

# DEVELOPMENT AND TESTING OF TELEPHONE SURVEY METHODS FOR HOUSEHOLD SURVEYS AT STATISTICS CANADA

J. Douglas Drew, Peter Dick, Karen Switzer  
Social Survey Methods Division, Statistics Canada, Ottawa, Ontario K1A 0T6

## 1. Introduction

Since 1985, Statistics Canada has been testing and developing telephone survey methods with a view to reducing costs and improving data quality for household surveys. Most of the effort has focused on the Canadian Labour Force Survey (LFS), due to its role as a central vehicle for household surveys in terms of sample design, data collection and processing. The LFS is a national sample survey of 46,000 households per month. It follows a rotating panel design in which households remain in the sample for six consecutive months, after which they are rotated out. It is based on a multi-stage area sample, with a decentralized interviewing staff of 800 local interviewers located across Canada and reporting to one of 5 Regional Offices (RO's). The data collection method is "warm telephoning" - that is, face-to-face interviews with households during their first month in the sample and telephone interviews in later months. Introductory letters are mailed to households prior to the initial face-to-face interview.

Given the existing survey design, potential advantages of more telephone intensive methods are seen in :

- (i) Adoption of more efficient telephone, list, or dual frame designs, to replace the clustered area sample, so that data reliability requirements can be met with smaller sample sizes.
- (ii) The replacement of more expensive face-to-face interviews with less expensive telephone interviewing.
- (iii) Shifts in the organization of data collection from decentralized (i.e. local interviewers) to a mixed organization where some local interviewers are retained for households requiring personal interviews, but where telephone interviewing is done by central interviewers working out of the RO's. Such a mixed organization has potential for productivity gains stemming from doing the telephone interviewing in a more structured and supervised office setting.
- (iv) The shift to a mixed organization for data collection is also seen as having the potential for improved data quality, again due to opportunities for closer supervision and observation of the centralized interviewers. In addition, further data quality gains are anticipated from eventual adoption of CATI for the centralized portion.

Section 2 briefly describes the designs of major tests carried out thus far under the program. Findings are presented in section 3-5 - not as they pertain to individual tests, but bringing together all the pertinent results as they pertain to certain key issues. Section 3 deals with coverage and frame issues. In it, telephone coverage in Canada, and characteristics of the telephone and non-telephone populations are described along with the implications on frame strategies. Section 4 deals first with mode of collection, and the impact of telephone versus face-to-face interviews. It then deals with organization of interviewing staff, and the impact of local versus centralized interviewing staffs versus a mixture of both. Section 5 highlights

results comparing CATI with paper and pencil interviewing. Finally section 6 concludes with development plans and likely time frames for implementation of new methods relating to each of these issues.

## 2. Summary of Individual Tests

The tests conducted at Statistics Canada are briefly described in this section in terms of objectives and main aspects of the methodology. The salient features of each test are summarized in Table 2.1.

### 2.1 Telephone First Interview Test

The Telephone First Interview Test was designed to determine the impact of using a telephone interview instead of the traditional face-to-face interview during the first month households are in the sample. The test has been embedded into the ongoing Labour Force Survey since October 1985 in Quebec, and from October 1985 - March 1989 in Ontario. The methodology is reported fully in Drew, Choudhry and Hunter (1988). In brief, newly sampled LFS dwellings are matched to lists purchased from the telephone companies on the basis of address information, which allows telephone numbers for 65% of the new sample to be determined. Test and control samples are selected from the matched dwellings such that each test dwelling is paired with a control dwelling within the same sampling unit (city block). For these dwellings telephone numbers were provided. Interviewers were instructed to attempt a telephone interview in the test dwellings, but to revert to personal visits if initial telephone contacts were unsuccessful. Most of the interviewers in urban areas of Quebec and Ontario were involved in the test, with typically 1-3 of the 10-12 new dwellings per interviewer per month being assigned to the test treatment. Interviewers were unaware of the existence of a control sample, and for all non-test dwellings, followed regular LFS procedures.

### 2.2 Telephone Sampling Tests

At an early juncture, studies into the telephone coverage (results of which are described in Section 3), ruled out the feasibility of basing the Labour Force Survey solely on a telephone frame. Rather, dual frame or other approaches would be needed.

Two telephone sampling tests have been conducted with the objectives therefore of (i) assessing the performance of the telephone frame portion of a dual frame design, and (ii) providing information on costs, variances and non-sampling errors for use in constructing models for optimum allocation of the sample between area and telephone frames.

In both telephone sampling tests, data collection was in a regionalized, office setting, using paper and pencil methods.

The first test, run out of the Toronto and Montreal Regional Offices from December 1985 to September 1986, involved examination of two sampling approaches - Random Digit Dialing, using a modified version of the Waksberg 1978 methodology, and telephone list sampling. The latter method used telephone/address lists purchased directly from telephone companies,

which permitted letters to be sent to households prior to their being telephoned.

One of the principal findings emerging from this test related to the rapid deterioration of telephone frames, and implications of this deterioration, especially for panel surveys. As described in section 3, telephone samples which are up to date and representative of the telephone universe when a panel is first selected, can be seriously biased at later time periods if special measures are not taken to update the sample.

This problem can be readily overcome for the case of telephone list sampling, provided updated frames containing new numbers are available on a regular basis. At each occasion of frame updating, the sample for each panel is purged of numbers which are no longer in-scope and the sample for each panel is augmented by a representative sample drawn from the new telephone numbers.

A second telephone sampling test was carried out from July 1988 to March 1989 in Alberta and Nova Scotia, as an operational test of such a sample updating methodology. The test also served to broaden the base of experience in conducting the LFS as a telephone survey.

### **2.3 Centralized Warm Telephoning Test**

As already mentioned, the current Labour Force Survey collection methodology can be characterized as decentralized, warm telephoning - that is local interviewers carrying out both face to face first month interviews, and telephone follow-up interviews. The initial tests of alternative collection methodologies just described concentrated on cold telephone methods - with promising results at first. For these tests, the organization was decentralized when sampling from an area frame, and centralized when sampling from a telephone frame.

However, data quality problems with cold telephoning identified in the Telephone First Interview Test during 1987, lead to an adjustment of the testing strategies. The Centralized Warm Telephoning Test was initiated to study a mixed mode, mixed organization collection methodology. The test methodology consists of face-to-face collection for first month in sample cases by local interviewers, and telephoning by central interviewing staff working out of the Regional Offices for the majority of non-first month cases. Where nonresponse follow-up for central cases is required, this is carried out by the local interviewers. This methodology was initially tested for the Labour Force Survey by Muirhead et al (1975) and is being studied for the Current Population Survey by the United States Bureau of the Census (1987) where in addition the centralized interviewing is being done by using CATI technology.

The principal issue under study in the Centralized Warm Telephoning Test was the desirability of centralizing telephone interviews when in a mixed mode situation where both face-to-face and telephone interviews are required. Aside from improved technology (CAPI or CATI), such centralization was seen as the only opportunity for data quality or cost improvements in collection, in the event that the data quality problems observed with cold telephoning could not be resolved. In the event that these problems were resolved, it was felt the test would shed useful light on the desirability of centralizing, with local follow-up,

the portion of the sample selected from a non-telephone frame which could be handled by telephoning.

The Centralized Warm Telephoning Test was carried out from January 1988 to March 1989 in two CMA's in which Regional Offices are located - Montreal and Halifax.

The test was embedded into the ongoing LFS, and it was designed strictly as a test of operations and costs. Plans for further testing related to data quality and extension to non-RO cities are described briefly in Section 6.

Five to six adjacent interviewer assignments (approximately 420 dwellings) were selected in each city selected for the test. This amounted to roughly half the LFS sample for Halifax and about one quarter for Montreal. All other assignments within the boundaries of the CMA's were denoted as the control group.

Within the test group, all birth interviews and all interviews for households which did not have a telephone or which had denied permission to telephone were conducted face-to-face by field interviewers. All other interviews were conducted from a central location (the regional office) by telephone. This resulted in approximately 80% of all interviews being initially attempted from the centralized location, and the remaining 20% which were comprised of first month households, vacants and households without telephone numbers, being conducted in the field. If, by late afternoon of the Wednesday of survey week, no contact had been made with a household designated for central telephoning, either due to repeated unanswered calls or the telephone being disconnected (or similar message), the interview was recycled to the field for follow-up. At this time the face-to-face interviewer would attempt to make a personal call. If this was unsuccessful however, both telephone and personal calls by the field interviewer would be attempted as was practical until the end of the survey week. At this point a nonresponse code would be assigned.

During the November 1988 survey, a change in the criteria for recycling was implemented in order to assess the impact on nonresponse rates. At this stage it was decided that recycling and personal follow-up would only take place for those nonresponding households for which there was an operator intercept and the telephone company confirmed that the number was no longer valid. All other dwellings would continue to be telephoned from the central location and, if no response was obtained by the end of survey week, the non-response code 'No One at Home' was assigned.

### **2.4 CATI Test**

This test (see Catlin, Ingram and Hunter, 1988) was a controlled study comparing Computer Assisted Telephone Interviewing (CATI) with paper and pencil interviewing. The survey chosen for the test was the LFS with off line telephone samples of 1,000 Ontario households per month per treatment. All interviewing took place from Statistics Canada headquarters in Ottawa.

The study was part of a collaborative research effort with the United States Bureau of the Census (USBC), and the CATI software used was developed by the USBC. The wording of the questions was purposely the same for both treatments. Features unique to the CATI treatment were automatic branching, some basic on-line edits, and automated call scheduling.

### 3. Coverage and Frame Issues

#### 3.1 Telephone Coverage

Telephone coverage and the extent to which characteristics of those without telephones differ from those with telephones are crucial to the implementation of telephone survey methods - particularly as regards frame strategies.

In an international review of telephone coverage, Trewin and Lee (1988) found telephone coverage in Canada to be one of the highest in the world at 97-98%. As is typical of the situation in most countries they surveyed, persons in non-telephone households in Canada tend to have lower incomes and higher rates of unemployment.

Table 3.1 gives the percentage of non-telephone households in Canada from 1976 to the present. Telephone coverage, while already high in 1976 has been steadily edging upwards, although it appears to have flattened out over the last few years at around 98.5%.

Table 3.2 compares the characteristics of the telephone and non-telephone universe for the September 1985 LFS. The non-telephone households, while accounting for only 1.6% of households, accounted for 4.5% of unemployed persons, leading to a two-tenths of a percentage point difference in the unemployment rates for telephone households versus for all households. Given these differences, and the 1.5 - 2% coefficient of variation required for the national unemployment rate, it is not feasible to base the LFS solely on a telephone frame.

#### 3.2 Deterioration of Telephone Frames Over Time

Another factor bearing on the suitability of telephone frames is their rate of deterioration over time.

Random Digit Dialing has the advantage over telephone list sampling that it yields a sample which is representative of the telephone universe at the point in time when the selection is carried out. This is ideal for a one-time survey. List frames, on the other hand, will suffer from out-of-datedness depending on the age of the list.

The deterioration of telephone frames over time is illustrated in Table 3.3, taken from the first Telephone Sampling Test. In this test, each month a new panel was selected. For the RDD approach, the sample for the panel was representative of the telephone universe when numbers were initially screened, which was done two weeks prior to commencement of interviewing. For the list sampling approach, sampling was done using the most recent list frame, with list frames being purchased on a quarterly basis from the telephone companies. Table 3.3 shows percent numbers which were no longer in scope when interviewing was carried out for each of the six panels. The 2.1% out of scope numbers for Panel 1 under the RDD treatments was a result of the advanced screening. The list frame, which on average was 3-4 months old when sampled, contained 5.2% numbers that were no longer valid in panel 1. This number increased to over 7% for later panels. These rates of no longer working numbers are illustrative of the rates of frame changes. They can be considered "deaths", and assuming overall frame coverage is relatively constant over short time intervals, there exists a roughly equal number of "births" or new numbers in the frame.

A problem with the sampling strategies used for the first Telephone Sampling Test is that individual panels over time become less and less representative of the current telephone universe, due to the failure to include new numbers. The implications of this undercoverage was studied by matching the regular Labour Force Survey samples for Ontario and Quebec for October 1986 and January 1987 with telephone files. Unemployment rates in Quebec for households with new telephone numbers were 20.2% vs 10.9% for households retaining the same telephone number, with comparable figures being 9.6% and 6.6% for Ontario. This demonstrates that for the sampling strategy adopted for the first test, failure of the telephone frames to adequately represent the telephone universe lead to a coverage problem of greater magnitude than the known undercoverage of the telephone universe.

In the second Telephone Sampling Test, lists were obtained monthly, and were no more than six weeks out-of-date when interviewing started. Also as mentioned earlier, the panel samples were updated each month to reflect frame changes. Table 3.4 shows the frame changes observed during the test for Nova Scotia. Operationally, the sample updating strategy was a success. More detailed analysis for individual panels is yet to be carried out.

#### 3.3 Frame Options for Household Surveys

We now briefly discuss frame options for household surveys, and their impacts on the telephone survey methods that can be employed. Options that have been identified are as follows: (i) telephone frame, (ii) area frame, (iii) Address Register, (iv) dual (area & telephone) frame. These are discussed in turn below.

Telephone Frame Although a telephone frame has been ruled out as the sole basis for the Labour Force Survey, the situation is different for other surveys. Catlin et al (1984) showed that coverage biases for general population characteristics were less than for Labour Force characteristics. Further, for smaller surveys (e.g. those with national sample sizes of 10,000 or less) small biases are less important in relation to the larger sampling variances involved.

Hence, an RDD household survey capacity was developed, which has been used since 1986 for several one-time surveys, and in addition, for the General Social Survey, which is an annual survey of 10,000 households.

Area Frame As the Telephone First Interview test demonstrates, it is possible in urban areas to match selected addresses from the area frame to telephone lists in order to permit cold telephone interviewing. Telephone numbers can be obtained in this fashion for approximately 60% of households. The drawback to this procedure is the relatively low rate of telephoning it permits, and the fact that clustering of the sample is still required.

On a small scale, the feasibility of extending this procedure to rural areas was looked into. For the test sites, the address information collected when sampled areas are listed was modified to include last name (where visibly displayed on the dwelling or mail box), rural route numbers and Postal Code. Matches of this information to the telephone company lists produced a low match rate in the 20-30% range, which lead to abandoning further work along these lines.

Address Register Statistics Canada recently decided to construct an Address Register in urban areas

of Canada to improve coverage in the 1991 Census (Drew, Royce and van Baaren 1989). The Address Register will be a machine readable list of addresses constructed by linkage of various administrative data sources, including lists of customers with published telephone numbers purchased from telephone companies. During the Address Register's use in the 1991 Census, its coverage will be updated to correspond to that of the 1991 Census.

It is planned over the next few years to conduct in-depth studies into the use of the Address Register as a frame for household surveys in urban areas. If the conclusion is that it should be adopted as a frame, the Address Register will be updated on an ongoing basis following the 1991 Census. If its coverage is not adequate, a dual frame design with the Address Register supplemented by a small area sample could be considered.

An advantage of the Address Register as a frame over the area frame is that telephone and non-telephone households are known ahead of time. Hence the two can be sampled as separate strata - with clustering restricted to the non-telephone stratum requiring face-to-face interviews. The non-telephone stratum would include those households with non-published numbers. Evidence from earlier studies showed the refusal rate to cold telephoning for households with non-published numbers was 12% compared to 2% for all households. Under warm telephoning there is no corresponding increase in refusal rates for households with non-published numbers and there is a good success rate in converting these household to respond by telephone in later months. This finding, and the desire to be sensitive to privacy concerns of individuals support the face-to-face initial interview of such households, which account for an estimated 10-15% of numbers.

Dual Frame While the Address Register has reasonable prospects for providing a complete frame which is amenable to efficient use of telephone survey methods, it is only being developed in urban areas. It is not feasible to extend the Address Register to rural areas at present, due to the general lack of good addressing systems in rural areas. There is, however, a trend towards improved address information in rural areas for fire and ambulance emergency services, which may change this situation in the long term.

Looking ahead to the 1990's the frame option that looks most attractive for rural areas is a dual frame method, combining area and telephone frames. Research to consider issues such as the allocation of sample to area and telephone frames, and the impact of reduced sample sizes on the efficient design of the area sample is beginning. Another issue for research relates to deriving subprovincial estimates from the telephone frame. The subprovincial areas of interest will in general be unplanned domains, and the geographic information available from the telephone frame - such as Postal code, telephone exchange, and community name - may not be sufficient to determine which domain sampled units belong to.

#### **4. Collection/Organization Issues**

##### **4.1 Test Results**

Perhaps the most basic question associated with the issues of data collection and organization is the determination of whether telephone interviews yield data comparable to those obtained with face-to-face or personal interviews. Although the cost implications are

also important, the use of any new data collection methodology must ultimately depend on obtaining "good" data.

The Telephone First Interview Test, which was described earlier, does not allow for a pure comparison of telephone versus face-to-face data collection. Approximately 80% of first month in sample cases were done by telephone for the test, versus approximately 20% for the regular LFS. Thus we are comparing a predominantly telephone procedure with a predominantly face-to-face procedure. Aside from the lack of purity however, this test does allow for the control of other factors. In this situation the same interviewers are used, the sampled areas are the same, and the organization of the data collection is the same.

A somewhat purer comparison of telephone versus face-to-face interviewing can be obtained by using the Telephone Sampling Tests, and comparing the results to those obtained from the ongoing Labour Force Survey. This allows for the comparison of pure telephone contact from the off-line test with the predominantly face-to-face contact observed in the Labour Force Survey. However, the results in this case are somewhat confounded by the different organization of the test, the use of different interviewers, and the smaller sample sizes found in the Telephone Sampling Tests.

Results from individual tests are described below, followed by a general discussion.

##### **4.1.1 Telephone First Interview Results**

For this test, no differences were observed in response rates for the test and control samples. In Quebec, the response rates were virtually identical at 96.1%, while in Ontario the response rate of 96.3% observed for the test treatment was marginally less than the rate of 96.5% found for the control.

When comparing the Labour Force estimates obtained in the test and control samples, certain estimates from Quebec early in the test for the period Oct. 85 - Feb. 87 were significantly different. In particular the estimates of Employed and Unemployed Males in households of three or more persons were being undercounted in the test sample (see Drew, Choudhry and Hunter 1988). Table 4.1 presents data over the full life of the test from October 1985 to March 1989. For Quebec, a few statistically significant differences still exist - which show up only due to the influence of the early time period. When the data are analyzed from March 1987 onwards, these differences are not significant. In Ontario, the other site for this test, no such problems were encountered and no major deviations have been seen to operate for a prolonged period of time that cannot be explained by sampling errors. The reasons why the perceived deviations were occurring in Quebec are still unknown, although they may have been due to factors external to the survey leading to decreased respondent trust of cold telephone interviews. No major changes were made to the conduct of the test since its inception. Although all recent data from this test are comparable to those obtained in the Labour Force Survey, this test does seem to point to the sensitivity of telephones for data collection.

##### **4.1.2 Telephone Sampling Tests**

We begin with results from the first Telephone Sampling Test. The standard against which the test treatments were compared was the ongoing Labour Force Survey, adjusted to exclude households without telephones.

Detailed analysis of nonresponse is presented in Drew, Chouldhry and Hunter (1988). In summary, over all months of the test nonresponse rates were 8.5% for the RDD sampling treatment, 6.9% for telephone list sampling, and 4.1% for the LFS. All these differences are statistically significant at a 1% confidence level.

Table 4.2 presents estimated unemployment rates and participation rates. The only statistically significant difference at a 5% level is in Quebec, between the unemployment rates under RDD versus the ongoing LFS. It is worth noting that this difference is consistent with the differences observed between cold and warm telephoning in Quebec during the same time period under the Telephone First Interviewing Test.

Table 4.3 presents average design effects, which were calculated as the ratio of variances for a given method to the variance under the assumption of srs sampling. Here, it is worth noting that the LFS design effects reflect those for an up-to-date area frame - since a new sample based on the decennial redesign of the survey had been introduced in 1985. Also, in the LFS design, the clustering effect is partly removed by post-stratified estimators - which perhaps explains the low design effects. While design effects based on the area frame would deteriorate over the 10 years between survey redesigns, those for a telephone sample, assuming the frame is continuously updated, would remain unchanged.

Interviewing costs per household were estimated to be \$3.53 for RDD, \$2.72 for telephone list sampling, and \$4.76 for the ongoing LFS. The costs for the test treatments include \$0.46 per household for long distance charges. This amount was estimated based on long distance rates and data on length of calls, since record keeping practices in the RO's did not permit the extraction of actual charges incurred. The difference in costs between the RDD and telephone list sampling treatments was due to the extra step required in RDD - that of screening to determine in-scope numbers.

The response rates for the second Telephone Sampling Test are reported in Table 4.4 for both Nova Scotia and Alberta. The response rate in Nova Scotia was generally constant at approximately 90%. In Alberta however, the response rates varied considerably from a low of 84.5% to a high of 93.3%. The main anomaly in the response rates occurred in November 1988 and was caused by operational problems of a nature that would not occur in an ongoing production environment. These problems resulted in the survey being cancelled in Nova Scotia, and the loss of two days of interviewing time in Alberta because the samples did not arrive in the offices in time. The late start to the survey week in Alberta resulted in fewer call-backs being attempted, and a large number of "No Answer" outcomes being obtained.

In the early months of the test, particularly in Alberta, a fair amount of overlap was noted between samples for the test and the regular Labour Force, as denoted by the relatively high occurrence of "other" nonresponse. This operational problem was addressed in later months, and the number of cases in this category decreased.

The response rates in the Labour Force Survey for approximately the same areas as the test sites for the same time period are 94.3% for Nova Scotia and 94.5% in Alberta. These rates are approximately 4-5% higher than those found in the Test.

Analysis of costs, estimates and variance estimates are not yet available for this test.

### 4.1.3 Centralized Warm Telephone Test

The Centralized Warm Telephone Test, as already mentioned, was designed as a test of costs and operations. Due to problems with the breakdown of the test design in one of the test cities, analysis of the data will only be presented for the Montreal site. As shown in Table 4.5, on average a cost saving of \$0.78 per household was obtained when using the test methodology as opposed to the local approach to interviewing used for the ongoing LFS. This represents a 14% reduction in interviewer costs. There appears also to be evidence of a learning curve, since the cost savings averaged only 5.9% over the first six months of the test, as compared to 19.6% for the last nine months. The largest component of these savings is due to a lower number of kilometers per household associated with the personal interviews done by the test group. The interviewers doing a large number of personal interviews seem to be able to plan their route so as to minimize the distance between each dwelling. The actual time spent on each interview was slightly less in the test group for both telephone and personal interviews. However, the difference for personal interviews was so small that it was not a factor in the cost savings for that group.

Although the test was not developed to measure data quality, it was possible within the design framework to make some comparisons of nonresponse rates between the test and control groups. As noted earlier the recycling criteria were changed in November 1988, so that telephone cases remained with the central interviewing staff unless the number was no longer valid, instead of being shipped to the local interviewers at mid-week. From January 1987 to October 1988, the nonresponse rates were 8.0% for the test treatment versus 6.1% for the control. Under the restricted recycling from November 1988 to March 1989, the gap in nonresponse rates narrowed to 7.3% versus 6.7%, respectively.

### 4.2 Discussion

It appears from an analysis of the test results to date that data quality is equivalent when we are comparing cold telephone contact with a warm telephoning where the initial interview is face-to-face (First Interview Test). The addition of a centralized component, while still allowing for face-to-face follow-up, (Centralized Warm Telephone Test) tends to decrease the response rates by approximately 1%. When using the centralized approach without allowing for face-to-face follow-up (Telephone Sampling Tests), the response rates decrease by 4% -5%.

Cost savings of \$0.78 per household were observed with the Centralized Warm Telephone methodology which allowed for face-to-face interviews for first month cases and for follow-up of nonresponse. Results from the first Telephone Sampling Test, where face-to-face follow-up was not allowed, show cost savings of \$2.00 per household. Further testing is proposed to shed light on the amount of face-to-face follow-up that is desirable given its costs and its impact on reducing nonresponse rates. (See Section 6).

## 5. Technology Issues

On the whole, findings from the 1987 CATI test were positive. A few key findings are included here. More details can be found in Catlin, Ingram and Hunter (1988).

The principal motivation for CATI is one of improved data quality. Three quality differences were discernable between CATI and paper and pencil methods, with CATI being favoured in each case. First, the overall rate of edit failures detected during post-collection data processing was 50% lower for CATI. Second, the virtual elimination of branching errors under CATI was important. This was particularly the case for certain portions of the questionnaire involving complex branching, which, although they are encountered infrequently, have a bearing on determination of labour force status, and which under paper and pencil are subject to high levels of branching error. Third, the average household size reported under CATI was 3% higher, which represents roughly a 50% reduction in the underenumeration in the LFS relative to the Census. This improvement seems to stem from the enforced probing built into the CATI instrument for additional household members and for persons temporarily away.

It is worth noting that there were no differences in overall costs of the methods.

## 6. Plans for Future Development

While clearly a lot of developmental work remains to be done, the results of testing to date, when pieced together, provide a view of a more telephone intensive design for large scale household surveys in the future. It is likely to be a mixed design with respect to the various design parameters we have considered in this paper - that is, mode of collection, organization of interviewing staff, technology, and choice of frames. The future design will be mixed to a far greater extent than had been envisaged at the outset of the development program.

We conclude with a brief look at developments underway or foreseen on each of these fronts.

Technology As noted in the previous section, the test results reported by Catlin, Ingram and Hunter (1988) were sufficiently positive to conclude that CATI should be adopted for household surveys whenever hardware and software are in place to support it. Operational plans call for acquisition of new hardware in the Regional Offices capable of supporting CATI by 1990. In parallel, as part of a major initiative at Statistics Canada to develop general survey systems, a Data Collection and Capture system with capability for CATI, CAPI and off-line data capture is being developed, with completion targeted for 1991 (Green 1988).

While this software development is taking place, it is planned to study methodological issues related to how to take full advantage of the "computer assisted" aspect of CATI to improve data quality for household surveys. Research is planned in three areas. The first is robust branching. Here it will be attempted to make the questionnaire branching less sensitive to response errors to a single question, by developing branching rules contingent on responses to multiple questions. Another area for research is the issue of dependent versus independent interviewing. Dependent interviewing - by which we mean merely the probing for changes as opposed to independently re-asking questions in later months - is currently used for items such as industry and occupation and name of employer. The research will address the question of whether dependent interviewing should be extended to other items, such as duration of job search. A third area of research is into the feasibility of on-line computer assisted coding of

industry and occupation. Coding these items on-line would seem to have potential for significant data quality gains. The system would be designed to probe the respondent to provide additional information or to choose between competing descriptions. The initial feasibility study would investigate whether the Automatic Coding and Text Recognition (ACTR) system developed at Statistics Canada (Wenzowski 1989) could be adopted to a computer assisted, prompting mode of operation. If not, systems developed elsewhere would be looked into.

The time frame for availability of hardware, software, and conclusion of the above research places a realistic time frame for implementation of CATI to be the mid 1990's - possibly as early as 1993. Its adoption in the LFS would likely coincide with the introduction of a redesigned sample following the 1991 Census - which would take place in 1995.

Frame The timing for implementation of frame and sample design changes in the LFS will be tied to its post-1991 Censal redesign. In urban areas, plans are being formulated for evaluating the use of a complete list frame of dwellings with telephone numbers for dwellings whose occupants have published numbers. This frame will be known as the Address Register (AR), and it is initially being developed for use in the 1991 Census of Population (Drew, Royce, Van Baaren 1989). An AR based design in urban areas would probably consist of an unclustered telephone sample and a clustered sample of dwellings for which telephone numbers are not available. Depending on the rate of undercoverage of dwellings on the Address Register, the design may call for a separate coverage improvement sample. This procedure could work following an approach of half-open intervals, or could be comprised of a small area sample matched to corresponding portions of the Address Register.

In rural areas, a dual frame design would appear to be needed, with an area frame and a telephone list frame, and research efforts will be focussed on this. Circumstances capable of changing this could be the development of equivalents to civic addresses in rural areas, or increases in telephone coverage rendering dual frame approaches unnecessary.

Collection and Organization Unlike CATI and the frame developments, changes with respect to mode and organization of data collection can take place within a much shorter time frame. As was described, testing to date has looked separately at cold telephone contact under local organization of data collection, and warm telephoning under mixed local and central organization. Consideration is being given to a further test to examine the implications again of a mixed organization, but for the case where a greater proportion of interviewing is done by the central staff - through replacement of warm telephoning with cold telephoning wherever possible (i.e. for the area frame, wherever telephone numbers can be obtained by matching to telephone company lists). Unlike in the Centralized Warm Telephone Test which was designed to measure costs, in the new test both the test and control samples would be representative of the same areas, in order to permit valid data quality comparisons.

Successful results from this test in terms of cost and data quality could lead to implementation of a mixed collection, mixed organization methodology as early as 1991.

## References

- Catlin, G., Choudhry, H., and Hofmann, H. (1984). Telephone Ownership in Canada. Internal Memorandum, Statistics Canada.
- Catlin, G., Ingrams, S., and Hunter, L. (1988). The Effects of CATI on Cost and Data Quality.
- Drew, J.D., Choudhry, H., and Hunter, L. (1988). Nonresponse Issues in Government Telephone Surveys, Telephone Survey Methodology, R. Groves etc. al Editors, Wiley, New York.
- Drew, J.D., and Jaworski, R. (1987). Telephone Survey development on the Canadian Labour Force Survey, ASA Proceedings, Section on Survey Research Methods.
- Drew, J.D., Royce, D. and Van Baaren, A. (1989). Address Register Research at Statistics Canada. ASA Proceedings, Section on Survey Research Methods to appear).
- Green, I.S., and Wade, M.J. (1988). Data Collection and Capture for Surveys, Symposium 88 - The Impact of High Technology on Survey Taking, Statistics Canada, Ottawa.
- Muirhead, R.C., Gower, A.R, and Newton F.T. (1975). The Telephone Experiment in the Canadian Labour Force Survey, Survey Methodology, Vol. 1, No. 1, pp. 158-180.
- U.S. Bureau of Census (1987). Computer Assisted Interviewing - Research, Development and Implementation, Internal Memorandum.
- Waksberg, J. (1978). Sampling Methods for Random Digit Dialing, Journal of the American Statistical Sssociation, Vol. 73, No. 361, pp. 576-579.
- Wenzowski, M.J. (1988) ACTR - A Generalized Automated Coding System, Survey Methodology, Vol. 14, No. 2, pp. 299-307

### Acknowledgements

The authors gratefully acknowledge the contributions made to the development and implementation of the tests described herein by other team members - Allison Hale, Eddie Ross, Ghislaine Diotte, and Brian Williams. We thank also the Telephone and CATI Steering Committee, chaired by Ray Ryan, for the direction it provides. Comments by Gordon Brackstone and Nanjanma Chinnappa were helpful in finalizing the paper.

**Table 2.1: Summary of Recent Telephone Tests Conducted at Statistics Canada**

Test	Test Site	Test Date	Purpose	Mode of Data Collection	Collection Organization	Frame
Telephone First Interview	Quebec and Ontario	October 85 to Mar 89	Data quality	Telephone (with minimum of personal interviewing)	Local	Area (with linked telephone numbers)
Warm Centralized	Montreal CMA	January 88 to Mar 89	Cost Study	Telephone and personal	Local and Central	Area
Telephone Sampling - 2	Alberta and Nova Scotia	July 88 to March 89	Data quality, Cost	Telephone	Central	List (with MO. sample updates
Telephone Sampling - 1	Quebec and Ontario	Oct 85 to Sept. 86	Data quality, cost	Telephone	Central	REE and List
Computer Assisted Telephone Interviewing (CATI)	Ontario	March 86 to November 86	Data quality, cost	Telephone	Central	Telephone (RDO)

**Table 3.1 Non-telephone Hhlds by Province (%)**

	1976	1981	1985	1987
Canada	3.5	2.4	1.8	1.5
NFLD	10.0	6.0	5.1	3.6
P.E.I.	--	--	--	--
N.S.	7.5	4.6	3.5	3.2
N.B.	5.8	5.3	5.3	3.3
Quebec	3.3	2.1	1.6	1.5
Ontario	2.5	1.9	1.0	1.0
MAN.	4.1	2.3	2.7	2.4
SASK.	3.6	2.5	2.3	2.4
ALTA	3.0	2.4	2.0	1.8
B.C.	4.2	2.8	2.4	1.3

Source: Estimates from Household Facilities & Equipment Survey

**Table 3.2: Labour Force Characteristics of Non-telephone Universe**

Characteristics	Non-Telephone Universe		Total
	Telephone Universe	Telephone Universe	
Unemployment rate (%)	9.1	29.0	9.3
Participation rate (%)	65.4	51.2	65.1
Employment/population (%)	59.5	36.3	59.1
Population 15+ (000's)	18,824	345	19,169

Source: Estimates from labour force survey (September 1985)

**Table 3.3: Telephone Sampling Test  
(October 1985 - September 1986)  
No Longer Numbers by Panel (%)**

Month in Sample	RDD Sampling	List Sampling
1	2.1	5.2
2	3.1	6.3
3	3.7	6.7
4	4.1	7.6
5	4.6	7.7
6	5.6	7.1
<b>Overall</b>	<b>3.9</b>	<b>6.8</b>

**Table 4.3: Telephone Sampling Test  
(October 1985 - September 1986)  
Design Effects**

Prov	Design	Unemployed	Employed
4	LIST	1.2	1.3
	RDD	1.1	1.4
	LFS	1.6	1.5
5	LIST	1.3	1.4
	RDD	1.1	1.2
	LFS	1.4	1.1

**Table 3.4: Telephone Sampling Test (July 1988 - March 1989)  
Nova Scotia Telephone List Population**

Month	Retained from Previous Month	Lost from Previous Month	Additions from Previous Month	Total Listed Numbers	Percent Loses	Percent Gains
Sep	--	4,842	5,036	301,493	--	--
Oct	296,942	4,551	5,352	302,294	1.5	1.5
Nov	296,739	5,555	7,414	304,153	1.8	2.5
Dec	279,657	4,496	6,741	306,398	1.5	2.2
Jan	302,036	4,362	5,881	307,917	1.4	1.9
Feb	303,933	3,984	5,073	309,006	1.3	1.7
Mar	306,230	2,776	3,307	309,537	0.9	1.1

**Table 4.1: Telephone First Interview Test (October 1985 - March 1989)  
Estimate for Test Treatment as  
Percent of Estimate for Control Treatments**

Characteristic	Ontario		Quebec	
	Percent	T-Statistic	Percent	T-Statistics
Employment	98.5	- 1.22	97.2	- 1.64
Unemployment	96.3	- 0.94	111.8	- 2.27*
Not in LF	101.1	0.88	98.2	- 0.63
Pop 15+	99.2	- 0.86	98.5	- 1.31
Pop in Hhld = 3+	97.8	- 0.76	94.1	- 1.26
Pop. in 1 person Hhlds	100.6	0.25	104.1	- 0.93
Pop. in 2 person Hhlds	100.1	0.26	101.0	0.43
Pop. in 3 person Hhlds	98.6	- 0.56	94.7	- 1.28
Emp. Male 15-24	95.2	- 0.75	88.7	- 2.00*
Emp. Male 25+	98.5	- 1.11	96.7	- 1.54
Emp. Female 15-24	99.0	- 0.11	111.5	- 1.44
Emp. Female 25+	99.2	- 0.65	97.1	- 1.11
Unemp. Male 15-24	105.6	0.53	119.0	1.53
Unemp. Male 25+	94.5	- 0.80	96.7	- 0.31
Unemp. Female 15-24	99.6	- 0.15	119.4	1.30
Unemp. Female 25+	90.9	- 1.22	123.9	2.71**
Not in LF Male 15-24	99.8	- 0.07	99.9	0.47
Not in LF Male 25+	105.6	1.57	101.3	0.07
Not in LF Female 15-24	101.9	0.56	95.3	- 0.56
Not in LF Female 25+	99.2	- 0.14	97.3	- 0.92
Pop. Male 15-24	97.2	- 0.49	95.2	- 0.73
Pop. Male 25+	99.9	- 0.09	97.7	- 1.49
Pop. Female 15-24	99.9	0.16	105.4	0.95
Pop. Female 25+	98.9	- 1.28	98.4	- 1.24

Notes 1. \* -> significance at the 5% level  
2. \*\* -> significance at the 1% level

**Table 4.2: Telephone Sampling Test  
(October 85 - September 86):  
Unemployment and Participation Rates**

Prov	Design	Unemployment Rate		Participation Rate	
		(S.D.)	(S.D.)	(S.D.)	(S.D.)
Quebec	LIST	12.3	(0.78)	64.1	(1.08)
	RDD	13.0*	(0.88)	62.8	(1.28)
	LFS	10.9	(0.27)	63.4	(0.29)
Ontario	LIST	7.3	(0.59)	69.0	(1.11)
	RDD	7.9	(0.63)	69.0	(1.18)
	LFS	6.9	(0.16)	69.0	(0.20)

\* Significant difference between RDD and LFS Unemployment rates for Quebec

**Table 4.4: Telephone Sampling Test  
(July 1988 - March 1989)  
Response Results**

Site	Month	Sampled Numbers	In-Scope Numbers (% of 3)	Refusal	Nonresponse Total (% of 4)
(1)	(2)	(3)	(4)	(5)	(6)
Nova Scotia	Jul 88	351	93.4	0.9	8.5
	Aug 88	585	91.8	1.7	9.9
	Sep 88	708	90.4	2.7	10.0
	Oct 88	711	90.6	3.0	10.6
	Nov 88	590	89.7	2.5	10.4
	Dec 88	710	90.4	2.5	10.0
	Jan 89	710	90.6	3.0	10.9
	Feb 89	714	91.2	2.8	11.1
	Mar 89	714	92.3	2.9	9.9
	<b>Total</b>	<b>5,793</b>	<b>91.0</b>	<b>2.5</b>	<b>10.2</b>
	Jul 88	282	92.6	2.3	9.2
	Aug 88	469	90.6	2.8	7.5
	Sep 88	572	90.0	3.7	11.7
	Oct 88	583	90.4	3.4	11.6
	Nov 88	576	90.8	4.4	15.3
	Dec 88	574	90.2	4.6	9.5
	Jan 89	574	90.4	4.2	9.2
	Feb 89	571	92.5	2.3	5.9
	Mar 89	564	93.1	3.4	5.9
	<b>Total</b>	<b>4,765</b>	<b>91.1</b>	<b>3.5</b>	<b>9.6</b>

**Table 4.5: Centralized Warm Telephone Test  
Cost Savings Per Household**

Survey	Tel Hhlds	Pers Hhlds	All Hhlds	% Saving Due to Test
Jan 88	.35	1.88	.66	12.98
Feb 88	-.17	1.65	.19	3.19
Mar 88	.52	1.43	.71	13.76
Apr 88	.35	-.94	.09	1.69
May 88	.00	.10	.02	.40
Jun 88	.17	.16	.17	3.53
Jul 88	.69	2.29	1.10	18.22
Aug 88	.87	3.73	1.41	27.48
Sep 88	.35	2.23	.74	13.47
Oct 88	1.21	5.57	2.05	33.97
Nov 88*	-	-	-	-
Dec 88	.17	2.64	.59	12.92
Jan 89	.35	1.02	.50	7.88
Feb 89	.87	2.06	1.11	19.23
Mar 89	1.04	2.86	1.38	23.46
<b>Total</b>	<b>.52</b>	<b>1.77</b>	<b>.78</b>	<b>14.16</b>

\* data not available