## THE IMPACT ON PERSONAL AND FAMILY INCOME OF ADJUSTING THE CURRENT POPULATION SURVEY FOR UNDERCOVERAGE

Robert Yuskavage and David Hirschberg, Bureau of Economic Analysis Frederick J. Scheuren, Social Security Administration

This paper presents the results of adjusting the Current Population Survey (CPS) for undercoverage, with attention focused on the impact of alternative adjustment procedures on the distribution of personal and family income. In addition, the impact on selected population characteristics and labor force estimates is reviewed.

The data base to which the coverage adjustments were made is the March 1973 CPS. This particular survey was selected because a special matching study conducted jointly by the Bureau of the Census and the Social Security Administration brought together information from the Current Population Survey, the Internal Revenue Service's personal income tax returns and Social Security's wage and benefit systems. 1/ As part of the reconciliation of the differences among these various sources, it was necessary to "correct" for the understatement of the population in the CPS.

Organizationally, the material has been divided into five sections. We begin in section 1 with some background on the nature of the March 1973 CPS' undercoverage and the alternative methods employed to deal with it. The next three sections examine the differential impact of the adjustments on income (sections 2 and 3) and other selected characteristics (section 4). Section 5 provides a few concluding remarks.

## 1. CPS UNDERCOVERAGE ERRORS AND ALTERNATIVE ADJUSTMENTS CONSIDERED

The papers by Bateman [3] and Korns [4] have already provided a detailed discussion of the nature and magnitude of Current Population Survey and Annual Housing Survey coverage errors. Some further points still need to be made, however, especially with regard to the March 1973 CPS. After sufficient background has been set, we will then describe the alternative coverage adjustments considered.

1.1 <u>Types of CPS undercoverage</u>.--As we have just seen [3, 4], undercoverage errors in a survey or census may be classified into omissions of two types. People can be missed if they live in households that are missed or they can be in an enumerated household but for one reason or another not be counted as members.

In the paper by Korns, estimates from the 1975 CPS were presented to show that for adults, more than two-thirds of the undercoverage was the result of persons missed in enumerated households. This pattern, which seems to have been typical since at least 1975, was not present for the March 1973 CPS due to a number of special circumstances. Instead, in 1973, we estimate that each of the two kinds of error accounted for about half of the undercoverage 2/--roughly the same ratio that was observed in the 1970 Census [7].

The undercoverage of households in the March 1973 CPS arises mainly from the following sources [8-9]:

- (a) Deficiencies inherited from the 1970 Census insofar as the CPS relied on Census addresses as one of the sampling frames (68 percent of the total CPS universe consists of address-list enumeration districts (ED's) drawn from the 1970 Census).
- (b) Incomplete listings in area segments; the failure to include established mobile homes in address ED's or to systematically include new mobile homes.
- (c) The omission of addresses converted from nonresidential to residential or homes moved to a site which was not a residential address in the 1970 Census; failures to include housing units completed after the Census for which permits were issued before January 1, 1970.
- (d) Failures to include addresses imputed in the Census or ones inadequately described in the address registers (The "Cen-Sup" or E6 portion of the CPS now corrects this inadequacy [3]; however, that sample was <u>not</u> put in place until shortly after March 1973.)

Altogether, the effect of these omissions was to understate the total number of occupied housing units by approximately 3.3 percent.3/ With the addition of the Cen-Sup sample, CPS household undercoverage has been reduced to a rate now less than 3 percent [9].

The sources of within household undercoverage in March 1973 can only be speculated about. Siegel [5] suggests causes of within household misses in the 1970 Census which may be applicable:

- (a) Deliberate concealment, carelessness, confusion, or apathy on the part of respondents.
- (b) A failure to adequately allow for persons who do not fit into any household according to the conventional rules of residence.

There are some differences between the CPS and the Census that could give rise to other possible sources of error which, when taken together, probably lead to additional within household undercoverage. These are:

- (a) Differences in enumeration (i.e., the replacement of self-enumeration with enumeration by highly trained, generally experienced CPS interviewers--possibly leading to some improvement in coverage).
- (b) Differences in counting rules. (In particular, college students living in dormitories are counted at their college in the Census. They are supposed to be counted at their parent's home [permanent address] in the CPS. We suspect this change leads to some loss of coverage.)
- (c) Longitudinal nature of CPS. (Coverage rates typically decline over the life of a CPS rotation panel. This may be partly due to new household members not being completely accounted for on household rosters in succeeding interviews.)

March 1973 was characterized by exceptionally good overall coverage, better than that for any other March survey in the period 1970-1977. This is true even though the survey's household coverage, as we saw above, was not as good as that for the March CPS surveys since then. One of the main reasons for this apparent paradox may be that in the 1973 survey, because of updates in the sample design, the mix of rotation panels was not the same as that which typically occurs in every other month. The average number of previous interviews each household had received was considerably less than normal in March 1973.

1.2 <u>Alternative coverage adjustments</u>--A coverage adjustment has been part of the CPS since its inception in the 1940's [10, p. 10]. Basically, the adjustment has only "corrected" the differential undercoverage of the CPS relative to the previous census. This has usually been done by ratioing the sample estimates to independently derived age-race-sex population totals obtained by carrying forward decennial census population estimates to account for subsequent aging of the population, births, deaths and net (legal) migration.

Some consideration has, of course, been given before today's session [e.g., 12-13] to what might happen if the survey were adjusted to "true" and not just census-level population totals. Our approach differs from these previous efforts in scope but not in purpose. We have, as a result of the 1973 Exact Match Study, much more information with which to attempt adjustments. Our principal goal is still, however, to examine the sensitivity of the survey estimates to the problem of coverage, not to make a definitive statement on what the CPS coverage adjustment should be. Three alternative survey estimates are compared in the remainder of this paper. A brief definition of each of these is given below: 4/

- (a) <u>Initial</u>.--This is the survey estimate before any adjustment for coverage. It is also known as the First Stage weight because it consists of all the estimation steps in the CPS up to and including the application of the first stage factors in effect for March 1973.
- (b) <u>Standard</u>.--This is the usual March Supplement estimate. It is obtained by inflating the first stage weighted sample results to the census-level agerace-sex population totals discussed earlier. Further adjustments are also made so that husbands and wives living together have the same sampling weight, while at the same time leaving unchanged the estimates for certain labor force categories.
- (c) Extended.--This is an estimate obtained by a combination of adjustments designed to yield a complete "correction" for all the March 1973 undercoverage, not just the differential undercoverage relative to the 1970 Census. It was derived as a byproduct of the 1973 Exact Match Study and consists of adjusting the first stage weighted sample to independent population totals based on Jacob Siegel's Preferred Series D population estimates corrected for the 1970 Census undercount, an independent estimate of the total number of U.S. occupied households, and extensive administrative data from social security and tax records for persons eligible for interview in the March CPS.

It should be noted that neither the standard nor extended coverage adjustments make any special allowance for aliens illegally residing in the United States. To the degree that such individuals are not included in Census Bureau estimates of the true population, they have been omitted from consideration. 5/

1.3 Overall CPS population coverage.--Table 1 displays CPS undercoverage rates in March 1973 for all persons and for persons 14 years and older by age, race and sex. The rates express the discrepancy between the initial and stan-dard estimates and the corrected total CPS-eligible population. 6/

 (a) <u>Initial</u>.--The March 1973 CPS before adjustment underestimated the population eligible for interview by 9.4 million or 4.4 percent. For adults 14 years or older the undercoverage was proportionately greater, about 5.2 percent. Table 1 shows that the undercoverage rates were much more severe for males of other races (23.2 percent) than they were for white males (4.7 percent). Even so, the absolute number of white males missed in the survey (3.2 million) far exceeds the number of other males missed (2.1 million). Women, as the table shows, were generally better covered than men; whites better covered than other races. The worst coverage was for males of other races 22 to 39 years of age where over 30 percent were missed.

(b) <u>Standard</u>.--The standard Census Bureau coverage adjustment (the March supplement weighting) reduces the amount of undercoverage quite substantially in nearly every age-race-sex group. Differential undercoverage still exists, however, and therefore could have an impact on estimates of characteristics which vary greatly from one group to another.

Table 1.--Undercoverage Rates by Age, Race, and Sex Before Adjustment and After Standard Adjustment

	All Races		White		Other Races		
Age and Sex	Before	After Standard	Before	After Standard	Before	After Standard	
OVERALL	4.4	2.6	2.7	1.9	15.5	7.1	
Under 14 years old 14 years or older	2.2 5.2	2.3 2.7	0.3 3.5	1.4 2.0	11.9 17.2	6.9 7.2	
MALES							
14 years or older, total	6.9	3.7	4.7	2.8	23.2	10.1	
14 to 21 years 22 to 39 years 40 to 64 years 65 years or older	3.6 11.6 5.6 3.2	2.0 5.0 4.2 1.4	1.5 8.9 3.3 2.6	1.6 3.8 3.0 1.7	15.9 30.7 24.4 8.3	4.0 13.2 14.2 -1.5 *	
FEMALES							
14 years or older, total	3.5	1.7	2.4	1.3	11.8	4.7	
14 to 21 years 22 to 39 years 40 to 64 years 65 years or older	2.1 4.6 2.8 5.0	1.4 2.2 1.2 2.4	0.8 3.4 1.6 4.0	1.1 1.8 0.6 2.2	9.3 12.9 12.0 15.0	3.2 5.1 5.5 4.5	

(Percent of corrected total CPS-eligible population)

\*In this case, the standard estimate exceeded the corrected population total.

It might be good to mention one more thing about the March 1973 CPS population undercoverage before going on to look at the impact of alternative adjustments on income distribution statistics. Insofar as we can tell, the basic demographic dimensions of the missed groups are roughly the same as those which have been observed historically [4, Chart 3]. The only difference of any importance is that the overall coverage is slightly better in 1973 than that in more recent years. This implies, among other things, that any sensitivity we might observe would in all probability be greater if we were doing the same study with, say, the March 1977 CPS.

# 2. INCOME STATISTICS FOR PERSONS

A considerable body of conjecture exists about the socio-economic characteristics of persons not covered by the decennial censuses or the CPS. For example, the supposition has been advanced earlier at this session [4] that persons missed in enumerated CPS households tend to have smaller incomes on the average (i.e., are poorer) than covered persons of the same sex, race and age. They may also have a weaker attachment to the employed labor force; that is, be more often unemployed.

There is some evidence from the 1970 Census-CPS Match Study supporting the hypothesis that a household's coverage in the Census was directly related to the amount of income received by its members. Median family income for persons missed in the 1970 Census was 73 percent of the median family income of the entire population [5, p. 8]. This pattern of difference applies, however, only to white families; no such relationship emerged for variation of coverage with respect to income among families of other races.

A natural question to ask of the CPS coverage adjustments presented in this paper is whether or not they yield results which conform to working hypotheses such as those just mentioned. In the remainder of this section we will try to provide some answers to this question for statistics on the income of persons.

2.1 Income recipiency.--The number of persons 14 years or older reporting money income in the CPS rose by 7.4 million from the initial to the extended estimates and by 3.5 million from the initial to the standard estimates. The percentage increases of 6.2 percent and 3.0 percent, respectively, are slightly higher than the corresponding percentage increases for all persons. The overal! income recipiency rate, consequently, has increased from 78.8 percent at the initial stage to 79.4 percent at the extended stage. Moreover, as table 2 shows, the slight increase in income-recipiency rates is broadly based. All four age groups and all four race-sex groups show increases from the initial to the extended estimates.

Table 2.--Income Recipiency Rates for Persons 14 Years or Older

(In	percent)
-----	----------

Item	Initial	Standard	Extended
Overall	78.8	79.0	79.4
AGE			
14 to 24 years	66.6	67.1	68.0
25 to 44 years	80.9	81.2	81.5
45 to 64 years	82.1	82.1	82.3
65 years or older	91.3	91.3	91.8
RACE AND SEX			
White males	92.5	92.6	92.7
White females	66.7	66.7	67.5
Males of other races	84.1	84.4	84.9
Females of other races.	71.5	72.1	72.1

These results are perhaps inconsistent with the working hypothesis just discussed that persons missed in the survey may have a weaker attachment to the employed labor force than do covered persons and hence smaller or no income from earnings.

Mean total money income reported in the survey falls, however, as a result of both the standard and extended coverage adjustments. (See Table 3.) This result is clearly consistent with at least the second part of our hypothesis since it implies a mean income among the persons missed in the survey which is smaller than that for covered persons.

Table 3.--Mean Income Amounts for Persons with Income 14 Years or Older  $% \left( {{{\left[ {{{T_{{\rm{B}}}} \right]}_{{{\rm{A}}}}}} \right.} \right)$ 

(In	do 1	lars)
-----	------	-------

Item	Initial	Standard	Extended
Overall	6,398	6,376	6,289
AGE			
14 to 24 years 25 to 44 years 45 to 64 years 65 years or older	2,798 8,223 8,405 3,941	2,847 8,182 8,377 3,932	2.837 8,089 8,246 3,909
RACE AND SEX			
White males White females	9,001 3,608	8,980 3,609	8,841 3,616
Males of other races Females of other races	5,744 3,349	5,615 3,362	5,653 3,336

Why, then, should income-recipiency rates rise, given that mean money income falls (on the average) and given what we believe to be the nature of CPS undercoverage? Two explanations can be advanced. The first is that while a weak attachment to the employed labor force clearly implies low earnings, it does not necessarily imply no earnings and certainly not a total lack of income.

Another way in which we have to modify the hypothesis relates to certain dependent groups often without income of their own: wives living with their husbands (especially if they have young children) and teenage children still living with their parents and younger siblings. These groups tend to be better covered than persons of the same age, race or sex who are living in different circumstances. The effect of trying to account for this difference (as is done in the extended procedure) results in decreasing the original weight of these groups relative to that of groups more likely to have some income. (For white women and persons 14 to 24 this effect is so strong that it not only increases recipiency rates, it also increases mean incomes for persons with income.)

2.2 <u>Income aggregates.--</u>Table 4 shows the Bureau of Economic Analysis (BEA) 1972 benchmarks for each type of money income collected in the March

Table 4.--March 1973 CPS Aggregrate Income by Type as Percent of BEA Benchmark Before and After Coverage Adjustment

	Revised BEA	CPS as a percent of Benchmark				
Type of Income	Benchmark (in billions of dollars)	Initial	Standard	Extended		
		_				
Total money income	867.0	87.5	89.7	91.3		
Wages and salaries	619.9	94.5	97.1	99.0		
Nonfarm self-employment.	56.5	94.2	95.9	95.4		
Farm self-employment	16.3	64.8	65.3	65.5		
Social security 1/	39.8	91.1	93.1	94.9		
Property income	75.0	44.2	45.1	45.8		
Public assistance	10.9	67.9	70.8	72.9		
Other transfers	27.2	66.3	68.1	70.6		
Other income	21.4	64.1	65.6	66.6		

1/ Includes Railread Retirement.

1973 CPS. Also shown is the percentage of each income type obtained from the CPS at the initial, standard, and extended stages of estimation.

Increasing the number of income recipients will, of course, raise the aggregate amount of money income estimated in the CPS. Note, for example, that the standard adjustment raises the aggregate amount of money income reported from 87.5 percent of the benchmark to 89.7 percent. The extended adjustment lifts the aggregate amount reported to 91.3 percent. The shortfall in the CPS reporting of personal income is reduced by 30 percent after an extended coverage adjustment, 18 percent after the standard adjustment.

The underreporting and nonreporting of income in the survey [16-17] are perhaps the chief causes of the remaining BEA-CPS differences. A full discussion of how those problems occur is beyond the scope of this paper. Other papers [e.g., 18-20] from the Exact Match Study have addressed this question and interested readers may wish to consult them for further information.

2.3 <u>Income distributions for missed persons</u> by race.--Table 5 compares the percentage distribution of CPS total money income for persons with income age 14 or older covered in the survey at the initial stage with that for persons missed in the survey as estimated by the standard and extended coverage adjustments. The results, at least for the extended estimates, are entirely consistent with the findings of Siegel presented earlier [5] for the 1970 Census.

Focusing on the distribution of white persons, it is clear that the extended coverage adjustment picks up persons who tend to have much lower incomes than covered persons. This is not the case for the standard adjustment. The mean income for white persons with income imputed as missed by the standard adjustment was \$6,506. This compares to a mean of \$6,620 for covered persons with income and a mean of \$4,513 for persons imputed as missed by the extended adjustment. (See table 5.)

Unlike those for white persons, the income distributions imputed to covered and missed persons of other races are more nearly identical. This is true of both the standard and extended adjustments and is entirely in keeping with [5].

Table 5Income	e Size Distribution of	Persons with	Income 14	Years or Olde	er, Covered an	d Missed, by Race
---------------	------------------------	--------------	-----------	---------------	----------------	-------------------

		All races			Whites			Other races		
Money Income	Initial Covered	Missed	Persons	Initial Covered	Missed	Persons	Initial Covered	Missed	Persons	
	Persons	Standard	Extended	Persons	Standard	Extended	Persons	Standard	Extended	
Total										
number (millions)	118.5	3.5	7.4	106.2	1.9	4.5	12.3	1.6	2.9	
PERCENT	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	. 100.0	
\$1-\$4,999 or loss	51.9	54.4	64.4	50.5	45.6	64.6	64.1	64.6	64.2	
\$5,000-\$6,999	12.2	15.5	12.4	12.1	16.5	12.8	12.9	14.3	11.6	
\$7,000-\$9,999	14.3	16.1	14.2	14.4	19.1	13.9	13.3	12.7	14.7	
\$10,000-\$11,999	7.2	6.1	4.5	7.4	7.4	4.4	4.8	4.6	4.6	
\$12,000-\$14,999	6.3	4.1	2.6	6.7	5.6	2.4	2.7	2.1	2.9	
\$15,000 or more	8.1	3.8	1.9	8.8	5.7	1.9	2.1	1.7	2.0	
Mean income										
(in dollars)	6,398	5,631	4,526	6,620	6,506	4,513	4,489	4,376	4,530	

2.4 <u>Overall distributional impact.--What impact</u> do the alternative coverage adjustments have on the entire distribution of personal money income? Very little, it would seem, given the small size of the adjustment in relative terms. (See figure 1.) Differences between the initial, standard, and extended distributions are extremely small. The differences do, however, follow the hypothesized [5] pattern: Both of the adjusted distributions are shifted slightly to the left of the initial estimates. The extended adjustment, again, shows the larger change, in effect shifting weight directly from the \$10,000 and above classes down to the lowest income class. There is an increase in the under \$5,000 group of 0.8 percentage points from the initial (51.9 percent) to the extended (52.7 percent); all of this increase is compensated for by declines in the size classes above \$10,000, there being no change in the proportion of individuals with incomes of \$5,000 to \$9,999.

Figure 1





CPS Total Money Income

Still another way to look at the distributional differences between the three estimates is to examine the following selected income percentile points.

Percentile	Income at	Percentile	(In dollars)
Points	Initial	Standard	Extended
20th	1,359	1,370	1,352
50th	4,710	4,703	4,609
80th	10,363	10,315	10,193
95th	17,566	17,476	17,278

Again we see the downward shift in the income distribution after adjusting for coverage. For the standard procedure it is fairly slight (\$7, for example, at the median or 50th percentile); for the extended adjustment, the shift is somewhat larger (from \$4,710 to \$4,609 at the median).

2.5 <u>Measures of Distributional Inequality</u>.--Mixed results have been obtained with respect to some measures of distributional inequality. The overall Gini concentration ratio or coefficient of inequality declines slightly from the initial to the standard and extended stages, from 0.5042 to 0.5029 and 0.5036, respectively. Since the Lorenz curves which underlie those Gini coefficients do not intersect, it can be said unambiguously that measured inequality in the total distribution has been reduced, however slightly, by each coverage adjustment. <u>7</u>/

A more revealing measure of changes in the relative distribution, (that is, of changes in

the shares of various quantiles relative to one another), is the mean income of any quantile divided by the mean income of the distribution as a whole [22, p. 247]. However, even this more sensitive "relative mean income" measure registers very little change as a result of coverage adjustments to the CPS. The mean income of the bottom quintile has increased 2.6 percent relative to the mean income of the distribution as a whole after the extended coverage adjustment; all other quintile share changes are less than 1 percent of the new mean.

# 3. INCOME STATISTICS FOR FAMILIES AND UNRELATED INDIVIDUALS

In this section we will continue our analysis of the impact of coverage adjustments on income statistics. Attention will be focused now on consumer units (families and unrelated individuals) rather than on persons.

3.1 <u>Number of families and unrelated indivi-</u> <u>duals</u>.--Adjusting for net CPS undercoverage errors has a marked impact on the relative number of families and unrelated individuals. (See table 6.) At the initial stage, unrelated individuals comprise 16.1 million or 23.2 percent of the 69.4 million consumer units. The group rose in importance to 23.6 percent after the standard coverage adjustment; and, after the extended coverage adjustment, unrelated individuals represent 25.2 percent of the total. The difference for unrelated individuals between the standard and extended estimates (about 1.5 million) is even more striking when one notes

Itom	Families and unrelated individuals		Families		Unrelated individuals				
I LEIII	Initial	Standard	Extended	Initial	Standard	Extended	Initial	Standard	Extended
Total	69.4	71.2	72.7	53.3	54.4	54.4	16.1	1 <b>6</b> .8	18.3
AGE OF HEAD									
4 to 24 years 5 to 44 years 5 to 64 years 5 years or more	6.3 25.6 24.0 13.4	6.7 26.2 24.5 13.8	6.9 27.0 24.9 13.9	4.0 22.3 19.6 7.4	4.2 22.7 19.9 7.6	4.1 22.9 19.9 7.5	2.3 3.3 4.4 6.0	2.5 3.6 4.6 6.2	2.8 4.1 5.0 6.4
RACE AND SEX OF HEAD									
/hite males /hite females	48.5 13.5	49.3 13.7	49.4 14.2	43.2 4.6	43.8 4.7	43.4 4.7	5.3 8.8	5.5 9.0	6.0 9.5
Males of other races Females of other races	4.7 2.8	5.1 3.1	5.9 3.2	3.7 1.8	4.0 1.9	4.3 1.9	1.0 1.0	1.2 1.1	1.6 1.3

Table 6.--Number of Families and Unrelated Individuals by Type of Estimate, Age, Race and Sex

(In millions)



Distribution of CPS Total Money Income for 1972 Before and After Coverage Adjustments

that the number of families is virtually identical (at 54.4 million) for both procedures.

Although the total number of families did not change between the standard and extended coverage adjustments, the race-sex composition of family heads changed significantly. Particularly noteworthy is that families headed by males of other races jumped 16.2 percent from the initial to the extended stage and 7.5 percent from the standard to the extended stage. Large relative shifts also occurred in the racesex and age composition of unrelated individuals as a result of the coverage adjustments.

3.2 <u>Distributional impact</u>.--Figure 2 shows the CPS money income distribution before and after adjustment. This is done separately for families, and families and unrelated individuals combined. What should be noted about these size distributions of income is, first, that the initial, standard and extended are quite close; the second thing to notice is that, as with persons, both coverage adjustments introduce a very small yet persistent downward shift in income.

Perhaps an even better picture of what is happening emerges when we look at median income. For families, median income falls by about 1 percent between the initial and extended estimates, from \$11,101 to \$10,990 respectively. The standard estimate yields a median family income which, at \$11,045, is about midway between the other two.

For unrelated individuals, there is virtually no significant change in the medians. The initial median is \$3,526 with the standard (\$3,538) and extended (\$3,540) being slightly larger.

For families and unrelated individuals combined we see the largest differences between the medians. There is a drop from the initial estimate of \$9,304 to \$9,225 for the standard and \$9,074 for the extended. The reason for larger declines in the combined distribution than in the components is due to the increase in importance of unrelated individuals commented on earlier.

With respect to distributional inequality, there is a tendency for the coverage adjustments to increase inequality among consumer units, in contrast to the decrease among persons. This is especially the case for families and unrelated individuals combined where the Gini ratios rise from .4085 for the initial estimate to .4094 and .4121 for the standard and extended estimates, respectively.

3.3 <u>Some limitations on family data</u>.--A problem exists with the current family estimates in that we are following the standard Census procedure of using the weight of the family head (primary or secondary) in deriving the estimates. Neither the standard nor extended coverage adjustment reflect the increasing probability of a family not being completely enumerated as household size increases.

By shifting weight from smaller-sized to largersized families in order to properly reflect the family size distribution, the income distribution will probably be shifted slightly upward. Such an elaboration on the extended coverage adjustment is currently being tested.

The total impact of such a change is expected to move the extended coverage-adjusted family distribution into closer alignment with the already published standard family distribution. Differences caused by the increased importance of unrelated individuals or certain types of families (e.g., those of other races) are expected to remain; however, distributional shifts within each family type will diminish, becoming even less significant than they are now.

### 4. SELECTED POPULATION CHARACTERISTICS AND LABOR FORCE STATISTICS

CPS characteristics, other than income, are also

sensitive to the nature of the coverage adjustment. This section briefly examines the impact of alternative adjustments on labor force data, poverty rates, educational attainment, and residence statistics.

4.1 Labor force and unemployment rates.--The effect of adding about 8.2 million persons age 14 and older to the CPS population (4.0 million more than the standard CPS coverage adjustment) naturally tends to increase aggregate labor force totals. Table 7 shows that, relative to the published estimate [23], the extended coverage adjustment increased the size of the labor force by three million persons or 3.4 percent. While the estimated employed segment of the labor force was increased 3.0 percent, the estimated unemployed segment increased 11.6 percent. This meant that we imputed an unemployment rate of 17.6 percent to the population not covered by the standard CPS coverage adjustment. The overall impact on the measured unemployment rate was to raise it from 5.17 percent to 5.57 percent.

Table 7.--Labor Force Estimates for March 1973 Before and After Coverage Adjustment

(Numbers in millions)

Labor force category	Type of estimate					
(16 years and older)	Initial	Standard <u>1</u> /	Extended			
Total civilian labor force	84.7	87.3	<b>90.3</b>			
Employed	80.3	82.8	85.3			
Unemployed	4.5	4.5	5.0			
Unemployment rate (percent)	5.3	5.2	5.6			

1/ Obtained from [23].

The unemployment rate grows from the standard estimate to the extended one in part simply because we have given certain groups (e.g., males of other races) more importance. Some increase in unemployment rates also occurred, though, for each race-sex group separately.

Johnston and Wetzel in a 1969 paper examined the possible impact of undercoverage on unemployment rates [13]. Their results are sharply different from ours. They prepared two estimates of the unemployed to look at the problem. The first was under a "comparability" assumption. Missed persons were assumed to have the same labor force characteristics as their peers (sex, race, and age cohort). The second was a "poverty neighborhood" assumption. This procedure assumed that the missed population had the same labor force characteristics as its peers in poverty neighborhoods. Under both assumptions the published unemployment rate for 1967 was raised only from 3.8 percent to 3.9 percent.

Differences in the results of the two studies

arise for a number of reasons, including changes in the nature of the CPS, demographic changes in the labor force and in the undercoverage itself. There is also a deficiency in the Johnston and Wetzel methodology which may account for the greater insensitivity they observed. They did not look at the total CPS coverage problem, only that portion of the shortfall which is accounted for by bringing the CPS up to undercount-corrected totals after it had already been adjusted to census-level population estimates.

4.2 <u>Poverty rates.</u>--Brief mention needs to be made of the sensitivity of poverty estimates to alternative coverage adjustments. The following summary comparison in table 8 may aid in this endeavor.

# Table 8.--1972 Poverty Estimates Before and After Adjustment

(Numbers in millions)

Persons 14 Years or Older	Initial	Standard	Extended
Total	150.5	154.5	158.6*
Poor	15.8	16.4	17.3
lonpoor	134.7	138.0	141.3
overty rate (in percent):			
Overal1	10.5	10.6	10.9
Imputed to missed	**	15.5	21.2

\*This is slightly smaller than the corrected CPS-eligible population 14 or older of 158.7 because the extended adjustment does not force exact agreement with all population totals [11].

#### \*\*Not applicable.

In keeping with the changes in the overall income distribution we found a slight apparent increase in poverty for all persons 14 or older from the initial (10.5 percent) to the standard (10.6 percent) and extended (10.9 percent). Poverty is a family characteristic and hence would be affected by our failure, so far, to correct the standard or extended estimates for within family undercoverage. As a result, it is likely that the above difference between the initial and adjusted estimates overstates somewhat the impact of the problem of undercoverage on poverty rates.

4.3 Other selected characteristics.--To complete this paper's brief discussion of the impact of alternative coverage adjustments, we will now look at changes which take place in statistics by Census Region, metropolitan residence and educational attainment. Some examination of table 9 will show that basically both the standard and extended adjustments are imputing missed persons in a way which would be roughly appropriate if the pattern of under-coverage described in [5] for the 1970 Census were applicable to the March 1973 CPS. In particular, missed persons are more likely to have only an elementary education and to live in the South. The central city population estimates experience the largest absolute and proportionate increases in the residence category as a result of the coverage adjustments. This is mainly due, however, to the heavy concentration of persons of other races in the central cities. The extended estimate, as usual, shows more of a shift than the standard but the direction is generally the same.

Table 9.--Selected Characteristics by Type of Estimate

(Numbers in millions)

Selected Characteristic	Initial	Standard	Extended
Persons 14 years or older by Census region	-	-	
Northeast	36.3	37.2	37.8
North Central	41.2	42.2	43.4
South	46.8	48.2	49.8
lest	26.2	27.0	27.6
Persons 14 years or older by residence			
In metropolitan areas:			
Central city	45.6	47.4	48.7
Suburban fringe	57.3	58.6	59.7
Outside metropolitan areas	47.6	48.5	50.2
Persons 25 years or older by educational attainment			
Elementary	26.2	26.9	27.9
ish school	57.5	58.9	60.1
2011ege	26.4	27.1	27.6

### 5. CONCLUSION

In this paper we have done some further exploration of the impact of coverage adjustments on income and other socio-economic characteristics. All of the earlier CPS studies of this type [6, 12, 13, 24], and ours as well, show that such adjustments have their major effect on aggregates, with the effects on overall percentages and rates being much less pronounced.

Especially important perhaps for users of CPS income data is the extent to which aggregate income is increased when coverage adjustments are made. We have seen, for example, that some 30 percent of the CPS understatement of income is eliminated by employing the extended coverage adjustment.

The income size distributions of both persons and families were much less affected by our adjustments. Even so, our results demonstrate that CPS coverage problems significantly limit the survey's usefulness in studying economic well-being.

We would like to urge, along with Bateman [3], that additional research be undertaken to improve the coverage adjustment procedures now being employed in the Current Population Survey. Certainly, as users of the CPS, we will be continuing our own studies in this area.

## ACKNOWLEDGEMENT AND FOOTNOTES

The authors owe thanks to a great many people for contributing to the work being presented in this paper. Jean Salter at the Bureau of Economic Analysis (BEA) prepared the Income Benchmark revisions used in table 4. H. Lock Oh, Linda DelBene and Faye Aziz at Social Security (SSA) made major contributions to the extensive background material [11, 14] provided as handouts at the session.

Special thanks also must be given to the other participants at this session for their

very helpful comments. Editorial and other assistance was provided by Wendy Alvey, Ben Bridges and Denton Vaughan at SSA and Gordon Green at the Census Bureau. The typing was done at SSA by Rubye Ellis and Helen Kearney.

- 1/ The 1973 CPS-IRS-SSA Exact Match Study has been the subject of numerous papers at previous American Statistical Association meetings in 1974, 1975 and 1976. For further details on its goals and content, see, for example, [1] or [2].
- 2/ Derived from the extended coverage adjustment. It should be noted that we imputed a smaller average household size to missed households than to enumerated ones. This does not seem to be consistent with the findings from coverage checks done in connection with the 1970 Census [5]. However, the two results may be reconcilable if one takes into account the impact of the (uncorrected) household size bias that is typical of CPS noninterviews [6].
- 3/ The CPS estimate of total occupied units before any coverage adjustment was 66.7 million. The total number of occupied units was estimated to be 69.2 million [11, pp. 30-32].
- 4/ In the companion paper [11] provided as a handout at the session each of these estimates was described in greater detail than space will permit here. It should also be mentioned that the extended estimator is the average of two different approaches to the correction problem on which results are available separately [14].
- 5/ For some preliminary estimates of the number of illegal aliens not included in Census Bureau population totals, see [15].
- 6/ The starting point in developing our estimates of the CPS eligible population was Siegel's Preferred 1970 undercount corrected population totals aged to April 1, 1973. To these an adjustment was then made to exclude the institutional population and that portion of the Armed Forces not eligible for interview (See [11, pp. 16-18] for full details.)
- 7/ All percentile estimates, Gini ratios and income shares shown in this paper were prepared from the grouped data in [14] using the procedures discussed in [21].

### REFERENCES

 [1] "The Role of the Social Security Number in Matching Administrative and Survey Records," <u>1974 American Statistical Association Pro-</u> ceedings, Social Statistics Section, pp. 126-156. See also, "The Reconciliation of Survey and Administrative Income Distribution Statistics through Data Linkage," <u>1975 American Statistical Associa-</u> tion Proceedings, <u>Social Statistics</u> <u>Section</u>, pp. 119-158.

- [2] U.S. Social Security Administration, <u>Studies from Interagency Data Linkages</u> (especially Report No. 4).
- [3] Bateman, D.V., "Analysis of Census Bureau National Housing Inventory Estimates," <u>1977 American Statistical Association Pro-</u> ceedings, Social Statistics Section.
- [4] Korns, A., "Coverage Issues raised by Comparisons between CPS and Establishment Employment," <u>1977 American Statistical</u> <u>Association Proceedings, Social Statistics</u> <u>Section</u>.
- [5] Siegel, J.S., <u>Coverage of population in</u> <u>1970 Census and some implications for</u> <u>public programs</u>, Series P-23, No. 56, 1975.
- [6] Scheuren, F., Kilss, B. and Oh, H.L., <u>Studies from Interagency Data Linkages</u>, <u>Report No. 2, 1973.</u>
- [7] Siegel, J.S., Estimates of coverage of population by sex, race and age: demographic analysis, 1970 Census of Population and Housing: Evaluation and research program, PHE(E)-4, 1974.
- [8] Montie, I.C. and Schwanz, D.J., "Coverage Improvement in the Annual Housing Survey," <u>1977 American Statistical Association Proceedings, Social Statistics Section.</u>
- [9] Brooks, C. and Bailar, B., "Nonsampling errors in the Current Population Survey as they affect the Employment Statistics" (Unpublished OMB working paper appearing in a condensed form in <u>1977 American Statistical Association Proceedings, Social</u> <u>Statistics Section</u>).
- [10] U.S. Bureau of the Census, <u>Concepts and</u> <u>methods used in labor force statistics</u> <u>derived from the Current Population Survey</u>, <u>Series P-23, No. 62, 1976.</u>
- [11] Scheuren, F.J., "Methods of Estimation for the 1973 Exact Match Study" (unpublished working paper distributed with [14] at the session; to appear in the series <u>Studies</u> from Interagency Data Linkages).
- [12] Siegel, J. (1968), "Completeness of coverage of the nonwhite population in the 1960 Census and current estimates, and some implications," <u>Social Statistics and the</u> <u>City</u>. Cambridge: Harvard University, pp. 13-54.
- [13] Johnston, D. and Wetzel, J. (1969), "Effect of the census undercount on labor force estimates," <u>Special Labor Force Report</u> <u>No. 105</u>. Bureau of Labor Statistics.

- [14] Yuskavage, R. and Oh, H.L., "Four alternative Estimates of CPS income data for 1972" (unpublished working paper distributed with [11] at the session).
- [15] Lancaster, C. and Scheuren, F., "Counting the Uncountable Illegals: Some initial Statistical Speculations employing Capture-Recapture Techniques," <u>1977 American</u> <u>Statistical Association Proceedings</u>, <u>Social Statistics Section</u>.
- [16] U.S. Bureau of the Census, <u>Current Population Reports</u>, Series P-60, No. 90, pp. 24-25.
- [17] Budd, E., Radner, D. and Hinrichs, J., "Size Distribution of Family Personal Income: Methodology and Estimates for 1964," <u>Bureau of Economic Analysis Staff Paper</u> <u>No. 21</u>, 1973.
- [18] Herriot, R. and Spiers, E., "Measuring the Impact on Income Statistics of Reporting Differences between the Current Population Survey and Administrative Sources," <u>1975</u> <u>American Statistical Association Proceedings, Social Statistics Section.</u>
- [19] Kilss, B. and Alvey, W., "Further explo-

ration of CPS-IRS-SSA Wage Reporting Differences," <u>1976 American Statistical</u> <u>Association Proceedings, Social Statistics</u> <u>Section</u>.

- [20] Vaughan, D. and Yuskavage, R., "Investigating Discrepancies between Social Security Administration and Current Population Survey Benefit Data for 1972," <u>1976</u> <u>American Statistical Association Proceedings, Social Statistics Section.</u>
- [21] Oh, H.L., "Osculatory Interpolation with a Monotonicity Constraint," <u>1977 American</u> <u>Statistical Association Proceedings</u>, <u>Statistical Computing Section</u>.
- [22] Budd, E., "Postwar Changes in the Size Distribution of Income in the U.S.," <u>American</u> <u>Economic Review Proceedings</u>, May 1970, pp. 247-260.
- [23] U.S. Bureau of Labor Statistics, Employment and Earnings, April 1973, p. 23.
- [24] Vaughan, D. and Ireland, C.T., "Adjusting for coverage errors in the March 1973 Current Population Survey," <u>1975 American</u> <u>Statistical Association Proceedings,</u> <u>Social Statistics Section</u>.