739	15415	15104	14642	14368			
7	18039	18151	17763	17492	16988	16400	15755		
8	20175	20353	19986	19655	19210	18632	18030	17877	
9	24269	24386	24062	23790	23343	22728	22171	22033	21185

1985

Fam Size	Number of Children								
	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	5593								
1 (> 65)	5156								
2	7199	7410							
3	8409	8653	8662						
4	11089	11270	10903	10940					
5	13373	13567	13152	12830	12634				
6	15381	15442	15124	14819	14366	14097			
7	17698	17808	17428	17162	16667	16090	15458		
8	19794	19969	19609	19284	18847	18280	17690	17539	
9	23811	23926	23608	23341	22902	22299	21752	21617	20785

1984

Fam Size	Number of Children								
	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	5401								
1 (> 65)	4978								
2	6951	7155							
3	8119	8355	8364						
4	10707	10882	10528	10564					
5	12912	13100	12699	12388	12199				
6	14851	14910	14603	14308	13871	13611			
7	17089	17195	16827	16571	16093	15536	14925		
8	19112	19281	18933	18620	18198	17650	17080	16935	
9	22991	23101	22794	22537	22113	21531	21003	20872	20069

1983

Fam Size	Number of Children								
	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	5180								
1 (> 65)	4775								
2	6668	6863							
3	7788	8014	8022						
4	10270	10438	10098	10132					
5	12385	12565	12180	11882	11701				
6	14245	14301	14007	13724	13304	13055			
7	16391	16493	16140	15894	15436	14902	14316		
8	18332	18494	18160	17859	17455	16930	16383	16244	
9	22052	22158	21864	21617	21211	20652	20146	20020	19250

1982

Fam Size	Number of Children								
	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	5019								
1 (> 65)	4626								
2	6460	6649							
3	7545	7765	7772						
4	9949	10112	9783	9817					
5	11999	12173	11801	11512	11336				
6	13801	13855	13570	13296	12890	12649			
7	15880	15979	15637	15399	14955	14437	13870		
8	17761	17917	17594	17303	16911	16402	15872	15738	
9	21365	21468	21182	20943	20550	20008	19518	19396	18650

16

1981

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	4729								
1 (> 65)	4359								
2	6087	6265							
3	7109	7316	7323						
4	9375	9528	9218	9249					
5	11306	11470	11119	10847	10681				
6	13004	13055	12786	12528	12145	11918			
7	14963	15056	14734	14509	14091	13603	13068		
8	16735	16882	16578	16303	15934	15455	14955	14829	
9	20131	20228	19959	19733	19362	18852	18390	18276	17572

1980

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	4284								
1 (> 65)	3949								
2	5515	5676							
3	6441	6628	6635						
4	8494	8633	8352	8380					
5	10243	10392	10074	9828	9677				
6	11782	11828	11585	11351	11004	10798			
7	13557	13641	13349	13146	12767	12325	11840		
8	15162	15296	15020	14771	14437	14002	13550	13435	
9	18239	18327	18083	17879	17543	17081	16662	16558	15921

1979

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	3774								
1 (> 65)	3479								
2	4858	5000							
3	5674	5839	5845						
4	7482	7604	7357	7382					
5	9023	9154	8874	8657	8525				
6	10378	10419	10205	9999	9693	9512			
7	11942	12016	11759	11580	11246	10857	10430		
8	13356	13474	13231	13012	12717	12334	11936	11835	
9	16066	16143	15929	15749	15453	15046	14677	14586	14024

1978

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	3392								
1 (> 65)	3127								
2	4366	4494							
3	5100	5248	5253						
4	6725	6835	6612	6635					
5	8110	8228	7976	7781	7662				
6	9328	9365	9172	8987	8712	8549			
7	10733	10800	10569	10408	10108	9758	9374		
8	12004	12110	11892	11695	11430	11086	10728	10637	
9	14440	14510	14317	14155	13889	13523	13192	13110	12605

1977

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	3151								
1 (> 65)	2904								
2	4056	4174							
3	4737	4875	4880						
4	6246	6349	6142	6163					
5	7533	7642	7409	7227	7117				
6	8664	8699	8520	8348	8092	7941			
7	9970	10032	9817	9667	9389	9064	8707		
8	11150	11249	11046	10863	10617	10298	9965	9880	
9	13413	13478	13299	13148	12901	12561	12253	12177	11709

1976

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	2960								
1 (> 65)	2728								
2	3810	3921							
3	4450	4579	4584						
4	5868	5964	5770	5789					
5	7076	7179	6960	6789	6686				
6	8139	8171	8003	7842	7602	7460			
7	9366	9424	9222	9082	8820	8515	8180		
8	10475	10567	10376	10205	9974	9673	9361	9281	
9	12600	12661	12493	12351	12119	11800	11511	11439	10999

Each block header reads "Number of Children" spanning the numeric columns.

1975 — Number of Children

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	2798								
1 (> 65)	2579								
2	3602	3708							
3	4207	4329	4334						
4	5548	5639	5455	5474					
5	6690	6788	6580	6419	6321				
6	7695	7726	7567	7414	7187	7053			
7	8855	8910	8719	8586	8339	8050	7734		
8	9903	9991	9810	9648	9430	9146	8850	8775	
9	11913	11970	11811	11678	11458	11156	10883	10815	10399

1974 — Number of Children

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	2564								
1 (> 65)	2363								
2	3300	3397							
3	3855	3967	3971						
4	5083	5166	4998	5015					
5	6130	6219	6029	5881	5792				
6	7051	7079	6933	6793	6585	6462			
7	8113	8164	7989	7867	7640	7376	7086		
8	9074	9154	8989	8840	8640	8380	8109	8040	
9	10915	10968	10822	10700	10499	10222	9972	9909	9528

1973 — Number of Children

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	2311								
1 (> 65)	2130								
2	2974	3061							
3	3474	3575	3578						
4	4581	4656	4504	4519					
5	5524	5604	5433	5300	5219				
6	6354	6379	6248	6122	5934	5823			
7	7311	7357	7199	7089	6885	6647	6385		
8	8177	8249	8100	7966	7786	7552	7308	7246	
9	9836	9884	9752	9642	9461	9212	8986	8930	8586

1972 — Number of Children

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	2175								
1 (> 65)	2005								
2	2800	2882							
3	3270	3365	3369						
4	4312	4383	4240	4255					
5	5200	5276	5115	4989	4913				
6	5982	6005	5882	5763	5587	5482			
7	6883	6925	6777	6674	6482	6257	6011		
8	7698	7766	7626	7499	7330	7109	6879	6821	
9	9260	9304	9181	9077	8906	8672	8459	8407	8083

1971 — Number of Children

Fam Size	0	1	2	3	4	5	6	7	8 or more
1 (< 65)	2107								
1 (> 65)	1942								
2	2712	2791							
3	3167	3259	3263						
4	4176	4245	4107	4121					
5	5036	5110	5954	4832	4758				
6	5794	5816	5967	5581	5411	5309			
7	6666	6707	6564	6464	6278	6060	5822		
8	7456	7522	7386	7263	7099	6885	6662	6606	
9	8969	9011	8892	8791	8626	8399	8193	8142	7829

Appendix B: The Income-Deficit Model

In order to decrease the number of cases for which the poverty measures discussed above could not be calculated due to item (particularly income) nonresponse, an effort was made to impute poverty status. The apparatus used to perform this was a simple OLS regression model, using 17 independent variables.

We regressed an uncollapsed calculation of INCDEF on this range of 17 variables for all the cases (appropriately weighted) that had income information. Following is a list of the variables used, accompanied by their definitions.

BLACK   A dummy variable, recoded from RACE (the GSS variable). BLACK equals 1 if RACE equals 2, and is 0 otherwise.

EARNRS   The GSS variable; the number of earners in the household.

EDUC   The GSS variable; years of education. If missing, EDUC is assigned a code of 0.

FULLTIME A dummy variable, recoded from WRKSTAT (the GSS variable). FULLTIME is equal to 1 if WRKSTAT is equal to 1, and is 0 otherwise.

KIDS   The number of children under age 18 in the household; the sum of BABIES, PRETEEN, and TEENS (the GSS variables).

MARRIED   A dummy variable; recoded from MARITAL (the GSS variable). MARRIED is equal to one if MARITAL is equal to one, and is 0 otherwise.

MIDCLS   A dummy variable, recoded from CLASS and CLASSY (the GSS variables). MIDCLASS is one if CLASS or CLASSY is equal to three, and is zero otherwise.

PRESTIGE The GSS variable; occupational prestige. If PRESTIGE is missing, it is set to zero.

RACOTH   A dummy variable, recoded from RACE (the GSS variable). If RACE is equal to 3, then RACOTH is equal to 1, and is 0 otherwise.

RETIRED   A dummy variable, recoded from AGE (the GSS variable). If AGE is greater than or equal to 65, then RETIRED is equal to 1, and is 0 otherwise.

SPEDUC   The GSS variable; spouse's years of education. It is treated just as EDUC above.

SPPRES   The GSS variable; spouse's occupational prestige. It is treated just as PRESTIGE above.

19

UNDER25   A dummy variable, recoded from AGE (the GSS variable). If AGE is less than 25, then UNDER25 is equal to 1, and is 0 otherwise.

UPCLS   A dummy variable, recoded from CLASS and CLASSY (the GSS variables). UPCLASS is one if CLASS is equal to 4, or CLASSY is equal to 4 or 5, and is zero otherwise.

WORKCLS   A dummy variable, recoded from CLASS and CLASSY (the GSS variables). MIDCLASS is one if CLASS or CLASSY is equal to three, and is zero otherwise.

YEAR   The unrecoded GSS variable; year of survey.

Following are some tables and information relating to the regression:

					Analysis of Variance			
.tiple R	.57463					DF	Sum of Squares	Mean Square
;quare	.33020	R Square Change	.33020					
justed R Square	.32970	F Change	653.21303		Regression	16	7521444898224.75700	470090306139.047
ndard Error	26826.45134	Signif F Change	.0000		Residual	21200	15256760024942.1000	719658491.74255

F = 653.21303   Signif F = .0000

riable	B	SE B	95% Confdnce Intrvl B		Beta	SE Beta	Correl	Part Cor	Partial	Tolerance	T
RNRS	7312.268526	210.477901	6899.715975	7724.821076	.230743	.006642	.309171	.195276	.232089	.716209	34.741
DCLS	8041.988814	710.984534	6648.405534	9435.572094	.121519	.010743	.207760	.063578	.077451	.273731	11.311
COTH	-4774.707918	1472.025053	-7659.987980	-1889.427855	-.018351	.005658	.000053	-.018232	-.022272	.987023	-3.244
AR	1341.402826	36.656440	1269.553441	1413.252212	.212935	.005819	.220205	.205690	.243748	.933106	36.594
DER25	-6418.335584	612.346515	-7618.580912	-5218.090256	-.064095	.006115	-.081428	-.058915	-.071802	.844903	-10.482
CLS	22952.503659	1231.926554	20537.834752	25367.172566	.122413	.006570	.123802	.104725	.126926	.731889	18.631
ACK	-4258.342065	559.836401	-5355.663611	-3161.020520	-.044482	.005848	-.124183	-.042755	-.052170	.923855	-7.606
LLTIME	5809.212349	428.528146	4969.264881	6649.159817	.088589	.006535	.223038	.076198	.092703	.739811	13.556
DS	-1826.036482	156.683459	-2133.147872	-1518.925092	-.072149	.006191	-.010565	-.065507	-.079787	.824358	-11.654
PRES	212.021586	16.118716	180.427687	243.615485	.141877	.010786	.306122	.073936	.089974	.271570	13.154
UC	838.134219	75.812025	689.536937	986.731502	.083080	.007515	.337675	.062141	.075711	.559457	11.055
TIRED	-1846.721407	617.137431	-3056.357290	-637.085524	-.020744	.006932	-.160675	-.016820	-.020548	.657437	-2.992
RESTIGE	144.285681	13.604582	117.619675	170.951687	.073755	.006954	.296775	.059613	.072648	.653275	10.606
RRIED	-12858.08911	1058.798501	-14933.41398	-10782.76423	-.190477	.015685	.193785	-.068260	-.083117	.128425	-12.144
ORKCLS	-2467.678989	696.846342	-3833.550346	-1101.807632	-.037484	.010585	-.198999	-.019905	-.024314	.281985	-3.541
PEDUC	1250.574695	88.045519	1077.998842	1423.150549	.243655	.017154	.299830	.079837	.097091	.107364	14.204
Constant)	-120298.6710	3131.804221	-126437.2433	-114160.0986							-38.412

Appendix C: Constructing a "Low-Income" Variable

In addition to the poverty definition formulated by the Social Security Administration, many arms of government also use the alternative, or supplemental, classification of "low-income" families or individuals. The definition of a low-income family is a family with an annual income of less than 125% of the poverty level.

One can easily construct a binary low-income variable from INCDEF, the GSS income-deficit measure, and the relevant family income variable. In order to do so, one could follow this procedure:

1) Recode INCDEF and the family income variables into dollars (e.g. recode a 3 for INCDEF to -$3000, the midpoint of the category).

2) Subtract the recoded INCDEF from the recoded income variable.

3) Multiply this result by 1.25 and subtract the result from the recoded income variable; this is the resulting income-deficit for the low-income measure.

4) If this low-income income-deficit is negative, the family is low-income.

Because most cases for whom poverty status is imputed are missing income information, these cases cannot be assigned low-income status without either imputing income, or deriving their already (implicitly) imputed income by subtracting the relevant poverty number (See Appendix A) from INCDEF.

In order to impute income, the researcher could easily adapt the income-deficit model (outlined in Appendix A) to his or her needs, or alternatively develop a different model. Imputing income by using the income-deficit model, however, would yield the same results as deriving income from the calculated poverty numbers and INCDEF, save for some increase in measurement error due to INCDEF's crudeness.

References

Citro, Connie and The Center for Census Use Studies, Bureau of the Census, *The Measure of Poverty*, Technical Paper IX, *Inventory of Federal Data Bases Related to the Measurement of Poverty*, Washington D.C.: U.S. Government Printing Office, 1976.

Davis, James and Tom W. Smith, *General Social Surveys, 1972-1988: Cumulative Codebook*, Chicago: NORC, 1988.

Friedman, Rose D., *Poverty: Definition and Perspective*, Washington D.C.: American Enterprise Institute, 1965.

King, Jill, *The Measure of Poverty*, Technical Paper V, *The Consumer Price Index*, Washington D.C.: U.S. Government Printing Office, 1976.

Ono, Mitsuo and Herman P. Miller, "Income Nonresponses in the Current Population Survey," in *American Statistical Association: Proceedings of the Social Statistics Section, 1969*, Washington D.C.: American Statistical Association, 1969.

Orshansky, Mollie, "Children of the Poor," in *Social Security Bulletin*, Washington D.C.: U.S. Department of Health, Education, and Welfare, July 1963.

_____, "Counting the Poor: Another Look at the Poverty Profile," in *Social Security Bulletin*, Washington D.C.: U.S. Department of Health, Education, and Welfare, January 1965.

_____, "Who's Who Among the Poor: A Demographic View of Poverty," in *Social Security Bulletin*, Washington D.C.: U.S. Department of Health, Education, and Welfare, July 1965.

_____, "Memorandum for Dr. Daniel P. Moynihan," reprinted in *The Measure of Poverty*, Technical Paper I, *Documentation of Background Information and Rationale for Current Poverty Matrix*, Washington D.C.: U.S. Government Printing Office, 1977.

Spiers, Emmett F. and Joseph J. Knott, "Computer Method to Process Missing Income and Work Experience Information in the Current Population Survey," in *American Statistical Association: Proceedings of the Social Statistics Section, 1969*, Washington D.C.: American Statistical Association, 1969.

Stephenson, C. Bruce, "A Comparison of Full-Probability and Probability-With-Quotas: Sampling Techniques in the General Social Survey," GSS Technical Report No. 5, published as "Probability Sampling With Quotas: An Experiment," in *Public Opinion Quarterly*, Winter 1979.

U.S. Department of Commerce, Bureau of the Census, *The Extent of Poverty in the United States: 1959 to 1966*, No. 54 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1969.

_____, *Characteristics of the Population Below the Poverty Level: 1974*, No. 102 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1976.

_____, *Characteristics of the Population Below the Poverty Level: 1975*, No. 106 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1977.

_____, *Characteristics of the Population Below the Poverty Level: 1976*, No. 115 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1978.

_____, *The Current Population Survey: Design and Methodology*, Technical Paper 40, Washington D.C.: U.S. Government Printing Office, 1978.

_____, *Characteristics of the Population Below the Poverty Level: 1977*, No. 119 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1979.

_____, *Characteristics of the Population Below the Poverty Level: 1979*, No. 130 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1981.

_____, *Characteristics of the Population Below the Poverty Level: 1980*, No. 133 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1982.

_____, *Characteristics of the Population Below the Poverty Level: 1981*, No. 138 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1983.

_____, *Estimates of Poverty Including the Value of Noncash Benefits: 1983*, Technical Paper 52, Washington D.C.: U.S. Government Printing Office, 1984.

_____, *Characteristics of the Population Below the Poverty Level: 1982*, No. 144 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1984.

_____, *Characteristics of the Population Below the Poverty Level: 1983*, No. 147 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1985.

_____, *Characteristics of the Population Below the Poverty Level: 1984*, No. 152 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1986.

_____, *Poverty in the United States: 1985*, No. 158 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1988.

_____, *Poverty in the United States: 1986*, No. 160 in series P-60, *Consumer Income*, Washington D.C.: U.S. Government Printing Office, 1988.

U.S. Department of Health, Education, and Welfare, *The Measure of Poverty: A Report to Congress as Mandated by the Education Amendments of 1974*, Washington D.C.: U.S. Government Printing Office, 1976.