

BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-86/13

CPS-Census Retrospective Study
Final Report

by

Gregg J. Diffendal
Undercount Research Staff
Statistical Research Division
U.S. Bureau of the Census

Room 3582, F.O.B. #3
Washington, D.C. 20233

(301)763-5926

This series contains research reports, written by or in cooperation with staff members of the Statistical Research Division, whose content may be of interest to the general statistical research community. The views reflected in these reports are not necessarily those of the Census Bureau nor do they necessarily represent Census Bureau statistical policy or practice. Inquiries may be addressed to the author(s) or the SRD Report Series Coordinator, Statistical Research Division, Bureau of the Census, Washington, DC 20233.

Recommended by: Howard Hogan

Report completed: March 26, 1986

Report issued: May 27, 1986

TABLE OF CONTENT

Executive Summary and Policy Implications.....	1
I. Introduction.....	3
II. Results.....	4
III. Sampling.....	27
IV. Tracing and Matching.....	30
V. Weighting and Nonresponse Adjustments.....	34
VI. Costs.....	35
VII. Comparisons.....	36
VIII. Conclusions.....	42
References.....	45
Appendix Final Match Codes.....	45

Acknowledgements:

I would like to thank David Altvater, Dan Childers, Judy Evaul, Sue Finnegan, Alan Heckert, Barry Hill and Howard Hogan who worked on conducting the CPS-Census Retrospective Study, and/or made comments on previous drafts of this paper.

CPS-Census Retrospective Study

Executive Summary and Policy Implications

The CPS-Census Retrospective Study took a sample from the 1977 CPS and traced and matched them to the 1980 Decennial Census. This study is the only attempt to trace and match people before the 1990 Census. Also this study can be viewed as a pre-enumeration survey taken 3 years before the census and tracing began two years after the census.

The CPS-Census Retrospective Study produced a nonmatch rate of 14% and a not traced rate of 4.8%. The nonmatch rate is comparable to the IRS/Census Direct Match Study but is about twice as high as the 1980 Post-Enumeration Program (see Section VII Comparisons). The not traced rate is slightly higher than the IRS/Census Direct Match Study (which used the 1979 IRS file). The 1976 Canadian Reverse Record Check (using 5 year trace period) had the same not traced rate as the CPS-Census Study while the 1960 U.S. Reverse Record Check (using a 10 year trace period) had a not traced rate over twice as high as the CPS-Census Study. The results from the Forward Trace Study (to be completed in 1986) will provide further evidence of tracing for census evaluations.

The nonmatch rates by race are particularly important because of the strong evidence of differences by race from demographic analysis and the PEP. Blacks and other races had a nonmatch rate about twice as high as the white race group. However, the nonmatch rates may be affected by recall bias due to the late start of the project-two years after Census day. The not traced rate was also twice as high for black and other as compared to whites. The relative nonmatch rates are comparable to the results from the 1980 PEP.

The not traced rate (4.8%), noninterview (refusal) rate after tracing (4.3%), and noninterviews (3 to 5% for each month of CPS) gives an incomplete match status for about 12 to 14% of the sample. Given the controversy of the imputation used in the 1980 PEP which produced 12 estimates and the level of unresolved match status (about 4.0%) for the April 3 series compared to the level of the undercount (.8% PEP 3-8), the large number of incomplete match status is a serious liability that alone may rule out using any technique of tracing for 3 to 5 year period.

The undetermined rate (not traced and refusals) for this study are considerably higher than would be acceptable for a decennial census evaluation. The not traced rate was fairly consistent with other studies involving tracing. The Forward Trace Study will provide a better understanding of the ability to trace people over time. Some difficulties were observed in deciding whether to classify a person as not traced versus nonmatched.

I. Introduction

The CPS-Census Retrospective study took one rotation panel (about one-eighth of the full sample) from the March 1977 CPS and searched the census to determine a match status for the sample persons. In order to determine a match status, the person had to be found in the 1980 census or the person had to be contacted to verify the person's address on census day. The main purpose of this study is to examine the nonmatch rates and the not traced rates for this measurement method. Tracing is the ability to find someone after the original contact.

There were five stages in determining a match status for each person. The first stage searched for the people at their March 1977 address in the census files. If they were not found in the census at their 1977 address, the second stage searched for a new address in the 1979 IRS/IMF. Only matched persons or persons out of scope can be determined from the first two stages, since a person not found may live at a different address April 1, 1980. The third stage was mail followup where each household was mailed a questionnaire and was asked to mail it back to the Census Bureau. Nonresponse and postmaster returns from mail followup were sent to telephone followup, the fourth stage. A person could be determined to be a match, nonmatch, refusal, or out of scope case at stages three or four when a 1980 address was obtained. Step five was field followup. For those contacted the person could be categorized as above. After the five stages there was still a group of people who the Census Bureau could not find and contact after field followup. These were the not traced cases.

Table 1 shows the number of sample persons assigned a code of matched, nonmatched, noninterviewed, not traced and out of scope for each of the five operations.

Table 1

Persons' Status by Operation Resolved

	77 CPS	IRS/IMF	Mail	Telephone	Field	TOTAL
Matched	9155	3511	1072	994	1259	15991
Nonmatched	0	0	211	700	1064	1975
Noninterviewed	0	0	53	663	154	870
Not Traced	0	0	0	0	958	958
Out of Scope	56	0	83	139	138	416
TOTAL	9211	3511	1419	2496	3573	20210

The five operations were conducted from 1982 when the census match was done until August 1983 when the field followup and matching was conducted. So for nonmovers from March 1977 until April 1980 who were correctly enumerated, and for movers who gave their current address in their IRS filing for 1979 (sent to IRS from January 1980 to April 1980) and were linked to a sample respondent (those who gave social security numbers), the CPS-Census Retrospective Study reflects a 3 year tracing period. For those persons not found at their 1977 address or not found in IRS files, or found in IRS files but not found in 1980 census, this CPS-Census Retrospective Study reflects a five to six year tracing period since the tracing would have been done two to three years earlier if the original matching was done in 1980. Consequently, the result from this study should be interpreted in this context.

II. Results

In this section, the nonmatch and not traced rates are presented and discussed starting with the national estimates and continuing with estimates broken into multiple characteristics. The variables examined are age, race, sex, marital status, education and availability of social security number. Other information on the sampling, noninterview adjustments used to derive the estimates

discussed in this section, and the detailed steps in the tracing and matching are presented in later sections.

The total sample size was 19,794 people (out of scope persons are removed from the calculations in this section). Of these there was a national nonmatched estimate of 14.0%. The not traced estimate was 4.8%. These estimates use subsampling weights and noninterview adjustments discussed in later sections. The nonmatch rate should not be interpreted as an undercount estimate since it does not account for erroneous enumeration (overcounts) and imputations in the census. The nonmatch rate is much higher than that obtained from the 1980 Post Enumeration Program which had a gross nonmatch rate of 5.4% for April (3-8) and 6.1% for August (5-8).

Although the national match rate is interesting, breakdowns by the variables of interest will show where easy and difficult tracing and matching occurred. The variables examined are the following: age (3-17, 18-24, 25-34, 35-44, 45-54, 55-64, 65+), race (white, black, other), sex (male, female), marital status (single, married, divorced, seperated, widowed), education (grade school, some high school, graduated high school, some college, graduated college, post graduate college), and social security number (obtained, not obtained). These variables were recorded in the 1977 interview so education and marital status may have changed by census time. Age is coded as the persons age in 1980. See section V "Weighting and Nonresponse Adjustments" for a description of the nonresponse adjustments used here.

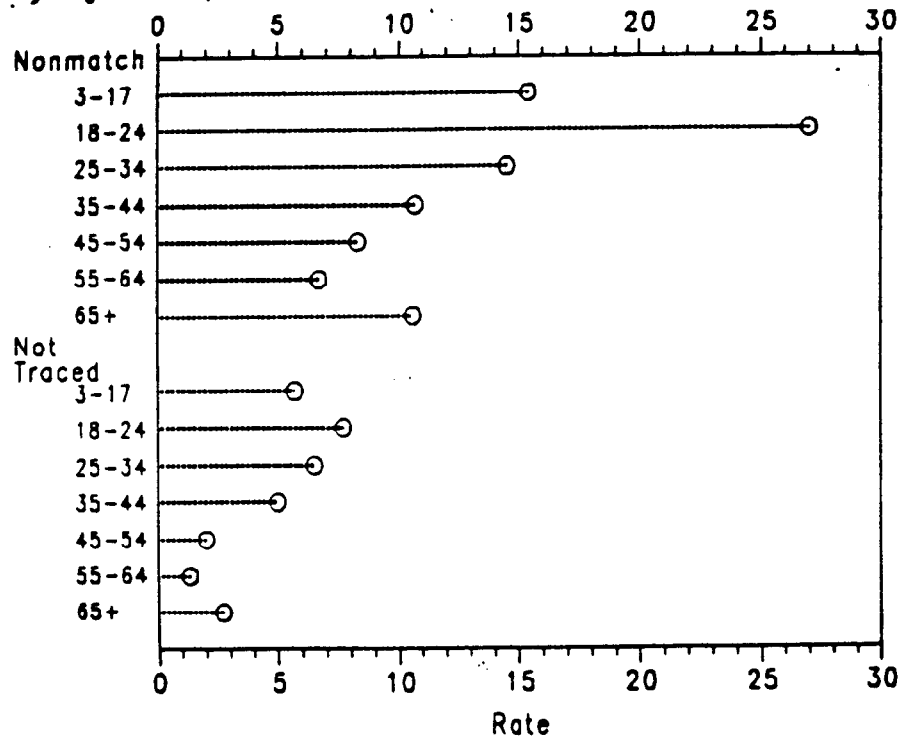
AGE

The seven age groups examined are: 3-17, 18-24, 25-34, 35-44, 45-54, 55-64, 65+. Note that this study does not cover the births or immigrants between 1977 and 1980. Table 2 shows the nonmatch, not traced and sample sizes for the seven age groups.

Table 2
Nonmatch Rate, Not Traced Rate and Sample
Size by Age

AGE	<u>3-17</u>	<u>18-24</u>	<u>25-34</u>	<u>35-44</u>	<u>45-54</u>	<u>55-64</u>	<u>65+</u>
Nonmatch Rate	15.4	27.0	14.5	10.7	8.3	6.7	10.6
Not Traced Rate	5.7	7.7	6.5	5.0	2.0	1.3	2.7
Sample Size	5270	2659	3129	2367	2011	1950	2408

Table 2: Nonmatch Rate and Not Traced Rate
by Age Group



The highest nonmatch and not traced rates are for the 18-24 years old age group. Being the most mobile part of the population, it is not too surprising to have large not traced rates. Mobility may also be part of the reason for an extremely high nonmatch rate. Failure at probing for college or alternate addresses may have led to the high nonmatch rate. The not traced rates for persons over 45 of less than 3 percent reflects the stability for this group. The ability to trace persons over 45 means this procedure may prove highly possible for this age group. It is surprising to see an increase in the not traced rate for the population over 65, perhaps due to movements after retirement. The very high nonmatch rates for person 65+ which were estimated at over 10 percent is troubling. Elderly people usually have a very high coverage rate in the census. The age groups 3-17, 25-34, and 35-44 are similar with not traced rates at 5 percent and nonmatch rates around 10 percent or higher.

RACE

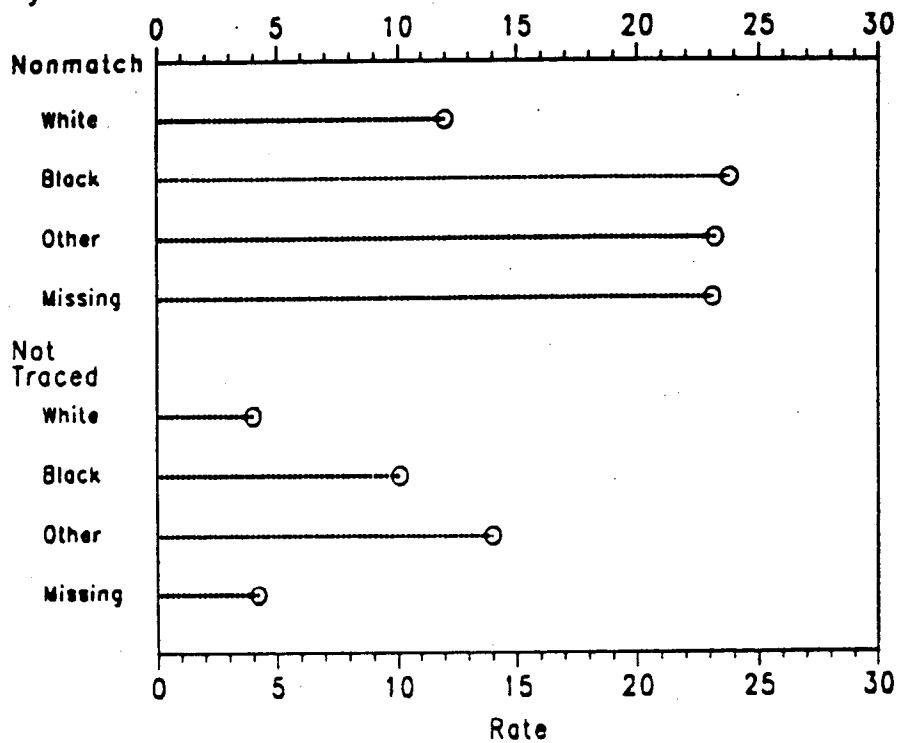
The race groups were examined using white, black, other, and missing (blank). Ethnicity was not asked on the 1977 CPS and is not available for analysis. Table 3 shows the nonmatch, not traced and sample sizes for the race groups.

Table 3

Nonmatch Rate, Not Traced Rate and Sample Size
by Race

RACE	<u>White</u>	<u>Black</u>	<u>Other</u>	<u>Missing</u>
Nonmatch Rate	12.0	23.8	23.2	23.1
Not Traced Rate	4.0	10.1	14.0	4.2
Sample Size	16411	1904	520	959

Table 3: Nonmatch Rate and Not Traced Rate
by Race



The white population had the lowest nonmatch and not traced rates. The blacks, others and missing race code persons have similar high nonmatch rates, about twice as high as white persons. However, persons with the race code missing have a not traced rate that is similar to white persons. Therefore it appears difficult to predict the missing race code by nonmatch and not traced since the missing race code resemble blacks and others for the nonmatch rate, but resemble whites for the not traced rates.

SEX

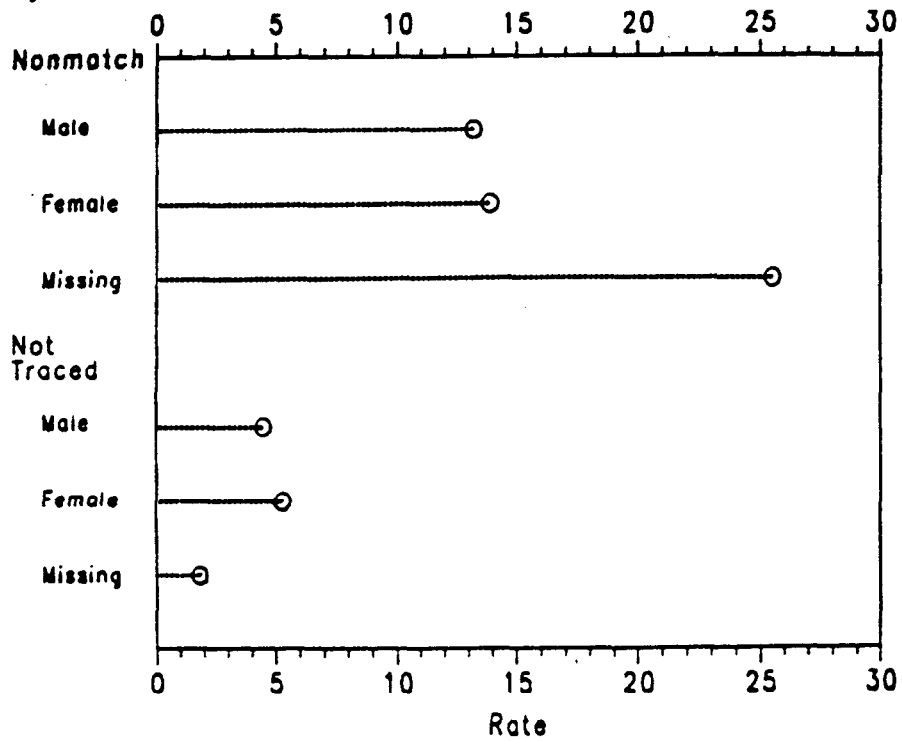
Next, sex of the sample person was examined to see its effect on nonmatch and not traced rates. Table 4 shows the nonmatch rates, not traced rates and sample sizes for males, females and missing categories.

Table 4

Nonmatch Rate, Not Traced Rate and
Sample Size by Sex

SEX	<u>Male</u>	<u>Female</u>	<u>Missing</u>
Nonmatch Rate	13.2	13.9	25.5
Not Traced Rate	4.5	5.3	1.8
Sample Size	9209	10035	551

Table 4: Nonmatch Rate and Not Traced Rate
by Sex



There are only small differences in the nonmatch and not traced rates for males and females, the females being slightly higher in both categories. When missing the sex code, there was a higher nonmatch rate but a lower not traced rate.

MARITAL STATUS

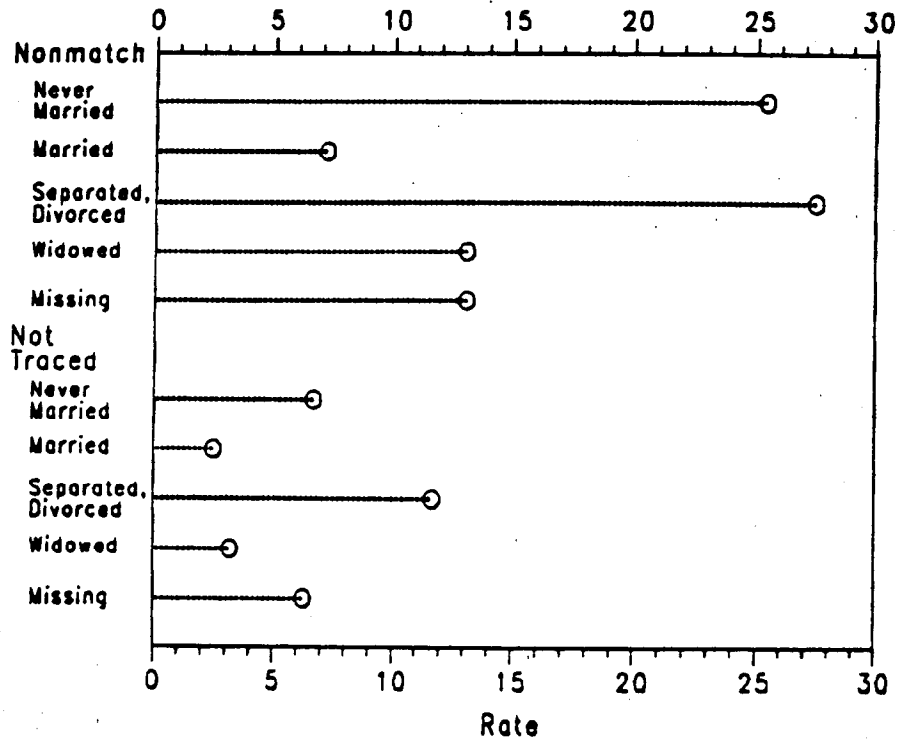
Marital status was examined using the codes never married, married, seperated divorced, widowed, and missing. The marital status was recorded in the 1977 interview and may have changed before the 1980 census. Table 5 shows the nonmatch, not traced and sample size by marital status.

Table 5

Nonmatch Rate, Not Traced Rate, and Sample Size by Marital Status

MARITAL STATUS	Never Married	Married	Separated, Divorced	Widowed	Missing
Nonmatch Rate	25.4	7.2	27.5	13.1	13.1
Not Traced Rate	6.7	2.5	11.7	3.2	6.3
Sample Size	4533	8671	1139	1042	4410

Table 5: Nonmatch Rate and Not Traced Rate by Marital Status



Care should be used in interpreting this table since most never married are under 20 years old, most married, seperated/divorced are over 21 years old and most widowed are over 65 years old. Table 9 later in this section examines marital status by age which gives a more complete picture of the effects of marital status on the nonmatched and not traced rates.

The not traced rates are the highest for seperated/divorced persons, about twice as high as any other marital status category. Never married and missing categories had similar high not traced rates. Married and widowed had low not traced rates. Married persons had the lowest nonmatch rates, almost one-half the level of the other categories. Widowed and missing marital status had the next lowest, near the national average, followed by the highest nonmatch rates for never married and seperated/divorced. These comparisons show difficulties in tracing and matching for single, seperated and divorced persons.

SOCIAL SECURITY NUMBER AVAILABLE

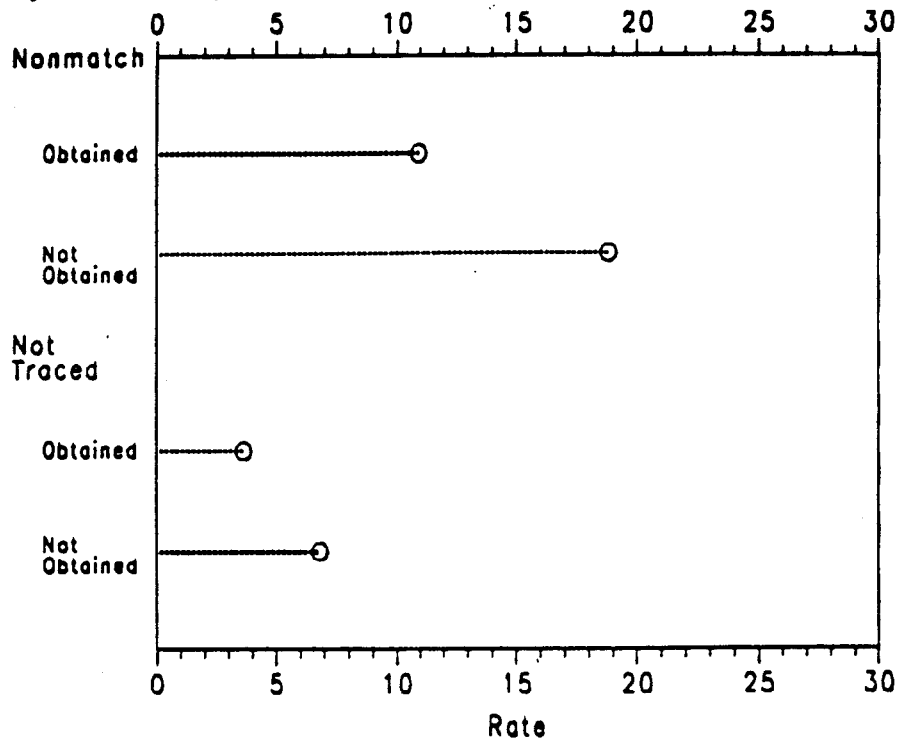
Availability of social security number was examined for its effects on matching and tracing. Table 6 shows the nonmatch, not traced and sample size by social security number obtained or social security number not obtained.

Table 6

Nonmatch Rate, Not Traced Rate and Sample Size
by Availability of Social Security Number

SSN	<u>Obtained</u>	<u>Not Obtained</u>
Nonmatch Rate	10.9	18.8
Not Traced Rate	3.6	6.8
Sample size	12287	7507

Table 6: Nonmatch Rate and Not Traced Rate
by Availability of Social Security Number



Lower nonmatch and not traced rates were obtained if a social security number was given in the interview. The reason for the lower rates is that the second step of this study used the IRS/IMF to get a new address for persons not originally matched to the census. If a social security number was not available then the person could not be searched in the IRS/IMF. However some persons without an obtained social security number were matched during the IRS/IMF look up since they may have lived in the same household as someone who gave their social security number in the 1977 interview.

EDUCATION

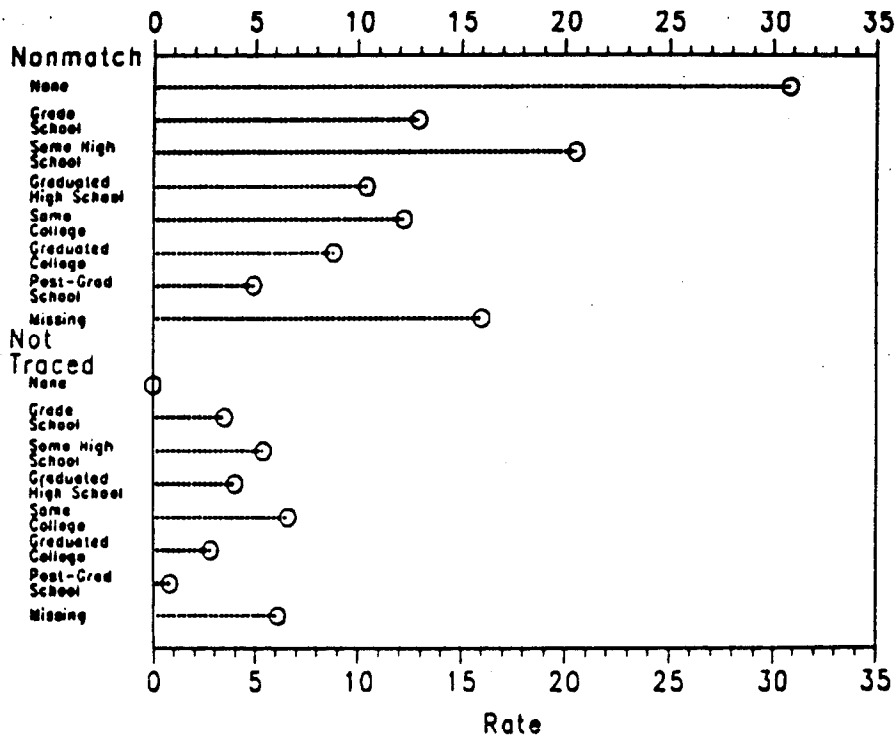
Education was examined using eight categories listed in Table 7. The categories for some high school and below are highly correlated with the younger age groups since they are too young to have finished high school. The nonmatch, not traced and sample sizes by education are shown in Table 7.

Table 7

Nonmatch Rate, Not Traced Rate and Sample Size
by Education Level

EDUCATION	None	Grade School	Some Graduated			Some Graduated College	College	Post-Grad.	
			High School	High School	High School			School	Missing
Nonmatch Rate	30.8	12.9	20.5	10.4	12.2	8.8	4.9	16	
Not Traced Rate	0.0	3.5	5.4	4.0	6.6	2.8	0.8	6.1	
Sample Size	78	2420	3576	4526	2680	997	745	4772	

Table 7: Nonmatch Rate and Not Traced Rate
by Education Level



The nonmatch rates are generally lower with increased education, except for persons who started and did not finish high school or college. This effect could be due to the mobility of these persons. At the time of the census, the high school and college who did not finish in 1977 may have moved to new jobs or other schooling which would be harder to match. The not traced rates show a similar trend as the nonmatch rates, decreasing rates for higher education. The grade school educated and no education are exceptions to the rule. This may be caused by the least educated have lower mobility and hence easier to trace.

CROSS TABULATIONS

Although the nonmatch and not traced rates are important for the variables examined, some important features may be hidden in the tables. For example, are grade school educated nonmatch rates the same for children who have only reached those grades or is it also true for adults with only a grade school or less education? Therefore three cross classified tables will be examined for their nonmatched rates and not traced rates to see if further insights are obtained for some groups. The tables examined are: age by race by sex; marital status by age; and education by age. Because of the small sample sizes some collapsing of categories was necessary to produce stable estimates.

AGE BY RACE BY SEX

Table 8 examines the nonmatch, not traced and sample size for age by race by sex. In order to keep the sample sizes at a reasonable level the age categories 25-34 and 35-44 were combined to 25-44, and 45-54 and 55-64 were combined to 45-64. The race categories used were collapsed to white and nonwhite.

Table 8

Nonmatch Rate, Not Traced Rate and Sample Size
by Age-Race-Sex

Nonwhite Females

	<u>3-17</u>	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	23.3	27.6	22.9	18.5	10.5
Not Traced Rate	11.2	6.0	11.4	6.7	6.6
Sample Size	534	199	417	267	136

Nonwhite Males

	<u>3-17</u>	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	19.4	38.1	24.9	16.5	10.0
Not Traced Rate	16.4	9.9	3.6	7.5	6.2
Sample Size	542	227	291	200	97

White Females

	<u>3-17</u>	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	13.4	22.9	10.6	8.9	10.5
Not Traced Rate	5.6	10.4	4.4	1.8	3.0
Sample Size	1779	1152	2372	1861	1318

White Males

	<u>3-17</u>	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	10.6	28.7	11.3	4.2	9.9
Not Traced Rate	2.6	4.7	6.0	0.0	1.2
Sample Size	1946	1066	2397	1616	827

Table 8: Nonmatch Rate by Age, Race and Sex

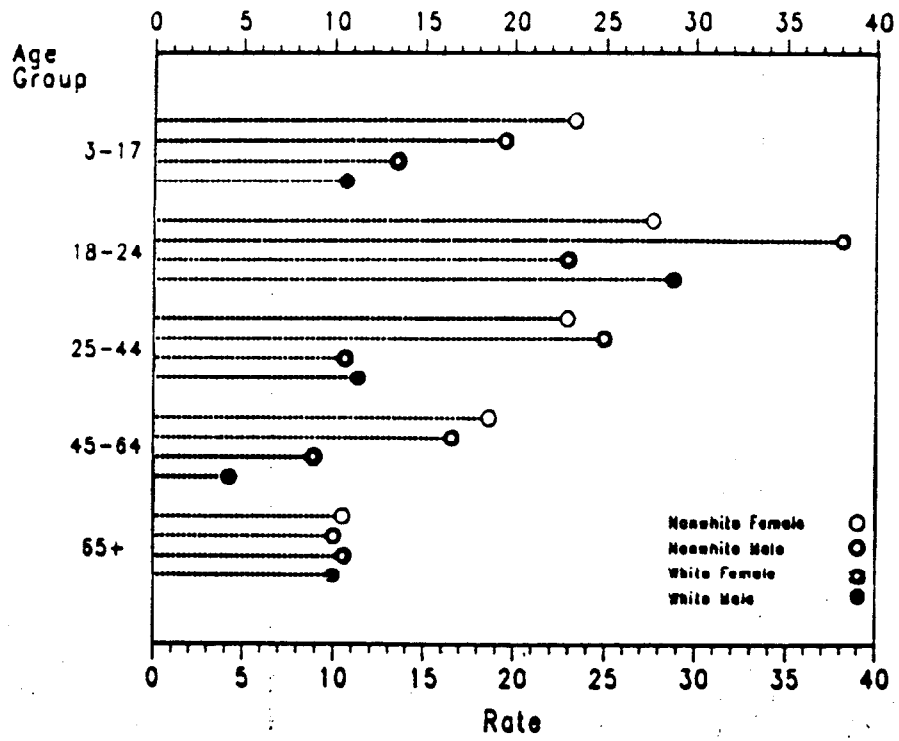
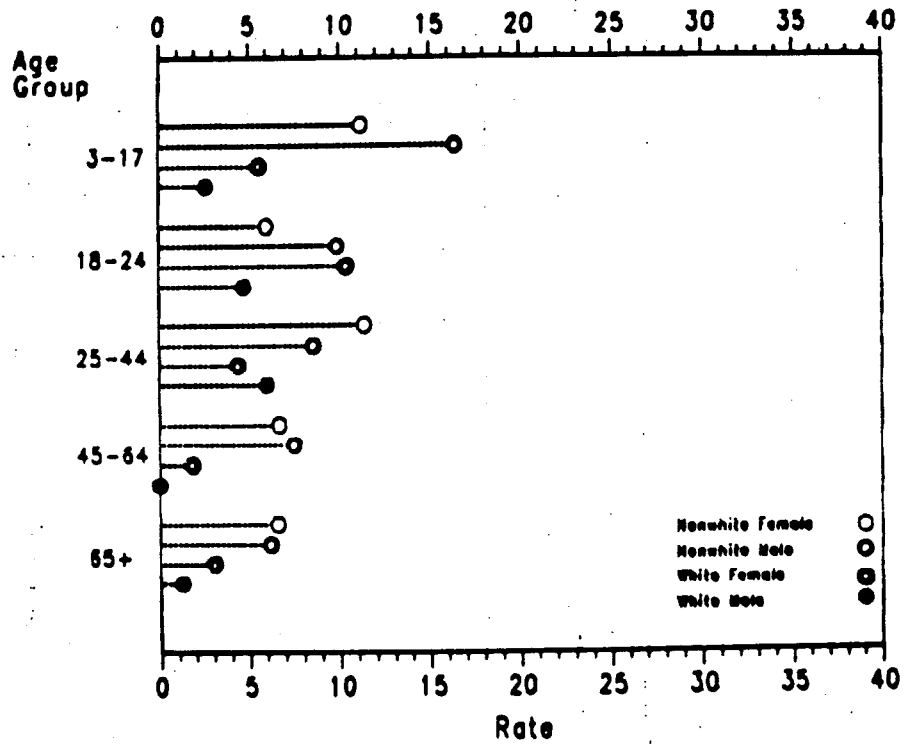


Table 8: Not Traced Rate by Age, Race and Sex.



There are many observations to be made from these tables. For the nonwhites, there is a decline in the nonmatch rates from 18-24 to 65+. This is not true for whites, where the 65+ has higher nonmatch rates than 45-64. Whites and nonwhites, males and females for 65+ have similar nonmatch rates. Perhaps this is due to higher mobility of white elders. There is always a lower nonmatch rate for 3-17 than for 18-24. For males 18-24 there is a higher nonmatch rates than for females 18-24. This coincides with the demographic analysis results.

One of the peculiarities observed are the lower nonmatch rates for the 45-64 white males versus white females.

The not traced rates are higher for nonwhite males and females than white males and females except for white females 18-24. Also puzzling are the high not traced rates for nonwhite 3-17 (males and females) and for nonwhite females 25-44. Evidently our inability to trace the nonwhite females 25-44 is why we failed to trace their children as well. For the age group 3-17, white females have higher not traced rates than white males, while nonwhite males have higher not traced rates than nonwhite females. The high not traced rates for nonwhite males 18-24 and all age groups was somewhat expected but still poses serious problems for this methodology.

MARITAL STATUS BY AGE

Table 9 examines the nonmatch, not traced and sample sizes for marital status by age. Marital Status is collapsed into three categories: divorced and seperated, married and widowed, and never married. Since almost everyone 3-17 was never married, this age group is not shown. Otherwise the age groups are the same as in table 9 on age-race-sex, which are 18-24, 25-44, 45-64, 65+. When examining these tables, remember that the marital status was recorded in

1977 and the age is for 1980. Because of the small sample size of divorced/separated for ages 18-24 and 65+, the nonmatch and not traced rates have very large variances and unstable point estimates.

Table 9

Nonmatch Rate, Not Traced Rate and Sample Size for Marital Status by Age

Divorced, Separated

	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	55.4	31.0	25.1	8.3
Not Traced Rate	16.4	16.4	7.1	6.1
Sample Size	61	569	409	99

Married, Widowed

	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	11.5	9.1	5.4	9.3
Not Traced Rate	4.0	3.9	1.0	2.2
Sample Size	321	3883	3377	2115

Never Married

	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	28.8	19.5	9.2	21.0
Not Traced Rate	8.1	7.4	1.8	6.3
Sample Size	2229	1031	168	158

Table 9: Nonmatch Rate for Marital Status by Age

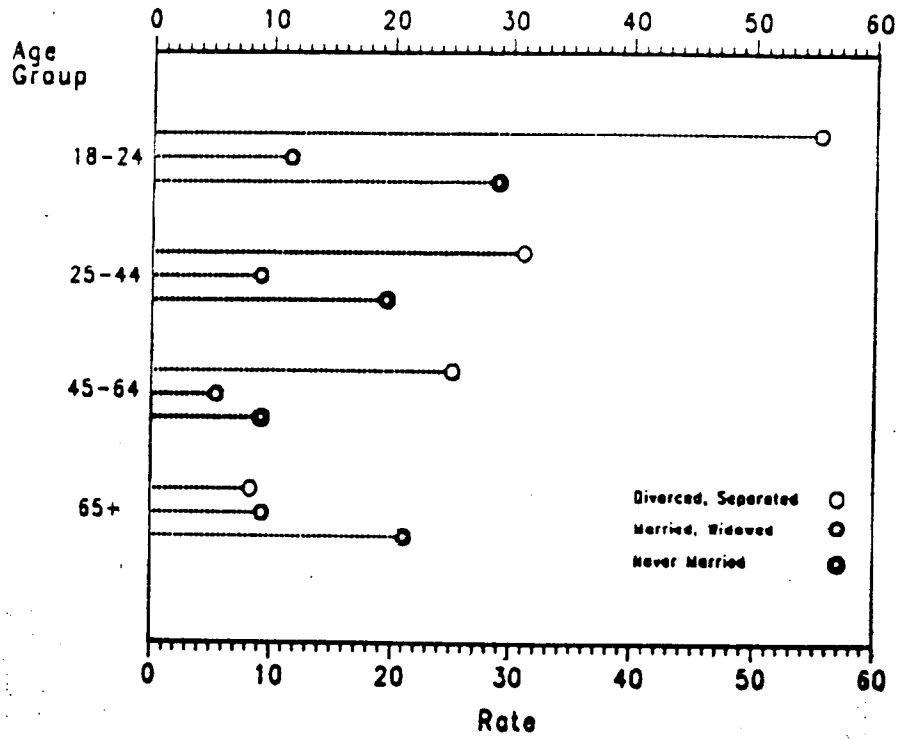
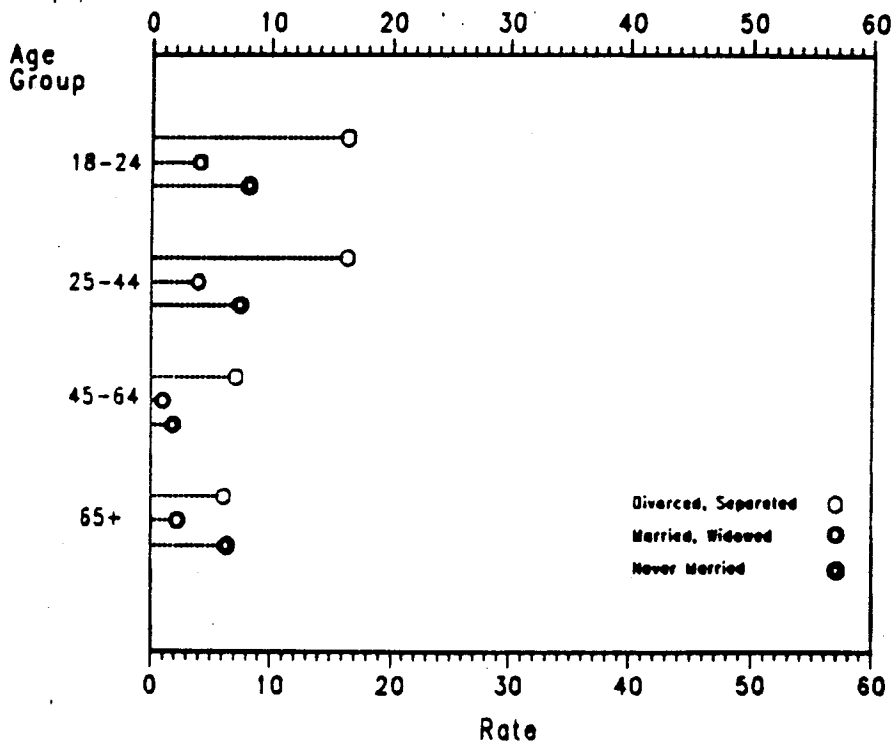


Table 9: Not Traced Rate for Marital Status by Age



The married/widowed group has the lowest nonmatch and not traced rates for all age groups. This is about half the nonmatched or not traced rates for all other marital status groups except divorced/separated 65+. The divorced/separated 65+ group has a slightly lower nonmatch rate than married/widowed. The divorced/separated and never married show similar nonmatch and not traced rates.

The high nonmatch rate for the age group 18-24 seen in table 2 is partly due to their being never married. The higher nonmatch and not traced rate for 65+ than for 45-64 holds for married/widowed and never married which comprise the majority of these age groups.

EDUCATION BY AGE

Table 10 examines the nonmatch and not traced rates for education by age. Education is coded as: not completed high school, graduated from high school, some college, and graduated from college and post graduate college training. The age group 3-17 is dropped from this analysis since almost everyone age 3-17 would be in the not completed high school category. The college and post college schooling variable is not shown for 18-24 since the sample size was very small (18) and may be misleading. The remaining age groups were used as presented for the age-race-sex and marital status by age table (25-44, 45-64, 65+).

Table 10

Nonmatch Rate, Not Traced Rate and Sample Size
for Education by Age

Not Completed High School

	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	30.4	20.0	10.5	11.3
Not Traced Rate	7.0	8.3	1.2	3.1
Sample Size	1606	1027	1446	1384

Graduated High School

	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	25.1	11.9	5.0	6.5
Not Traced Rate	10.1	4.5	2.5	0.0
Sample Size	495	2070	1426	535

Some College

	<u>18-24</u>	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	17.3	12.5	8.0	10.6
Not Traced Rate	7.9	8.3	2.3	4.8
Sample Size	524	1316	571	269

Completed College and Post Graduate College

	<u>25-44</u>	<u>45-64</u>	<u>65+</u>
Nonmatch Rate	6.1	3.4	7.5
Not Traced Rate	3.2	0.0	0.0
Sample Size	1062	500	160

Table 10: Nonmatch Rate for Education by Age

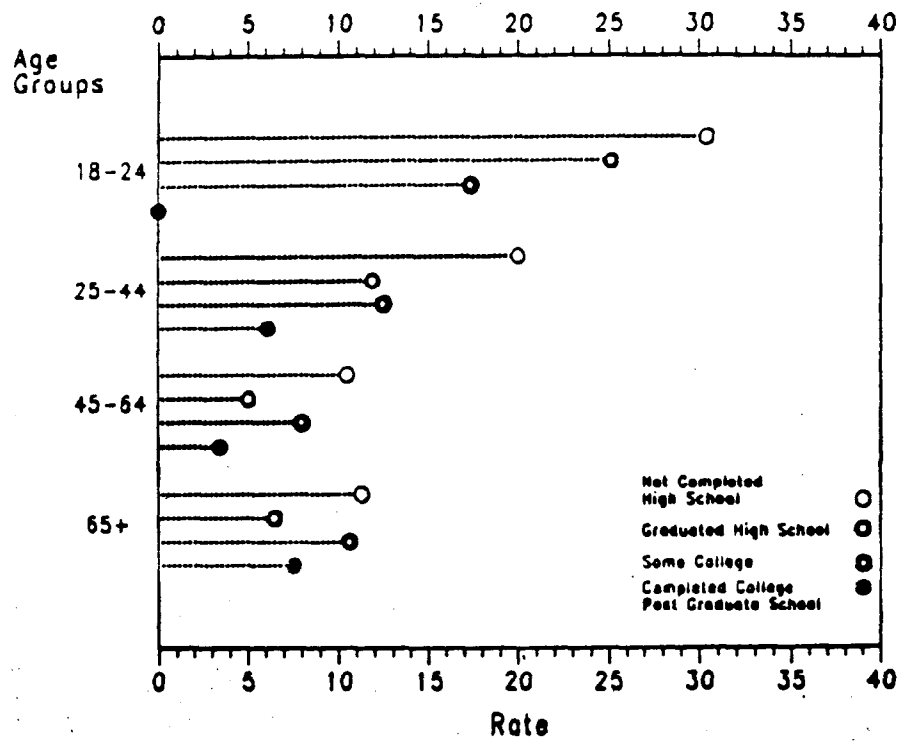
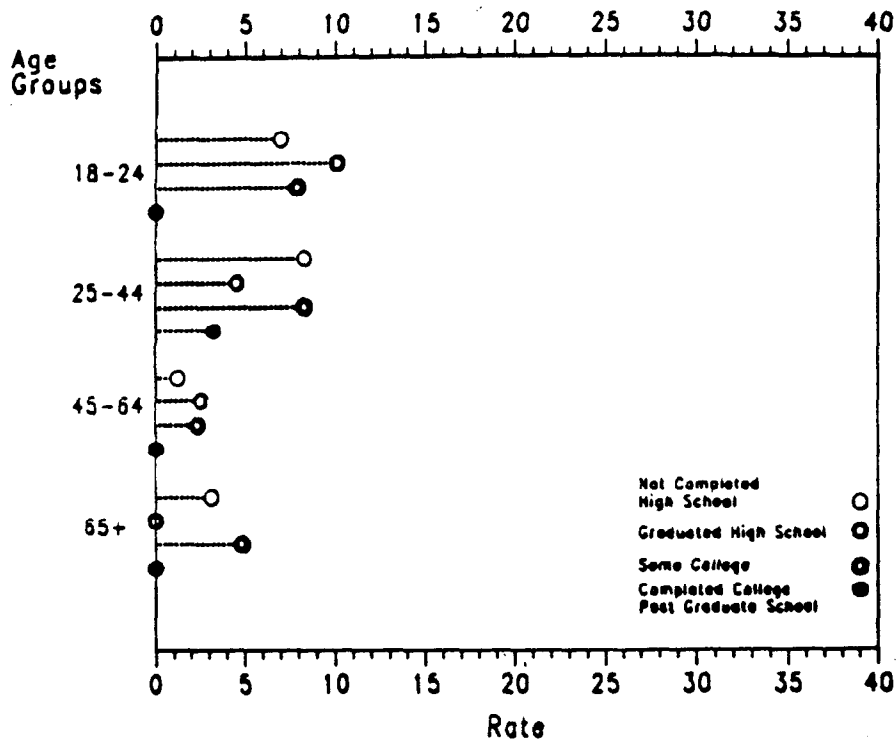


Table 10: Not Traced Rate for Education by Age



In general we would expect smaller nonmatch and not traced rates for higher education and the older the person (although for age 65+ this is not true from just looking at age). Some exceptions may be possible for those who start college and do not finish. However the 18-24 age group may be difficult to interpret since they may cut across the education variable. The 18-20 years old probably did not finish high school in 1977. Those 21-24 probably would have finished high school and may have gone on to college, but would not have finished college by 1980. The 18-24 age group show high nonmatch and not traced rates for all education levels. High school graduates have the highest not traced rate, but those not completed high school have a higher nonmatch rate for this age group. Although the nonmatch rate for some college 18-24 is still high overall, it is lower than other education variables in this age group. These people would have probably been in college in 1980. The lower nonmatch rate may be due also to their incorrect inconclusion on their parents census form.

The college graduate category had the lowest nonmatch and not traced rates except for the nonmatched rate for 65+ which was slightly higher than the high school graduate (the not traced were the same for 65+ at .0). The 25-44 and 45-64 age categories vary for the three other education variables. The lowest not traced rates for 25-44 is the high school graduate but for 45-64 is for not completed high school. The high school graduates have the lowest nonmatched rates for 25-44 and 45-64. For 65+ the nonmatch rates are higher for every education level than the 45-64 age groups which in turn the 45-64 are lower than 25-44 across all education categories. The 65+ high nonmatch rate may be partly attributed to their lack of a high school degree, but the consistently higher nonmatch rates for all education levels implies that it is partly their age as well.

III. Sampling

This section will describe the sampling design and the subsampling done in telephone and field followup. The sample used in this study is the eighth rotation group from the March 1977 CPS. See U.S. Bureau of the Census (1978) for a complete description of the CPS design. Only those cases that were data defined (i.e., had name and some characteristics) were chosen from the eighth rotation group so that only matchable person were included. That is, person with insufficient information were not included in this analysis since the person could never be matched or not matched with certainty. No adjustments were made to any results reported in this paper for the exclusion of this group.

In order to cut cost, subsampling was done on the telephone and field followups. For nonresponse and postmaster returns only one-half were sampled in telephone followup. For mail returns that needed additional information, no subsampling was done. For incomplete telephone interviews, the sampled cases were split into black and nonblack and were further subsampled. Two-thirds of the blacks were subsampled and sent to field followup; one-fifth of the other races were subsampled and sent to field followup. The different subsampling rates helped insure sufficient sample sizes for inference of the black population. Since the subsampling for field followup was stratified by black and nonblack, the proportion nonmatched and not traced and their standard errors were calculated on this breakdown. The results are shown in Table 11 for proportion nonmatched and in Table 12 for proportion not traced. These tables are also used in Section VII for comparisons with similar type studies.

Table 11

Proportion Nonmatch
(Standard Errors in Parenthesis)

	<u>3-17</u>	<u>18-24</u>	<u>25-34</u>	<u>35-44</u>	<u>45-54</u>	<u>55-64</u>	<u>65+</u>
Nonblack Males	.109 (.016)	.286 (.024)	.142 (.019)	.100 (.020)	.052 (.009)	.032 (.008)	.101 (.017)
Black Males	.234 (.031)	.463 (.044)	.300 (.051)	.020 (.020)	.270 (.055)	.144 (.046)	.079 (.036)
Nonblack Females	.135 (.017)	.229 (.024)	.113 (.016)	.106 (.021)	.098 (.018)	.093 (.019)	.107 (.017)
Black Females	.259 (.034)	.290 (.046)	.311 (.043)	.217 (.046)	.237 (.045)	.079 (.031)	.093 (.032)

Table 12

Proportion Not Traced
(Standard Errors in Parenthesis)

	<u>3-17</u>	<u>18-24</u>	<u>25-34</u>	<u>35-44</u>	<u>45-54</u>	<u>55-64</u>	<u>65+</u>
Nonblack Males	.040 (.012)	.046 (.017)	.079 (.018)	.042 (.017)	.000 (.000)	.000 (.000)	.012 (.011)
Black Males	.130 (.026)	.125 (.032)	.098 (.036)	.000 (.000)	.156 (.046)	.041 (.029)	.073 (.034)
Nonblack Females	.065 (.015)	.101 (.021)	.034 (.013)	.061 (.019)	.014 (.011)	.021 (.013)	.030 (.012)
Black Females	.118 (.027)	.082 (.030)	.175 (.036)	.050 (.027)	.152 (.042)	.030 (.022)	.000 (.000)

The formula for the estimate of the proportion nonmatch and the proportion not traced are given in equation 1 and 3 respectively, and the formula for the variance are given in equation 2 and 4 respectively. These formulas were derived in the IRS/Census Direct Match Study. For completeness the necessary notation is included here.

$$p = w_1 p_1 + w_2 p_2 + w_3 p_3 \quad (1)$$

$$v(p) = \frac{p_1 q_1}{n} + w_2^2 (k_2 - 1) \frac{p_2 q_2}{n_2} + w_3^2 k_2 (k_3 - 1) \frac{p_3 q_3}{n_3} \quad (2)$$

where

p = proportion nonmatched (or not traced)

$v(p)$ = variance of the proportion nonmatched (or not traced)

p_i = proportion nonmatched (or not traced) in stratum i
 $i=1, 2, \text{ or } 3$

q = $1-p$

q_i = $1-p_i$, $i=1, 2, \text{ or } 3$

n_i = number of persons in stratum i , $i=1, 2, \text{ or } 3$

n = $n_1 + n_2 + n_3$ = total sample size

w_i = $\frac{n_i}{n}$, $i=1, 2, \text{ or } 3$

k_i = the inverse of the subsampling rate $i=2 \text{ or } 3$

k_2 = 2

k_3 = 3/2 for Blacks
= 5 for Nonblacks

For the not traced rated, $p_1 = p_2 = 0$ since the not traced category is only defined after field followup.

Consequently equations 1 and 2 simplify to

$$p = w_3 p_3 \quad (3)$$

$$v(p) = \frac{p_3 q_3}{n} + w_3^2 k_2 (k_3 - 1) \frac{p_3 q_3}{n_3} \quad (4)$$

The estimates of the proportion nonmatched and not traced and their standard errors were calculated within each age, race, sex cell.

IV. Tracing and Matching

In this section the matching and tracing are discussed. The sample size and percents reported in this section are the same as reported in table 1, but different from those reported in the tables in section II. The differences are due to the out of scope cases which are not included in section II but are included in table 1 and the discussion in this section. Table 1 gives the breakdowns of the tracing and matching determination for the five phases of the CPS-Census Retrospective Study. This table includes all weighting due to subsampling for telephone and field followup. The five phases of the study will be examined separately.

The first stage was the match to the Census. This was conducted in spring of 1982 with most being matched to the census. A few people were coded as deceased when the spouse is matched to a census questionnaire and is widowed. The total resolved with the census search at the 1977 address was 45.6% of the total sample. This seems a little low when compared to the figures from the long form which showed that 53.6% lived in the same house 5 years ago. Noninterviews and nonmatches were not allowed from the census search.

The second stage was the IRS/IMF search for a new address of all unresolved persons. When a new address was found, the persons were searched in the census. Only matches were allowed during the IRS/IMF search. Of the total sample, 17.4% were resolved and matched to the census. Of the remaining cases (Total minus resolved during census stage) 31.9% were resolved. During the IRS/IMF search, new

addresses could be found only if a social security number was available from the 1977 interview. However in section II over 30% of the resolved cases from the IRS/IMF search were from persons without social security numbers. This is because the sample is a household sample and if a social security number is available for only one member of the household then a new address could be found and other member of the household without social security numbers would have been matched during the IRS/IMF search. A slightly higher match rate was obtained if a social security number was available during the IRS/IMF search. (See Table 6 in Section II).

The next stage was mail followup which used the latest available address, the 1977 address or the IRS/IMF address. Only 7% of the cases were resolved during mail followup which is 19% of the unresolved cases before mail followup. Mail followup was conducted in the fall of 1982, 5 years after the original interview and 2 years after census day. Although the small response rates were disappointing, they were probably caused by a large number of moves in the five year span. Also mail followup response is always low in comparison to field followup, but the cost is minimal. So all resolved cases from the mail followup phase may be viewed as a bonus.

The fourth stage was telephone followup. In order to cut costs, the telephone interviewing was broken into five categories according to the nonresponses from mail followup. The five unresolved categories were: postmaster return, whole household unmatched; postmaster return, partially matched household; no response, whole household unmatched; no response partially matched household; and mail returned, needing additional information. During the telephone followup 12.4% of the total sample was resolved, that is, 41.1% of the total cases left to complete. The telephoning was conducted in the spring of 1983. An

important observation is the large percentage of noninterviews from the telephone followup. The results shows that 76.3% of all noninterviews occurred during telephone followup. Table 13 shows the results for the telephone followup broken down by the three main types of noninterview; refusals, unable to geocode, and unresolved (no unique address). Possible matches were also classified as noninterviews but only one person was coded as a possible match. It appears that insufficient care was taken during telephone followup. The problem may have been that the questionnaire did not clearly define the address, there was inadequate training for the interviewer, or perhaps there was difficulty in using telephone interviewing for tracing. If further tracing and matching studies are done, telephone interviewing should be more carefully examined to obtain correct and geocodeable addresses.

Table 13

Type of Noninterview for Telephone Followup

	<u>From Telephone</u>	<u>Percent of Total Noninterviews</u>
Refusals	86	9.9
Unable to geocode	444	51.1
Unresolved	133	15.3
Total	663	76.3

The telephone interviewing was conducted by taking the 1977 CPS phone number and trying to contact the person. If that failed, then the interviewer called the directory assistance at any of the possible addresses, especially at the last known address, IRS/IMF address or address from mail followup.

The fifth and final phase of the CPS-Census Retrospective Study was field followup. Field followup resolved 12.9% of the total sample or 73.2% of the remaining cases. The cases not resolved were assigned the code tracing failed which amounted to 4.7% of the total sample. The field interviewing was conducted during August 1983, three years after census day and over six years after the original CPS interview. The field followup began searching at the latest available address in order to find the person. The searching involved looking at the latest address, asking neighbors if they knew the sampled person's whereabouts, and searching the local telephone directories for a listing of the person. The interviewers used their own initiative to locate the sample person for example, they asked at local bars, departments of motor vehicles and police departments. If the person was not found then the previous address(es) were searched.

Similar to telephone followup, further subsampling was conducted during field followup to cut costs. The same five categories used in telephone followup were used in field followup but they were further divided into black and nonblack with different subsampling rates for each to insure sufficient sample size for blacks. Within the five categories, two out of three blacks were subsampled. Field and telephone subsampling imply that blacks sent to field followup were given a weight of 3.0 for all nonmail returns and postmaster returns and a weight of 1.5 for mail returns needing additional information. The subsampling rate for nonblacks was one out of five during field followup. Field and telephone subsampling imply that nonblacks sent to field followup were given a weight of 10.0 for all nonmail returns and postmaster returns and a weight of 5.0 for mail returns needing additional information. There were lower weights for the mail returns needing additional information because no

subsampling was done in the telephone followup for this category.

V. **Weighting and Nonresponse Adjustments**

Two different weights could be used, the probability of selection and subsampling for telephone and field followups. The probability of selection was not used in this analysis because when the primary sampling unit (PSU) and segment were matched to the 1977 March CPS, not all PSU's matched. It is unknown why but it may be keying errors. Consequently all results reported were not weighted by their base weights.

The second weights were applied so that the results account for the subsampling. The nonrespondents from the mail followup were subsampled by one half within the codes H1, H2, H3, and H4 (See Appendix). Mail returns with incomplete information were not subsampled. The cases that were resolved in telephone followup, except the mail returns which needed additional information, were given a weight of two.

The incomplete interviews from telephone followup were further subsampled during the field followup. The subsampling for field followup used the codes H1, H2, H3, and H4 and other (incomplete information mail followup) but used different subsampling rates for Blacks and others. For Blacks, the subsampling rate was two-thirds. The weights for field followup cases for blacks are three ($3/2$ from field x 2 from telephone) for categories H1, H2, H3, H4 and one and one-half for other (incomplete mail followup). For the other races the subsampling was one-fifth. Therefore the weights for field followup for other races are ten (5 (from field) x 2 (from telephone) for categories H1, H2, H3 and H4 and five for other (incomplete mail followup). The codes H1, H2,

H3, H4 and other for blacks were coded J1, J2, J3, J4 and J6 and for other races were coded K1, K2, K3, K4 and K6.

VI. Costs

Besides evaluating the quality of the data and the estimates of the nonmatched rate and not traced rates, this study needs to examine the costs of performing the operations. Even if good quality data was obtained, if the costs are prohibitive, the study will not be done on a large scale.

Table 14 Cost by Operations shows the costs for four operations: geocoding, keying, telephone, and field. There is some error in these figures since the IRS/Census Direct Match Study and the CPS-Census Retrospective Study occurred at the same time and some costs were incorrectly recorded in the other project. Other costs not show here are not easily estimated. Most operations were done in the Data Preparation Division in Jeffersonville. The personnel there are highly trained and experienced which may help reduce the costs.

Table 14

Costs by Operations

	<u>Total Cost</u>	<u>Persons</u>	<u>Forms</u>	<u>Cost Per Unit</u>	<u>Hours</u>	<u>Hours Per Unit</u>
Geocoding	24912	--	6794	3.70	3363	.50
Keying	2815	19068	--	.15	440	.023
Telephone	3053	2947	--	4.45	2018	.70
Field	9700	497	--	19.50	--	

Geocoding is the process of converting the address to census geography. For vague and rural addresses, geocoding is difficult and error prone operation. The cost per unit (form) is shown at \$3.70. Estimates of the cost (including

overhead) for geocoding during the 1977 address and 1979 IRS/IMF address were \$10.00 to \$15.00, so the cost per unit for geocoding may be too low.

Keying converts the CPS information to a computer file. The costs here are very small. The cost is 15 cents per person and the timing was about 2 minutes per person.

Telephone followup was the fourth stage in the tracing operations. The costs listed do not include the telephone bills or equipment. The costs were \$4.45 per person and the tracing took close to 45 minutes per person. Only about one half of the persons sent to telephone followup were resolved so the cost per resolved case is about twice as high as the cost listed, estimated to be \$9.60 per person resolved.

Field followup was the fifth and last stage in the tracing operation. The costs were estimated at \$19.50 per person. No hour estimates were available for the field followup.

VII. Comparisons

There were two other studies, 1980 PEP and IRS/Census Direct Match Study, that produced nonmatch rates for the 1980 Census which are shown in table 15. The IRS/Census Direct Match Study also involved tracing. The not traced rates are compared in table 16. Two other studies that involved tracing and matching are the 1960 US Reverse Record Check and the 1976 Canadian Reverse Record Check. These studies use the previous census augmented by births, immigrants and persons missed in test census as the sampling frame. The Canadians have conducted reverse record checks since 1961 and their censuses are five years apart. The results from these two studies are not directly comparable to this study since their coverage is for a

different group of people and they were conducted in a different time, but may give indications of potential improvements.

Table 17 shows the nonmatch rate, not traced rate and tracing period for these four studies.

Table 15 shows the proportion nonmatched for the CPS-Census, IRS/Census and the 1980 PEP Studies for race by sex by age. The 1980 PEP was the April 1980 CPS sample. The results show that the CPS-Census nonmatch rates are usually twice as high as the PEP. The only instance of the CPS-Census nonmatch rates being lower is for black males 35-44 which only had 50 sample persons and an extremely low proportion nonmatched (.02). The IRS/Census nonmatch rates usually fall between CPS-Census and PEP. For nonblack males, the nonmatch rates for the IRS/Census Match Study are much closer to the CPS-Census Study than for any other race-sex group.

Table 15

Proportion Nonmatched CPS-Census Retrospective Study,
IRS/Census Direct Match Study, and PEP

	<u>3-17</u>	<u>18-24</u>	<u>25-34</u>	<u>35-44</u>	<u>45-54</u>	<u>55-64</u>	<u>65+</u>	
Nonblack	.109	.286	.142	.100	.052	.032	.101	CPS
Males	NA	.201	.128	.060	.049	.039	NA	IRS
	.046*	.092**	.054	.039	.031	.026	.025	PEP
Nonblack	.135	.229	.113	.106	.098	.093	.107	CPS
Females	NA	.168	.058	.059	.057	.042	NA	IRS
	.045*	.075**	.042	.034	.023	.023	.027	PEP
Black	.234	.463	.300	.020	.270	.144	.079	CPS
Males	NA	.189	.228	.125	.144	.063	NA	IRS
	.118*	.131**	.117	.104	.096	.099	.052	PEP
Black	.259	.290	.311	.217	.237	.079	.093	CPS
Females	NA	.177	.098	.032	.100	.032	NA	IRS
	.110*	.123**	.098	.091	.065	.060	.057	PEP

Note: The numbers in the table are from the survey listed directly across in the right hand column.

NA - not applicable

* PEP age group reported is 5-19

** PEP age group reported is 20-24

These comparisons clearly show the much higher nonmatch rates for the CPS-Census Study versus the 1980 PEP for almost all age-race-sex groups examined. Some possible explanations for the higher nonmatch rates for the CPS-Census Study are presented. The CPS-Census sample may more closely reflect the true nonmatch rates than the PEP. Some of the extreme nonmatch rates (eg .463 for black males 18-24) may be due to sampling variability, but the overall trends may be better estimated from the CPS-Census Study than the PEP. The assumption of independence between the Census and the CPS-Census Study may be more nearly met than in the PEP. Since the CPS-Census interviews were conducted two to three years after the census, there may be recall bias of the respondent's census day address. This can be especially troublesome for persons moving around census

day. The PEP's original interview was taken in April or August 1980 and the followup in February 1981. The time between interview and census day is less than a year and should have lower recall bias of the respondent's census day address. Also, the interviewers may have found a person with the same name and mistakenly listed the person as an interview. Since the person's characteristics would not match, the person would be recorded as a nonmatch. There is some anecdotal evidence of this occurring. It is not possible to measure the effects of these errors on the nonmatch rate, but clearly these effects would be much smaller the closer the study is to census day.

Table 16

Proportion Not Traced CPS-Census Retrospective Study
and IRS/Census Direct Match Study

	<u>3-17</u>	<u>18-24</u>	<u>25-34</u>	<u>35-44</u>	<u>45-54</u>	<u>55-64</u>	<u>65+</u>	
Nonblack	.040	.046	.079	.042	.000	.000	.012	CPS
Males	NA	.046	.061	.027	.021	.000	NA	IRS
Nonblack	.065	.101	.034	.061	.014	.021	.030	CPS
Females	NA	.051	.040	.014	.002	.000	NA	IRS
Black	.130	.125	.098	.000	.156	.041	.073	CPS
Males	NA	.034	.071	.166	.000	.000	NA	IRS
Black	.118	.082	.175	.050	.152	.030	.000	CPS
Females	NA	.000	.039	.161	.000	.000	NA	IRS

Note: The numbers in the table are from the survey listed directly across in the right hand column.

NA - not applicable

Table 16 shows the proportion not traced for the CPS-Census and IRS/Census Studies. Nonblack males show similar not traced rates for both studies. Nonblack females have higher not traced rates for ages 18-24 and 35-44 for the CPS-Census study. The IRS/Census Study has not traced rates that are extremely low for age groups 45-54 and 55-64 for all race-sex groups while the CPS-Census not traced rates are low for nonblacks but high for 45-54 and moderate for 55-64 for blacks. The IRS/Census Study had high proportion not traced for blacks 35-44 (male and female) while the CPS-Census Study had moderate to low for these groups. The CPS-Census Study had high not traced rates for young blacks male and female (15-24 and 25-34) while the IRS-Census Study had moderate to low rates except for black males 25-34 which was about equal to the CPS-Census Study's value. In general our tracing ability for blacks is not encouraging.

Table 17

Comparison of Proportion
Nonmatched and Not Traced
with Other Studies

	<u>Nonmatched</u>	<u>Not Traced</u>	<u>Time from sample to census (in years)</u>
CPS-Census	.140	.048	3
IRS/Census ¹	.126	.031	0
1976 Canadian R R C	.028 ²	.048	5
1960 US R R C	.037 ³	.099 ⁴	10

¹ Coverage for ages 18-64 only

² Base is cases in scope minus not traced

³ Estimate 3 which approximately is imputation done in CPS-Census study

⁴ Not traced for 1950 census only, higher not traced rates were recorded for births (.149) and missed in 150 census (.185), but aliens were listed as zero due to special circumstances.

Table 17 shows the proportion nonmatched and not traced for the CPS-Census Retrospective Study, IRS/Census Direct Match Study, 1976 Canadian Reverse Record Check and the 1960 US Reverse Record Check (RRC). The time frame from the sample to the census and matching are also shown in this table. For both reverse record checks, the matching would begin within a year of the census. The CPS-Census and IRS/Census studies were started about 2 years after the census. The proportion nonmatched for the CPS-Census Study are extremely high compared to the other reverse record checks, but comparable to the IRS/Census study. The CPS-Census proportion not traced are very comparable to all the other studies listed especially accounting for the time of sample to census. The CPS-Census study results have the same proportion not traced as the 1976 Canadian Reverse Record Check. The much higher proportion not traced for the 1960 U.S. Reverse Record Check may be the extremely long time span between the sample and the census. If a full reverse record check was to be conducted for the 1990 census, the proportion not traced would most likely be at least .05 to .10 and perhaps even higher.

VIII. Conclusion

The CPS-Census Retrospective Study is the only test of matching and tracing before the 1990 Census. The results show a much higher nonmatch rate than the 1980 PEP. The not traced rates are comparable to other major studies using tracing, about 5% of the sample not traced. Results from the Forward Trace Study will provide further evidence on the ability to reduce the not traced rate for a coverage measurement survey.

The high nonmatch rates relative to the 1980 PEP is an area of major concern. Some possible reasons for the higher nonmatch rates are: greater independence between the Census and the CPS-Census Study than the Census and the PEP, recall bias from the '83 interview on '80 census address, and finding people with same name but different characteristics. If a study like the CPS-Census Study was performed in 1990, the followup interview would occur within a year of census day which would reduce recall bias and lowering the nonmatch rate. The tracing would need to confirm the respondent's address at time of the original sample to make certain the correct person was found. Clearly reducing the time from the original sample to the census would reduce the tracing workload since there would be fewer movers.

The nonmatch rates generally followed the expected patterns, persons 18-24 had the highest nonmatch rate, married and widowed persons had low nonmatch rates while divorced and separated had high nonmatch rates, people with higher education had lower nonmatch rates (except those who started but did not finish high school or college), and minorities had higher nonmatch rates than whites. Some of the unexpected results were the higher nonmatch rate for females than for males and the slightly higher nonmatch rate for age 65+ over the age groups 45-54 and 55-64. The not traced rates followed the same pattern as the nonmatch

rates. The unusual result was for the race and sex categories which had missing data, they had high nonmatch rates but low not traced rates.

The traced rates were moderate (good) given a six year lapse in searching. Actually only those who moved from 1977 to 1980 and those who filed tax return from these 1980 census addresses only had a three year time lapse. So movers and nonmatches have a different tracing time than those matched in steps 1 and 2 (matched at 1977 address and matched at IRS/IMF address). After processing it was discovered that some followup interviewers did not have the most up to date address. This probably caused a larger not traced rate. Conversely if a person with a similar name was contacted and thought to be the correct person, this would have caused a lower not traced rate and a higher nonmatch rate.

The noninterview rate for this study was about average at around 4.5%. This does not include the noninterviews in the original March 1977 sample. These cases were not chosen for this study because they could not be traced or matched. About 3 to 5% of a CPS sample are noninterviews also. Since the not traced rate of 4.5% also needs to be imputed into a match or nonmatch category, the not traced and the noninterviews together account for all the nonresponse adjustment needed to produce a nonmatch rate. Therefore the nonresponse adjustment amounts to around 12 to 15% of the sample, a significant amount of nonresponse adjustment.

A difficulty of this study is the high nonresponse from telephone followup. Over 75% of the noninterviews occurred in telephone followup. Clearly if we want to minimize all forms of nonresponse, telephoning as carried out in this

study is not the answer. One possibility is to send all refusals to field followup rather than using a nonresponse adjustment. This would take advantage of the low costs of telephoning and minimize its disadvantages.

Finally this study does not cover the entire U.S. population. Since the sample was taken three years before the census, there is no coverage of births or immigrants after March 1977. This needs to be changed if the study is to be used for coverage evaluation.

References

Childers, Danny "CPS/Census Retrospective Match Study Interim Report" PERM No. 48.

Childers, Danny and Howard Hogan "Census Experimental Match Studies" Proceedings of the ASA Section on Survey Research Methods (1983) p 173-176.

Childers, Danny and Howard Hogan, "The IRS/Census Direct Match Study Final Report" Statistical Research Division Report Series No. Census/SRD/RR - 84/11.

U.S. Bureau of the Census, (1978) "The Current Population Survey: Design and Methodology" Technical Paper 40.

U.S. Bureau of the Census, (1964) "Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: Record Check Studies of Population Coverage." Series ER60, No. 2, Washington, D.C.

Statistics Canada, "Coverage Error in the 1976 Census of Population and Housing", 1976 Census Quality of Data Series, Catalogue 99-840, March 1980.

Appendix

Final Match Codes

Final match codes are assigned to each sample person in the CPS household. These codes indicate the enumeration status in the 1980 Decennial Census. A description of the final match codes follows.

1. Matched (coded match) - A census questionnaire was located with the CPS sample person listed. This occurred after matching to the address obtained from:

Code Stage

M-1 1977 CPS control card
M-2 1979 IRS/IMF file
M-3 mail followup
M-4 telephone followup
M-5 field followup

2. Not enumerated (coded nonmatch) - A search of census records failed to find the person listed at the reported census day address. The address was reported during:

Code Stage

N-3 mail followup
N-4 telephone followup
N-5 field followup

3. Linked to a close-out case on the census questionnaire (coded nonmatch) - The census questionnaire for the reported census day address for the sample person was a close-out case in census. The census day address for the sample person was reported during:

Code Stage

L-3 mail followup
L-4 telephone followup
L-5 field followup

4. Census questionnaire not on microfilm (coded nonmatch) - The census questionnaire for the reported census day address for the sample person was not on the microfilm. The census day address for the sample person was reported during:

Code Stage

Q-3 mail followup
Q-4 telephone followup
Q-5 field followup

5. Possible match (coded noninterview) - We were not able to convert the "PM" in the initial match status because the name on the census questionnaire was blank or because there was not enough information on the census questionnaire to assign a final match status. (This match code is used only for data defined persons). The code was assigned during the processing of the address obtained from:

Code Stage

P-3 mail followup
P-4 telephone followup
P-5 field followup

6. Refused (coded noninterview) - The respondent refused the interview during:

Code Stage

R-3 mail followup
R-4 telephone followup
R-5 field followup

7. Unable to geocode (coded noninterview) - A census day address was given that could not be geocoded, (i.e., we could not determine an accurate DO and ED). The address resulted from:

Code Stage

G-3 mail followup
G-4 telephone followup
G-5 field followup

8. Unresolved (coded noninterview) - There was no unique address given to geocode. The information resulted from:

Code Stage

U=3 mail followup
U=4 telephone followup
U=5 field followup

9. Tracing failed (coded not traced) - No one could be located who could give any information about the sample person after mail, telephone, and field followup. The code is T=5.

10. Deceased before April 1, 1980 (coded out of scope) - The sample person was reported to have died before census day from:

Code Stage

D=1 IRS/IMF file
D=2 Classified deceased when the spouse is matched to a census questionnaire and is widowed.
D=3 mail followup
D=4 telephone followup
D=5 field followup

11. APO/FPO address (coded out of scope) - The sample person was reported to have been in the military, out of the country on April 1, 1980 from:

Code Stage

S=1 Zip code in IRS/IMF file
S=3 mail followup
S=4 telephone followup
S=5 field followup

12. Emigrated before April 1, 1980 (coded out of scope) - The sample person was reported to have moved out of the country before April 1, 1980 from:

Code Stage

E-3 mail followup
E-4 telephone followup
E-5 field followup

13. Nonsample telephone followup cases (coded subsampling adjustment) - If the respondent did not return the mail followup questionnaire or if the questionnaire was returned by the post office (PMR), a sample of the IRS cases were sent to telephone followup and a sample of them were coded as nonsample IRS cases. The nonsample cases will be given a noninterview adjustment. The final match codes for these cases are as follows:

Code Subsampling Categories

H-1 PMR, whole household unmatched
H-2 PMR, partially matched household
H-3 No response, whole household unmatched
H-4 No response, partially matched household

14. Nonsample field followup cases (coded subsampling adjustment) - if the telephone followup was unsuccessful in locating a telephone number for the sample person or anyone who could give any information about the sample person, the cases were stratified by race of CPS line 1 and subsampled further. The sample cases were sent to field for extensive followup. The nonsample cases were given final match codes as follows:

Code	Subsampling Categories
J=1	Black, H=1
J=2	Black, H=2
J=3	Black, H=3
J=4	Black, H=4
J=6	Black, Other
K=1	Other races, H=1
K=2	Other races, H=2
K=3	Other races, H=3
K=4	Other races, H=4
K=6	Other races, Other

The M code indicates that the sample person was counted in the 1980 Decennial Census. The codes N, L, and Q indicate that the sample person was missed in the census. The codes R, G, U, P and T will require a noninterview adjustment or imputation. The codes D, S, and E will be out-of-scope since they were not eligible to be counted in the census.