

REVERSE RECORD CHECK:  
TRACING PEOPLE IN CANADA<sup>1</sup>

J.-F. Gosselin, Statistics Canada

1. INTRODUCTION

The Reverse Record Check (RRC) method is generally recognized as one of the best procedures to evaluate the level of undercoverage in the census. A frame containing all persons who should be enumerated in the current census is built up from the returns of the previous census and intercensal birth and immigration registrations. A random sample is selected from each source and each selected person is traced to his/her current census address. Census documents are then checked to determine whether or not the selected person was enumerated.

The main advantage of this method lies in the fact that it does not involve any form of re-enumeration which generally leads to under-estimates of coverage errors because of the strong tendency for persons missed in the census also to be missed in the re-enumeration process. However, one potential problem with this method is that there will always be a nucleus of selected persons who cannot be traced to their current census address and for which the enumeration status cannot be determined. Since the level of undercoverage among these persons is probably higher than average, it becomes extremely important to keep this group as small as possible in order to reduce the potential bias in the undercoverage estimates (this is a form of non-response).

The RRC method has been used successfully in Canada and is the main vehicle for assessing the level of undercoverage in the current Censuses of Population and Housing. It was first introduced in the 1961 Census using a small sample selected from the previous census and from which only national estimates could be produced reliably. In 1966 the method was used on a larger scale with a sample size of about 26,100 persons which included a sample of intercensal births and immigrants, and a sample of persons missed in the 1961 Census.

Both the 1971 and 1976 Reverse Record Checks were developed along lines very similar to the 1966 Check with improvements being made to the design and methodology. The main difference in 1976 was a further increase in and a redistribution of the sample to allow population undercoverage to be estimated at the province level. Currently, a study of similar size is being planned for 1981. Reports on the 1966, 1971 and 1976 RRC may be found in [1], [2], [3] and [4]. In addition, [5] gives a comprehensive report on the 1976 Coverage Measurement Programme.

This method has been used successfully in Canada because of our ability to keep the proportion of untraced persons at a very low level through extensive use of administrative records as well as follow-up from the regions. The tracing operations are by far the most complex, costly and time-consuming operations associated with this study. This paper describes the various tracing methods used as well as the success obtained from each of them. A brief description of the methodology of the study will first be presented.

2. OVERVIEW OF THE METHODOLOGY

This study involves five major steps:

- 1 - The construction of a frame of persons who should be enumerated in the census based on sources independent of the current census.
- 2 - The selection of a random sample from each of these sources.
- 3 - The tracing of selected persons to determine the address of their usual place of residence on census day.
- 4 - The searching of census forms to determine whether or not the selected persons were enumerated at the address traced and a follow-up operation for cases not found.
- 5 - The weighting of the sample data and the production of final results.

Each will now be discussed briefly.

The frame is constructed from four different sources (also called frames):

Census frame: all persons enumerated at their usual place of residence in the previous census;  
Birth frame: all intercensal births;  
Immigrant frame: all intercensal immigrants;  
Missed frame: all persons missed in the previous census.

For the first three frames, records are available from which a sample can be selected. For the missed frame, no exhaustive list is available. However those persons classified as missed in the census by the previous RRC are taken as a random sample from this frame.

These frames combined cover practically all persons who should be enumerated in the census<sup>2</sup>. The major groups which are not covered are illegal immigrants, and persons missed in the previous census and not given a chance of selection in all of the previous RRCs. The latter group is probably becoming very small as more RRCs are conducted.

The sample design varies from frame to frame depending mostly on the nature of the lists or records available. The 1976 sample sizes are given in Table 7. Further details may be found in [4].

Since the address obtained at the time of selection is usually out of date, a Tracing operation must be undertaken to determine the address of each selected person as of Census day. The tracing methods used vary from frame to frame and include: registered letters sent to the last known address; searches of administrative files such as Family Allowance and Old Age Security records; telephone and field tracing conducted from the Regional Offices.

All cases traced to a possible census address undergo a Head Office Searching operation whereby the current census documents are checked to determine whether or not the selected person has been enumerated at the address at which he/she was traced. Those who are found are automatically classified as 'Enumerated' and are considered finalized. For those cases not found, a Follow-up operation is undertaken from Regional Offices

whereby each selected person is contacted to verify his/her address as of Census day or to obtain other possible addresses, and to collect some basic data on characteristics of those missed in the census.

The sample data are then weighted to produce final estimates of population and household undercoverage.

### 3. TRACING OPERATIONS

Since the addresses obtained at time of selection are generally out of date, a Tracing operation has to be undertaken to determine the place of residence of each selected person (SP) on Census day. The tracing system used consists of a series of individual operations or 'traces', carried out sequentially in a pre-determined order (i.e., cases not traced at a given stage undergo the next stage of tracing).

The system used in 1976 was a slight variation of the one associated with the 1971 study. The tracing methods can be classified into three general types:

1 - Operations that are conducted in or directed from Statistics Canada's eight Regional Offices (RO). These include a match with current census records built into the Census Regional Office processing operations (RO Match), as well as telephone and intensive field traces (Telephone Trace and Main RO Trace).

2 - Operations that make use of administrative records such as Health and Welfare Canada's Family Allowance and Old Age Security records, Unemployment Insurance Commission records, and Revenue Canada/Taxation records.

3 - Mail-out operations from Head Office.

As a general rule, attempts are made to maximize the use of existing administrative records because to do so often eliminates the need to contact the selected person and in general is much cheaper. However, complete and up-to-date records unfortunately tend to exist only for certain special groups of the population such as persons 65 or over. Hence there was a need to resort to the other two types of operations noted above.

Diagram 1 illustrates the sequence of the various tracing operations for each frame. Individual traces will now be discussed.

#### 3.1 R.O. Match

The first tracing operation undertaken for the census frame was the Regional Office Match. Its purpose was to look in the 1976 census records for nearly 28,000 persons selected from the 1971 census returns, to see whether they were still residing at, and had been enumerated at, their 1971 census address.

This operation was built into the census RO Processing operations in order to provide early access to census forms. In broad terms, this operation consisted of a search of the 1976 census records for those Enumeration Areas (EA) corresponding to the selected 1971 EAs to locate the selected persons.

As a result of this operation, cases were classified into five categories. Table 1, which presents these categories, indicates that 42% of the Census frame sample was matched during this

operation, which makes it very effective. This is particularly so, given that this operation made use of current census records, which eliminates the need for any further searching in Head Office as was required for other tracing operations. The small number of cases in category 2 had to be later re-checked since the 1971 information could not be matched exactly with the 1976 data (e.g., different initial or a small difference in the year of birth). Most of these were later re-classified as enumerated.

#### 3.2 R.O. Telephone Trace

The purpose of the telephone trace was to contact selected persons who were not located during the RO Match or to obtain leads as to the whereabouts of SPs through the use of the telephone from each Regional Office. Specifically the cases for which this operation was undertaken correspond to SPs classified in categories 3, 4 or 5 in the RO Match (see Table 1).

For cases where the address was found but not the SP (category 3), a telephone call was made to the number at that address (which was noted during the RO Match from the census form) to get information on the whereabouts of the SP. If this was unsuccessful, an attempt was made to locate a telephone number by searching telephone directories either for the SP or for the head of his/her household in 1971. This latter procedure was also used for cases classified in category 4 or 5 in the RO Match. All leads provided by this initial contact were followed-up by telephone unless they were outside the Regional Office area jurisdiction.

Over 15,500 cases were sent to the telephone trace and the results are summarized in Table 2. Almost 59% of these cases were traced while a lead was obtained for an additional 32.5%. These were used in subsequent phases of tracing. In only about 9% did the telephone trace lead to a dead end. The telephone trace success rates for categories 3, 4, and 5 are shown in Table 3. It is worth noting that in the 1971 Reverse Record Check, the telephone trace was only undertaken for cases in category 3 in the R.O. Match. However considering the tracing rate obtained, the extension of the operation to categories 4 and 5 was an overall success.

#### 3.3 Tracing Using Health and Welfare Canada Records

Health and Welfare Canada (H.&W.C.) administers two programmes whereby payments are made to families with children aged 0 to 18 and to persons 65 years or over. These programmes are called Family Allowance (FA) and Old Age Security (OAS) respectively. Because monthly payments are involved, one would expect these files to be relatively complete and up-to-date for the population that is covered by these programmes and hence these records are ideal for tracing purposes.

These files were used in the tracing of persons selected from all frames but at different stages of the operation. For the census frame, about 4,100 cases were sent to H.&W.C. These were cases which had not been located by the Telephone Trace and for which the selected persons either themselves were eligible for FA or OAS, or whose

family contained an eligible member in the 1976 census. A search was undertaken only in the last known province of residence and produced a success rate of slightly over 50% (see Table 5).

A very similar procedure was carried out for the immigrant (539 cases) and the missed (177 cases) frames with the exception that for these frames this search was actually carried out earlier, around census time using FA records only. The success rates obtained for these frames were about 64% and 75% respectively.

For the birth frame, the FA records were the primary source of tracing information. This was carried out in two phases. In the first phase, a search of records in the province of birth was undertaken. The few cases that were not found during the first phase were sent to neighbouring provinces. This was the second phase and was intended to cover cases that either had changed province before the census or for which the province of birth and residence differed. As one might expect the tracing rate for births was over 90%.

The overall success rate of this operation was 68%. One problem that was experienced is the presence of a relatively large number of cases (13%) which were found on the records but which could not be considered as traced since the information provided was a postal address such as rural route or postal box number. This type of address could not be used for searching purposes, although they were used as leads in further stages of tracing.

### 3.4 Matching with Unemployment Insurance Commission and Revenue Canada/Taxation Records

The Unemployment Insurance Commission (UIC) maintains records of persons who have applied for a social insurance number (SIN) which is required for anyone entering the work force. Although this file is not kept up to date it was felt that the information could at least be used as a starting point particularly for immigrants for whom no precise address is available at time of selection. In addition this had the added feature of providing a SIN which could then be used to match with Taxation records to obtain a more recent address.

The following cases were sent to UIC where a manual search of their records was undertaken:

- (a) all persons selected from the immigrant and missed frames;
- (b) persons from the census frame not traced by the Telephone Trace and who did not qualify for the Health and Welfare Canada searches.

For those cases found, a computer match was carried out with the 1974 Taxation records (i.e., the latest available file at that time) using the SIN obtained from UIC.

The results are summarized in Table 4. These indicate that a UIC record was found for nearly 75% of cases sent. Of those, more than 2/3 could be found on the Taxation files giving an early 1975 address. Since neither source gave addresses which were current to the census, cases found on either of these records were not considered traced. Rather, the information was used as leads in subsequent tracing stages (e.g. in a mail-out operation).

### 3.5 Mail-out

A mail-out operation was undertaken for persons selected from the immigrant and missed frames who were not traced by Health and Welfare Canada. This was undertaken soon after Census day and involved about 1600 persons.

This operation was carried out in two phases. For the first phase, a registered letter was sent to the last known address. For a large proportion of cases, this address had been obtained from the UIC/Taxation match, but in some cases it was the reported intended address at time of immigration. The latter was often very unreliable. The second phase of mailing was undertaken four weeks after the first mail-out. Reminder letters (also registered) were sent to persons who did not respond to the first letter. Also there were letters sent on the first mail-out that were undeliverable and therefore were returned by the Post Office. Of those which had a proper mailing address, a letter was sent to the householder at the address in an attempt to determine the whereabouts of the selected person.

The results of this operation are summarized in Table 6. The success rates for the initial and reminder letters were about 47% and 37% respectively. The householder letter produced rather poor results with less than 3% traced.

When taken as a proportion of the number of persons involved in this operation, the reminder letter added 8% of cases traced to the initial 47% produced by the first letter while the householder letter produced less than 1% cases traced. This resulted in an overall success rate of about 56% which is a definite improvement over the same operation in 1971. This may be attributable to the introduction of the use of Taxation records.

It should be noted that a mail-out operation was also used for a very small number of cases from the census and birth frames where postal addresses were obtained from the Health and Welfare trace.

### 3.6 Main Regional Office Trace

The Main Regional Office Trace was an extensive telephone and field tracing operation carried out from the Regional Offices for about 5,500 cases from all frames not previously traced. This was the first tracing procedure common to all frames.

First, attempts were made to resolve most of the cases by telephone contact using city and local telephone directories to locate selected persons, or possible acquaintances, in order to obtain leads. Whenever possible, all leads were followed up by telephone. As a last resort, problem cases were sent to the field where all leads were to be explored.

In addition to relatives and neighbours, the following were used as sources of information: former landlords or employers, school or university files, social clubs, union offices, files of various government departments and agencies.

Nearly 80% of cases were traced by this operation. The variation by frames is shown below.

<u>FRAME</u>	<u>%</u>	<u>FINAL STATUS</u>	<u>%</u>
Census	80.1	Enumerated	88.2
Birth	71.0	Missed	2.5
Immigrant	72.8	Deceased	3.2
Missed	84.8	Emigrated	1.3
TOTAL	79.6	Tracing Failed (final)	4.8

This operation was very successful considering that, by that stage, we were down to a relatively small nucleus of persons for whom other means of tracing had already failed. It is also quite expensive as one might expect. Although no detailed cost figure could be obtained from the 1976 operation separate from other RRC field operations, the cost based on 1971 experience was probably of the order of \$15-\$20 per case sent.

### 3.7 Final Health and Welfare Canada Trace

This was the final tracing operation which involved cases not traced by the Main RO Trace and which were eligible for a FA or OAS search. The procedures were identical to the Main Health and Welfare trace (see section 3.3) with the exception that cases were sent not only to the province of their last known address but also to neighbouring provinces. A total of about 650 cases were involved in this trace of which about 42% were traced.

### 3.8 Overall Results of the Tracing

The overall results of the tracing operations are given in Table 7. The initial tracing rate was 96.4% and this rate ranged from a low of 90.2% for the immigrant frame to a high of 96.9% for the census frame.

This table also indicates the percentage of the original sample traced at various stages of the tracing operations. For the census frame, about 75% of the sample was traced in the R.O. Match or the Telephone trace, while for the birth frame, almost 89% were located through the Health and Welfare search. For the other two frames, the mail-out operation was the most effective in tracing about 45% of the original sample.

As mentioned earlier, once a selected person was traced, a search of census documents was undertaken to determine whether or not he/she was enumerated at that address. Cases not found underwent a follow-up operation to verify their census address. As one might expect there were a number of cases that were traced but could not be contacted again during follow-up. These cases, albeit small in number, were re-classified as 'untraced' which explain the fact that the final tracing rate was lower than the percentage initially traced as shown in Table 7. The unweighted percentage distribution of the sample in the final status categories is given as follows:

### 4. SUMMARY

The Reverse Record Check method has been used successfully in Canada to assess the level of undercoverage in the census. This is made possible by the implementation of a very effective but also very expensive and time-consuming tracing operation which results in a very high percentage of the original sample being traced. Currently a study of similar size is being planned for the 1981 census.

### FOOTNOTES

1. The content of this paper is based largely on internal Statistics Canada reports prepared by G.J. Brackstone, J.-F. Gosselin and G. Thérout.
2. The frame also includes some persons who legitimately could not be enumerated in the census, essentially emigrants and deaths since the previous census. However these persons can be eliminated from the sample during the tracing operation and hence would not bias the results.

### REFERENCES

- [1] Muirhead, R.C., "1966 Census Evaluation Programme-Reverse Record Check", Technical Report, Sampling and Survey Research Staff, Dominion Bureau of Statistics, Canada, June 1969.
- [2] Brackstone, G.J., Gosselin J.-F., "1971 Census Evaluation Project MP-1: 1971 Reverse Record Check Results Memorandum", CDN 71-E-23 (Part 1), Census Field, Statistics Canada, October 1974.
- [3] Gosselin, J.-F., Thérout, G., "1976 Census Parametric Evaluation Programme, Reverse Record Check: Basic Results on Population and Household Undercoverage in the 1976 Census", Census and Household Survey Methods Division, Statistics Canada, January 1978.
- [4] Gosselin, J.-F., Brackstone, G.J., "The Measurement of Population Undercoverage in the 1976 Canadian Census Using the Reverse Record Check Method", American Statistical Association Proceedings of the Social Statistics Section, 230-235, 1978.
- [5] Statistics Canada, "Coverage Error in the 1976 Census of Population and Housing", 1976 Census Quality of Data Series, Catalogue 99-840, March 1980.

DIAGRAM 1: THE 1976 REVERSE RECORD CHECK TRACING SYSTEM

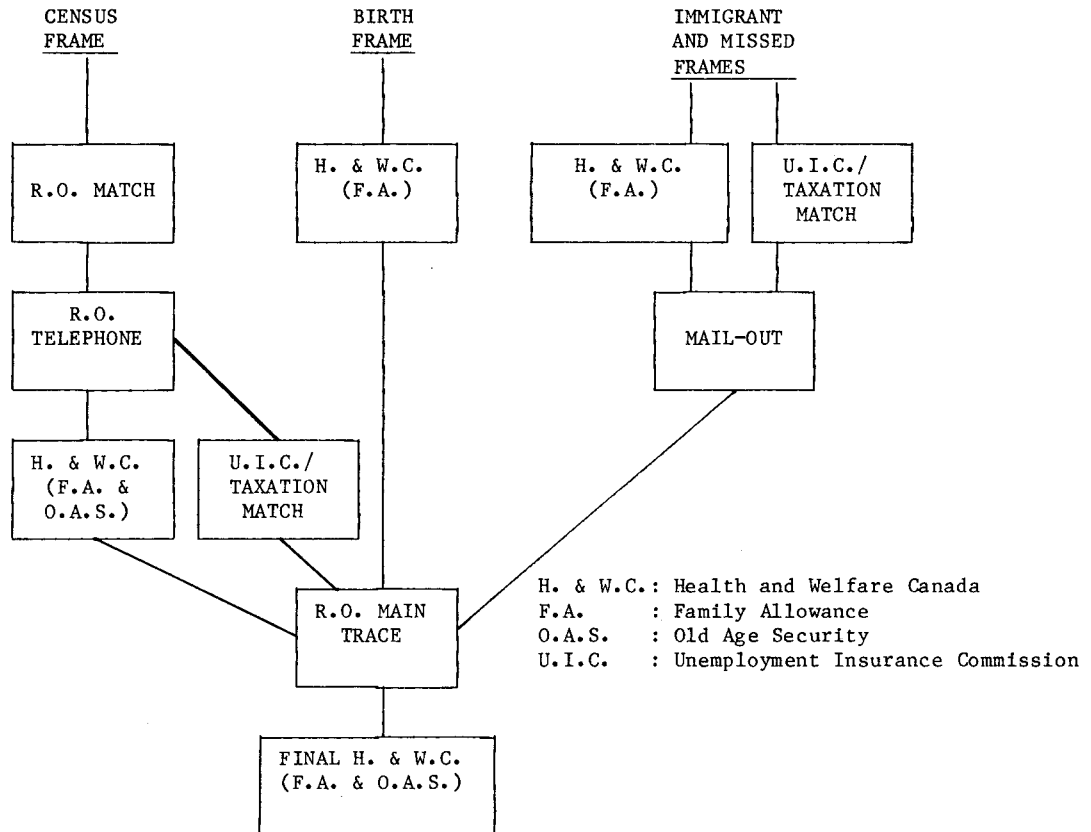


TABLE 1: RESULTS OF THE R.O. MATCH

Categories	%
1. Perfect match: SP found	42.0
2. Partial match: SP probably found	0.7
3. 1971 address/head of household found, SP not found	42.3
4. 1971 address found vacant in 1976	1.5
5. 1971 address/head of household not found in 1976	13.5

TABLE 3: TELEPHONE TRACE SUCCESS RATE FOR CATEGORIES 3, 4 and 5 OF THE R.O. MATCH

Results of the RO Match	% Traced in the Telephone Trace
3. Address found, SP not found	58.1
4. Address found vacant in 1976	46.0
5. Address not found in 1976	62.5

TABLE 2: RESULTS OF THE TELEPHONE TRACE

Results	%	
TOTAL TRACED		58.8
- Address found	54.0	
- Deceased	4.3	
- Emigrated	0.5	
TOTAL UNTRACED		41.2
- Lead obtained	32.5	
- Tracing failed	8.7	

TABLE 4: RESULTS OF THE UIC/TAXATION SEARCH

Results	%
UIC record found	
- Taxation record matched	55.4
- Taxation record not matched	18.8
UIC record not found or match not unique	25.8

TABLE 5. RESULTS OF HEALTH AND WELFARE TRACE

Results	Census Frame	Birth Frame	Immigrant Frame	Missed Frame	TOTAL
	%	%	%	%	%
Total traced	50.5	90.8	64.2	75.1	68.0
Total untraced	49.5	9.2	35.8	24.9	32.0
- lead obtained or address incomplete*	22.8	2.4	3.2	5.6	13.0
- SP not found	26.5	6.8	32.6	19.3	19.0

\*Rural route or post office box number only.

TABLE 6: RESULTS OF THE MAIL-OUT TO THE IMMIGRANT AND MISSED FRAMES

First Letter	- Total letters sent	1575
	% traced	47.2%
Reminder Letter	- Total letters sent	336
	% traced	37.5%
Householder Letter	- Total letters sent	416
	% traced	2.6%
Summary Results - Total selected persons		1575
	- Traced - First letter	47.2%
	- Traced - Reminder letter	8.0%
	- Traced - Householder letter	0.7%
	- Total traced by Mail-out	55.9%

TABLE 7: OVERALL RESULTS OF THE TRACING OPERATIONS - 1976

STAGES OF TRACING	FRAMES				
	Census	Birth	Immigrant	Missed	Total
SAMPLE SIZE	27,913	3,262	1,169	767	33,111
% TRACED BY STAGES <sup>1</sup>					
- R.O. Match	42.3	n.a.	n.a.	n.a.	35.7
- R.O. Telephone	32.9	n.a.	n.a.	n.a.	27.7
- Health and Welfare	7.5	88.9	29.6	17.1	16.4
- Mail-out	(-)	2.0	44.9	46.4	2.9
- Main R.O. Trace	13.5	4.4	24.5	30.5	13.3
- Health and Welfare (final)	0.9	1.6	(-)	(-)	0.8
% INITIALLY TRACED <sup>2</sup>	96.9	95.4	90.2	93.0	96.4
% TRACED-FINAL <sup>3</sup>	96.0	92.4	89.4	90.4	95.2

- Less than 0.1%

- The UIC/Taxation operation is not listed since cases found are not treated as traced but are passed on to subsequent tracing operations.
- The sum of the percentage traced by various stages is higher than the Total Initially Traced since a number of cases were traced by two different operations (because of overlap) and therefore are included more than once in the individual percentages.
- The difference between the final and the initial percentage traced are cases that were traced, not found in census documents, but that could not be contacted again during follow-up. These are treated as untraced.