6. TECHNICAL ANNEX

The General Social Survey (GSS) was initiated by Statistics Canada in order to reduce gaps in the statistical information system, particularly in relation to socio-economic trends. Many of these gaps cannot be filled through existing data sources or vehicles because of the range or periodicity of the information required, or the lack of capacity of relevant vehicles.

OBJECTIVES

The GSS has two principal objectives, first to gather data on trends in Canadian society over time, and second, to provide information on specific policy issues of interest. To meet these objectives, the GSS was established as a continuing programme with a single survey cycle each year.

The GSS gathers a wide variety of data to meet different kinds of unmet needs for a very broad spectrum of users. To achieve the objectives outlined above, the GSS has three components: Core, Focus and Classification. Core content is directed primarily at monitoring long-term social trends by measurement of temporal changes in living conditions and well-being. Main topics within Core content include health, time use, personal risk, education and work, and family and social support. As all Core content topics cannot be treated adequately in each survey cycle, a single cycle covers a specific topic which recurs on a periodic basis. The Core content of the 1990 General Social Survey, the fifth cycle, was family and friends.

CONTENT

Focus content is aimed at meeting the second objective of the GSS, namely to provide information touching directly on a specific social problem or policy issue such as retirement. In comparison to Core content, Focus is more specific to immediate policy issues. For the fifth cycle of the GSS, there was no Focus content.

Classification content provides the means of delineating population groups and is used in the analysis of Core and Focus data. Examples of classification variables are age, gender, education and income.

The target population of the GSS-90 consisted of all people aged 15 years and over living in the ten provinces of Canada, with the exception of full-time residents of institutions.

SAMPLE DESIGN

The population was sampled using random digit dialling techniques and interviewed by telephone, thus excluding from the sample those people living in households without telephones. These households account for less than 2 per cent of the target population. The sample was allocated to provinces in proportion of the square root of the size of their populations, and to strata within provinces in proportion to their population. As well, Health Canada sponsored a supplementary sample of the elderly (aged 65 and over) which was derived from the Labour Force Survey and the Province of Ontario sponsored an increase in the sample in that province.
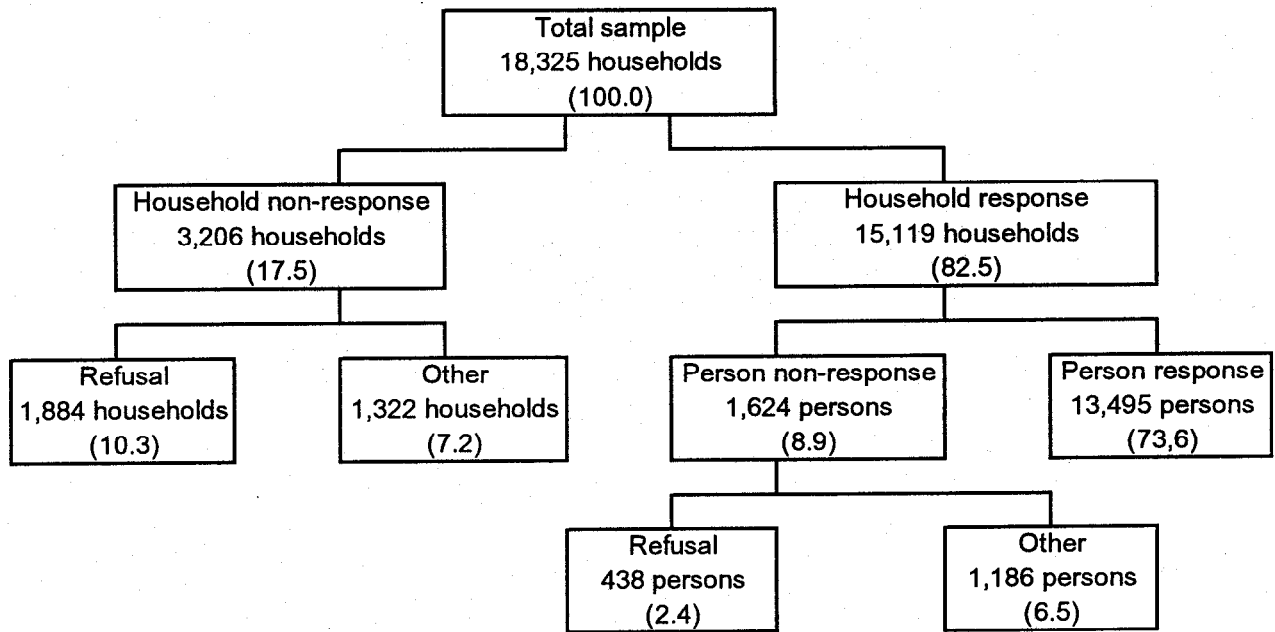
The total sample size of 13,495 people is large enough to allow extensive analysis at the national level, some analysis at a regional

level, but only very limited analysis at a provincial level.

DATA COLLECTION AND FORMS

Data collection took place between January and March 1990. Data were collected from 13,495 respondents aged 15 and over. There were 4,830 non-responses, on a total sample size of 18,325. Non-response at the household level was 17.5 per cent, at the person level 8.9 per cent (figure 6.1).

Figure 6.1
Total sample
Response magnitudes and rates
(in per cents)

```
                        ┌─────────────────────┐
                        │    Total sample     │
                        │  18,325 households  │
                        │      (100.0)        │
                        └─────────────────────┘
           ┌──────────────────────┴──────────────────────┐
┌─────────────────────────┐              ┌─────────────────────────┐
│ Household non-response   │              │   Household response    │
│    3,206 households      │              │    15,119 households    │
│        (17.5)            │              │        (82.5)           │
└─────────────────────────┘              └─────────────────────────┘
    ┌───────────┴───────────┐                ┌───────────┴───────────┐
┌──────────────┐ ┌──────────────────┐ ┌──────────────────────┐ ┌──────────────────┐
│   Refusal    │ │     Other        │ │  Person non-response │ │  Person response │
│1,884 households│ │1,322 households  │ │    1,624 persons     │ │  13,495 persons  │
│   (10.3)     │ │     (7.2)         │ │        (8.9)         │ │      (73,6)      │
└──────────────┘ └──────────────────┘ └──────────────────────┘ └──────────────────┘
                                          ┌───────────┴───────────┐
                                    ┌──────────────┐ ┌──────────────────┐
                                    │   Refusal    │ │     Other        │
                                    │  438 persons │ │  1,186 persons   │
                                    │    (2.4)     │ │      (6.5)       │
                                    └──────────────┘ └──────────────────┘
```

Source : Statistics Canada, 1994.

Data were collected on two forms. The Control Form was used to ensure that the telephone number reached belonged to an eligible household, to record some demographic data for each household member (age, sex, marital status and relationship to a reference person) and to randomly select a respondent aged 15 and over. Only one respondent was selected per household. The Family and Friends Questionnaire, composed of the Core content questions and the Classification content questions, was then administered. No proxy responses to the questionnaire were accepted.

DATA PROCESSING AND ESTIMATION

Data capture personnel in the Statistics Canada regional offices keyed data directly from the survey questionnaires into mini computers. Following the interviews, all questionnaires were captured and put through a computer edit allowing the interviewers to resolve problems (e.g. improper skip patterns or key punch errors). These data were then sent electronically to Ottawa. All survey records were subjected to an extensive computer edit. Partial non-responses, flow pattern errors and abnormally high or low responses were identified. Missing or incorrect data were recoded as « not stated » or, in a very few cases, imputed from other areas in the same questionnaire.

Each person in a probability sample can be considered to represent a number of others in the surveyed population. In recognition of this, and utilising sample design information, each survey record was assigned a weight that reflected the number of individuals in the population that the record represented. These weights were adjusted for non-response and for the differences between the target and the survey population using population counts for the target population. The estimates presented in this report were calculated using the adjusted weights.

It is important to recognise that the figures which appear in this report are estimates based on data collected from a small fraction of the population (roughly 1 person in 2,000) and are subject to error. The error can be divided into two components: sampling error and non-sampling error.

DATA LIMITATIONS

Sampling error is the difference between an estimate derived from the sample and the one that would have been obtained from a census that used the same procedures to collect data from every person in the population. The size of the sampling error can be estimated from the survey results (Statistics Canada, 1994).

All other types of errors such as coverage, response, processing and non-response are non-sampling errors. Many of these errors are difficult to identify and quantify.

Coverage errors arise when there are differences between the target and the survey population. Households without telephones represent a part of the target population that was excluded from the surveyed population. To the extent that this excluded population differs from the rest of the target population, the estimates will be biased. Since these exclusions are small, one would expect the biases introduced to also be small. However, since there are correlations between a number of questions asked on this survey and the groups excluded, the biases may be more significant than the small size of the groups would suggest.

Individuals residing in institutions were excluded from the surveyed population. The effect of this exclusion is greatest for people aged 65 and over, for whom it approaches 9 per cent.

In a similar way, to the extent that the non-responding households and persons differ from the rest of the sample, the estimates will be biased. The overall response rate for the survey was 74 per cent. Non-response could occur at several stages in this survey. There were two stages of information collection: at the household level and at the individual level.

While refusal to answer specific questions was very low, accuracy of recall and ability to answer some questions completely can be expected to affect some of the results presented in the preceding chapters.

Since the survey is cross-sectional, caution is required in making causal inferences about the association between variables. Observed

associations may be a reflection of differences between cohorts, period effects, differences between age groups or a combination of these factors.