Statistics Canada | Statistique Canada

Methodology Branch

Social Survey Methods Division

Direction de la méthodologie

Division des méthodes d'enquêtes sociales

Canada

[redacted]

WORKING PAPER NO. SSMD-94-002-E
METHODOLOGY BRANCH

CAHIER DE TRAVAIL N0. DMES-94-002-E
DIRECTION DE LA MÉTHODOLOGIE

NONRESPONSE RATES AND TRENDS IN HOUSEHOLD SURVEYS

by

H. Bérard and S. Kumar
February 1994

ABSTRACT

The study of nonresponse trends and reasons for it leads to a better understanding of the phenomenon thereby facilitating the identification of remedial measures to control or reduce it. In this paper these trends are examined for several household surveys. Components, seasonal patterns, and longitudinal evolution of nonresponse for the Labour Force Survey (LFS) are analyzed. The annual nonresponse rates for LFS and other households surveys are compared and possible explanations for these differences are presented. Some of the current research activities at Statistics Canada on this topic are described.

RÉSUMÉ

L'étude de la non-réponse (et des raisons qui la justifient) permet une meilleure compréhension du phénomène de la non-réponse et facilite ainsi l'identification de mesures qui peuvent être prises pour la réduire ou la contrôler. Dans cet article, les patrons de non-réponse de plusieurs enquêtes ménages sont examinés. Les composantes de la non-réponse, ses patrons saisonniers ainsi que la non-réponse des ménages longitudinaux de l'Enquête sur la population active (EPA) sont étudiés. Les taux de non-réponse annuels de plusieurs enquêtes ménages, dont l'EPA, ont été comparés. Des raisons ont été proposées afin d'expliquer les différences observées dans les taux de non-réponse. Des travaux de recherche récents dans le domaine de la non-réponse à Statistique Canada sont présentés.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The phenomenon of nonresponse is common to all sample surveys and censuses. However, its impact on the estimates and their reliability varies from survey to survey. The magnitude of nonresponse and reasons for it are diverse and complex. It may be due to the subject matter of the survey, data collection methodology, complexity of the survey instrument, an individual's inability to provide the data and his attitude to the survey sponsoring organization, concerns about confidentiality, and sometimes outright refusals. Various survey organizations have developed techniques to deal with this problem at various stages including the data collection and estimation stages. The information on nonresponse (complete or partial) is generally produced for all surveys. This compilation has included the production of nonresponse rates by various categories, and generating profiles of nonrespondents using partial responses (if available), and other sources of data. The overall objective of nonresponse related activities, besides informing users of data quality, has been to undertake and develop procedures that minimize the degree of complete nonresponse at the data collection stage. Various efforts in this regard have included the development of policies on reducing response burden, publicity campaigns aimed at specific groups, and the use of incentives. Attention has also been focused on the development of better data collection instruments, improvement of interviewer training procedures, strengthening and monitoring operational procedures. Another important aspect of work for this problem has been directed at making meaningful use of partial responses. This has led to the development of various imputation techniques for item nonresponse. However, some of the nonresponse definitions, on which these policies and research are based, have lacked conceptual consistency. These inconsistencies are not only between different survey organizations but also sometimes among various surveys conducted by the same organization. Thus the survey practitioners are unable to exploit fully some of the research results due to these problems. Recently the importance of this topic has led to an increased interest in collaboration and a more integrated research approach among several government statistical agencies resulting in the conduct of international workshops on nonresponse at which Statistics Canada has been an active participant.

The resolution of conceptual differences in reporting nonresponse has also received considerable attention at Statistics Canada in recent years [3]. We describe the development of uniform concepts, standards and guidelines for reporting nonresponse data at Statistics Canada in Section 2. These have resulted in Statistics Canada, adopting as policy, standards and guidelines for reporting nonresponse information for units about which data are collected directly. All Statistics Canada surveys are covered by this policy thereby ensuring conceptual consistency of various nonresponse reports. However, surveys based on data collected from administrative sources are excluded from the framework of these standards. In the case of mixed mode surveys (e.g. combining a sample survey and use of administrative records), our standards apply to the non-administrative component of the survey.

In later sections, we analyze nonresponse trends for some important household surveys at Statistics Canada. These are the Labour Force Survey (LFS), the Family Expenditure Survey (FAMEX), the Food Expenditure Survey (FOOD), the Survey of Consumer Finances (SCF), and the Household Facilities and Equipment Survey (HFE). A brief description of LFS is presented in Section 3. Section 4 examines the trends in LFS nonresponse rates. These are examined and discussed from various perspectives, i.e. comparing overall rates, analyzing rates by reasons for nonresponse and sample structure, and investigating seasonal behaviour as well as longitudinal evolution of nonresponse. Nonresponse trends for the remaining surveys are presented in Section 5. Section 6 describes future direction of some of the research work to be done on nonresponse at Statistics Canada. Conclusions are presented in section 7.

2. NONRESPONSE REPORTING STANDARDS AT STATISTICS CANADA

Recent research work has resulted in the formalization of concepts for reporting nonresponse information for various surveys and censuses conducted by Statistics Canada. Its Methods and Standards Committee (a policy-making body on statistical methods) has approved these concepts and their utilization by all survey programs for reporting nonresponse information. The definitions and other related concepts are contained in a Statistics Canada document entitled "Standards and Guidelines for Reporting of Nonresponse Rates" [11]. The nonresponse information about sampled units is required to include details such as the reasons for nonresponse. Furthermore, the counts are to be weighted and unweighted, so that weighted and unweighted rates can be produced at both the data collection and estimation stages depending on whether interest focuses on the effect on estimates or on the control of nonresponse at the collection stage. For additional details see [11].

This policy is being implemented in the following manner. Program managers are responsible for providing, on a regular basis, nonresponse data based on the definitions in [11]. It also contains the degree of detail and the format in which these data are to be provided. These data are to be stored in one of two Statistics Canada's internal databases. These are: Central Nonresponse Information Database (CNID), and Statistical Data Documentation System (SDDS). CNID is a new repository of detailed information on nonresponse for major recurring surveys. The starting date for data input into CNID is the survey reference year 1993. The 1993 reference year was chosen in order to give project managers time to set up coding and computer systems if those in place did not conform to these standards. The development of CNID is already in progress and it is being set up in a SAS environment. Although several database management systems were available, it was decided to use SAS mainly because of its extensive use at Statistics Canada. This is also a good choice because it will allow extensive research and analysis with varying objectives to be carried out by different analysts.

While CNID establishes a detailed nonresponse database for major recurring surveys and censuses, SDDS will continue to maintain basic nonresponse rate information for all surveys. SDDS is Statistics Canada's repository of documentation on all statistical surveys. It includes information about content, methodology, frequency, and size as well as summary quality indicators. The basic nonresponse measures that it contains are now being brought in line with the new standards.

Any unforeseen problems in reporting nonresponse data based on the guidelines will be documented. We will subsequently review the guidelines in the light of these problems to see if any refinements are necessary.

3. LABOUR FORCE SURVEY (LFS)

The LFS is conducted monthly and provides estimates for labour force characteristics such as employment, unemployment, participation levels and rates (for more survey design and methodology details see [12]). The sample follows a rotating sample design. It is composed of six rotation groups that are of equal size. The selected dwellings remain in the sample for a period of six consecutive months. Every month, one sixth of the sample is dropped and is replaced by an approximately equal number of new dwellings. To keep track of the nonresponse status of LFS households, every nonrespondent household is coded according to its reason for nonresponse. These four mutually exclusive and exhaustive reasons, along with their codes in parentheses, are:

(1) temporary absence of dwelling occupants during the interview week (T);

(2) contact with dwelling could not be made though occupants were present during the interview week (N);

(3) refusal (R), and;

(4) other reasons (O). Examples of other reasons are: (a) interview not conducted due to bad weather, (b) no interviewer available, (c) survey form arrived too late for processing or other operational problems, and (d) interview not conducted due to special household circumstances, such as sickness, death or language problems.

4. LFS NONRESPONSE RATE

Overall or total LFS nonresponse rate is the ratio of the number of LFS nonrespondent households to the total number of households in the sample. We examine these rates from various perspectives. These are: (1) overall nonresponse rate and its composition, i.e. which components contribute to the nonresponse and their relative importance; (2) seasonal pattern of nonresponse rates; (3) longitudinal evolution of nonresponse; and (4) annual trends in nonresponse.

4.1 Nonresponse Rate Components

The analysis of the Canada level LFS nonresponse rates, presented in this section, is based on ninety-six months (May 1985 to April 1993) of data. The total nonresponse rate (the average of ninety six monthly nonresponse rates) and its breakdown by "number of months in the sample" and by reason for nonresponse (code) is presented in the Table 1.

Table 1. Nonresponse Rate Components (Rate in %).[1]

Code	Number of Months in the Sample						Total
	1	2	3	4	5	6	
T	0.41	0.29	0.27	0.26	0.25	0.23	1.71
N	0.48	0.26	0.21	0.20	0.19	0.17	1.51
R	0.20	0.18	0.20	0.22	0.23	0.23	1.26
O	0.09	0.06	0.06	0.06	0.06	0.06	0.39
Total	1.18	0.79	0.74	0.74	0.73	0.69	4.87

[1] Based on averages of nonresponse rates from May 1985 to April 1993.

Based on the above table, we observe that:

(1) the overall nonresponse rate is 4.87%;

(2) T's account for (1.71/4.87) or 35% of the total nonresponse; similarly N and R respectively account for 31% and 26% of the nonresponse while the contribution of O to the total nonresponse is only 8%;

(3) births households (households in the sample for the first time) which represent roughly 17% of the sample account for (1.18/4.87) or 24% of total nonresponse; the nonresponse in each of the following months is virtually the same with each month accounting between 14% and 16% of total nonresponse;

(4) the relatively high contribution to total nonresponse by households that are in the sample for the first time seems to be mostly due to T's and N's; the nonresponse components, attributable to T and N, decrease by about 40% after the first month (this may be due to success in establishing contact with households after the first month);

(5) there is almost no difference in the refusal rates (R) between the households in the sample for the first month and others in subsequent months.

4.2 Seasonal Patterns

Some factors that can cause variations in nonresponse are seasonal in nature. Two such factors are climate and vacation patterns. We examined nonresponse rates to determine if these follow any seasonal pattern. Since 1982, July has been the month with the highest total nonresponse rate for the LFS. Generally, higher levels of nonresponse are observed between February and May while lower levels are observed in October (Figure 1).

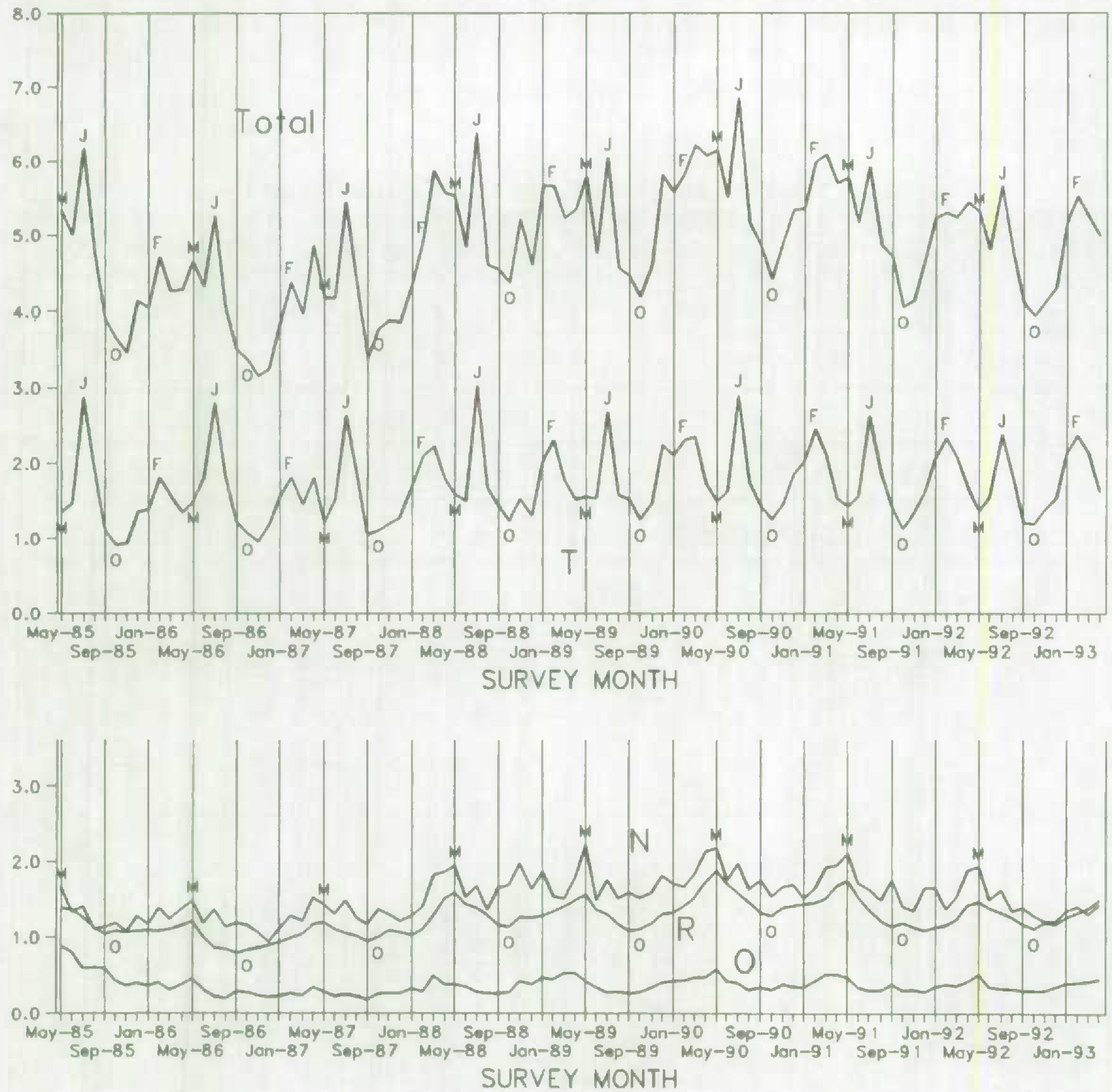

Figure 1. Seasonal Patterns of LFS Total Nonresponse and Components T, N, R and O.

Various contributions based on nonresponse codes are presented in Figure 1. It is interesting to note that in recent years the percentage of T's in February has increased while the percentage of T's in July has decreased. They both appear to have reached a similar level in 1992 (Figure 1). The fluctuations observed in the nonresponse components indicate that the major cause of nonresponse varies from month to month (Figure 1). In order to determine the relative contribution of each component to total nonresponse, averages of nonresponse rates were computed for each of the twelve calender months over an 8 year period (Table 2). A graphic representation of total nonresponse rate is presented in Figure 2.

Table 2. Nonresponse Rates (%) by Survey Month and Code.

Code	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	AVG
T	1.88	2.19	1.97	1.64	1.44	1.58	2.74	1.79	1.31	1.16	1.29	1.56	1.71
N	1.45	1.46	1.54	1.74	1.87	1.47	1.61	1.35	1.45	1.39	1.40	1.43	1.51
R	1.19	1.24	1.32	1.46	1.51	1.37	1.27	1.18	1.09	1.10	1.16	1.17	1.26
O	0.38	0.41	0.44	0.46	0.53	0.42	0.35	0.34	0.35	0.33	0.35	0.36	0.39
Tot	4.90	5.30	5.27	5.30	5.35	4.84	5.97	4.66	4.20	3.98	4.20	4.52	4.87

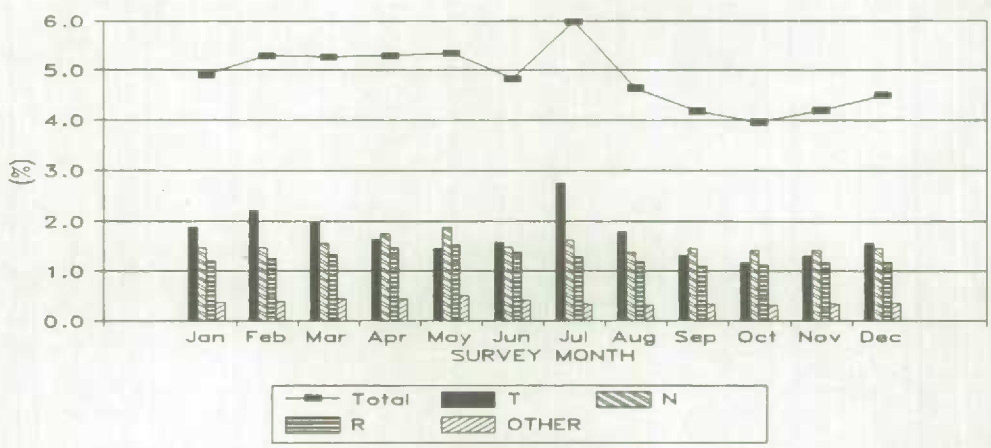

Figure 2. Nonresponse Rates (%) by Survey Month and Code.

We observe that total nonresponse rates are less than 5% for the month of June and the period from August to January. For all other months, the rate is higher than 5%. The highest nonresponse rate is in the month of July, perhaps because of a peak in vacations at that time. The higher nonresponse for July can be attributed in large part to an increase in the number of temporary absent households (T) where the rate attributable to T is 2.74% as compared to an annual average (for T based codes) of 1.71%. High rates for T are also seen in January, February and March when a large number of people travel south to avoid the harsh weather.

For "not at home" (N) category, nonresponse appears to fluctuate from month to month. However, there is a small increase in N's in April and May, where respectively 1.74% and 1.87% of sampled households were not at home during interview week. In other months this value range from 1.35% to 1.61% and the average over all months is 1.51%.

The lowest refusal (R) rates are in September where 1.09% of the households refused to be interviewed. The refusal rate increases slowly from September to May where it reaches a maximum value of 1.51%, then it decreases from May to September. However, we have no explanation for this phenomenon.

In summary, the peak in total nonresponse in July appears to be mostly due to T while the May maximum seems to be due to a mix of N, R and O. Although there is a decrease in T's from February to May, total nonresponse remains above 5% due to an increase in N, R and O during that period.

4.3 Longitudinal Evolution of Nonresponse

The longitudinal evolution of nonresponse can be investigated by examining the nonresponse behaviour of the same group of households (single rotation group) during its life in the sample. The term "longitudinal households" will also be used to refer to the households in the same rotation (birth) group. There are 12 distinct rotation groups in a calendar year. We will evaluate the potential seasonal effect of a group's sample entry month on nonresponse. For this purpose, we introduce the following terminology and notation.

(i) The term "b group" will refer to the rotation group that entered the sample for the first time in the calender month b. Here b=1 or January, 2 or February,.., 12 or December.

(ii) $Q_b(i,c)$ will denote the average nonresponse rate (averaged over the 8 years study period), attributable to code c (c= T, N, R, O) for the "b group" that is in the sample for the i-th time, i=1,2,..,6. For example $Q_{10}(4,T)$ denotes the average nonresponse rate, attributable to "T", for October birth (b=10) group during its 4-th (i=4) month in the sample, i.e. survey month January.

(iii) The average total nonresponse rate (average over the years for the study period) accounted for by the "b group" during its i-th time participation in the sample is denoted by $Q_b(i,.)$.

It was pointed out in section 4.1 that total nonresponse for the full sample generally decreases as the sample gets older. It was also shown in section 4.2 that nonresponse rate reaches relatively higher level in February, May and July. The longitudinal households nonresponse trends were analyzed to determine their contribution to these nonresponse patterns. The various $Q_b(i,.)$ values, for different rotation group (b values) are presented in Figure 3.

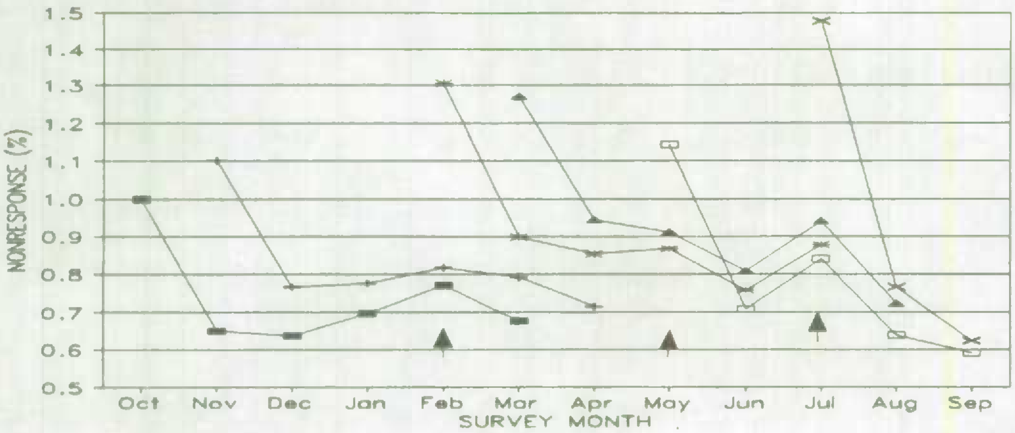

Figure 3. Total Nonresponse $Q_b(i,.)$ for October, November, February, March, May and July groups, for $i=1,2,...,6$.

The arrows indicate increase in LFS total nonresponse rates observed in February, May and July attributed to different nonresponse components (section 4.2). Based on these longitudinal nonresponse patterns we observed that:

(1) "July group", $Q_7(1,.)$, has the highest nonresponse rate of 1.48% during its birth month (July survey) followed by "February group" with a rate of 1.31% in its birth month;

(2) total nonresponse for longitudinal households does not necessarily decrease from the birth month to the 6-th month in the survey; nonresponse rate increased between June and July survey months for the four rotation groups that were introduced into the sample from February to May (note that only 6 rotation groups are shown in Figure 3);

(3) nonresponse for the July group dropped considerably in the final survey month, i.e. December survey; however, the drop in nonresponse from first time to last time is much smaller for the February group; this group shows the highest nonresponse rate in its final month in the sample, i.e. in July;

(4) the average nonresponse for all groups in their final month in the sample, $Q_b(6,.)$, is 0.69%; groups introduced into the survey from April to October have $Q_b(6,.)$ values

below 0.69%, while groups introduced into the survey from November to March have $Q_b(6,.)$ values higher than 0.69%.

Based on these observations it appears that the nonresponse rates $\{Q_b(i,.), i=1,2...,6\}$ are also dependent on the month b in which it was introduced into the survey. In order to examine this nonresponse pattern further, the evolution of nonresponse by reasons (T, N and R) were analyzed.

Figure 4 represents the percentage of sampled households, coded T, for November, December, February, April, June and July groups.

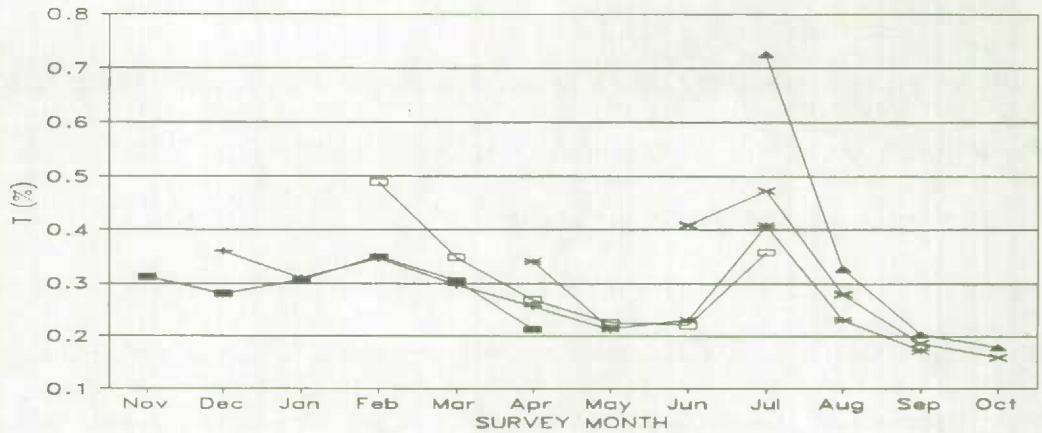

Figure 4. Percentage of "T" households in November, December, February, April, June and July groups.

From this figure we observe that:

(1) the usual decrease in "T", associated with the age of the group (see section 4.1), is disrupted in February and July survey months;

(2) four of the five rotation groups that were introduced into the sample before February show an increase in T's between January and February; all five rotation groups introduced into the sample before July show an increase in T's between June and July (note that only 6 rotation groups are shown in Figure 4);

(3) the July group shows the highest level of T's in its first month in the sample.

It was mentioned earlier that February and July are popular vacation months; this is reflected in larger than usual proportion of T's found during those months in all rotation groups. It is interesting to note that households introduced to the sample in February show a relatively high level of T's in both their birth month (February survey) and final month (July survey) in the sample.

Figure 5 represents the percentage of sampled households, coded N, for December, January, February, March, May and July groups.

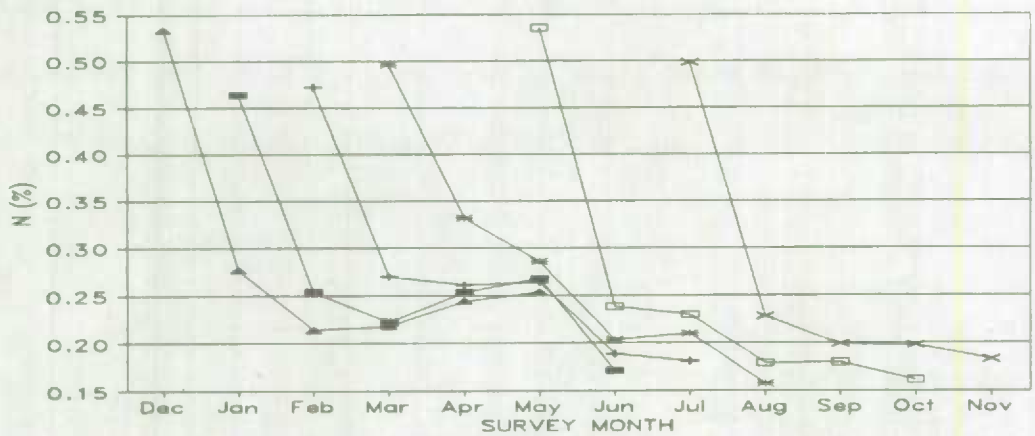

Figure 5. Percentage of "N" households for December, January, February, March, May and July groups.

The nonresponse patterns for households coded N are as follow:

(1) there is a general decrease in N's as the sample get older for groups introduced into the sample between February and November;

(2) for the December and January groups the percentage of N's is decreasing (as expected, see section 4.1) from survey month 1 to 3, however, there is a reversal of this trend in April and May;

(3) in May, all the rotation groups show relatively higher levels of N's; furthermore all

rotation groups, except the birth group, shows similar proportions of N's (around 0.27%);

(4) the May group had the highest percentage of sampled households coded N in its first month in the sample, i.e. May.

The relatively higher level of N's observed in the full sample in May appears to be due to a relatively higher level of N's in all rotation groups. We have no explanation for this phenomena, however, it is interesting to note that all nonbirth rotation groups are equally affected.

When considering an overall sample (section 4.1) refusal rates (percentage of sampled households coded R) remain constant as the sample gets older. In section 4.2 it was shown that refusal rates reach relatively higher values in May. Figure 6 presents the percentage of households coded R for rotation groups introduced to the sample between December and May.

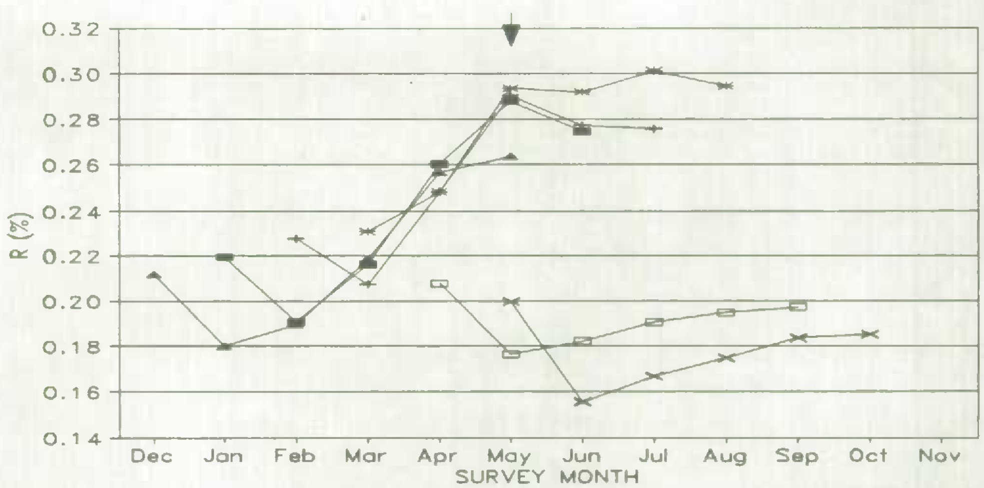

Figure 6. Percentage of Sampled Households Coded R Belonging to Rotation Group Introduced Between December and May.

Refusal rates for longitudinal samples have distinctive patterns:

(1) the refusal rates generally decreases form the birth month to the second month, then increases from the second month to the final month in the sample;

(2) the average refusal rates for the 6-th month in the survey, $Q_b(6,R)$, is usually higher than the rate observed in the first month $Q_b(1,R)$; there are two exceptions to this observation, namely, households introduced in April and May;

(3) households that were introduced into the sample in December, January, February and March show the highest increase in refusals as the sample gets older.

The peak in refusal rates in May in the full sample appears to be due to the relatively high refusal rates of four rotation group, i.e. 4/6 of the sample. It may be pointed out that a 1975 study indicated that around 48% of households which refused the first survey remained refusals in the subsequent months [10].

4.4 LFS Annual Nonresponse Rates: 1966 - 1993

In this section we consider historical trends in LFS nonresponse by comparing annual nonresponse rates. Figure 7 presents these trends.

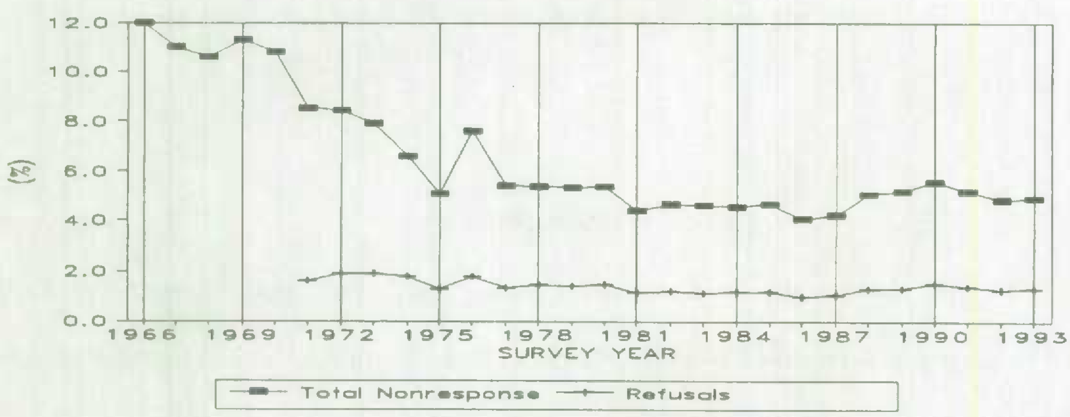

Figure 7. LFS Nonresponse and Refusal Rates (%) by Year.

The LFS annual response rate decreased substantially in the late sixties and early seventies. This decrease in nonresponse is possibly associated with increased emphasis on interviewer training and controlling the quality of interviewers' work. In the early seventies the interviewers manual was revised to make it more detailed and to include explicit procedures for various situations [7]. For example, it included specific and detailed instructions on how to do call backs and how to make contacts. Also, a special Monday follow-up procedure was introduced in an effort to contact No-one-at-home (N) households for July and August surveys.

Telephone interviewing was successfully introduced in almost all Canadian cities by 1975 to replace non births personal interview. Muirhead *et al.* suggested that in agreeing households the risk of nonresponse due to absence are minimized by a more flexible timetable of phone backs.

There was a noticeable increase in the nonresponse rate to 7.6% in 1976, but thereafter the annual nonresponse rate has remained around 5%. The increase in 1976 is coincident with doubling (almost) of the LFS sample necessitating hiring and training of a large number of new interviewers. It is interesting to note that refusal rates have remained between 1% and 1.9% since 1971. Hence, the decrease in nonresponse rate is mostly due to a decrease in other components (T, N, O).

5. OTHER SURVEYS

We will discuss Family Expenditure Survey (FAMEX), Food Expenditure Survey (FOOD), Survey of Consumer Finances (SCF), and Household Facilities and Equipment Survey (HFE) in this section. These four surveys and LFS are based on a common area frame.

The FAMEX and FOOD are intended to generate estimates of the distribution of household expenditures over several goods and services categories [16]. FAMEX has the same reference year as the FOOD but it is conducted in the following year. It is based on recall or memory. It takes several hours for a respondent to complete the FAMEX questionnaire. Use of income/expenditure data records may be necessary in some cases. Although the samples are based on the same frame as LFS, their samples differ completely from that of the LFS. This is done to minimize the response burden on households in the samples. For details see [13, 16].

SCF and HFE are conducted annually as LFS supplements in April and May respectively. SCF provides estimates for annual family and individual incomes whereas HFE provides data on housing and other indicators of the standard of living [14, 15]. Collection of SCF income information sometimes requires the use of financial statements. Four rotation groups of the LFS April sample are used for these two surveys. The reasons for this choice are to link data to LFS and to minimize operational costs.

5.1 FAMEX and FOOD Nonresponse Rates

These surveys have had different geographic area, coverage and content from time to time. The two surveys to date have operated on a four year cycle. In years 1 and 2 the survey is conducted in seventeen cities only. It is followed by a national sample in years 3 and 4. In order to study the trends in a meaningful way, one must take into consideration this cyclic feature of the survey. Table 3 presents nonresponse rates at the Canada level.

Table 3. FAMEX and FOOD Nonresponse Rates (%), 1969-1990*

CITIES			NATIONAL		
REFERENCE	FAMEX	FOOD	REFERENCE	FAMEX	FOOD
1974	24.7	28.4	1969	30.8	34.9
1976	23.0	23.6	1978	27.7	
1978		23.8	1982	18.6	18.6
1984	25.5	16.8	1986	23.4	21.1
1990	28.0	32.0	1992	26.4	20.3

*1980 and 1988 surveys were cancelled

The reference year is that to which the survey data pertain. FAMEX is always conducted in the year following the reference year whereas the FOOD is conducted in the same year as the reference year. The nonresponse rates have varied from year to year, and generally the rates in larger cities tend to be lower than those for smaller cities or rural areas.

It was observed that for the 1990 FOOD, the nonresponse rate had increased considerably as compared to the previous years. It is not easy to identify factors that might have resulted in this increase of nonresponse rate. Some possible explanations for this increase may be the changes in public attitudes to government or changes in the sampling method. It was felt that the 1990 FAMEX might experience similar increases in nonresponse rate. As a possible solution to reverse the trend of increase in nonresponse rate, we considered giving a gift to the sampled households at the time of first contact. Before formulating a policy on the use of gifts for response improvement, it was decided to conduct an experiment and evaluate the results. The sample was divided into three groups, control group (no gift), group with gift 1 (a Statistics Canada Publication " A Portrait of Canada", a book), and group with gift 2 (a clipboard bearing the Statistics Canada logo). The study concluded that these gifts did not have statistically significant effect on the national FAMEX response rate. However, incentives did have a significant effect in some cities [5]. The conclusion, based on this test, was that it was not worth introducing gifts to improve response.

For the 1992 Food Survey, several steps to improve the response rate were taken. A colourful vinyl portfolio, a mechanical pencil and a fridge magnet were added to the survey material provided to households included in the sample.

5.2 SCF Nonresponse rates

SCF produces income related data collected from the household individuals. However, the estimates are produced for different aggregations such as census families, economic families, or unattached individuals. For the definition of these aggregations see [14]. The nonresponse rates presented in Table 4 are computed at the economic family level.

Table 4. SCF Nonresponse Rates (%) 1979-1992

Year	1979	1980	1981	1982	1983	1984	1985
Rate	28.2	34.1	20.0	21.3	20.3	21.6	21.0

Year	1986	1987	1988	1989	1990	1991	1992
Rate	21.9	21.2	27.0	24.5	26.1	25.5	24.1

The average of these nonresponse rates is 24.1%. The rates are about 21% for the period 1981-1987 and above 24% for all other years. Some increase in nonresponse rates in recent years may be attributed to a number of supplementary surveys that have the same sample as SCF thereby increasing the household response burden. Recently a number of steps have been taken with a view to increasing the SCF response rate. These include the introduction of improved follow-up procedures, and simplified manuals for the interviewers.

5.3 HFE Nonresponse Rates

Table 5 presents the HFE nonresponse rates.

Table 5. HFE Nonresponse Rates (%) 1984-1992

Year	1984	1985	1986	1987	1988	1989	1990	1991	1992
Rate	8.0	11.5	10.7	11.3	13.7	14.6	14.2	15.3	13.8

The average of these nonresponse rates is 12.6%. The lowest nonresponse rate (8.0%) was achieved in 1984, the highest (15.3%) in 1991. Note that the rates have been below the average from 1984 to 1987, and above the average in the past five years. Noticeable increase in nonresponse rate were observed in 1985, 1988 and 1989. The increase in nonresponse might be due to the increase number of supplements conducted on the same sample [13].

Note that the nonresponse rates for HFE are lower than those for SCF although both surveys use the same sample. Part of the differences may be due to the differences in the subject matter content of the two surveys.

6. RESEARCH ACTIVITIES

Although there are many nonresponse research related activities at Statistics Canada, we will mention below only a few of them. The development of CNID is receiving high priority and system development is being undertaken. Conceptual problems, if any, will be evaluated and resolved. Household surveys are undergoing redesign following the 1991 Census and are also changing their data collection methodologies. The research on the impact of these changes on nonresponse behaviour will be undertaken and may involve models combining linear and seasonal effects [1, 9].

Recently longitudinal surveys have received considerable attention at Statistics Canada.

Three such surveys are: (1) National Population Health Survey, (2) National Longitudinal Survey of Children, and (3) Survey of Labour and Income Dynamics (SLID).

Nonresponse from the pilot test for SLID was examined to identify its major components. One indicator was to determine the number of follow-up attempts required to obtain a full or partial nonresponse. It was found that, for complete respondents, 31% of the households provided data on first attempt while 19% did so on the second attempt. The average number of attempts to achieve a complete response was about 3.1, while 6.2 attempts were necessary for partial responses. Also it was observed that smaller households tended to have higher nonresponse rates [17].

Research will be undertaken to attempt to understand the underlying nonresponse phenomena for longitudinal surveys. Appropriate steps will be taken to reduce its effect at the data collection and imputation stages.

Another area of ongoing activity has been the work on imputation of missing or inconsistent data. The work continues to involve the introduction and evolution of new techniques. The question of generating variance estimates taking account of imputation is also being addressed [4, 6]. Current research work also includes the modelling of the autocorrelative structure in the original nonresponse data and applying control charts to the residuals [2].

7. CONCLUDING REMARKS

The Labour Force Survey (LFS) nonresponse rate for the last fifteen years has been about 5%. The group of households in the sample for the first time has a relatively high nonresponse as compared to the groups that were sampled in earlier months. The two major reasons for the nonresponse are: (i) temporary absence of households occupants during the interview week, and

(ii) inability to establish contact with the occupants. These two reasons account for about 66% of the nonresponse. These rates also have a seasonal pattern, and are relatively higher in the months of February, May and July. Possible explanations of this phenomenon are the severity of cold weather in February and vacations during the months of May and July. Both of these result in the households occupants being away from their home. Roughly, 1% of sampled households are refusals, this monthly percentage generally does not change.

The reduction in the LFS nonresponse from 12% in 1966 to its current level of 5% can be attributed to several factors. Some of these factors are better interviewer training, introduction of improved operational procedures and controls, and implementation of new data collection methodologies such as telephone interviewing. The introduction of Computer Assisted Interviewing (CAI) is expected to further improve the monitoring and management of data collection operations. It will also be possible to keep track of the number of interview attempts and the status of nonresponse conversion.

The Survey of Consumer Finances (SCF) always has a relatively higher nonresponse rate than for LFS. Some of the possible reasons may be attributed to the nature of data being collected (i.e. income and its sources) and collection and operational procedures. Some individuals are not willing or are unable to provide this data. Recently improved SCF follow-up procedures for contacting "LFS No-one at home" have been introduced. It will help in reducing SCF nonresponse rate. Furthermore the future introduction of CAI will facilitate the monitoring and improvement of follow-up procedures.

The Family Expenditure Survey (FAMEX) and the Food Expenditure Survey (FOOD) also have relatively high nonresponse rates. The length of the Famex questionnaire and maintenance of a Food expenditure dairy may be possible reasons for it. A study on giving gifts with the objective of decreasing Famex nonresponse rate concluded that it did not achieve the desired goal.

ACKNOWLEDGEMENTS

The authors would like to thank G. Brackstone, G.J.C. Hole, F. Mayda and M.P. Singh of Statistics Canada for their constructive suggestions during the progress of the work, as well as Anis Ashraf and K.P. Srinath for their valuable comments that helped improve the quality of this paper.

REFERENCES

1. Binder, D.A.(1991). A Framework for Analyzing Categorical Survey Data with Nonresponse. Journal of Official Statistics, 7, 393-404.

2. Hapuarachchi, K.P., and Wronski, A. (1992). Control Charts for Nonresponse Rates in the Canadian Labour Force Survey. Technical Report, Statistics Canada.

3. Hidiroglou, M.A., Drew, J.D., and Gray, G.B. (1993). A Framework for Measuring and Reducing Nonresponse in Surveys. Survey Methodology, 19, 81-94.

4. Kovar, J., and Chen, E. (1993). Jackknife Variance Estimation Under Imputation: An Empirical Investigation, Working Paper No. METH 93-007 E, Methodology Branch, Statistics Canada.

5. Kumar, S., and Durning, A. (1992). The Impact of Incentives on the Response Rates for FAMEX 1990: An Evaluation, Working Paper No. SSMD 92-001 E, Methodology Branch, Statistics Canada.

6. Lee, H., Rancourt, E., and Särndal, C.E. (1991). Experiments with Variance Estimation

from Survey Data with Imputed Values. Proceedings of the Survey Research Section, American Statistical Association, 690-695.

7. Manuel de l'interviewer de l'enquête sur la population active.

8. Muirhead, R.C., Gower, A.R. and Newton, F.T. (1975). The telephone experiment in the Canadian Labour Force Survey. Survey Methodology, 1, 158-180.

9. Platek, R., Singh, M.P., and Tremblay, V. (1977). Adjustment for Nonresponse in Surveys. Survey Methodology, 3, 1-24.

10. Sugavanam, R. (1975). Non-interview patterns in the Canadian Labour Force Survey. Survey Methodology, 1, 44-58.

11. Statistics Canada (1993). Standards and Guidelines for Reporting of Nonresponse Rates, Methods and Standards Committee.

12. Statistics Canada (1990). Methodology for the Canadian Labour Force Survey, Catalogue 71-526.

13. Statistics Canada (1992). Annual Program Report 1990-1991, Household Surveys Division.

14. Statistics Canada (1991). Income Distribution by Size in Canada, Catalogue 13-207.

15. Statistics Canada (1991). Household Facilities and Equipment, Catalogue 64-202.

16. Statistics Canada (1992). Family Expenditures in Canada, Catalogue 62-555.

17. Statistique Canada (1993). Aspects qualitatifs de la collecte du test 3A de l'enquête sur la dynamique du travail et du revenu, Catalogue 93-07.