

BUREAU OF THE CENSUS
STATISTICAL RESEARCH DIVISION REPORT SERIES
SRD Research Report Number: Census/SRD/RR-86/24

OVERVIEW AND SUMMARY OF THE
FORWARD TRACE STUDY

by

Mary Mulry-Liggin
Statistical Research Division
Bureau of the Census

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, D.C. 20233.

Recommended by: Howard Hogan
Report completed: September, 1986
Report issued: November 11, 1986

Overview and Summary
of the
Forward Trace Study

Mary Mulry-Liggan
Undercount Research Staff
Statistical Research Division

September, 1986

OVERVIEW AND SUMMARY OF THE FORWARD TRACE STUDY

1. Introduction

The Bureau of the Census conducted the Forward Trace Study to test tracing strategies necessary in an evaluation of the decennial census with a reverse record check. In a reverse record check, a sample is drawn from the population sometime before the census traced forward to the census, and matched to the census. A sample was drawn from the 1980 census and supplemented by a sample of those missed by the census, a sample of immigrants and a sample of births. The people in the four samples were traced over the years 1980 to 1985.

The reverse record check has been used effectively in the evaluations of the Canadian censuses since 1961. The underlying logic in a reverse record check is that, with time, a person's chance of being interviewed changes. For example, some people are very mobile during their late teens and early twenties but are less mobile as children and older adults. They are more likely to be interviewed during these more stable periods of their life.

The Forward Trace Study does not consider the feasibility of the reverse record check for evaluating the census because there was no census in 1985 available for matching. However, the study focuses on an essential ingredient, the tracing techniques. The three tracing strategies that are considered are:

Treatment A, periodic tracing with intermediate personal contact,

Treatment B, periodic tracing with one initial contact,

Treatment C, periodic tracing without personal contact.

The results of the study indicate that a reverse record check does not appear workable on the scale that would be required to produce estimates of census coverage error suitable for census adjustment. The estimates of the tracing rates obtained in the closeout interview in 1985 are not high enough to recommend a reverse record check as a method of evaluating the census in 1990. Although there is some evidence that the tracing rates might be higher with intensive tracing techniques, tracing people over time is difficult to manage and control. Since even experienced interviewers needed a month or more to locate the harder cases, a sufficient number of qualified personnel probably would not be available at the time of the census.

This report contains the major findings of the study. Section 2 describes the various tracing procedures and discusses the cost. Section 3 contains the estimates of the tracing rates. Section 4 summarizes the conclusions that can be drawn from the study.

2. Description of Operations

The sample for the Forward Trace Study is derived from four sources: (1) the 1980 Census, (2) persons missed in the 1980 census, (3) people immigrated since the 1980 census and, (4) children born since the 1980 census. The sample persons are traced, in time, from location to location. Record is kept of the new locations of movers.

The sources are referred to as the C,M,I and B samples. The C or census sample is a sample of the 1980 Census Post Enumeration Program (PEP) E-Sample. The E Sample was a sample of households Enumerated in the 1980 Census.

The M or missing sample is a portion of the PEP person or P sample. The P sample resulted from the matching of April and August 1980 Current Population Survey (CPS) cases to census questionnaires. A sample of the persons determined missed in the census by that 1980 Census CPS Match is the Forward Trace M sample.

The I or immigrant sample is a sample of immigrants to the United States from April 1, 1980 to March 31, 1984. The sample was selected for each year by the Immigration and Naturalization Service through seventeen of their files control offices.

The B or birth sample is a sample of births in the United States from April 1 to December 31, 1980. The B Sample was selected in conjunction with the National Center for Health Statistics (NCHS) and 43 states, the District of Columbia and City of New York vital statistics offices. Seven states were not able to participate in the study because of regulations controlling the disclosure of confidential birth record information. These states were Idaho, Louisiana, Maryland, Michigan, New Jersey, Oklahoma and Rhode Island.

The C, M, I, and B Samples are divided into three treatments. The treatments are defined by the one tracing procedure particular to itself and different from the others. Treatment A cases are defined to be periodically traced, including periodic personal contact tracing. Treatment B is periodically traced but the initial contact was the only personal contact. Treatment C is defined by only periodic tracings, but with no personal contact.

Because of limited fiscal year '85 funds, a subsample of the original C, M, I, and B Samples was selected. The final sample sizes were the following:

<u>Sample</u>	<u>Treatment</u>	<u>Number of</u>	
		<u>Households</u>	<u>Persons</u>
C	A	1,373	4,119
	B	1,341	3,977
	C	1,310	3,854
C Sample Total		<u>4,024</u>	<u>11,950</u>
M	A	489	967
	B	504	1,071
	C	437	958
M Sample Total		<u>1,430</u>	<u>2,996</u>
I	A	677	677
	B	787	787
	C	885	885
I Sample Total		<u>2,349</u>	<u>2,349</u>
B	A	316	316
	B	315	315
	C	314	314
B Sample Total		<u>945</u>	<u>945</u>
All Samples	A	2,855	6,079
	B	2,947	6,150
	C	2,946	6,011
Sample Totals		<u>8,748</u>	<u>18,240</u>

Seven types of tracing techniques were used during the study. The ones performed on a case depended on its treatment and sample. Because of timing restrictions, the same techniques were not included in each sample.

The tracing techniques were as follows:

- 1) An initial interview of cases assigned to A and B treatments was conducted.
- 2) The Post Office was asked to confirm the address we had for a sample person or provide a forwarding address.
- 3) Letters explaining that the sample person had been selected for a research study were mailed to each person in the sample. An address correction was

requested on the envelope, and the clerks recorded the new address when these were returned in addition to mailing a letter to the new address.

4) An interim interview of cases assigned to treatment A was conducted.

5) An administrative records match (ARM) of Forward Trace records to Internal Revenue Service records was performed to obtain new addresses. The ARM was done for sample people in treatments A and B when we had their social security numbers.

6) The closeout consisted of mailing out questionnaires to all sample people. If the sample person did not return the questionnaire, an interviewer was sent to the last address we had on file for the person.

7) The Super Trace was an intensive field trace of a sample of the people not found in the closeout.

One way of comparing the three methods of tracing, the three treatments, is by the cost. Cost records were not kept in a way that permits determining the cost for individual techniques or treatments. However, the cost were kept in such a way that allows for a relative comparison on a case basis.

The estimates of the cost for each treatment and each sample are contained in the tables below. Each contains an estimate of the cost to select the sample for the Forward Trace Study. The cost of the initial interview is not included for the C and M samples. The initial interview for these cases was the interview for the 1980 PEP.

The assessment of the cost of the C and M sample cases does not include the cost of the selection of the 1980 PEP. The Census Bureau's 1980 PEP cost slightly under \$17 million which includes both the E and P sample operations. The E sample was selected clerically from boxes of census questionnaires using a list of questionnaire numbers. However, the \$17 million does not include the selection of the P sample because the P sample was the sample for the Current Population Survey.

Since the Super Trace samples are subsamples of those not traced in the closeout, the cost for a person in the Super Trace samples is determined by adding the estimated cost of the Forward Trace and the estimated cost of the Super Trace itself. The Super Trace samples contained 760 of 2890 people not traced in the closeout. The amount of \$44.50 is the estimated cost of field work and clerical processing for each of the 760 people selected for the Super Trace.

**Cost Per Person for the C and M
Samples by Treatment**

	A	B	C
Interview	-	-	
FTS Selection	3.00	3.00	3.00
PO Check	5.00	5.00	5.00
Letter	3.50	3.50	3.50
ARM	5.10	5.10	
Interview	19.00		
PO Check	5.00	5.00	5.00
Closeout	<u>13.50</u>	<u>13.50</u>	<u>13.50</u>
Total	<u>54.10</u>	<u>35.10</u>	<u>30.00</u>

**Cost Per Person for C and M
Super Trace Samples by Treatment**

	A	B	C
Forward Trace	54.10	35.10	30.00
Super Trace	<u>44.50</u>	<u>44.50</u>	<u>44.50</u>
Total	<u>98.60</u>	<u>80.70</u>	<u>74.50</u>

**Cost Per Person for the I Sample
by Treatment**

	Years 1, 2 and 3		Year 4
	B	C	
FTS Selection	13.00	13.00	13.00
PO Check	5.00	5.00	
Interview	19.00		
Letter	3.50	3.50	3.50
Closeout	<u>13.50</u>	<u>13.50</u>	<u>13.50</u>
Total	<u>54.00</u>	<u>35.00</u>	<u>30.00</u>

**Cost Per Person for I Sample
Super Trace Sample by Treatment**

	Years 1, 2 and 3		Year 4
	B	C	
Forward Trace	54.00	35.00	30.00
Super Trace	<u>44.50</u>	<u>44.50</u>	<u>44.50</u>
Total	<u>98.50</u>	<u>79.50</u>	<u>74.50</u>

Cost Per Person for the B Sample

FTS Selection	11.20
Letter	3.50
Closeout	<u>14.00</u>
Total	28.70

**Costs Per Person for B Super Trace
Sample by Treatment**

Forward Trace	28.70
Super Trace	<u>44.50</u>
Total	73.20

3. Trace Rates

Estimates of the trace rate were made for each sample as a whole and for each treatment within the samples. Estimates of the trace rate also were made for demographic subgroups within each sample. A sample person was considered to have been traced if the person was found during the final closeout operation either by returning the mail questionnaire or by field interview or the person was identified as deceased or emigrated during one of the tracing operations. When the Super Trace results were included, sample persons were considered traced if they were traced during the final closeout or the Super Trace.

Tables 1-6 contain the trace rates with and without the Super Trace results for the C, M, I and B samples by race and by treatment when appropriate. The I sample did not contain race information, and the B sample people were not divided into treatments. In the regular tracing operation, the estimates of the trace rates are 91.1 percent for the C sample, 83.4 percent for the M sample, 70.6 percent for the I sample and 73.7 for the B sample. The estimated standard errors for these trace rates are 0.6 percent, 1.6 percent, 1.1 percent, and 1.4 percent, respectively. When the Super Trace results are added the estimates of the trace rates increase to 92.2 percent for the C sample, 85.2 percent for the M sample, 72.3 percent for the I sample, and 76.4 percent for the B sample. The estimated standard errors for these trace rates are 0.6 percent, 1.5 percent, 1.1 percent and 1.4 percent, respectively. The Super Trace increased the overall trace rate in each sample from 1.1 to 2.7 percentage points.

Table 1 C-Sample Percentage Estimates of Trace Rates and Standard Errors by Race and Treatment

<u>Treatment</u>	<u>White</u>	<u>Black</u>	<u>Other</u>	<u>Combined</u>
A	95.4 (1.0)	86.8 (1.6)	90.3 (1.4)	94.1 (0.8)
B	90.8 (1.5)	81.7 (1.9)	83.3 (1.9)	89.3 (1.3)
C	90.7 (1.2)	81.7 (1.7)	80.9 (3.2)	89.7 (1.1)
Combined	92.3 (0.7)	83.7 (1.0)	85.3 (1.2)	91.1 (0.6)

Table 2 C-Sample Percentage Estimates of Trace Rates and Standard Errors by Race and Treatment When Super Trace is Included

<u>Treatment</u>	<u>White</u>	<u>Black</u>	<u>Other</u>	<u>Combined</u>
A	96.2 (0.8)	88.7 (1.4)	91.6 (1.3)	95.1 (0.7)
B	91.3 (1.5)	82.9 (1.8)	84.1 (1.9)	89.9 (1.3)
C	92.3 (1.1)	83.1 (1.7)	84.5 (2.7)	91.3 (1.0)
Combined	93.3 (0.7)	85.2 (1.0)	87.1 (1.1)	92.2 (0.6)

Table 3 M-Sample Percentage Estimates of Trace Rates and Standard Errors by Race and Treatment

<u>Treatment</u>	<u>White</u>	<u>Black</u>	<u>Other</u>	<u>Combined</u>
A	87.8 (2.8)	82.5 (3.3)	89.6 (7.2)	86.8 (2.2)
B	84.0 (3.7)	77.4 (3.7)	77.7 (9.4)	82.4 (2.9)
C	83.7 (3.3)	80.8 (3.5)	48.8 (14.2)	81.2 (2.8)
Combined	85.0 (2.0)	80.1 (2.0)	73.9 (6.9)	83.4 (1.6)

Table 4 M-Sample Percentage Estimates of Trace Rates and Standard Errors by Race and Treatment When Super Trace Is Included

<u>Treatment</u>	<u>White</u>	<u>Black</u>	<u>Other</u>	<u>Combined</u>
A	90.0 (2.5)	84.3 (3.2)	89.6 (7.2)	88.7 (1.9)
B	85.3 (3.7)	78.6 (3.7)	77.7 (9.4)	83.6 (2.0)
C	85.8 (3.2)	82.1 (3.4)	57.2 (15.0)	83.5 (6.8)
Combined	86.8 (2.0)	81.5 (2.9)	76.4 (2.8)	85.2 (1.5)

Table 5 I-Sample Percentage Estimates of Trace Rates and Standard Errors by Treatment With and Without Super Trace, All Four Years Combined

<u>Treatment</u>	<u>Regular</u>	<u>With Super Trace</u>
B	72.9 (0.9)	74.3 (0.9)
C	66.1 (1.9)	68.2 (3.7)
Combined	70.6 (1.1)	72.3 (1.1)

Table 6 B-Sample Percentage Estimates of Trace Rates and Standard Errors by Race With and Without Super Trace

<u>Race</u>	<u>Regular</u>	<u>With Super Trace</u>
White	76.5 (1.6)	79.3 (1.5)
Black	62.6 (3.2)	64.0 (3.2)
Other	61.6 (6.9)	63.9 (6.6)
Combined	73.7 (1.4)	76.4 (1.4)

Although the increase in the trace rates due to the Super Trace is not significant at the 90 percent level of confidence for any of the samples, the estimates of the trace rates for the Super Trace alone indicate that it was a successful operation. Table 7 contains estimates of the trace rates for the Super Trace. All the sample people in the Super Trace were not traced during the final closeout operation. These tracing rates are the result of a longer tracing period and more intensive effort by the interviewers.

Table 7 Percentage Estimates of Trace Rates for the Super Trace by Sample

<u>Sample</u>	<u>Trace Rate</u>
C	45.6
M	55.0
I	26.8
B	46.9

In the C sample treatment A was more successful than treatments B and C. The result is significant at the 90 percent level of confidence. The estimated trace rate for the A treatment in the C sample is 94.1 percent with an estimated standard error of 0.8 percent. The estimated trace rate for the B and C treatments is 89.3 percent and 89.7 percent, respectively, with estimated standard errors of 1.3 percent and 1.1 percent. The addition of the Super Trace results increases these percentages 1.5 to 2.6 points but bears the same conclusions.

The A treatment in the M sample was also more successful the B and C treatments. The result is not significant at the 90 percent level of confidence. However, the method of variance estimation used for the M-sample probably tends to be conservative. The estimated trace rates for treatments A, B, and C are 86.8 percent, 82.4 percent and 81.2 percent, respectively. Their estimated standard errors are 2.2 percent, 2.9 percent and 2.8 percent, respectively. Including the Super Trace increases the trace rates for treatments A, B, and C to 88.7 percent, 83.6 percent and 83.5 percent, respectively. These increases are not significant at the 90 percent level of confidence.

In the I sample only the B and C treatments are shown because time did not permit an interim interview for the A treatment. The B treatment is more successful than the C treatment. The result is significant at the 90 percent level of confidence. The estimated trace rate for the B treatment is 72.9 percent with an estimated standard error of 0.9 percent while the estimated trace rate for the C treatment is 66.1 percent with an estimated standard error

of 1.9 percent. Including the Super Trace results increases the trace rates 1.4 to 2.1 percentage points but the same conclusions hold.

3.2 Trace Rates by Race

The Forward Trace Study was more successful in tracing whites than blacks and other races in the C, M and B samples. The results are significant at the 90 percent level of confidence in the C and B samples, but not in the M sample. The race of the I sample people was not available.

The estimated trace rates for whites is higher in the C sample than in the M sample. The result is significant at the 90 percent level of confidence. However, this not the case for minorities. The difference in the estimated trace rates for blacks for the C sample and the M sample is not significant at the 90 percent confidence level. The same conclusion holds for races other than whites and blacks. The same pattern is evident within each of the three treatments for each of the two race categories with the exception of the C treatment for others.

The estimated trace rates for whites in the C, M and B samples are 92.3 percent, 85.0 percent, and 76.5 percent, respectively. Their estimated standard errors are 0.7 percent, 2.0 percent, and 1.6 percent, respectively. The addition of the Super Trace results increases the rate by 1.3 to 2.8 percentage points. The estimated trace rates for whites in the C and M samples traced with the A treatment are 95.4 percent and 87.8 percent, with estimated standard errors of 1.0 percent and 2.8 percent, respectively.

The estimated trace rates for blacks in the C, M and B samples are 83.7 percent, 80.1 percent, and 62.6 percent, respectively. Their estimated standard errors are 1.0 percent, 2.0 percent, and 3.0 percent, respectively. The trace rates for blacks traced with the A treatment in the C and M samples are 86.8 percent and 82.5 percent, respectively. Their estimated standard errors are 1.6 percent and 3.3 percent, respectively. The inclusion of the Super Trace results increases these rates slightly.

The estimated trace rates for races other than white and black are 85.3 percent for the C sample, 73.9 percent for the M sample and 61.6 percent for the B sample. The estimated standard errors are 1.2 percent, 6.9 percent, and 6.9 percent, respectively. The trace rates for others traced with the A treatment are 90.3 percent, with an estimated standard error of 1.4 percent, in the C sample and 89.6 percent, with an estimated standard error of 7.2 percent, in the M sample.

3.3 Comparisons With Other Studies

Trace rates in Table 8 for reverse record checks in the United States in 1960 and in Canada in 1976 and 1981 provide a basis of comparison. The samples for the reverse record check were drawn at the time of the 1960 census and traced retrospectively from the address available from the source. The sources were the 1950 Census, the 1950 Post Enumeration Survey, state birth records and alien registration records. The trace rate for immigrants was 100 percent because at that time legal aliens were required to register their address with the government every year, and the sample was drawn from this list. The trace rate of 70.6 percent for the I sample can not be compared with the 1960 results. The estimated trace rates for the C and M samples differ only by 0.1 percent and 0.3 percent, respectively, from the trace rates from the 1960 study. When the Super Trace results are included, the trace rates from the Forward Trace Study are higher. The A treatment was more successful in the C and M samples than the 1960 study.

Table 8 Percentage Trace Rates in Reverse Record Checks in the U.S. and Canada

	U.S.	CANADA	
	1960	1976	1981
Census	91.0	96.0	97.1
Missed	83.2	91.4	96.1
Births	85.6	92.4	92.3
Immigrants	100.0	89.4	96.1
Combined	87.8	95.2	96.6

The Canadians have an advantage over the Forward Trace Study in that they are able to start with a sample that has addresses and names for most all the sample people. The percentage of sample persons in each sample that did not have sufficient information to attempt tracing is shown in Table 9. For five percent of the sample persons selected for the I sample and over six percent of those selected for the B sample no attempts at tracing were made. With the I sample the problem was usually that the address was not complete. The main problem with the B sample was that confidentiality laws in some states, particularly California, permitted parents to request that all information not be released. With these cases, not even a name was received. The sample persons in the C sample without sufficient information for tracing were often enumerations without any names.

Table 9 Percentage of Sample Persons Without Sufficient Information to Attempt Tracing

<u>Sample</u>	<u>Percentage</u>
C	0.5
M	0.1
I	6.5
B	5.2

4. Summary

The estimates of the trace rates from the Forward Trace Study do not merit a recommendation that a reverse record check be used to measure census coverage in 1990. More importantly, a reverse record check does not appear workable on a scale that would be required to produce estimates of census coverage error suitable for census adjustment.

The estimates of the trace rates are comparable to those achieved in the Census Bureau's 1960 reverse record check. The estimates of the tracing rates in the 1960 reverse record check are shown in Table 8. However, the Super Trace techniques would be necessary to ensure the highest trace rates possible. Since census adjustment would require high trace rates, the Super Trace techniques should be used for all personal interviews.

The cost estimates illustrate that tracing is expensive. Treatment A which had the highest trace rates also was most expensive because of the periodic personal contact.

A significant finding was that the Forward Trace Study was also a challenge to manage and control. The records for the original sample filled 22 file cabinets. File folders were constantly being pulled and refiled for recording results of the tracing techniques. Any future study requiring tracing would be advised to consider controlling their records with computers.

Tracing is not a quick operation that can be accomplished with an inexperienced staff. The final closeout operation took four months to complete the clerical and field work. The Super Trace required an additional two months. All the interviewers were experienced, dedicated, and motivated. The letter in the Appendix illustrates how motivated and interested the interviewers were. These time and staff requirements would be difficult to meet in a census adjustment environment.

The primary methodological advantage a reverse record check has over a post-enumeration survey is that there is no response correlation between the independent record frames and the census being checked. This type of correlation introduces bias in the estimates of census coverage error.

However, which of the two methods achieves better coverage of the population, particularly for minority subgroups, is not obvious. The effective coverage rate can be viewed as the product of the frame coverage rate and the response rate,

$$\begin{array}{rcccl} \text{Effective} & & \text{Frame} & & \text{Response} \\ \text{Coverage} & = & \text{Coverage} & \times & \text{Rate} \\ \text{Rate} & & \text{Rate} & & \end{array}$$

The response rate for a post-enumeration survey will probably be higher than the estimates of trace rates observed for treatment A persons in the Forward Trace Study. However, how well the sampling frames for the two methods cover the population and their respective strengths is not clear, especially for subgroups that are hard to enumerate in the census. The estimates of the trace rates for minorities are lower than those for whites. The post-enumeration survey method has also shown weak response rates in these groups.

The Forward Trace Study shows that a reverse record check is more complicated to execute than a post enumeration survey. When estimates of census coverage are required for census adjustment, no evidence was produced to indicate that a reverse record check would be more effective.

Appendix

The Chief of Field Division, Stanley Matchett, sent a letter to the interviewers inviting their comments. The portion of the response from Ida Reiter of Smyrna, Georgia, that pertains to the Forward Trace Study follows:

Hello!

9/9/85

I've just finished the Supertrace survey for Linda Leier at 2900 and I'm feeling pretty proud of myself on being able to find 4 of 4 sample persons. I'm a good sleuth but I'd like to call whoever's attention to the fact that a little more time to follow up on it really helped (plus a bit of overtime)--even the passage of time helped locate one person!

At this end what you need is a person like me, experienced interviewer WHO KNOWS THE TERRITORY, knows how to LISTEN and ask questions with a happy enough disposition to disarm people so they'll tell you stuff. Too many people don't want the old girlfriend, the landlord, the ex, the cops, the bank, the WORLD to know where they are that will simply try to disappear and they do a good job of it, too! At "your" end give me please a slightly warmer trail to follow (5 years is too long!) CORRECT information if you can please. (Well I worked on Decennial too, I know how it was.) I hope this feedback is of some value, I may be dead in 1990.