RESEARCH INTO A REGISTER OF RESIDENTIAL ADDRESSES FOR URBAN AREAS OF CANADA

J. Douglas Drew, John B. Armstrong and Ruth Dibbs, Statistics Canada

I. INTRODUCTION

As part of its programme of research and testing in preparation for the 1991 Census of Population and Housing, Statistics Canada has initiated research into the creation of an automated list or register of residential addresses in urban areas of Canada, and into the investigation of uses of such a list in the conduct of the census.

Currently in taking the Canadian Census, address lists are compiled manually by approximately 40,000 Census Representatives (CRs) each responsible for an Enumeration Area (EA) containing 200-300 dwellings. This listing of addresses in a control document called the visitation record (VR), is carried out co-incident with drop-off of census questionnaires. In urban areas, the questionnaires are mailed back to the local CR.

The CR is responsible for basic edits ensuring completeness and consistency of the data, for telephone or personal follow-up of edit failure cases, and for personal follow-up of nonresponding households. When completed, work for the EA is returned via a supervisory network to regional sites where data capture takes place. While address ranges for block faces are data captured, individual addresses are not.

The initial steps in census taking in Canada may be characterized as being manual, highly decentralized, and heavily reliant on the quality of work by a large staff of temporary employees. An automated address register (AR) would be a pre-requisite to increased automation and centralization, which on the surface would appear to have good potential for reducing costs and increasing quality. An AR, either in combination with increased automation and centralization of front end census steps or on its own, also has the potential for reducing the dependence of census coverage on the temporary field staff. Finally, an AR has potential to yield better coverage of dwellings than the current methodology.

The topic of AR research at Statistics Canada is not a new one. As part of the research programme for the 1971 Census, Fellegi and Krotki (1967) constructed and evaluated address registers for two medium sized cities: Kitchener-Waterloo and London. The approach followed in the Kitchener-Waterloo test was similar to that currently under investigation. It involved creation of a register from the merging and unduplication of address information from multiple administrative data sources. Sources included municipal assessment rolls, the 1961 Census lists of households, and electricity billing lists. Due to technological limitations of the day, construction of the register was largely a manual process.

During the 1970's a series of studies, summarized by Booth (1976), were conducted concerning creation and maintenance of an AR for use in mailing out the 1981 Census. The approach considered involved data capturing address lists from a previous census, with updates based on information supplied by Canada Post. It was found that such a register yielded dwelling coverage comparable to that of traditional census methods. However, the high initial data capture costs were viewed as problematic despite estimated longer term savings, and the research was terminated.

It was decided to re-examine the feasibility of an AR for the 1991 Census, since it was felt that environmental factors are now more conducive to such an undertaking than was the case in previous decades. These factors include the increased availability of machine readable administrative record systems with address information, the widespread use of postal codes, the increased power and decreased cost of computers and the availability of improved record linkage methods and software.

Remaining sections of the paper include: the methodology being used for AR construction (Section 2); the evaluation of pilot registers in Ottawa (Section 3); a description of possible uses of an AR in the census (Section 4); and directions of future work (Section 5).

2. ADDRESS REGISTER CONSTRUCTION METHODOLOGY

In this section the methodology used to construct pilot address registers for test sites in Ottawa is described. Two pilot registers were constructed – one for comparison to the results of the 1986 Census and another to enable us to evaluate coverage of an address register relative to that of the Canadian Labour Force Survey (LFS). Before the methodology used to construct the pilot registers is reviewed the issue of address register contents is examined. Subsections 2.3 and 2.4 contain more detailed discussions of address standardization and record linkage, two methodological problems involved in address register construction. The section concludes with a list of possible enhancements to the construction methodology, suggested by the pilot study experience, that may be included in future address register work.

2.1 Address Register Contents

Dwelling address is of course the most important item that should appear on an addresss register. Depending on the eventual uses of a register it may be appropriate to include both physical location and mailing addresses. Failure to differentiate between these two addresses or appropriately link them would lead to significant overcoverage. From the operational point of view, obtaining information needed to link location and mailing addresses is difficult. Among the administrative files used in the pilot tests, only the municipal assessment files included both location and mailing addresses corresponding to the same dwelling. However, the mailing address referred to the property owner rather than the resident.

Only location addresses were included on the pilot registers. Each address was stored on the register in two forms. In addition to the free format version obtained from the original administrative source an address search key was stored. The search key is a standardized version of the address produced by Statistics Canada software and used for exact matching and record linkage purposes. Other contents of the pilot registers included the name and creation date for the administrative file from which the free format version of the address was obtained and information about other files that included addresses for which the same search key was generated. Telephone number and dwelling type were also stored when available. The Federal Electoral District – Enumeration Area number, providing a link to census geography, was stored on those pilot registers constructed for census test sites.

2.2 Steps in Address Register Construction

Before the address register construction process began, lists of postal codes corresponding to the test areas were obtained. For the census tests, this was done using Statistics Canada's files linking postal code and census geography. For the Labour Force Survey tests, postal code information obtained by LFS interviewers was verified by clerical staff.

Construction of the pilot registers using these postal code lists involved five steps: (i) selection of addresses that should correspond to test areas from administrative files using exact matching on postal code; (ii) address standardization (creation of search keys); (iii) elimination of exact duplicates; (iv) postal code verification; and (v) elimination of probable duplicates using record linkage techniques.

Given that exact matching on postal code was used in the first step of this procedure, a high incidence of correct postal codes on the administrative files is of obvious importance for the methodology. On most of the administrative files used in the study, a postal code appeared on over 99% of the addresses. Accuracy studies for various files have indicated rates of correct postal code use considerably higher than 90%. In addition, most errors occur in the final three digits of postal code, which determine the exact location of an address within a relatively small geographic area. Cases in which an address contains a postal code corresponding to another city are relatively uncommon.

The second step of address register construction involved the use of address standardization software developed at Statistics Canada. The software identifies the components of a free format address (street number, street name, etc.) and constructs an address search key using these components.

In the third address register construction step, the addresses that had been selected from the various files were merged and those with search keys that agreed exactly for certain important components were eliminated. In the next step, postal codes for addresses on the merged file were validated using Statistics Canada's postal coding software and those that were assigned a postal code outside the test area were eliminated.

After the elimination of exact duplicates and addresses with postal codes outside the test area, an address register may still contains some address pairs that refer to the same dwelling due to minor errors on the administrative files. An example of such a pair is "Apt 1604 1591 Riverside Dr Ottawa Ont K1G 4B7", and "Apt 1604 1591 Rivesride Dr Ottawa Ont K1G 4B6". The final step of address register construction involved elimination of such duplicates using record linkage techniques. For each pair of duplicate addresses identified by record linkage, the address that appeared on the largest number of administrative files was the one that was retained.

2.3 Address Standardization

The address standardization software used for construction of the pilot registers was developed at Statistics Canada during the 1970s. It accepts free format address information and identifies various components. The initial step of the algorithm involves scanning the input string and breaking it into "tokens". Each token is either a numeric or an alphabetic string that does not include blanks. For example, the address "Apt 1604 1591 Riverside Dr Ottawa Ont K1G 4B7", includes 9 tokens. Tables are searched in a attempt to identify the alphabetic tokens. In this case Dr is a street designator, Apt is an abbreviation of "Apartment", Ont is the province and the other strings cannot be identified from the tables. In the second stage, the numerics and the alphabetic tokens that were not found in the tables are identified by matching a series of address patterns to the available information.

A new address standardization package is currently under development. The new software, which is written in MPL, a compiler generator, offers a number of advantages over the package now available, including improved alphabetic token tables and a parsing algorithm that uses approximately 100 production rules to define valid address syntax. Standardization success rates for urban addresses are considerably higher with the new package than with the old software.

Unlike many other syntax parsing problems, such as mathematical equation parsing, address standardization involves an unlimited syntax. Consequently, certain cases will cause difficulty for any parsing algorithm. An example is "10 Main St Fort St John BC V6E 2G7". The street designator St appears in the city name "Fort St John". The new software identifies addresses like this one as ambiguous and consults tables of municipality names in an attempt to resolve them.

2.4 Record Linkage

Record linkage work involved in address register construction was done using Statistics Canada's Generalized Iterative Record Linkage System (Hill and Pring-Mill 1985), which is a record linkage package based on the methodology proposed by Fellegi and Sunter (1969). The Fellegi-Sunter methodology involves assigning weights to each pair of records according to the probability that the addresses both refer to the same dwelling. The total weight, W, for the record pair (a,b) is obtained by adding weights for individual linkage fields. That is, $W=w_1+w_2+...w_k$, where each w_i is a log-odds ratio: $w_i=\log (P(O_i) \mid (a,b)\in M) / (P(O_i) \mid (a,b)\in U)$.

M is the set of address pairs that refer to the same dwelling (true matches) and U is the set of pairs that do not refer to same dwelling (true non-matches). O_i is the outcome of the comparison of linkage field[1] i for addresses a and b. It may be agreement, disagreement or some form of partial agreement.

In practice, since the sets M and U are unknown for the file of addresses of interest, so are the conditional probabilities needed to calculate weights. Estimates of the weights can be obtained if a calibration file is available for which the sets M and U have been determined manually. Creation of such a calibration file is a process that is both expensive and somewhat error-prone. Recently Jaro (1986) has experimented with a method of weight calculation suggested by Fellegi and Sunter that does not need a calibration file, but instead, involves solution of nonlinear equation systems.

The Generalized Iterative Record Linkage System (GIRLS) uses another procedure that avoids the calibration file requirement. Initial weights are calculated assuming that the linkage fields on the file of interest are free of errors. Under this assumption weights can be computed using the frequency of occurrence of various values of each linkage field. These weights are used to rank address pairs that show some evidence of agreement, that is, might be true matches. This list of pairs is examined by the user who makes a judgement concerning a "cut-off" weight. Pairs with a weight higher than the cutoff are temporarily considered to be true matches. Estimates of error rates for the linkage fields are calculated using this set of pairs. These estimates are used to adjust the initial weights and the process of ranking pairs and selecting a cut-off is repeated. This procedure can be continued until differences between consecutive iterations are small.

In order to reduce the number of comparisons between address pairs, GIRLS allows the user to specify one or more blocking fields. Each set of values for the blocking fields defines a "pocket". During the iterative process only addresses in the same pocket, that is, addresses for which all the blocking fields agree, are compared.

In pilot AR construction, address search key fields were used as input to GIRLS. Street number and apartment number were used as blocking fields. Partial agreement outcomes were defined for street and city names using a simple character string comparison algorithm that allows for character transposition and as well as random insertion and deletion of characters. For postal code, partial agreement outcomes were defined based on the number of characters that agreed and the positions of those characters. According to these definitions, the addresses "Apt 1604 1591 Riverside Dr Ottawa Ont K1G 4B7", and "Apt 1604 1591 Rivesride Dr Ottawa Ont K1G 4B6", partially agree on both street name and postal code.

2.5 Future Methodological Improvements

Future address register construction work will involve some enhancements to the methodology described here, such as: (i) use of the new address standardization package currently under development; (ii) use of telephone number in record linkage (names may also be used but they will not be included on the final register); (iii) matching on phonetically encoded versions of street and city names; and (iv) imputation for apartment buildings (if a building includes units 101, 201 and 401 we have strong evidence that 301 has been missed).

3. EVALUATION OF PILOT ADDRESS REGISTERS

In this section results of the evaluation study for the pilot registers are reported. Some information concerning the test sites and the administrative sources are given in subsection 3.1. The methodology used to evaluate the pilot registers is described in subsection 3.2. In addition overall coverage, coverage by structural type of dwelling, and the potential for improving LFS and census coverage using an address register are examined.

3.1 Ottawa Test Sites

Two pilot registers were evaluated in Ottawa. The first register was constructed for areas covered by active Labour Force Survey clusters. Clusters, which consist of small numbers of adjacent block faces, are the penultimate LFS sampling units (households are the ultimate units). A cluster is considered active if households from the cluster are included in the current LFS sample. When a cluster becomes active it is visited by an LFS interviewer, who produces a dwelling list without making contact with residents.

While the area covered by the LFS pilot register represents a probability sample of block faces, the second Ottawa pilot register was constructed for a purposive sample of four tracts from the June 1986 Census. Three of the census tracts selected were from older parts of the city, and the fourth was from a very affluent area. They were chosen because they represented types of areas in which census coverage has traditionally been below average.

Five administrative files were used to construct the pilot registers - Revenue Canada's file of taxfiler addresses, files containing addresses of family allowance and old age security recipients from Health and Welfare Canada, the telephone company's billing list for published residential numbers and the municipal assessment file.

3.2 Coverage

The first step in evaluation of each pilot register involved matching the register to dwelling lists from the alternative source. Both exact matching and record linkage techniques were used. For each test area, it was assumed that all matched dwellings were valid. A field check was conducted to resolve unmatched addresses. For each test site a list of street names and numbers that corresponded to unmatched addresses from either the address register or the alternative source was determined. Head office personnel were asked to visit each of these street numbers and determine the total number of associated dwellings without making contact.

Coverage information for the pilot registers is given in Table 1 (LFS test) and Table 2 (census test). The tables also include information about the coverage of the LFS and the 1986 Census for the test areas. All cases in which more than one dwelling is associated with a street number are classified as multiples so this category includes some townhouse developments as well as apartment buildings. The coverage rates in these tables are calculated relative to our "best guesses" at the true lists of dwellings, based on the sets of matched dwellings and the field check results.

Before we comment on the coverage of the address registers, some remarks concerning LFS and census coverage rates are appropriate. The LFS single family undercoverage rate of 0.5% is similar to results obtained using a listing check incorporated in the LFS re-interview program (Cyr 1984). The multiple undercoverage rate of 5.3% reported in Table 1 should be carefully interpreted, since the rate of LFS multiple detection at the time of interview is substantial.

The census undercoverage rate of 3.7% appears somewhat high when compared to the 1981 reverse record check estimate of national undercoverage for urban households, which was 1.71% with a standard error of 0.13% (Burgess 1983). However, one must remember that the census tracts included in the test were selected because relatively poor coverage was expected. Over 47% of the dwellings in the test area were either apartments in buildings with less than five storeys, townhouses, or duplexes. The 1981 reverse record check undercoverage estimates for apartments in small buildings and for duplexes were 5.81% and 3.02%, respectively.

By far the largest factor contributing to the overall LFS overcoverage rate of 5.5% was cases in which the cluster lists indicated more dwellings at an address than did the field check. Such situations accounted for 124 of the 195 units incorrectly listed by the LFS. Discrepancies of this magnitude underline the difficulties associated with determining the number of dwellings in certain types of small multiples using a no-contact field check. The census overcoverage rate of 3.9% is surprising. It is important to note, however, that the field check was conducted in March 1987, ten months after the 1986 Census. In addition, each address listed on census visitation records was considered a dwelling covered by the census, regardless of whether or not a census questionnaire was returned from that address.

The overall address register coverage rates of 96.0% and 95.5% for the LFS and census tests, respectively, are comparable to the LFS and 1986 Census coverage rates of 97.7% and 96.3%. Address register coverage for multiple dwellings was not as good as single family dwelling coverage. Booth (1976) reports similar results. Much of the undercoverage and overcoverage for multiples was due to problems with the address standardization software used to construct the pilot registers. The software rarely recognized words like "basement", "upper" and "lower" when they were used to identify multiple households at the same address. A small study involving manual verification of unstandardized addresses from administrative files suggests that use of the new address standardization software currently under development would have improved address register coverage by at least one percent for both test areas.

Use of an address register as a coverage improvement tool in conjunction with the current field operations is one of the options that is currently under consideration for both census and the LFS. In view of this possibility, the information reported in Table 3 is quite promising. For the Ottawa LFS test area, use of an address register would have reduced undercoverage on initial no-contact dwelling lists from 2.3% to 1.4%. For the census test area, the coverage improvement potential is greater. Use of an address register would have reduced undercoverage in the test areas from 3.7% to 1.5%, assuming a comparable success in a production mode. This represents almost a 60% decrease in census undercoverage. It must be cautioned that with full scale implementation in a census, the reduction in coverage could be less. For example, if implementation consisted of matching an independently derived VR with an AR, the matching would be a manual operation with resolution of discrepancies by temporary staff, unlike the test which involved automated matching and use of experienced staff to resolve discrepancies.

4. POTENTIAL USES OF AN ADDRESS REGISTER IN THE CENSUS

In this section we enumerate possible uses of an address register in the conduct of the census. As these uses are examined in detail later this year we will be assessing factors such as cost and quality implications, requirements for development and testing, and the degree to which individual uses can stand alone or are interdependent. To be recommended for the 1991 Census, uses will have to satisfy criteria of low risk, adequate time to develop and test, and better or equivalent dwelling coverage at reduced or equivalent cost with a net overall gain from a combined cost and quality perspective.

Potential census uses are described below as they relate to delivery of questionnaires, follow-up of nonresponse and edit failures, control of census documents in processing, and geography operations. More details can be found in Dibbs et al (1987).

4.1 Use of an Address Register In Delivery of Census Questionnaires

On the previous occasions when AR research was carried out at Statistics Canada, the primary use envisioned was as a vehicle for conducting a mail-out census. We decided this time to investigate as well other options for use of a register in collection that could be integrated more easily into existing processes, and hence implementation would hence pose less risk to already functioning systems. The main alternatives we have identified for use of an AR in delivery of questionnaires are:

(i) To permit a mail-out census

A cost comparison for urban areas of the current list/drop-off methodology versus mail-out from an AR was conducted by Gamache-O'Leary, Nieman, and Dibbs (1987). In comparing costs, the largest unknown at this time is the cost of creating a national AR with dwelling coverage comparable to that obtained with the traditional method. The analysis revealed that if the AR created directly from administrative sources achieved comparable coverage, then there would be a clear cut cost advantage with a mail-out census. If one field check of the register was required, this would represent a rough break-even point in terms of cost. If two field checks were needed, say one by Statistics Canada personnel and another by Canada Post letter carriers, then a mail-out census would be more costly.

It was decided at an early juncture to rule out a general mail-out for 1991 and to concentrate on other uses of an AR in collection. Nevertheless, if an AR can be created and maintained at a high level of accuracy intercensally at a reasonable cost, mail-out may become an option worth reconsidering for 1996.

(ii) Pre-list with automated VR

Under this alternative, an address list would be pre-printed from the AR for each Enumeration Area. The Census Representative would update the list at drop-off. Any additions or deletions would be recorded by the CR for data capture.

An automated VR with capture of CR changes would fit best with scenarios for greater centralization of follow-up activities and an automated control file. The capture of changes made by the CR's would improve the coverage and accuracy of the AR, which could benefit other applications such as household surveys or the next census. On the other hand, there is some concern that if a CR is given a nearly complete list in advance a poor job of finding additions could translate into worse coverage than that with the current methodology.

(iii) Post drop-off coverage check

An alternative use of an AR in collection with lower risk would be to use the AR after the fact as an independent check on the quality of the manual VR's. Three procedures which are not mutually exclusive are possible. First, there could be a macro check of drop-off totals against the number of dwellings on the AR to determine if coverage at drop-off is acceptable. Second, a sample of addresses from the register could be matched against addresses on the VR as acceptance quality control of the CR's drop-off work. Third, there could be a detailed comparison of the AR against the VR to identify missed dwellings. A small scale test of the third option in the 1986 Census, reported by McCririck (1986), indicated that this could be a more cost effective method of coverage improvement than a similar check performed in the 1981 Census using post office lists and staff (Copp 1982).

4.2 Follow-up for Nonresponse and Edit Failures

Currently follow-up for nonresponse and for certain edit failures is carried out locally by CR's. In 1986, 13% of returns required nonresponse follow-up. In nonresponse follow-up the CR is encouraged to use the telephone if a name was obtained at drop-off that can be found in the local telephone directory. However, since contact rates at drop-off have fallen to as low as 40% in urban areas, this information is often not available and personal visits are required. The availability of telephone numbers on the AR for households with published numbers (roughly 80% of households) would greatly facilitate nonresponse follow-up by telephone.

The pre-list automated VR delivery alternative would lend itself to centralization of nonresponse and edit follow-up for those cases that can be handled by telephone. Potential benefits of such centralization would appear to be improved quality due to the use of experienced staff with more direct supervision, and possibilities for automation of processes; for example, in the longer term the use of Computer Assisted Telephone Interviewing.

4.3 Automated Control File

An automated control file (ACF) would be a file containing information on the status of each dwelling to which a census questionnaire is delivered. An ACF would go hand-in-hand with mail-back of census questionnaires to central sites, and centralization of nonresponse and edit follow-up. Such a scenario would permit greater control and monitoring of those early steps from drop-off up to data capture of questionnaires that are now decentralized and less amenable to close control.

Under the pre-list automated VR delivery alternative, the starting point for the ACF would be the data capture of additions and modifications to the VR carried out by the CR at the time of drop-off. This activity would take place right on the heels of drop-off, to ready the ACF for the check-in of questionnaires as they are mailed back. Check-in would be streamlined by the use of bar-coded labels. These labels would be generated from the AR for pre-listed dwellings, and affixed to questionnaires at the time of drop-off. Extra labels could be pre-printed to handle additional dwellings. The ACF would be updated to reflect the current status as individual census documents pass through the processing steps of check-in, edit, and follow-up (where necessary).

4.4 Applications in Geography Processes

An address register would impact on ongoing geography maintenance activities, on pre-censal activities such as delineation of Census Tracts and Enumeration Areas, and on processing of census data to add geographic coding.

One of the major ongoing geography maintenance activities is the updating of the Area Master File (AMF). The AMF is a machine readable file which covers urban areas of Canada, and contains the geographic identification down to the block face level required to generate automated maps showing street networks and standard geographical boundaries. Updating consists of capturing new and changed street patterns. A regularly updated AR could increase the cost efficiency of AMF updating efforts by signalling areas of change.

The delineation of EA's and CT's prior to the census would benefit from up-to-date counts, since number of dwellings is one of the principal criteria used in delineation. Currently, for many areas nothing more recent than counts from the previous census is available. A regularly updated AR providing up-to-date dwelling counts would eliminate the need for special field checks in known growth areas. The improved delineation that would result would greatly reduce the frequency of EA splits in the field at census time such splits are both costly and disruptive.

Gains would also result in geocoding of census data, provided the AR is used to generate labels which are affixed to questionnaires at drop-off. Geocoding is currently a manual operation that links household numbers to their block face centroids (x and y co-ordinates) to permit retrieval of census data for non-standard geographical areas. The capture of a unique address key appearing on the label coincident with capture of the census questionnaire would link census data to the AR, where geocodes would already be present for the majority of addresses.

5. DIRECTIONS OF FUTURE WORK

Decisions on what uses, if any, to make of an AR in the 1991 Census will have to be made towards the end of 1987 or early in 1988, when development and implementation plans will be formalized.

In order that supportable recommendations can be made, the major thrust of the project in the coming months will be in the design, conduct and analysis of a field study to evaluate the cost and quality (coverage) implications of use of an AR in census collection as a pre-list to be updated by CR's at drop-off (delivery alternative (ii) from the previous section), versus use of an AR in a post drop-off coverage check (alternative (iii)). The test is planned for five cities representative of Canada's various regions, the Regional Office cities, in November 1987.

Other components of the work plan include:

(i) Pilot register for pre-censal geography applications
One of the first production activities associated with the 1991 Census will be the delineation of EA's and CT's, which will begin early in 1988. To test the benefits of an AR to pre-censal geography operations, and other geography operations such as AMF updating, it is proposed to construct a pilot register for a sample of areas within each of the five Regional Office cities, and to maintain the registers over the period leading up to the 1991 Census.

(ii) Proposal to test household survey use of an AR
This test would use the pilot AR constructed in (i) above. For a collection of Labour Force Survey strata in each of the Regional Office cities, the LFS sample would be converted to an AR based design featuring separate strata and data collection methodologies for households with and without telephones on the AR. The study would also include investigation of approaches to dealing with undercoverage of the AR.

(iii) Non-Statistics Canada uses of an AR
During the current research and development stage, Statistics Canada can declare any pilot registers to be confidential, thereby exempting them from requests for release outside the agency. However, in the event of implementation of an AR, considerations relating to confidentiality would be different. The confidentiality of information pertaining to individual addresses would hinge on factors such as whether the information is already in the public domain and the conditions under which the information is acquired by Statistics Canada. Clearly it would be in Statistics Canada's interest to set a price for any portion deemed to be non-confidential.
In addition to information on individual addresses, summary data from a register, which by its nature would be non-confidential, may also be of potential benefit to others. Dwelling counts by postal code are an example of such summary data.
While the benefits of an AR to Statistics Canada

programmes should be of primary importance in decisions on implementation, consideration should also be given to the desirability of and demand for non-Statistics Canada uses.

REFERENCES

Booth, J.K. (1976). A summary report of all address register studies to date. Internal Report, Statistics Canada.

Burgess, R. (1983). Reverse record check tabulations. Internal Report, Statistics Canada.

Copp, J. (1982). Postal check evaluation. Internal Report, Census and Household Survey Methods Division, Statistics Canada.

Cyr, A. (1984). Analysis of dwelling coverage based on re-interview reports. Internal Report, Statistics Canada.

Dibbs, R., Tallon, P., Proulx, P., and Yan, J. (1987). Implications of an address register on census processes. Internal Report, 1991 Census, Statistics Canada.

Fellegi, I.P. and Krotki, K.J. (1967). The testing programme for the 1971 Census in Canada. Proceedings of the Social Statistics Section, American Statistical Association, 29-38.

Fellegi, I.P. and Sunter, A.B. (1969). A theory for record linkage. Journal of the American Statistical Association, 64, 1183-1210.

Gamache-O'Leary, V., Nieman, L., and Dibbs, R. (1987). Cost implications of mail-out of census questionnaires using an address register. Internal Report, 1991 Census, Statistics Canada.

Hill, T. and Pring-Mill, F. (1978). Generalized iterative record linkage system. Proceedings of the Workshop on Exact Matching Methodologies, Arlington, Virginia, 327-333.

Jaro, M. (1986). Matching the 1985 test census of Tampa, Florida, Proceedings of the Survey Research Section of the American Statistical Association, (forthcoming).

McCririck, D. (1986). Coverage check report. Internal Report, 1986 Census, Statistics Canada.

Table 1: Coverage of LFS and Pilot Address Register - Ottawa Test

	Single Family		Multiple		Overall Coverage	
	No.	%	No.	%	No.	%
Total Dwellings	2228		1312		3540	
Total LFS Dwellings	2231	100.1	1424	108.5	3655	103.2
LFS Overcoverage	14	0.6	181	13.8	195	5.5
LFS Undercoverage	11	0.5	69	5.3	80	2.3
Total AR Dwellings	2269	101.8	1401	106.8	3670	103.7
AR Overcoverage	64	2.9	206	15.7	270	7.6
AR Undercoverage	23	1.0	117	8.9	140	4.0

Table 2: Coverage of 1986 Census and Pilot Address Register - Ottawa Test

	Single Family		Multiple		Overall Coverage	
	No.	%	No.	%	No.	%
Total Dwellings	1898		4044		5942	
Total Census Dwellings	1906	100.4	4048	100.1	5954	100.2
Census Overcoverage	51	2.7	178	4.4	229	3.9
Census Undercoverage	43	2.3	174	4.3	217	3.7
Total AR Dwellings	2092	110.2	4551	112.4	6643	111.8
AR Overcoverage	276	14.5	690	17.1	966	16.3
AR Undercoverage	82	4.3	183	4.5	265	4.5

Table 3 : Pilot Address Registers - Coverage Improvement

	LFS Test		Census Test	
	Number	Percentage	Number	Percentage
Single Family	11	0.5	43	2.3
Multiple (Number of Dwellings)	20	1.5	87	2.2
Total	31	0.9	130	2.2