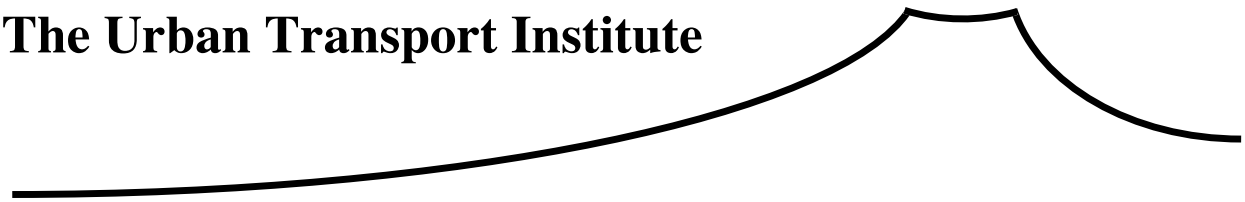
The Urban Transport Institute



TUTI Report 9-2002

A Study of the Potential for Developing an Integrated Vehicle-Use and Fuel Consumption Survey

for Transport Canada

A.J. Richardson and R.K. Seethaler

March 2002

Table of Contents

Abstract

This project considers the problem of obtaining estimates of vehicle use and fuel consumption from a single survey of Canadian vehicles. It reviews previous Canadian and foreign studies in similar areas, and highlights the major design issues to be addressed in the design of the new surveys (tentatively named CIVAFUS – the Canadian Integrated Vehicle and Fuel Use Survey). It then proposes a recommended survey procedure and a set of associated pilot studies to test the applicability of the method.

CIVAFUS builds on some of the methods already used in the CVS, but also contains some significant differences from the current CVS. The similarities are that it would:

- be conducted by StatCan,
- be a survey of vehicles,
- use vehicle registers as the sampling frame, and
- use a trip diary for provinces and quarterly odometer readings for territories.

The differences are that CIVAFUS would:

- be managed by an independent project manager on behalf of the two clients (TC and NRCan),
- use a mailout/mailback survey with full reminder regime and follow-up surveys,
- use a 1-day trip diary,
- survey up to three vehicles in selected households and companies,
- obtain demographic details of passengers in vehicles,
- use imputation more sparingly,
- use fuel consumption models in conjunction with the trip diary data in order to obtain estimates of total fuel consumption, and
- use the same modeling techniques to estimate a range of emissions.

The recommended method is proposed after a detailed consideration of the alternative methods. It is considered that the recommended method allows the current objectives of Transport Canada and Natural Resources Canada to be achieved most efficiently, while allowing scope for future developments and broadening of the environmental objectives of the survey.

1 Introduction

The Canadian federal government has commissioned a number of surveys to estimate vehicle use and on-road fuel consumption over the past 20 years. During the 1980s, Statistics Canada (StatCan or STC) conducted the Fuel Consumption Survey (FCS) for Transport Canada (TC), which used a one-month, or multiple-fill, fuel purchase log to estimate fuel consumption and vehicle-kilometres for the car and light-truck fleet. During the mid-1990s, Statistics Canada updated the Fuel Consumption Survey for Natural Resources Canada (NRCan), through the National Private Vehicle Use Survey (NaPVUS).

Since 1999, the primary means for collecting vehicle use data has been the Canadian Vehicle Survey (CVS), undertaken by Statistics Canada for Transport Canada. The CVS is a continuous survey of motor vehicles of all types (except motorcycles, off-road vehicles, and utility vehicles). Samples of vehicles are drawn quarterly from vehicle registration files, and the registered owner of each sampled vehicle (20,000 samples per year) receives a trip diary and is asked to record all trips and fuel purchases over a one-week period. The separate trip logs are aggregated by Statistics Canada to form estimates of vehicle-kilometres, passenger-kilometres, and fuel purchases for the entire on-road motor vehicle fleet (i.e. cars, trucks and buses).

The 7-day fuel purchase information collected in CVS does not permit fuel consumption to be estimated for each sampled vehicle, which was the intention of the previous FCS and NAPVUS. Conversely, the one-month, or multiple-fill, fuel purchase logs obtained in those earlier surveys did not provide trip characteristics, but merely aggregate vehicle-kilometres for the period.

Recently, discussions between Transport Canada, Statistics Canada and Natural Resources Canada (NRCan) have taken place to consider the potential of expanding or replacing CVS to provide vehicle-specific fuel consumption as well as trip characteristics. One set of options might still undertake two surveys, either in alternating periods, or simultaneously, but using the same sampling methods and sampling frame – these particularly are being examined by Statistics Canada. The other set of options, to combine the surveys in some manner to obtain both trip characteristics and vehicle-specific fuel consumption from each sampled vehicle, are the object of this project.

2 Study Objectives

The objective of this project is to advise Transport Canada on survey methods to obtain from the same respondents the trip characteristics (or good substitutes) currently obtained through CVS and the vehicle's fuel consumption as previously obtained through FCS or NaPVUS. All classes of vehicle covered in CVS are to be considered, though the priority would be for cars and light trucks. Advice on survey design and instruments is required in sufficient detail to enable them to be field tested in subsequent work (probably by Statistics Canada).

Without limiting the generality of the objectives, the proposed work should include the following elements:

- A review of previous vehicle-use and fuel consumption surveys in Canada, including the Fuel Consumption Survey, NAPVUS and CVS. This review should address methodologies employed for each survey, effectiveness of the data collection effort, and the advantages and disadvantages of each approach.
- A review of experience with vehicle-use and fuel consumption surveys in other countries.
- Consideration of means for combining fuel consumption and vehicle-use questions into a single survey. Leading options should be described, with comparative assessments of the relative strengths and weaknesses of the options.
- Recommendations of the options to be field-tested, including preliminary design of the survey methods and instruments.

To distinguish the proposed survey from previous surveys or other currently planned surveys (such as CVS-Fuel and CVS-Trip currently under investigation by STC), the survey options proposed in this report will be referred to as the Canadian Integrated Vehicle and Fuel Use Survey (CIVAFUS). There is no intention that this be the final name of the survey (more creative minds will need to be employed to design a more memorable and pronouncable acronym!), but it will serve to distinguish this survey from others.

3 Review of Previous Canadian Surveys

3.1 Fuel Consumption Survey (FCS)

FCS Name

Fuel Consumption Survey (FCS)

FCS Conducted for

Transport Canada

FCS Conducted by

Statistics Canada

FCS Duration of survey

October 1979 – December 1988

FCS Survey Objectives

The FCS was designed to provide a data base that would allow for the monitoring of average fuel consumption in personal-use passenger cars, light trucks and vans.

FCS Population

When the FCS started in 1979, the target population consisted of "personal-use" passenger cars operated in Canada. Between October 1981 and December 1987, the survey was expanded to include personal-use light trucks and vans operated in Canada. In 1998, the target population was reduced to include only personal-use passenger cars. The target population does not include rental vehicles, taxis, fleet operated vehicles, driver-training vehicles and antique vehicles. "Operated in Canada" means registered in Canada and used sometime during the survey period inside or outside Canada. It does not include vehicles registered outside Canada, but used within Canada sometime during the survey period.

FCS Sample frame

The survey population was defined by the Provincial vehicle registration lists created three months before each quarter.

FCS Sampling unit

The sampling unit from the sampling frame were individual vehicles.

FCS Sample size

The final sample of respondents included 83,115 vehicles (from October 1979 to December 1988). Until December 1985, the number of diary mailouts per month was approximately 1000 for passenger cars and 800 diaries for light trucks and vans. These numbers were cut between

January 1986 through December 1987 to approximately 333 passenger cars and 260 light trucks and vans. In 1988, only the 333 passenger cars were sampled monthly.

FCS Sample design

The FCS used a variable-fraction, stratified random sampling design. The sample was stratified by Province of registration, model year and vehicle weight. The samples were allocated proportionally to the population size of each Province (where population is defined in terms of registered vehicles), subject to minimum sample requirements for each strata. Since the number of telephone interviews conducted each month varied, depending on the number of unsuccessful telephone interviews and on the number of out-of-scope vehicles encountered, sufficient numbers of vehicles had to be selected each month to ensure the daily mailout requirements were met.

FCS Survey technique

The method of data collection used a combination of telephone interview and mailout/mailback Fuel Purchase Diary. Respondents were initially contacted by telephone and asked some screening questions concerning vehicle type and use. Ineligible vehicles (e.g. vehicles not operated during the survey month, vehicles not used for personal use, scrapped vehicles) encountered at this stage were dropped from the sample. Fuel purchase diaries were mailed for each vehicle included in the screened sample. The principal driver of each vehicle was asked to ensure that all fuel purchases were recorded for the survey month.

FCS Survey recording period

Respondents were asked to complete a 1-month Fuel Purchase Diary for the selected vehicle.

FCS Unit of Analysis

The “unit of analysis” about which information was provided was the vehicle selected from the registration lists. The person providing the information was the principal driver of the vehicle.

FCS Information collected

In the Fuel Purchase Diary, the following information is sought about the vehicle

- make of vehicle,
- model of vehicle,
- year of manufacture of vehicle,
- licence plate number of vehicle,
- number of cylinders of vehicle,
- transmission type of vehicle,
- does vehicle have air-conditioning,
- year and month of purchase,
- odometer reading when purchased,
- does odometer register in miles or kilometres.

In the Fuel Purchase Diary, the following information is sought about use of the vehicle

- odometer reading at start of survey period,
- odometer reading at end of survey period.

In the Fuel Purchase Diary, the following information is sought about fuel purchases

- date of fuel purchase,
- odometer reading at the time of the fuel purchase,
- type of fuel purchased,
- price per litre,
- number of litres purchased,
- was tank filled up,
- total price paid for fuel.

FCS Travel recording method

No detailed information was gathered about trips made during the period. Only the odometer readings give any idea of the total amount of travel.

FCS Fuel consumption recording

Fuel purchases were recorded in the Fuel Purchase Diary during the month. When a diary shows at least two fill-ups, it is possible to calculate the precise amount of fuel used between these fill-ups (assuming that both "fill-ups" were indeed complete fill-ups). Since the difference in odometer readings provides the distance between the two diary entries, it is therefore possible to calculate the fuel consumption rate of the vehicle between the fill-ups in litres/100km. It is then assumed that this rate of travelling and fuel consumption applies to the rest of the month, and the results are expanded to a full 31 days.

When there are less than two fill-ups in the month, it is not possible to calculate a fuel consumption rate directly, and the fuel consumption rate must be obtained by imputation, as described below.

FCS Imputation methods

While little detailed information on imputation methods is now available, the STC reports on the FCS state that "incomplete [fuel purchase] diaries were imputed. For example, to obtain estimates on fuel consumed and distance driven, the diaries that were incomplete for part of the survey month were adjusted upwards to reflect the complete month, unless it was known that the vehicle was not operated for part of the month". Furthermore, STC state that "due to the incompleteness of motor vehicle registration files for certain provinces, it was impossible to survey some model year cars in certain quarters. Whenever possible, data related to these quarters were derived from a subsequent quarter."

FCS Consideration of non-sampling errors

While some recognition of the possibility of non-response bias was given in various reports, no specific measurement of the extent of the bias and no specific corrections appear to have been

implemented. For example, with respect to possible non-response bias, STC state that "the estimates were derived on the assumption that the characteristics of the non-respondents do not differ from those of the respondents"

A potentially significant bias is the exclusion of vehicles with less than two fuel purchases (on the grounds that neither distance traveled nor fuel consumed could be calculated). This will result in an underestimation of the vehicle fleet and an underestimation of total fuel consumed. However, since these vehicles are likely to be less intensively used than vehicles that make multiple fuel purchases in the month, it will also result in an overestimation of average fuel consumption rates.

FCS Expansion to population totals

Expansion of the sample data to provide national and provincial population estimates is performed by a series of weights included in the data files. The weighting factors allow the production of annual, quarterly and monthly estimates (depending on the disaggregation level required) which are representative of the Canadian population of private vehicles.

FCS Response rates

The response rates varied from 55% to 75% across the provinces, with an overall average of 65% across Canada. Care should be taken in interpreting these response rates, however, since they apply only to the number of Fuel Purchase Diaries mailed out. They do not include non-contacts and refusals in the telephone survey stage of the overall survey.

3.2 National Private Vehicle Use Survey (NaPVUS)

NaPVUS Name

National Private Vehicle Use Survey (NaPVUS)

NaPVUS Conducted for

Natural Resources Canada

NaPVUS Conducted by

Statistics Canada

NaPVUS Duration of survey

October 1994 – September 1996

NaPVUS Survey Objectives

The objective of NaPVUS was to obtain, from Canadian households, socio-economic information on the household members, their driving habits, the composition of the private vehicle fleet of the household (passenger cars, light trucks and vans), as well as data on the distance traveled and the fuel consumption of a specific vehicle from the household fleet over a one month period.

The survey specifically sought to determine the following:

- the number of vehicles used for personal reasons;
- distances traveled by these vehicles;
- fuel consumption;
- vehicle fleet characteristics;
- household vehicle use;
- vehicle maintenance behavior;
- driver profile;
- the key factors affecting the decision to purchase a vehicle;
- on-road fuel consumption ratio; and
- comparison of on-road versus laboratory-tested fuel consumption ratios.

NaPVUS Population

The population was all Canadian households who own vehicles (passenger cars, light trucks, vans).

NaPVUS Sample frame

The initial sample was drawn from the households who participated to the Labour Force Survey (LFS) during the previous 9-18 months. The NaPVUS sample design is similar to that of the LFS. This latter sample design is briefly described below.

Next, a vehicle was selected randomly from the list of household vehicles eligible to participate to the second phase of the survey.

NaPVUS Sampling unit

Households were used as the sampling unit from the LFS sample frame.

NaPVUS Sample size

During the two year period (Oct 94 – Sept 96), a total of 39,287 households were selected to participate in NaPVUS. Among these, 25,572 accepted the telephone interview, for a response rate of 65 percent.

NaPVUS Sample design

The Labour Force Survey (LFS) is a monthly household survey whose target population is composed of Canadian residents aged 15 years and over; it excludes Yukon and Northwest Territories residents, persons living on Indian Reserves, full-time members of the Canadian Armed Forces, as well as the inmates of institutions.

The LFS used a multi-stage, variable-fraction, stratified random sampling design.

Primary Stratification: by province and then, by economical region within a province.

Secondary Stratification: the economic regions are divided into three categories: the self-representing areas (SRU), the non self-representing areas (NSRU) and the special regions. The

SRU are divided into regions of similar socio-economic situation (using the 1981 Census). The NSRU are divided into urban and rural strata, then according to a number of characteristics of the labour force. The special regions refer to the military bases, the institutions and the remote areas. Note that the last category require special administrative sampling procedures.

Household Selection: For each category of economic regions, a number of geographically contiguous groups of dwellings, or clusters, is selected (generally city blocks or block faces), then a sample of dwellings is selected within them.

NaPVUS Survey technique

The NaPVUS survey consists of two phases. In the First phase, a CATI survey allows the collection of information on the household and its vehicles.

In the Second phase, a diary to note the fuel purchases and odometer readings at the time of the purchases is sent to the household. The diary has to be completed over a one month period for a randomly selected vehicle from the list of household eligible vehicles.

No non-response questionnaire was sent, since by using the Labour Force Survey (LFS) a detailed profile of the non participants to the NaPVUS survey could have been obtained to study the non-response bias. In the end, such a non-response analysis was not conducted.

NaPVUS Survey recording period

Respondents were asked to complete a one-month diary of fuel purchases for the selected vehicle.

NaPVUS Unit of Analysis

The "unit of analysis" about which information was provided was the household selected from the LFS and a random vehicle selected from within the sampled household. The person providing the information in the CATI survey was any knowledgeable adult who could answer the questions about household and vehicle characteristics. The person providing the information in the Fuel Purchase Diary survey was the driver of the vehicle on that trip.

NaPVUS Information collected

In the CATI survey, the following information is sought about the household:

- Age, sex and marital status of household members,
- Relationships between household members,
- Educational status of household members,
- Employment status of household members,
- Annual household income,
- Ownership and availability of vehicles.

In the CATI survey, the following information is sought about all vehicles in the household:

- Make, model and year of manufacture of vehicle,
- Body type of vehicle,
- Number of cylinders of vehicle,

- Is vehicle turbocharged,
- Transmission type of vehicle,
- Does vehicle have air-conditioning,
- Is vehicle front-wheel, rear-wheel or four-wheel drive,
- Was the vehicle purchased new,
- Year of purchase,
- Reason for purchase of vehicle,
- Seasonal usage of vehicle,
- Frequency of use of vehicle,
- Reasons for using vehicle,
- Approximate kilometres driven in past 12 months.

In the CATI survey, the following additional information is sought about the vehicle selected for the fuel purchase diary survey:

- Use of vehicle for business purposes,
- Who uses the vehicle,
- Maintenance of vehicle.

In the Fuel Purchase Diary survey, the following information is sought about the fuel purchases made during the 1-month period:

- date of fuel purchase,
- odometer reading, at time of fuel purchase,
- type of fuel purchased,
- price per litre,
- number of litres purchased,
- was tank filled up,
- total price paid for fuel.

NaPVUS Travel recording method

No detailed information was gathered about trips made during the period. Only the odometer readings give any idea of the total amount of travel. Even then, these reading are only taken at the time of each fuel purchase. Unlike the FCS, odometer readings were not obtained at the start and end of the recording period, so vehicles with less than two fuel purchases have no objective measurement of the distance traveled within the survey period.

NaPVUS Fuel consumption recording

Fuel purchases were recorded in the Fuel Purchase Diary during the month. When a diary shows at least two fill-ups, it is possible to calculate the precise amount of fuel used between these fill-ups (assuming that both "fill-ups" were indeed complete fill-ups). Since the difference in odometer readings provides the distance between the two diary entries, it is therefore possible to calculate the fuel consumption rate of the vehicle between the fill-ups in litres/100km. It is then

assumed that this rate of travelling and fuel consumption applies to the rest of the month, and the results are expanded to a full 31 days.

When there are less than two fill-ups in the month, it is not possible to calculate a fuel consumption rate directly, and the fuel consumption rate must be obtained by imputation, as described below.

NaPVUS Imputation methods

When there are less than two fill-ups in the month, it is not possible to calculate a fuel consumption rate directly, and imputation methods must be used. From the 22,231 households with private vehicles, total of 9,391 (42 percent) diaries with usable information were returned. Among these, 898 vehicles (ten percent) were declared as not having being used during the one-month diary period; these were considered as nil distance traveled and nil fuel consumption. Among the remaining 8,493 vehicles, 7,178 cases had at least two fill-ups recorded in the diaries, allowing the estimation of the fuel consumption rate. For the 1,315 remaining usable cases, the availability of at least two fuel purchases made it possible to derive, through statistical modeling, an estimate of their fuel consumption rate. The final sample does not contain any vehicles with only one fuel purchase; these cases were discarded due to the lack of distance and fuel consumption estimation. The number of vehicles with only one fuel purchase is unknown.

If a vehicle does not make two fill-ups, but does make at least two fuel purchases, then it is possible to calculate the distance traveled between these purchases (and expand this to the month), but it is not possible to calculate the amount of fuel used (as opposed to the amount of fuel purchased). In this situation, an imputation model was developed to relate average fuel consumption rate to characteristics of the vehicle. This imputed rate could then be applied to the calculated distance to obtain the estimated fuel consumption in the month.

Where vehicles make less than two fuel purchases in the month, the problem is more complex since neither distance traveled nor fuel used is able to be calculated (the distance estimation problem could have been solved by asking for odometer readings at the start and end of the month, as well as when fuel was purchased). Work is continuing on imputation methods for this situation, since the omission of such vehicles from the data base will bias the fuel consumption estimates, since these vehicles are likely to be less used than those who make at least two fuel purchases in the month.

NaPVUS Consideration of non-sampling errors

While some recognition of the possibility of non-response bias was given in various reports, no specific measurement of the extent of the bias and no specific corrections appear to have been implemented.

A potentially significant bias is the exclusion of vehicles with less than two fuel purchases (on the grounds that neither distance traveled nor fuel consumed could be calculated). This will result in an underestimation of the vehicle fleet and an underestimation of total fuel consumed.

However, since these vehicles are likely to be less intensively used than vehicles that make multiple fuel purchases in the month, it will also result in an overestimation of average fuel consumption rates.

NaPVUS Expansion to population totals

Expansion of the sample data to provide national and provincial population estimates is performed by a series of weights included in the data files, depending on whether population estimates are required for household, person, fleet or selected vehicle variables. Different weights are included for expansion to yearly (or part-yearly) totals or to quarterly totals.

NaPVUS Response rates

During the 24-month period, a total of 39,287 households were selected to participate in NaPVUS. Of these households, 25,572 accepted the telephone interview, for a response rate of 65 percent. Among these participating households, 3,341 (13 percent) declared that they did not own any private vehicles. In the remaining 22,231 households, a total of 35,918 vehicles were enumerated, for which information on the uses and characteristics was collected. In each of the 22,231 households with private vehicles, one vehicle was randomly selected and specific questions were addressed: number and characteristics of the drivers, the usage of this specific vehicle and the maintenance habits. Drivers of these selected vehicles were then asked to fill out a fuel purchase diary for a period of one month and to return it to Statistics Canada. A total of 9,391 (42 percent) diaries with usable information were returned. The overall response rate was therefore 27.3% (=42% x 65%).

3.3 Canadian Vehicle Survey (CVS)

CVS Name

Canadian Vehicle Survey 2000 (CVS-2000)

CVS Conducted for

Transport Canada

CVS Conducted by

Statistics Canada

CVS Duration of survey

January 1, 2000 – December 31, 2000

CVS Survey Objectives

The objectives of CVS-2000 have been stated by Transport Canada as providing timely annual estimates of characteristics of the complete Canadian vehicle population by Province and Territory. In doing so, the survey attempts to satisfy the following data needs:

Physical Characteristics

- Total road vehicle-km in a year
- Vehicle-km by construction class of vehicle
- Vehicle-km by body type of vehicle
- Vehicle-km for trucks by configuration
- Vehicle-km by fuel type
- Vehicle-km by age of vehicle, by construction class

Operational Characteristics

- Vehicle-km by truck carrier type
- Vehicle-km by truck driver type
- Vehicle-km by type of vehicle and month
- Vehicle-km by hour of day
- Vehicle-km by day of week
- Vehicle-km by commodity

Demographic Characteristics

- Vehicle-km by driver age-group
- Vehicle-km by driver sex
- Vehicle-km by trip length
- Vehicle-km by trip purpose

CVS Population

The population was all registered-vehicle-days in 2000 in Canada. Most registered vehicles would have 365 registered-vehicle-days in 2000, although vehicles not registered for the entire year would have less than 365 registered-vehicle-days in 2000. The population of vehicles for the survey was defined as all vehicles registered in Canada, excluding motorcycles, off road vehicles (e.g., snowmobiles, dune buggies, amphibious vehicles) and special equipment (e.g., cranes, street cleaners, snowplows and backhoes).

CVS Sample frame

The survey population was defined by the 13 jurisdiction vehicle registration lists (ten Provincial and three Territorial Governments) created three months before each quarter. This population differs from the population of interest; e.g. vehicles that were registered less than three months before the quarter begins (or during the quarter) are not included in this quarter's sample.

The incoming lists underwent thorough preparation procedure:

- First, out-of-scope vehicles are removed (trailers, motorcycles, construction equipment, parade vehicles, etc.).
- Second, vehicles with expired registration are removed.

- Then, records with duplicate Vehicle Identification Numbers (VIN) within each list are removed leaving the one updated most recently.
- Next, records with duplicate Vehicle Identification Numbers (VIN) among all lists are removed leaving the one with the most recent update.
- Last, records with irregular data are verified.

CVS Sampling unit

The sampling unit from the sampling frame were individual vehicles. There are four vehicle types.

- Buses
- Light vehicles with gross vehicle weights below 4.5 tonnes,
- Trucks with gross vehicle weights of 4.5 tonnes or more and less than 15 tonnes, and
- Trucks with gross vehicle weights of 15 tonnes or more.

CVS Sample size

A total of 21,662 vehicles out of 17,742,186 from the survey population (0.122%) were drawn for the ten provinces. Another 9,742 vehicles out of 47,278 (20.6%) were included in the sample for the three territories.

CVS Sample design

The CVS-2000 used a multi-stage, variable-fraction, stratified random sampling design. All vehicles from the survey population were stratified (grouped) into 104 strata. First, the vehicles were stratified into four vehicle types (buses, light vehicles, and two groups of trucks) and 13 jurisdictions (ten provinces and three territories). Then, for efficiency of estimates, they were further divided into two vehicle-age strata of newer and older vehicles.

Next, a sample of vehicles (first stage sample) was selected from the survey population for each stratum. To minimize respondent burden, no vehicle was selected more than once during any consecutive four quarters for provinces (two consecutive quarters for territories) and the three characters of the postal code were used to spread the sample over all regions.

Subsequently, seven consecutive days starting within the quarter were randomly assigned (second stage) to each vehicle selected at the first stage. Within each stratum, the first reporting day was evenly spread over the quarter to ensure a uniform number of responses over time and for each day of the week. This step was not applied to the vehicles registered in the three territories since only odometer readings are collected from these vehicles at the start and end of the quarter.

CVS Survey technique

The data collection for the vehicles sampled in the ten provinces is different from the one for the vehicles sampled in the territories.

Provincial collection

The registered owners of the sampled vehicles were telephoned and interviewed (CATI). During the CATI interview, information was collected about each sampled vehicle. Then the respondent was asked to complete a seven-day trip log. If the respondent agreed to complete a trip log, personal information such as name and address were obtained in order to mail out a trip log for the vehicle. The log type depended on the type of vehicle. In all cases, the respondents were requested to record information about all the trips made in the selected vehicle over the assigned seven-day period. If the respondent could not be contacted by phone, a trip log with a short additional questionnaire (to collect some of the information normally collected during the CATI) was mailed out.

Territorial collection

The registered owners of the selected vehicles were mailed postcards and asked to provide two odometer readings, one at the beginning of the quarter and another at the beginning of the next quarter and information about the vehicle status (owned, sold, scrapped).

CVS Survey recording period

Respondents in the Provinces were asked to complete a 7-day trip diary for the selected vehicle.

Respondents in the Territories were asked to provide an odometer reading at the start and end of the quarter for the selected vehicle.

CVS Unit of Analysis

The "unit of analysis" about which information was provided was the vehicle selected from the registration lists. The person providing the information in the CATI survey was any knowledgeable adult who could answer the questions about vehicle characteristics. The person providing the information in the Trip-Log survey was the driver of the vehicle on that trip.

CVS Information collected

In the CATI survey, the following information is sought about the vehicle:

- vehicle type,
- fuel type used,
- distance driven last week,
- some information about anticipated vehicle usage during the following six weeks,
- current odometer reading at the time of the CATI survey, and
- passenger capacity for buses.

In the Trip-Log survey, the following information is sought about the trips made during the 7-day period:

- time and date of the beginning and the end of each trip,
- trip length,
- trip purpose (for the vehicle),

- number and age group of passengers,
- sex and age group of the driver,
- fuel purchases,
- whether dangerous goods were carried,
- number of kilometres traveled on roads with posted speed limit of 80 km/h or more, and
- for trucks, their configuration.

CVS Travel recording method

The definition of what delimits a trip depends on the vehicle type:

For light vehicles, if any of the following events happened:

- a stop of more than 30 minutes
- a change of driver
- a change in the main trip purpose

For heavy vehicles (trucks) weighing 4.5 tonnes or more if any of the following events happened:

- a stop of more than 30 minutes
- a change of driver
- a change of purpose or use
- a change in the truck configuration
- a change in the status of the load from loaded to unloaded or the reverse

For buses, if any of the following events happened:

- a stop of more than 30 minutes
- a change of driver
- a change in the type of bus service
- all the passengers have been dropped off and another passenger trip begins (does not apply to scheduled urban buses)

Using the definitions of a trip given above (which are non-standard definitions of a trip in the travel survey literature), the information described above was requested for each trip made during the seven day period. Space was provided in the trip logs for 40 trips for light vehicles, 49 trips for heavy vehicles and 32 trips for buses during the 7-day period. Apart from the trip purpose (of the vehicle) no information was gathered about activities at the end of trips.

CVS Fuel consumption recording

No specific questions were asked about fuel "consumption". Rather, a question was asked about any fuel purchases made during a trip. From this information, an aggregate estimate of total fuel purchases could be made for the population, by expanding the purchases made within the sample of vehicle-days up to the population of vehicle-days. However, no estimates of fuel consumption rates could be made for any specific vehicle.

CVS Imputation methods

Once all necessary information for the survey was collected, a series of verifications took place to ensure that the records were consistent and that collection and capture of the data did not

introduce errors. Reported data were examined for completeness and consistency using automated edits coupled with manual review. Outliers, i.e. respondents reporting extremely large values, were processed manually. Missing values and data found in error were imputed by another automated system. The system imputed the data using different imputation rules depending on the vehicle, available information and the type of data to be imputed. For example, the data can be imputed based on other responses for the same vehicle or by using data from a similar vehicle. The imputed data were then again examined for completeness and consistency. At the end of this process, every vehicle had seven days of trips. Whole trips or days of trips were imputed for units reporting distance driven but no usable trip-log, using the respondent's approximation of distance driven in a week (obtained in the CATI survey).

The imputation rates in CVS-2000 are relatively high. The following table shows the percentage of vehicle days reported and the percent imputed for each type of vehicle.

	Vehicle days reported			Vehicle days imputed		
PROVINCES	All	0 km	Non 0 km	All	0 km	Non 0 km
Light vehicles	55%	20%	35%	45%	6%	40%
Trucks 4.5t – 15t	69%	49%	20%	31%	8%	22%
Trucks 15t or more	64%	42%	23%	36%	9%	27%
Buses	96%	58%	38%	4%	0%	4%

It can be seen that 45% of all vehicle days for light vehicles were imputed. 41% of vehicle days were imputed where the respondent did not return a trip–log. This imputation is based on the respondent's approximation of travel distance given in the CATI interview and all seven days were imputed for these respondents. The next 3% were imputed from the postcard sent to respondents who refused to accept a trip-log (the postcard collects 8 consecutive odometer readings, one at the start of the week and one at the end of the next 7 days) The last 1% were imputed from respondents that returned their trip-logs but did not provide all seven days.

It can also be seen that the proportion of days with zero travel is much higher in the reported vehicle days than in the imputed vehicle days. One explanation for this is that the vehicles for which trip-logs were imputed were more heavily used than those vehicles for which trip-logs were actually provided, and hence they had less zero travel days.

Alternatively, it could be a result of the imputation process used to impute the trip-logs. The total amount of travel in imputed trip-logs was based on the respondent's estimate of the kilometres traveled in the week prior to the CATI survey. However, by comparing these estimates with the actual distance traveled, as recorded in the trip-logs for those respondents who returned trip logs, STC concluded that the respondent's estimates were an under-estimate of the true distance traveled (as recorded by odometer readings). For example, for drivers of light vehicles in Ontario, they conclude that odometer-reading estimates are 19% higher than respondents' estimates.

However, this result should be viewed with caution since there are several other studies that show the opposite result. For example, the Australian Survey of Motor Vehicle Use (SMVU) showed a 9% over-estimate (i.e. respondent's estimate greater than odometer estimate), the USA Residential Transportation Energy Consumption Survey (RTECS) showed a 11% over-estimate, while the USA National Personal Travel Survey (NPTS) showed a 4% over-estimate. The STC results mean that the respondent's estimate must be increased by 19% to match the odometer estimate. If this result is incorrect (and the SMVU, RTECS and NPTS results are more correct) then this would result in imputed distances that are higher than they should be, resulting in less zero travel days in the imputed data. The results from the imputation process, given that it constitutes such a large part of the final data set, should be re-checked to ensure that no arithmetic and logical errors have been made.

CVS Consideration of non-sampling errors

Four types of non-sampling error were identified by STC; coverage error, response error, non-response error and processing error.

Coverage errors arise when the survey population does not adequately cover the population of interest. As a result, certain units belonging to the population of interest are either excluded (undercoverage), or counted more than once (overcoverage). In addition, out of scope units may be present in the survey population (overcoverage).

The following sources of coverage errors for CVS were observed:

- Errors in the classification variables on the survey may result in either under- or overcoverage of the registered vehicles;
- The sample is drawn from the list created three months prior to the beginning of the quarter. Thus the vehicles registered after the list was created and before the end of the quarter cannot be drawn into the sample;
- A vehicle list from any jurisdiction that was not created on time or did not arrive at all results in even larger undercoverage since an older list has to be used for sampling;
- A vehicle list created early causes overcoverage;
- A vehicle that has been scrapped or salvaged and remained on the list causes overcoverage;
- A vehicle that was registered and subsequently unregistered between two consecutive registration lists causes undercoverage.

Response errors occur when a respondent provides incorrect information due to a misinterpretation of the survey questions or lack of correct information, gives wrong information by mistake, or is reluctant to disclose the correct information. Large response errors are likely to be caught during editing. However, others may simply go through undetected. Few response errors were discovered during editing of the data.

Non-response errors can occur when a respondent does not respond at all (total non-response) or responds only to some questions (partial non-response). These errors can have a serious effect if the non-respondents are systematically different in survey characteristics from the respondents and/or the non-response rate is high. No detailed examination of this differential effect was evident in the STC documentation of CVS-2000.

Processing errors can arise in data capture, coding, transcription, editing, imputation, outlier detection and treatment, and other types of data handling. A coding error occurs when a field is coded erroneously because of a misinterpretation of the coding procedures or a bad judgment (e.g. errors in commodity coding). A data capture error occurs when the data are misinterpreted or keyed incorrectly. Once data are coded and captured, they are subject to editing and imputation of missing or erroneous values. The quality of the data used in the estimation depends on the amount of imputation and the difference between the imputed and the true, but unknown, values. The imputation system could result in bias of the estimates. This can happen due to wrong assumptions or due to inability to impute. For example in the CVS it is impossible to detect missing or entered in error fuel purchases for vehicles that travel only a small distance during the reported week.

CVS Expansion to population totals

Since the survey population differs from the population of interest, several corrections were done to assure that the estimates correspond (as closely as possible) to the population of interest. The sampling weights derived from the sample design were adjusted and improved using updated registration lists. This was possible because, during the passage of time since the sample was selected, a set of prepared vehicle lists was obtained for the beginning and for the end of the reference quarter. To improve the estimates for the vehicles registered in the ten provinces, all the days were further stratified into working days and holidays (or non-working days, including weekends). Second stage sampling weights were adjusted so that every day of vehicle activity within the same stratum contributed with equal weight to the total estimate. The final set of weights reflected as closely as possible the characteristics of the vehicle population during the reference period.

CVS Response rates

Response rates are often seen as an indication of survey quality, since low response rates provide more opportunity for non-response bias to affect the results. Note, however, that low response rates can only have a serious effect if the non-respondents are systematically different in survey characteristics from the respondents.

In CVS-2000, the response rate is defined as the number of vehicle-days for which respondents gave complete or partial (vehicle-kilometers only) answers to the survey divided by the total number of in-sample and in-scope vehicle-days.

Given that CVS-2000 in the Provinces is a two-stage survey (CATI followed by trip-logs), there are opportunities for non-response at each stage. In the table below, the first three columns refer to complete responses to both CATI and trip-logs, while the second three columns refer to complete response to CATI but non-response to the trip-log (thereby requiring imputation of the trip characteristics). Note again that these responses refer to vehicle-days, and hence one vehicle could contribute to both sets of columns. In addition, the "vehicles out of scope" actually refer to "vehicle-days" out of scope, in that one vehicle out of scope contributes 7 vehicle-days out of scope. The same comment applies to "contact made but no data".

PROVINCES	Vehicle-kilometres and trip characteristics reported			Only vehicle-kilometres reported (trip characteristics imputed)			Vehicle out of scope	Contact made but no data
	All	0 km	Non 0 km	All	0 km	Non 0 km		
Light vehicles	38%	14%	24%	32%	4%	28%	4%	4%
Trucks 4.5t – 15t	35%	25%	10%	15%	4%	11%	7%	9%
Trucks 15t or more	38%	24%	13%	21%	5%	16%	7%	12%
Buses	42%	25%	16%	2%	0%	2%	5%	34%

In terms of calculating response rates, separate response rates should be given for each stage, after due allowance is made for sample loss (vehicles out of scope), as shown below. It can be seen that an overall response rate of about 40% is obtained for the trip-logs from the initial sample of in-scope vehicles.

PROVINCES	Response to CATI	Response to Trip-Log after CATI	Response to CATI and Trip-Log
Light vehicles	73%	54%	40%
Trucks 4.5t – 15t	54%	70%	38%
Trucks 15t or more	63%	64%	41%
Buses	46%	95%	44%

The Territories did not have a CATI survey, and only required the provision of two odometer readings to estimate mileage by means of a postcard survey at the start and end of the 3-month recording period. Despite this minimal requirements the response rates in the Territories were quite low, as shown below.

TERRITORIES	Vehicle-kilometres and trip characteristics reported			Vehicle-kilometres reported			Vehicle out of scope	Contact made but no data
	All	0 km	Non 0 km	All	0 km	Non 0 km		
Light vehicles	N/A	N/A	N/A	15%	1%	15%	8%	8%
Trucks 4.5t – 15t	N/A	N/A	N/A	13%	2%	11%	15%	7%
Trucks 15t or more	N/A	N/A	N/A	15%	2%	13%	11%	5%
Buses	N/A	N/A	N/A	13%	1%	12%	14%	6%

4 Review of Related Overseas Surveys

4.1 Australian Survey of Motor Vehicle Usage (SMVU)

SMVU Name

Survey of Motor Vehicle Use 2000 (SMVU)

SMVU Conducted for

Australian Bureau of Statistics (ABS)

SMVU Conducted by

Australian Bureau of Statistics (ABS)

SMVU Duration of survey

1 November 1999 to 31 October 2000. The SMVU has been conducted about every three years since 1963.

SMVU Survey Objectives

The main purpose of the SMVU is to satisfy the information needs of Australian Commonwealth and State government agencies responsible for the allocation of funds for road development, the design and construction of highways, the regulation of road transport operators and accident exposure and energy use analysis. It collects vehicle usage data across all major vehicle types (cars, trucks, buses, etc.) and States, thus providing the direct comparisons of patterns between vehicle types and States that are vital for the above purposes. It is the only source of such information and a time series has now been established. The data are often used as proxies for road use and provide the broad context for project evaluation and measurement of growth rates in road use, mobility and traffic congestion.

The major users of SMVU data are Commonwealth, State and Territory transport authorities, other government departments, academics and private transport research organizations. The data provides a broad context for project evaluation and measurement of growth rates in road use, mobility and traffic congestion and the freight task. It is also used for environmental studies into greenhouse emissions and for a variety of economic modeling purposes.

SMVU Population

The population for the survey contains all registered vehicles, except caravans, trailers, tractors, plant and equipment, vehicles belonging to the defence services and vehicles with diplomatic or consular plates.

SMVU Sample frame

The sampling frame for the SMVU is all vehicles that were registered with a motor vehicle authority for road use at some stage during the 12 months ended 31 October 2000, except caravans, trailers, tractors, plant and equipment, vehicles belonging to the defence services and vehicles with diplomatic or consular plates. Where they were registered as such, vintage and veteran cars were also excluded from the survey. The population was identified using information obtained from the State and Territory motor vehicle registration authorities.

Selections for quarters one and two of SMVU 2000 were made from a population (or frame) of 12.1 million vehicles registered at 31 October 1998. For quarter three, in addition to selections from this population, selections were taken from new motor vehicles registered between November 1998 and the end of December 1999. Vehicles registered at 31 October 1999 and new motor vehicles registered between November 1999 and the end of April 2000 provided the population from which vehicles were selected for the fourth quarter.

SMVU Sampling unit

The sampling unit from the sampling frame were individual vehicles.

SMVU Sample size

For the 2000 SMVU, a sample of approximately 16,000 vehicles was selected to report on vehicle use over a three month period within the reference year 1 November 1999 to 31 October 2000. Of these, 26% were passenger vehicles and motor cycles, 58% were freight vehicles, 11% were buses and 5% were other non-freight carrying vehicles. The sample size was chosen to give a suitable level of precision for estimates of total distance traveled for each State/Territory of registration by type of vehicle category.

SMVU Sample design

The SMVU used a variable-fraction, stratified random sampling design. The population of vehicles was stratified within each State or Territory according to the vehicle description recorded by the registration authority. Each type of vehicle category was further stratified by other characteristics to take account of different usage patterns. These were:

- passenger vehicles according to whether taxis or other passenger vehicles;
- other passenger vehicles for most States, to capital city or rest of State;
- motor cycles according to age;
- buses according to size;
- light commercials and articulated trucks according to age and for most States, to capital city or rest of State;
- rigid trucks according to age and size and for most States, to capital city or rest of State; and

- non-freight carrying trucks according to whether ambulance, hearse, fire engine or tow truck, mobile crane etc.

SMVU Survey technique

The survey methodology is described as 'pre-advice', where owners of vehicles selected in the survey received early advice about their inclusion to encourage record keeping and minimize reliance on recall. These owners were asked to complete two mail questionnaires tailored to their vehicle type. The first, at the beginning of each quarterly survey period, asked for selected vehicle characteristics and the vehicle's odometer reading. Owners were also advised that they would receive a follow up questionnaire at the end of the quarter seeking details about the use of the vehicle over the quarter and a second odometer reading. Examples of the main items requested in the second questionnaire were included with the first questionnaire, together with an optional, simple worksheet to help compile the data during the period. However, most of the questions in the second questionnaire simply ask for the owners estimate of percentage breakdown of the use of the vehicle into various categories. There is no evidence whether the worksheets were used by respondents to provide better responses to the questions, or whether they simply relied on recall.

These methods were developed following an extensive review of alternative collection methodologies for the SMVU conducted during 1995 and 1996 with a view to improving the quality of SMVU statistics.

Procedures also have progressively been instituted to encourage providers to phone through their data on a free-call number, if they are having any difficulties with the questionnaire.

Several aspects of different methodological options were tested during 1995 and 1996. There were effectively two competing methodologies that were tested. The first was a Logbook methodology, comprising several interrelated methods. This involved Trip Logbook, Fuel Logbook and Odometer Reading surveys. The alternative is a Pre-advice methodology. Respondents were advised at the beginning of the period of their inclusion in the survey, the data items required, and were given advice on how to compile records to assist them. This also incorporated an odometer readings component. “Fleet Data Capture” was also investigated as an adjunct to both main methodologies. Straight Recall, the traditional SMVU methodology, was tested as a control and to help bridge between the old and new methodologies.

ABS analyzed the data from the first two quarters pilot tests, to establish whether there were statistically significant differences in key variables between methods. This work identified no significant differences but the sample sizes for the tests were small, rendering the findings inconclusive rather than indicative of there being no difference. The bridging tests will help resolve this issue. However, there is clear evidence that a Logbook methodology is not viable. Pre-advice proved to have high data quality and, relative to logbooks, a low cost. Also, Fleet Data Capture proved highly successful for urban bus fleets but not for other fleets. Logbooks had low response rates and very poor data quality. Trip and Fuel Logbooks would both require substantial

clerical intervention to raise the quality, but are unlikely to achieve an acceptable level even with intensive effort. They were also very burdensome on providers and require complex processing systems. Forms analyses of errors provided evidence that data quality from the Pre-advice tests is significantly better than for Straight Recall; and is far superior to Logbooks. It is also more robust for estimates of total distance traveled than a pure Odometer Readings method, because it obtains odometer readings and a separate respondent estimate of distance traveled. Further, it is simpler and cheaper than Logbooks, potentially providing all data items with one method. Finally, the number of reported data items per response were far higher than for Logbooks or Straight Recall. Combined with response rates that were substantially higher than Logbooks and comparable to Straight Recall, the effective reporting rates were clearly superior.

SMVU Survey recording period

The data were recorded for a quarter. At the start of the quarterly survey period, the provider is asked to return a form reporting their contact details and the vehicle's odometer reading. The provider is also alerted about a follow up form seeking details about the use of the vehicle over the quarter and a second odometer reading. Examples of the main items that will be requested are included, together with an optional, simple worksheet to help compile the data during the period.

SMVU Unit of Analysis

The “unit of analysis” about which information was provided was the vehicle selected from the registration lists. The person providing the information was the owner of the vehicle.

SMVU Information collected

In the2002 SMVU, the following information is sought about the vehicle and its usage:

- Vehicle type
- Total and average kilometres traveled
- Total fuel used
- Average rate of fuel consumption
- Fuel type
- State/Territory of registration
- State/Territory of operation
- Business and private use
- Area of operation
- Driver characteristics
- For Freight vehicles only:
  - Total and average tonnes carried
  - Total and average tonnes-kilometres
  - GVM/GCM
  - Commodity carried
- For Buses only:
  - Kilometres traveled by type of bus and main type of service.

SMVU Travel recording method

No detailed information was gathered about trips made during the quarter. Only the odometer readings give any idea of the total amount of travel, while the owner provides estimates of the breakdown of the kilometres traveled in various categories. While these estimates have proved to be consistent over many years of the SMVU, this does not necessarily mean that they are accurate estimates. All it means is that owners of vehicles are consistent in their estimations or mis-estimations. For an example of the difference between accuracy and precision in the SMVU, see Richardson (2000).

SMVU Fuel consumption recording

No detailed fuel consumption or fuel purchases were recorded. Rather, owners were asked to estimate the average fuel consumption rate of the vehicle. This appears to be a highly unreliable way to estimate fuel consumption.

SMVU Imputation methods

Non-respondent vehicles to the survey are excluded from the estimation process, and hence stratum weights are calculated based on the number of respondent vehicles, rather than the number of selected vehicles. This procedure is equivalent to imputing the stratum respondent mean for the non-respondent vehicles.

Partial respondent vehicles to the survey are imputed using an automatic imputation system. The imputation system for the survey is based on a strategy using stratum by class respondent means for the key data items.

The need for imputation of unfilled items on the returned questionnaires, as for previous SMVU surveys, remained quite high. Of the questionnaires returned, 14% of those reporting some vehicle use needed imputation of one or more items apart from the average rate of fuel consumption. The imputation for average rate of fuel consumption was 25%.

SMVU Consideration of non-sampling errors

Concerns about non-sampling error in previous SMVU surveys, most notably recall bias by providers, led the ABS to the introduce a new collection methodology for the 1998 SMVU. While the new collection methodology has resulted in an overall improvement in SMVU estimates, some data quality issues remain and these are described below.

The ABS introduced a "pre-advice" methodology in the SMVU from August 1997 to improve the quality of its estimates of motor vehicle use. The methodology, in which vehicle owners receive early advice about their inclusion in the survey, encourages (according to the ABS) a higher degree of record keeping about the use of the vehicle during the survey period, either within owners' systems or by using the worksheet provided. This reduces the reporting errors arising from inaccurate recollection of use identified as a deficiency in the previous collection methodology. There is however no available documentation about the extent of record-keeping within the quarter, and it is possible that owners continue to rely on recall of vehicle activity over

the past three months. While this may be better than previous SMVU surveys, where the recall period was 12 months, it is still far from ideal.

In addition, the reporting of odometer readings taken at the start and end of the survey periods (approximately three months apart) provides more reliable estimates of total distance traveled without the recall bias inherent in the previous methodology, which was totally dependent on recall of usage over the past 12 months. For the 2000 SMVU, 76% of providers reported two odometer readings. This compares with 67% for the 1998 SMVU and 70% for the 1999. Where odometer readings were not provided, the total distance traveled was based on the reported distance traveled.

A further indicator of improved data quality has been the reduction in the reporting of 'rounded' data for total distance traveled for the 1998, 1999 and 2000 SMVU. Such rounding could cause significant errors, especially with the prevalence of certain distances that could be seen as arbitrary guesses on the part of the provider. Distances considered to be rounded are every 1,000 km in the range 1,000km up to 10,000km and every 5,000km for distances over 10,000km. The proportion of 'rounded' responses for total distance traveled for both the 1998 and 1999 SMVUs was 6%. For SMVU 2000, 4% of responses for total distance traveled were rounded. This is a significant improvement on the 1991 and 1995 SMVUs where the comparable figures were 50% and 23% respectively.

A potentially important factor relating to non-sampling error is the response rate achieved. The ABS states that it makes all reasonable efforts to maximize response rates. Where appropriate, mail reminders and telephone follow-up are used to attempt to contact initially non-responding vehicle owners. SMVU non-response predominately occurs because the ABS is unable to trace the vehicle selected or the owner will not or cannot complete the form.

A large non-response increases the potential for non-response bias, which occurs if the usage patterns of the non-responding vehicles differ significantly from those of the responding vehicles. For the SMVU, however, it is assumed that the characteristics of non-responding vehicles including the proportion of de-registered, out of scope and nil use vehicles are the same as for responding vehicles. Adjustment occurs to the 'weights' (the factors which expand the sample data to obtain estimates for the population) allocated to the respondent vehicles to allow for non-responding vehicles.

SMVU Expansion to population totals

Expansion of the sample data to provide national and state population estimates is performed by a series of weights included in the data files.

SMVU Response rates

When vehicles found to be de-registered or out of scope are removed, the live response rate for the 2000 SMVU is 79%.

Of the total sample, 6% were un-registered vehicles, 7% were untraceable, while 12% were unusable because of unresolved queries or where the vehicle was sold during the reference quarter and the reported data covered less than 14 days, or were non-response where no listing could be found to enable contact by telephone, or the owner was contacted by telephone but response was still not secured.

4.2 USA Residential Transportation Energy Consumption Survey (RTECS)

RTECS Name

Residential Transportation Energy Consumption Survey 1994 (RTECS)

RTECS Conducted for

Energy Information Administration (EIA)

RTECS Conducted by

Energy Information Administration (EIA)

RTECS Duration of survey

October 1993 (staggered starting dates) through March 1995 (staggered ending dates). The 1994 RTECS was the fifth such survey covering a calendar year conducted by EIA; previous surveys were collected in 1983, 1985, 1988 and 1991. Prior to the 1983 RTECS, monthly surveys were conducted from June 1979 to September 1981. Beginning with 1985, the surveys were conducted triennially. The 1994 RTECS was the last in the series.

RTECS Survey Objectives

The Energy Information Administration (EIA) is mandated by the US Congress to collect, analyze, and disseminate impartial, comprehensive data about energy: how much is produced, who uses it, and the purposes for which it is used. To comply with this mandate, EIA collects energy data from a variety of sources covering a wide range of topics.

The purpose of this survey is to provide information on the use of energy in residential vehicles in the 50 States and the District of Columbia. Included are data about the number and types of vehicles in the residential sector, the characteristics of those vehicles, the total annual vehicle-miles traveled, per-household and per-vehicle vehicle-miles traveled, vehicle fuel consumption and expenditures, and vehicle-fuel economy.

RTECS Population

The Residential Transportation Energy Consumption Survey (RTECS) was designed by the Energy Information Administration (EIA) to provide data about the population of vehicles that are used for personal transportation in the United States.

RTECS Sample frame

As of July 1994, the target population for the RTECS was estimated at 97.3 million households, based on adjusted estimates of households from the U.S. Bureau of the Census, Current Population Survey (CPS). The universe for the RTECS is comprised of all housing units occupied as the primary residence in the 50 States and the District of Columbia. The sample of households selected for the 1994 RTECS was based on the 1993 RECS (Residential Energy Consumption Survey) multistage area probability sample. The RECS incorporates a rotating panel that allows the observation of changes in energy use over time when the same households are in two successive surveys. The original RECS sample consisted of 8,753 units, of which 1,642 were not completed. The RTECS sample consisted of 3,020 housing units selected from the 7,011 available 1993 RECS housing units for which data were successfully collected.

The fraction of RECS housing units selected for RTECS was 59.8 percent. At the beginning of the data collection period in October 1993, 2,842 (93.3 percent) of the 3,020 housing units were identified as housing units that could potentially be contacted by telephone, and 200 housing units (6.6 percent) were identified as households that could not be contacted by telephone, either because they did not have telephones, had unlisted numbers, or refused to provide a telephone number during the RECS interview. This group was classified as mail households and data were collected from these households via a mailed questionnaire rather than a telephone interview. Contact was not attempted for an additional 3 households. By the end of the RTECS survey cycle (March 1995), the percent of households considered mail households had increased to 485, or 16 percent, because of an increased number of households with unlisted numbers or disconnected telephones.

RTECS Sampling unit

The sampling unit from the sampling frame was households.

RTECS Sample size

The sample size of the 1994 RTECS was 3,020 households, similar to the 1988 and 1991 RTECS which had 2,986 and 3,045 households respectively.

RTECS Sample design

The sample design for the 1994 RTECS consisted of a self-weighting national sample of households.

RTECS Survey technique

The RTECS was divided into three data collection phases. The first beginning of year data collection was combined with the RECS personal interview. During this interview, the household's vehicle stock was enumerated, and the VIN and the odometer reading for each vehicle were recorded. Household characteristics were also collected. Phases two and three: Mid-Year (M-Y) data-collection, and End-of-Year (E-O-Y) data collection, were conducted via telephone interviews. For households that could not be contacted by telephone, the data were collected via a mail questionnaire.

M-Y Data-Collection Phase: The M-Y mailing consisted of a letter from the Director of the Office of Energy Markets and End Use and a vehicle update worksheet for the respondents to complete. At this time, no vehicle characteristic data were obtained; only an inventory update was collected. The respondent was instructed to either keep the worksheet by the telephone for the telephone interview or return the worksheet by mail, if the household was classified as a no-telephone household. Any respondent who had refused at the B-O-Y interview was not contacted during the M-Y data collection phase. These households were not contacted at the M-Y interview, in order to increase the probability that the household would respond to the E-O-Y data collection. During the telephone interview, data were collected using the RTECS questionnaires.

One week prior to each M-Y and E-O-Y data collection, a mailing was sent to the RTECS respondents. The E-O-Y mailings consisted of the following: (1) Odometer Reading Cards; (2) VIN cards; (3) a page of instructions; (4) a letter from the Director of the Office of Energy Markets and End Use of the EIA explaining the survey; and (5) a letter from the survey contractor explaining their role in the survey.

E-O-Y Data-Collection Phase: Data collected during the E-O-Y phase consisted of an update of the vehicle stock and the collection of vehicle characteristics and vehicle use data.

The data-collection instruments for the RTECS consisted of four types:

- The 1993 RECS questionnaire,
- Odometer Reading Card,
- Vehicle Identification Number Card, and
- RTECS questionnaires.

1993 Residential Energy Consumption Survey Questionnaire - This form was used during the RECS personal interview. During this interview, the household's vehicle stock was enumerated, and the VIN and the odometer reading for each vehicle were recorded. Household characteristics were also collected.

Odometer Reading Card - This form was used by the RECS data collector during the combined RECS/RTECS data collection and was mailed to the respondent prior to the E-O-Y data collection and was used as a reference by the respondent during the telephone interview. The card was used to record, on an assigned date, the odometer readings for each reported vehicle during both the B-O-Y and E-O-Y data collections. A computer-generated label attached to the card identified each vehicle by make, model, and year of the vehicle. The respondents were instructed to enter the vehicle's odometer reading on the card after the last use of the vehicle on the specified assigned date. For the B-O-Y data collection, the date of the RECS data collection was used if the vehicle was available. If the vehicle was not available, the card was left with the respondent to be filed out and dated when the vehicle was recorded. For the E-O-Y data collection, cards were mailed to respondents to be filled out. No odometer reading cards were provided during the M-Y data collection. An additional odometer reading card without a computer-generated label was included for the respondent to record odometer readings for any vehicle acquired since the last contact.

VIN Card - This form was mailed to the respondent and used as a reference during the telephone interview. The card was used to record the VIN for each reported vehicle. Each VIN card had a computer-generated label identifying the specific vehicle assigned to the card. A thorough explanation of the VIN and where to locate it was provided on the card. A blank VIN card was also provided to record the VIN for any vehicle acquired since the last contact. For the E-O-Y data collection, the VIN was collected only from households that had acquired a new vehicle at the M-Y data collection phase. Households were not mailed a VIN card if they had refused to provide a VIN at the time of the RECS contact, or if they were previously classified as a nonrespondent household.

The odometer and VIN cards were mailed to the respondents approximately 1 week prior to the telephone interview. The respondents were requested to keep both the odometer and VIN cards by their telephones so they would be readily available when the interviewer telephoned. If the household was classified as a no-telephone household, the respondent was requested to return the cards in a self-stamped, business reply envelope that was provided.

Residential Transportation Energy Consumption Survey Questionnaire - This form was used by the telephone interviewer to record information gathered during the telephone interviews. Vehicle data obtained with this questionnaire included: verification of the stock of vehicles; motor vehicle characteristics for each vehicle, such as transmission type, drive type, fuel system type, engine size, and number of cylinders; vehicle fuel characteristics such as fuel type, fuel grade and type of pump service; odometer readings; and VIN. The questionnaire consisted of two types: the M-Y mail and telephone questionnaire and the E-O-Y telephone questionnaire.

RTECS Survey recording period

The vehicle usage data were recorded for a full year.

RTECS Unit of Analysis

The "unit of analysis" about which information was provided was all vehicles in the sampled households.

RTECS Information collected

The first Beginning-of-Year data collection was combined with the RECS personal interview. During this interview, questions were asked about the number of vehicles in the household, and for each vehicle: the VIN; the vehicle type; the vehicle make, model, and model year; the odometer reading; and estimated miles traveled during the past year or since the vehicle was acquired, if the vehicle was acquired within the previous 12 months. Household characteristics questions included the number of household members, and for each household member: their age, gender, employment status, and relationship to the head of household. The education level and ethnic background were collected for the head of the household only. The number of drivers, annual family income, and income assistance were collected for the household.

The Mid-Year mailing consisted of a letter from the Director of the Office of Energy Markets and End Use and a vehicle update worksheet for the respondents to complete. At this time, no vehicle characteristic data were obtained; only an inventory update was collected.

Data collected during the End-of-Year phase consisted of an update of the vehicle stock and the following vehicle characteristics for each vehicle recorded: the make, model and model year, engine size, fuel system type, and transmission type; vehicle fuel characteristics such as the fuel type, fuel grade and type of pump service; odometer readings; and VIN.

Respondents were asked to provide the VIN for each vehicle. The decoded VIN was used to enhance the accuracy of reported vehicle characteristics. The collection of the VIN from the respondent, rather than from information obtained from a registration list of vehicles, is an unusual feature of the RTECS and deserves special comment.

The VIN is a sequence of numbers and letters that, when decoded, provides vehicle characteristics that range from the nation of origin to the individual assembly plant where the vehicle was manufactured. The first three characters of the standard VIN format, designated as the World Manufactures Identification (WMI), identify the nation of origin, the manufacturer, and the vehicle make. The next five characters are the Vehicle Description Section (VDS). These characters identify the vehicle model; the body type such as sedan or station wagon; the engine type, which includes characteristics like the number of cylinders, cubic inch displacement and net brake horsepower; the restraint system found in the vehicle; and a model change code. There is no fixed format or standard codes within this five-character field. The ninth field contains a check digit. The check digit is an internally consistent number computed from the other identification numbers according to a mathematical formula. It is used during the decoding process to verify the accuracy of the other identification numbers. The next section is the Vehicle Identification Section (VIS) and contains eight characters. The first character in this section is the vehicle model year, the second character is the assembly plant name and/or location. The last six letters in this final section represent the sequential production number for a specific vehicle. To protect the confidentiality of the respondents, the sequential production number for a specific vehicle was not included on the RTECS public-use tape.

In most passenger cars the VIN is attached to the left side of the dash or instrument panel and is visible through the outside of the windshield. In some instances, the VIN is located on the inside of the door panel on the driver's side. Imported automobiles often attach the VIN to the windshield pillar post or on top of the steering column. Also, the VIN can be inscribed on the following documents: insurance cards, vehicle registrations, vehicle titles, safety or emission certificates, insurance policies, and bills of sale.

The initial collection of the VIN occurred during the RECS. Since the RECS data are collected in personal interviews, this survey would allow the interviewer to actually record the VIN from either the vehicle or a document.

The survey design allowed for the collection of the VIN during several different stages of the RTECS, thus eliminating the need to rely on a one-time effort. The VIN could be obtained during the RECS interview or during the E-O-Y interview. Only the RTECS households that did not provide a VIN at the time of the RECS interview, or for which the VIN was incorrectly transcribed during the RECS, were asked to provide the VIN again during the RTECS E-O-Y data collection. A household that had refused to provide a VIN at any time was not asked for the VIN a second time. RTECS precollection mailings to the households contained an explanation of the VIN and how to locate and record it.

The importance of obtaining an accurate VIN for a successful RTECS was emphasized during both the RECS and RTECS interviewer training sessions. Considerable interviewer training time was allocated to describing the VIN and providing the interviewers with thorough directions for locating and recording the VIN.

Overall, the collection of the VIN was regarded by EIA as a highly successful endeavor that yielded quality data. VIN's were obtained for 4,617 (76 percent) of the 6,084 total sampled vehicles. Of the 4,617 obtained VIN's, 3,842 (83 percent) were considered "good" VIN's. In summary, "good" VIN's were obtained for 3,842 (63 percent) of the 6,084 sampled vehicles (a good VIN was one that did not require correcting for common transcription error before it could be decoded).

RTECS Travel recording method

No detailed information was gathered about trips made during the year. Only the odometer readings give any idea of the total amount of travel.

RTECS Fuel consumption recording

No detailed fuel consumption or fuel purchases were recorded in the 1994 RTECS. Rather, fuel consumption and expenditures were estimated using the Environmental Protection Agency (EPA) laboratory-test miles per gallon (MPG) adjusted for on-road fuel efficiency shortfall. The estimation of these statistics occurred in several steps. First, for each RTECS vehicle, the VMT were determined from two actual odometer readings or imputed using data from the 1993 RECS (the respondents estimate of VMT in the previous 12 months). When possible, VMT were determined for a sample vehicle by taking the difference between two odometer readings, which spanned a period of time. This method was used to determine VMT for 2,648 (48 percent) of the 5,553 RTECS sample vehicles. Attempts were made to obtain odometer readings during the RECS interviews, the End-of-Year (E-O-Y) RTECS interview and any time a vehicle was acquired or disposed. A "span" of odometer readings was the difference between two odometer readings. In most cases, this span was a B-O-Y to E-O-Y span, although due to an occasional nonresponse, only shorter spans were obtained, such as RECS to B-O-Y. Odometer spans of less than a full year were also obtained for vehicles that were either acquired or disposed of during the survey year. The VMT that were assigned to each RTECS vehicle corresponded to the period of time that the vehicle was in possession by the sample household. In most cases, however, this period of possession did not correspond exactly with the beginning and ending dates for the odometer span. This was true even for vehicles with a complete B-O-Y to E-O-Y odometer span; because odometer cards were mailed to respondents in several distinct waves at the beginning and end of the RTECS survey; and because the exact dates of odometer readings were often left to the convenience of the respondents. Therefore, all VMT obtained from odometer spans were adjusted to correspond to the period of time that the vehicle was in possession by the sample household. A 2-step adjustment procedure was used. STEP 1 adjusted the odometer-span VMT to a standard annualized mileage covering 365 days, and STEP 2 readjusted the annualized VMT to correspond to the exact period of time that the vehicle was in possession by the household. These adjustments took into account a typical distribution of VMT fractions among the different months of the year. Step 2 was performed only for vehicles that were not in the possession of the household for the entire calendar year 1994.

For 1,927 sample vehicles (35 percent), no odometer span was available, although an estimate of annual VMT had been obtained from the respondent during the RECS interview. VMT for these vehicles were imputed from a regression on the estimated VMT obtained from the RECS. For another 978 sample vehicles (17 percent), no odometer span was available and a VMT estimate was not obtained during the RECS interview. VMT for these vehicles were imputed using a multiple linear regression model, where the independent variables were number of drivers, household income, age of household head, type of vehicle, and use of vehicle on the job. This regression was also used for imputing VMT for vehicles that were imputed as being acquired or disposed. Both of the regression models described above yielded estimates of annualized VMT. The STEP 2 adjustment described previously was then used to adjust this VMT to correspond with the time the vehicle was in the possession of the household.

Second, the annual on-road fuel efficiency, given in terms of MPG, was estimated using the questionnaire responses, decoded Vehicle Identification Number (VIN) data, EPA fuel efficiency test results, and the months that the vehicle was in use. The MPG were adjusted to account for the difference between EPA test values and on-road values. Third, estimated vehicle fuel consumption was derived by dividing the VMT by the estimated MPG. Finally, the estimated vehicle fuel expenditures were derived by multiplying the vehicle fuel consumption by the fuel price. The 1994 RTECS did not collect vehicle fuel prices via fuel purchase diaries. Instead each RTECS vehicle was assigned a price based on reported fuel type used in each vehicle. Gasoline prices were obtained from the BLS 1994 Retail Gasoline Pump Price Series. Diesel fuel prices were obtained from the Lundberg Survey, Inc.

RTECS Imputation methods

In the instances when a RECS respondent refused to participate in one or more of the RTECS data collection phases, a decision was made to impute the missing data rather than readjust the weights to account for the RTECS nonresponse because of the availability of RECS data for these households. Missing data items were imputed using the following: RECS data files on vehicle characteristics; decoded VIN; hot- and cold-deck procedures; linear and logistic regression; predictive mean matching; and EPA data files.

RTECS Consideration of non-sampling errors

Nonresponse bias is one type of nonsampling error that contributes to the total error of a survey. Other nonsampling errors include population undercoverage during sampling, interviewer error, coding and/or key punching errors, and response bias. The wording and format of the survey questionnaires, the procedures used to select and train interviewers, and the quality-control procedures built into the data collection and processing operations were all designed to minimize these sources of error.

It was recognized in the early planning stages of the 1994 RTECS that special attention would have to be given to minimizing nonresponse bias, since the RTECS households were contacted several times a year, in addition to the initial RECS personal interview. The following steps were taken to minimize the nonresponse:

- If possible, the VIN was collected during the RECS interview, thus, reducing the need to ask for the VIN during the RTECS, if it was successfully collected during the RECS
- The M-Y data collection instrument was streamlined with the primary emphasis placed on updating the vehicle stock and obtaining the odometer readings for any disposed of or acquired vehicle. The vehicle characteristics and VIN for newly acquired vehicles were collected during the E-O-Y data collection instead of during the M-Y update
- No M-Y data collection was attempted for households declared as legitimate refusals at the initial interview. All households were recontacted for the E-O-Y data collection; however, only households that previously had a valid odometer reading were asked to provide the E-O-Y odometer reading. None of the households that previously refused were asked any VIN questions
- A letter describing the survey and its importance was mailed to the households approximately one week prior to the B-O-Y and E-O-Y telephone interviews

- The interviewers were requested to attempt a minimum of eight telephone calls before the household was classified as a noncontact and in many cases made up to 16 or more attempts to contact the household.

RTECS Expansion to population totals

All the statistics from RTECS 1994 are estimates of population values, such as the total number of households in the United States. These estimates are based on a subset of the entire population of households chosen according to multistage probability sample selection rules. The universe includes all households in the 50 States and the District of Columbia, including households on military installations. Survey estimates expand the RTECS sample results to represent the target population. This required the development of weights for each sample household using a multistage weighting procedure. The weights for the RTECS were developed from the weights that originally had been used in the 1993 RECS. These original weights were divided by the probability that a RECS household was selected into the RTECS sample. The probability that an individual 1993 RECS household was selected for the 1994 RTECS varied by the Secondary Sampling Unit (SSU) and the estimated vehicle annual mileage for the household as obtained from the RECS interview.

These RECS weights were appropriate for estimates of U.S. households as of November 1993 (the midpoint of the RECS data-collection time period). Since the midpoint of the RTECS data-collection period was July 1994, the RTECS weights were adjusted so that RTECS household counts were estimated at their presumed July 1994 levels. This was accomplished by use of post-stratification. In post-stratification, the survey weights in RTECS (and RECS) were adjusted by factors so that, within certain population subgroups, RTECS estimates of household counts would agree with those estimated from the Current Population Survey (CPS). Within each population subgroup or post-stratification cell, the weight adjustment factor was computed as the CPS household count estimate divided by the RTECS household count estimate. (RTECS household count estimates are produced by summing RTECS survey weights.) The CPS estimates within the subgroups are called "control totals," and they are considered to be more reliable than the corresponding estimates from RTECS.

The post-stratification cells were defined by a two-way contingency table. One margin represented metropolitan status within the Census regions. This margin had 12 categories comprised of 4 Census regions (Northeast, Midwest, South, and West) and 3 metropolitan statuses (metropolitan in center city, metropolitan outside of center city, and non-metropolitan). The other margin of the post-stratification table contained three categories (one-person-male households, one-person-female households, and all other households). Thus, the post-stratification table had a total of 36 cells. However, 36 separate weight-adjustment factors were not computed. Rather, the RTECS weights within these cells were adjusted by a limited "raking" procedure. The weights were first adjusted to agree with CPS totals for the Census region by metropolitan status margin, using 12 cells. Next the weights were adjusted to CPS totals for the household type margin using 3 cells. Finally, the weights were readjusted to CPS totals for the Census region by metropolitan status margin. Raking allows for the use of more sample units in computing each weight-adjustment factor, by not distributing them around too many cells.

However, the method is based on the assumption that there is no significant interaction between the margins of the post-stratification table.

RTECS Response rates

EIA claim that the response rate for the RTECS was effectively 100%. They state that "Unit nonresponse is the type of nonresponse that occurs when no data are available for an entire sampled household. Most unit nonresponse cases are caused by the respondent being unavailable or the respondent's refusal to cooperate. Unit nonresponse for the 1994 RTECS must be addressed in the context of the unit nonresponse for the 1993 RECS, since the 1994 RTECS sample was drawn from households that responded to the 1993 RECS. Thus, in all cases, at least the RECS data were available for every RTECS household, therefore, no RTECS household was a total nonrespondent. Generally, weight adjustment was the method used to reduce unit nonresponse bias in the RECS statistics and that adjustment carried over automatically to the RTECS subsample."

This interpretation of non-response is considered to be over-generous, and inflates the response rate actually achieved in RTECS 1994. However, no published information is available to quantify the response rates actually achieved.

Change in Method of Estimating Fuel Consumption Rates in RTECS

The 1994 method for collecting data on vehicle fuel economy for the Residential Transportation Energy Consumption Survey is to use the Environmental Protection Agency's fuel-economy ratings. The previous method, abandoned because of budgetary restraints in 1988, was to use fuel-purchase logs. Although the two methods have to date yielded similar estimates of fuel economy, the current method may yield less reliable estimates over time.

For the 1983 and 1985 RTECS, fuel-economy estimates were based on data from fuel-purchase logs kept by selected respondents. During a particular month of the survey year, each respondent recorded the odometer reading and the amount of fuel purchased at the time of each fuel purchase. That information was then used to compute the fuel economy of the vehicles in miles per gallon. Computing fuel economy in that manner gave accurate on-road monthly estimates. Those monthly estimates were converted to annualized estimates based on the observed way the fuel economy varied for different months of the year. The response rates for completion of the fuel purchase logs were 46 percent (3,526) for the 1983 RTECS and 40 percent (3,413) for the 1985 RTECS. For those vehicles for which no log data were available, the annualized fuel economies were imputed by using the medians of categories defined by vehicle type, make, model, and model year.

Beginning with the 1988 RTECS, EIA stopped collecting the fuel-purchase data, due to budget limitations. Instead of fuel-purchase logs, EIA developed a new approach to estimate fuel economies based on EPA's fuel-economy ratings. After applying EPA fuel-economy ratings (in miles per gallon) to RTECS sample vehicles, EIA multiplied the derived fuel-economy ratio by

the annual vehicle miles traveled to derive annual consumption. EPA's fuel-economy ratings are part of EPA's annual certification files, which contain extensive information on vehicles. The EPA files provide three fuel-economy ratings: city, highway, and composite. The composite rating is formed by combining the city and highway ratings, assuming a "typical" vehicle-use pattern of 55-percent city driving and 45-percent highway driving. EPA bases its ratings on the fuel economy of test vehicles under simulated driving conditions, adjusted for on-road use. This adjustment uses discount factors, which EPA developed in the early 1980's. The factors reduce the highway ratings by 22 percent and the city ratings by 10 percent. They were developed based on vehicles of late 1970's and early 1980's vintage.

EIA further adjusts the EPA ratings for individual driving circumstances before reporting them in the RTECS. This adjustment takes into consideration such factors as urban versus rural driving patterns, traffic congestion, seasonal temperature variations, humidity levels, geographic variations, altitude, wind, and road gradient and surface conditions. This adjustment uses a regression model unique to each vehicle, based on the vehicle's average miles per day and its geographic location. Although these adjustments are the best available for this estimation methodology, they are still unlikely to capture many of the inherent differences among vehicles and drivers that may affect fuel economy.

To compare the two methods for estimating fuel economy, estimates from both methods were needed. Therefore, either log-based fuel-economy estimates for the 1988 and 1991 RTECS vehicles or adjusted EPA fuel-economy ratings for the 1983 and 1985 RTECS vehicles had to be obtained. Accordingly, three steps were taken:

- Base EPA fuel-economy ratings for the 1983 and 1985 RTECS vehicles were obtained by matching them with vehicles in EPA's certification files, which list extensive information on each make and type of vehicle of model years 1975 and later. Some of the RTECS vehicles had to be excluded, either because they were of model years earlier than 1975, or because their make or model names could not be determined. As a result, 2,799 of the 1983 RTECS vehicles and 2,297 of the 1985 RTECS vehicles were excluded from the analysis, leaving 4,866 and 6,194 usable vehicle records in the 1983 and 1985 RTECS databases, respectively.
- The EPA fuel-economy ratings for these vehicles were adjusted for on-road use and individual driving circumstances, using the same method as had been used in the 1988 and 1991 RTECS
- Estimates of average fuel economy for the U.S. vehicle stock and various subpopulations were computed using both methods, and differences were analyzed.

A comparison of the results of the two methods showed that for the total vehicle stock, the EPA-based estimates were slightly higher than log-based estimates in both 1983 and 1985 (1 percent higher for 1983 and 4 percent higher for 1985). However, the EPA-based estimates were noticeably lower than the log-based estimates for vehicles in certain categories:

- For vehicles that traveled fewer than 5,000 miles per year, the EPA-based estimates were 6 percent lower in 1983 and 3 percent lower in 1985.
- For vehicles whose primary drivers were 66 years of age or older, the EPA-based estimates were 4 percent lower in 1983 and 3 percent lower in 1985.
- For vehicles with model years of 1975 or 1976, the EPA-based estimates were 3 percent lower in l 983 and 2 percent lower in 1985.
- For vehicles with rotary engines, the EPA-based estimates were 3 percent lower in 1983 and 8 percent lower in 1985. However, this difference was likely due not to the method used but to the fact that rotary engine vehicles were older and relatively rare (only 12 out of 4,866 sample vehicles in 1983 and 9 out of 6,194 in 1985).

4.3 USA Vehicle Inventory and Use Survey (VIUS)

VIUS Name

Vehicle Inventory and Use Survey 1997 (VIUS)

VIUS Conducted for

U.S. Census Bureau

VIUS Conducted by

U.S. Census Bureau

VIUS Duration of survey

The survey was conducted on vehicles registered in the US on July 1, 1997 and estimated the use of vehicles in the 12 months of 1997.

VIUS Survey Objectives

The economic census, of which VIUS is a part, is the major source of facts about the structure and functioning of the Nation's economy. It provides the framework for such composite measures as the gross domestic product, input/output measures, production and price indexes, and other statistical series that measure short-term changes in economic conditions. VIUS data are of considerable value to government, business, industry, academia, and the general public. Data on the number and types of vehicles and how they are used are important in studying the future growth of transportation and are needed in calculating fees and cost allocations among highway users. The data also are important in evaluating safety risks to highway travelers and in assessing the energy efficiency and environmental impact of the Nation's truck fleet. Businesses and others make use of these data in conducting market studies and evaluating market strategies; assessing the utility and cost of certain types of equipment; calculating the longevity of products; determining fuel demands; and linking to, and better utilizing, other data sets representing limited segments of the truck population.

VIUS Population

The 1997 VIUS is a probability sample of private and commercial trucks registered (or licensed) in the United States as of July 1, 1997. This survey excludes vehicles owned by Federal, state, or local governments; ambulances; buses; motor homes; farm tractors; unpowered trailer units; and trucks reported to have been sold, junked, or wrecked prior to July 1, 1996.

The results of the 1997 VIUS are comparable to the 1992 and prior Truck Inventory and Use Surveys (TIUS). The survey name was changed in 1997 from TIUS to VIUS to account for areas of future expansion, including the possible addition of automobiles and buses. However, while survey forms were produced for cars and buses, it appears that the 1997 VIUS was again limited to trucks of various sizes.

VIUS Sample frame

The sampling frame for the initial universe was constructed from files of active truck registrations for the 50 states and the District of Columbia. All registrations on the sampling frame were identified as being active as of July 1, 1997. The frame was stratified by geography and truck characteristics. The 50 states and the District of Columbia made up the 51 geographic "state" strata for the initial sample. Body type and gross vehicle weight (GVW) determined the following five truck strata: (1) pickups; (2) vans, minivans, and panels; (3) light single-unit trucks (GVW < 26,000 lb); (4) heavy single-unit trucks (GVW >= 26,000 lb); and (5) truck-tractors. Therefore, the sampling frame for the initial sample was partitioned into 255 geographic-by-truck strata.

VIUS Sampling unit

The sampling unit from the sampling frame was individual vehicles.

VIUS Sample size

The 1997 VIUS is a probability sample of about 131,000 trucks registered in the United States as of July 1, 1997.

VIUS Sample design

This survey consists of an initial and a supplementary sample of truck registrations.

Initial Sample. Within each of the 255 strata in the sampling frame, a simple random sample of truck registrations was selected, resulting in an initial sample of approximately 128,000 truck registrations.

Supplementary sample. Each truck represented on the sampling frame for the initial universe had associated with it a state of registration and a state as part of its mailing address. In the 1997 VIUS initial sampling frame and in previous TIUS sampling frames, a truck registered in a state other than the one listed on its mailing address was excluded. (For example, a truck registered in Florida that had a mailing address in Wisconsin was excluded from the sampling frame.) Therefore, previous TIUS estimates did not include data from these trucks. In order to represent

these trucks in the 1997 VIUS estimates, a supplementary sample was selected independently from the initial sample. The sampling frame for the supplementary universe was also stratified by geography and truck characteristics. The truck characteristics were the same ones used to stratify the frame for the initial sample. However, the geographic strata were defined differently. The 50 states and the District of Columbia were partitioned into four mutually exclusive geographic strata. Each "state" was included in one, and only one, geographic stratum. These geographic strata were used instead of those used for the initial sample because of cost and processing constraints. Within each geographic-by-truck stratum, a simple random sample of truck registrations was selected. Thus, an additional 20 geographic-by-truck strata were used to stratify the sampling frame for the supplementary universe, resulting in a total of 275 geographic-by-truck strata from which to select truck registrations for the entire 1997 VIUS sample. The size of the supplementary sample for the entire United States was approximately 3,000 truck registrations.

VIUS Survey technique

For each selected truck, a questionnaire was mailed to the registered owner. The registered owner was requested to provide data about the truck identified by the vehicle registration information imprinted on the questionnaire, regardless of whether or not the truck was still in his or her possession. The information provided by each respondent was subjected to extensive computer edits. Questionable responses were reviewed and corrected when necessary.

VIUS Survey recording period

The usage data was estimated for a complete year (1997).

VIUS Unit of Analysis

The "unit of analysis" about which information was provided was the vehicle selected from the registration lists. The person providing the information was the owner of the vehicle.

VIUS Information collected

The information obtained from the vehicle owner varied somewhat, depending on the type of vehicle. For pickups, panel trucks, vans, minivans, sport utility vehicles, jeeps, and station wagons built on truck chassis, the following information was obtained:

- Changes of ownership or disposal of vehicle during the year
- Body type of vehicle
- Axle and drive-train configuration
- Use of trailers
- Weight of vehicle
- Weeks in use during 1997
- Miles driven during 1997
- Miles driven since manufacture
- Home base of vehicle

- Percent of mileage driven outside home state
- Percent of miles within various distances of home base
- Estimate of fuel consumption rate (mpg)
- Type of fuel used
- Vehicle accessories
- Maintenance procedures
- Type of use of vehicle (private, business etc)
- Type of business for which vehicle was used
- Commodities carried
- Number of vehicles owned by owner or owner's company.

For larger trucks, the following questions were asked:

- Changes of ownership or disposal of vehicle during the year
- Leasing arrangements, if applicable
- Usual configuration of vehicle
- Trailer configuration
- Body type of truck
- Axle and drive-train configuration
- Cab type
- Length of truck and trailers
- Width of trailers
- Empty, average and gross weight of truck/trailer configuration
- Weeks in use during 1997
- Miles driven during 1997
- Miles driven since manufacture
- Home base of vehicle
- Percent of mileage driven outside home state
- Percent of miles within various distances of home base
- Estimate of fuel consumption rate (mpg)
- Type of fuel used
- Usual place of re-fuelling
- Type of brakes
- Vehicle accessories
- Maintenance procedures
- Type of for-hire operations
- Type of use of vehicle (private, business etc)
- Type of business for which vehicle was used
- Commodities carried
- Carriage of hazardous goods
- Number of vehicles owned by owner or owner's company.

VIUS Travel recording method

No detailed information was gathered about trips made during the year. Only the owner's estimate of total distance traveled gives any idea of the total amount of travel, while the owner also provides estimates of the breakdown of the miles traveled in various categories.

VIUS Fuel consumption recording

No detailed fuel consumption or fuel purchases were recorded. Rather, owners were asked to estimate the average fuel consumption rate of the vehicle. This appears to be a highly unreliable way to estimate fuel consumption.

VIUS Imputation methods

For item nonresponse, a missing value is replaced by a predicted value obtained from an appropriate model for nonresponse. To impute annual miles and lifetime miles, the sample is divided into a finite number of mutually exclusive cells based on state of registration, truck type, and model year. For each cell, estimates of average annual miles and aver-age lifetime miles are computed based on those trucks in the cell for which annual miles and lifetime miles have been reported. Missing values are then replaced with the appropriate average values. A slightly different imputation procedure is used for length and average weight (empty weight plus cargo weight). For these data items, a missing value is replaced with data from a truck with similar characteristics for which length and average weight have been reported. For all other data items, no imputation is performed. Instead, separate estimates are published in a "Not reported" category. For example, a respondent who did not indicate the major use of his or her truck would be included in the estimate for the "Not reported" category. Users of the estimates were instructed to exercise caution when allocating the estimate for the "Not reported" category to the estimates for the reported categories in the proportions observed for the reported categories. This is because the characteristics of the trucks for which information was obtained may differ significantly from those trucks for which no information was obtained.

VIUS Consideration of non-sampling errors

In the 1997 VIUS, non-sampling error can be attributed to many sources: (1) inability to obtain information about all trucks in the sample; (2) response errors; (3) differences in the interpretation of the questions; (4) mistakes in coding or keying the data obtained; and (5) other errors of collection, response, coverage, and processing. Although no direct measurement of the potential biases because of non-sampling error has been obtained, precautionary steps were taken in all phases of the collection, processing, and tabulation of the data in an effort to minimize its influence.

A potential source of bias in the estimates is non-response. Non-response is defined as the inability to obtain all the intended measurements or responses about all the selected trucks. Two types of non-response are often distinguished. Unit non-response is used to describe the inability to obtain any of the substantive measurements about a sampled truck. In most cases of unit non-

response, the questionnaire was never returned to the Census Bureau after several attempts to elicit a response. Item non-response occurs either when a question is unanswered or the response to the question fails computer edits. The procedures used to account for unit non-response are discussed below.

Unit non-response is handled in the estimation procedure by reweighting. To apply this method of non-response adjustment, the assumption is made that the population of trucks can be divided into a finite number of mutually exclusive adjustment cells so that within each cell, all the population elements possess similar characteristics and share a similar probability of responding, if selected in the sample. The adjustment cells for the 1997 VIUS are identical to the sampling strata. A non-response adjustment factor is computed for each adjustment cell and is equal to the ratio of the number of truck registrations selected into the sample to the number of responses received within each adjustment cell. In this sense, reweighting allocates characteristics to the non-respondents in proportion to the characteristics observed for the respondents within each adjustment cell. The amount of bias introduced by this non-response adjustment procedure depends on the extent to which the non-respondents differ, characteristically, from the respondents in each adjustment cell.

VIUS Expansion to population totals

National estimates of the number of trucks were computed in the following manner. For the initial and supplementary samples, weighted estimates of the number of trucks were computed for each geographic-by-truck stratum. The weight for a given truck was the product of two factors: the reciprocal of the truck's probability of selection and a non-response adjustment factor. The geographic-by-truck stratum estimates from the initial and supplementary samples were summed to form the national estimates. National estimates of truck miles were computed using the same method. National estimates of average miles per truck were computed by dividing the national estimates of truck miles by the national estimates of the number of trucks.

VIUS Response rates

The Census Bureau defines unit non-response as the inability to obtain any of the substantive measurements about a sampled truck. In most cases of unit non-response, the questionnaire was never returned to the Census Bureau after several attempts to elicit a response. Approximately 84.5 percent of the questionnaires were returned with substantive data.

4.4 German Mobility Panel (Fuel Purchase Sub-sample) (GMP)

GMP Name

German Mobility Panel (GMP) – Das Deutsche Mobilitätspanel

Fuel consumption and mileage sub-sample survey.

GMP Conducted for

German Federal Ministry for Transport, Infrastructure and Housing

GMP Conducted by

NFO Infratest, München

GMP Duration of survey

Pilot-test was conducted in 1994-1996. Since 1996, the survey has been an ongoing panel in the Fall of each year, with approximately 50% of the households owning at least one car responding to a separate survey on fuel consumption and mileage in the Spring of each year. The fuel consumption sub-sample is therefore also a panel survey, with previous respondents to the fuel consumption survey being augmented each year by approximately 50% of the car-owning households in the new wave of the fall mobility panel survey.

GMP Survey Objectives

The objective of the overall German Mobility Panel is to obtain

- actual data on travel behavior,
- regular and comparable information over time,
- indications for the changes in travel behavior
- the possibility to enter cross-sectional surveys (such as Kontiv) into time-serial assessment
- a reference base for before/after analyses.

The panel on mobility behavior sought to determine socio-economic information on the household members, daily mobility patterns over 7 days for all household members older than 10 years of age (trips, trip purpose, mode choice, trip distance and duration), and the composition of the private vehicle fleet of the households.

In addition to the daily travel behavior of the persons included in the GMP, the GMP also records mileage and fuel use of cars used by 50% of the panel households in the Spring of each year. The panel on mileage and fuel consumption specifically sought to determine the following:

- the vehicles used for different reasons (personal, business related)
- type of vehicle ownership;
- distances traveled by these vehicles;
- fuel consumption;
- vehicle fleet characteristics;
- household vehicle use;
- on-road fuel consumption ratio; and
- comparison of on-road versus laboratory-tested fuel consumption ratios.

GMP Population

The population of the panel was all private households from "old" West Germany (since 2000, all Germany) and for the mileage and fuel consumption survey, households with car availability with the GMP.

GMP Sample frame

The sample frame for the overall GMP was the German household register. These households were kept in the GMP for only three years, after which time they were rotated off the panel. For the survey on mileage and fuel consumption, a sub-sample of approximately 50% households owning at least one vehicle was selected randomly from the panel households of the Mobility Panel in that year.

GMP Sampling unit

Households were used as the sampling unit for the Mobility Panel. Households owning at least one vehicle were used as the sampling unit for the mileage and fuel consumption survey.

GMP Sample size

Since 1995, the sample size of the mileage and fuel consumption survey, in households and cars, has been as follows:

Year	Sample Size – households	Sample size - cars
1995	271	376
1996	368	483
1997	311	417
1998	336	460
1999	332	451
2000	337	458
2001	376	525

GMP Sample design

Because of the small sample size for the overall GMP, a representative sample on geographic grounds is not possible for this household-based survey. Therefore a non-uniform stratified random sample has been chosen according to the following stratification variables: household structure, spatial typology and car ownership.

A spatial typology has been defined in five categories according to residential density, job density, commuting quotas, revenue quotas of different economic sectors, which covers the entire range of German spatial areas.

A household typology in four categories has been defined according to household size, number of children in the household and professional status of the household members.

The sample for the mileage and fuel consumption was chosen randomly from the GMP households owning at least one car.

GMP Survey technique

Contacting procedures, household details and trip recordings for each household member are recorded in the wave of the mobility panel conducted in the previous Fall (e.g. information on respondents in the Spring 2000 fuel survey were recorded in the Fall 1999 mobility panel survey). A telephone call was used for the first contact of the household and regular mail was used for all further contacts, including distribution and collection of household questionnaire and trip diaries, and contact letters for successive waves.

The mileage and fuel consumption survey is conducted with a self-administered tanking diary (German: Tankbuch) extending over the period of approx. 2 months (end of April to End of June 2000), in which all tanking episodes and odometer readings had to be entered (including odometer readings at the start and end of period). Additionally, the level of the fuel gauge was recorded at the start and end of the recording period.

Two instruments are used to remind drivers several times of the necessity to fill in the tanking diary:

- Stickers on the speedometer and on the fuel cap for each car; and
- Reminder postcards to the household.

GMP Survey recording period

In the households sampled for the GMP, all persons older than 10 years were asked to complete a 7-day trip diary for all their trips and to start their recording on a specific day. In order to limit a number of adverse effects of the panel survey (e.g. attrition, learning effects), the selected households are rotated and only participate for the limited number of 3 waves. Respondents to the mileage and fuel consumption survey were asked to complete a two-month (8 weeks) diary of fuel purchases for all vehicles used by the household.

GMP Unit of Analysis

The "unit of analysis" about which information was provided was the household selected from the Mobility Panel and all vehicles used by the sampled household. The persons providing the information in the tanking diary were the drivers of each vehicle.

GMP Information collected

In the mileage and fuel consumption survey, the following information is sought:

<u>Vehicle characteristics</u>

- Make, model and year of manufacture of vehicle,
- Body type of vehicle,

- Number of cylinders of vehicle,
- Size of tank (liters)
- Type of fuel (super lead free, super plus, diesel, normal, other)
- Occurrence of special events (car break down, holiday trips, etc.)
- Approximate kilometres driven per year.

Vehicle ownership categories

- Privately owned, company car of self-employed, company car of employer

Type of vehicle use (km)

- Only private use, private & commercial, only commercial use
- Road type where vehicle was used, (on urban, on highway, on rural, on mix of road types)

Fuel consumption & Odometer readings

- Odometer readings are reported at the start of the reporting period, at every tanking episode, at the end of the reporting period.
- The level of the fuel gauge is reported at the start and end of the reporting period.
- For every fuel purchase during the reporting period the date of purchase, litre, cost (Euro) are recorded
- Whether the tank was filled up or not.

Household related information is already known from the Panel on daily mobility (from the wave in the previous Fall).

GMP Travel recording method

Detailed recording of trips made by each household member was undertaken during the mobility panel survey in each Fall. This survey consisted of a full stage-based travel survey of urban and long-distance trips over a 7-day period from all members of the household. These recordings were made in the Fall of each year.

During the mileage and fuel consumption survey in the following Spring, no detailed information was gathered about trips made during this 8 weeks period. Only the odometer readings give any idea of the total amount of travel. Since the mobility panel and the fuel consumption survey were performed in different seasons, it is not possible to link the fuel consumption reports with the detailed trip records.

GMP Fuel consumption recording

Fuel purchases were recorded in the Tanking Diary (the Tankbuch) during the 8-week period. When a diary shows at least two fill-ups, it is possible to calculate the precise amount of fuel used between these fill-ups (assuming that both “fill-ups” were indeed complete fill-ups). Since the difference in odometer readings provides the distance between the two diary entries, it is therefore possible to calculate the fuel consumption rate of the vehicle between the fill-ups in litres/100km. It is then assumed that this rate of travelling and fuel consumption applies to the rest of the month, and the results are expanded to a full 31 days.

When there are less than two fill-ups in the month, it is still possible to estimate a fuel consumption rate. The total mileage is known from the difference in odometer readings at the start and end of the recording period. The total fuel consumed is estimated as the sum of the fuel purchases plus the amount in the tank at the start of the period minus the amount in the tank at the end of the period.

GMP Imputation methods

When there is missing data on odometer readings, or fuel gauge readings at the start and end of the recording period, it may be impossible to calculate a fuel consumption rate. In these circumstances, it is necessary to impute information to enable the estimation of a fuel consumption rate. While such imputation methods are used in this survey, insufficient details are provided in the survey documentation to describe them here in detail.

GMP Consideration of non-sampling errors

There was some recognition of non-sampling errors that were accounted for in the following ways:

- Bias of non-mobile and highly-mobile households was reduced by limiting the participation to 3 waves only
- Correction for trip recording fatigue within the 7 day period (not specified in reports)
- Correction for design-change from 1996 onwards that had an effect on the average trip distance (not specified in reports).

GMP Expansion to population totals

The Mobility Panel data was expanded to population totals using data from the German register of population. The expansion factors were based on post-stratification at the level of the household, the person and the trip as shown below:

- Weighting on household level: Household size, Size of the residential community, Car ownership of the household (from census).
- Weighting on Person level: Age and gender, person-day
- Weighting on trip level: trip distance and reporting day.

For the sub-sample survey on mileage and fuel purchase, the classification of the official vehicle register (stratified by age of vehicle and number of cylinders) was used for expansion.

GMP Response rates

The response rates to the fuel consumption survey are shown below, in terms of both households and cars responding. It is also shown for households that were new to the mobility panel in the previous quarter and those who had been in the mobility panel for more than one wave. It can be seen that the household response rate was slightly higher than the cars response rate, indicating

that responding households had lower car ownership than non-responding households. This is probably because all cars in the household were being surveyed in the fuel consumption survey. If only one car in the household was being surveyed, this difference between responding and non-responding households may have disappeared. It can also be seen that new and old households had similar response rates, indicating that time spent on the panel neither increased nor decreased the likelihood of response to the fuel consumption survey. Overall, the response rates to the fuel consumption survey are around 75%, which is quite good for a mailback survey. However it must be realized that the sample frame for this survey was households that were already participating in the three-wave mobility panel and were therefore more predisposed to respond to such surveys.

	Response Rates			
Survey	Households	Cars	Old Households	New Households
1995	57%	54%	--	--
1996	65%	60%	61%	68%
1997	78%	77%	77%	78%
1998	79%	77%	80%	78%
1999	80%	76%	80%	82%
2000	70%	65%	73%	67%
2001	82%	77%	74%	75%

	FCS	NaPYUS	CYS	SMYU	RTECS	YUS	GMP
Conducted by	StatCan	StatCan	StatCan	ABS	EIA	US Census Bureau	Infratest Munich
Timing	1979-1988	1994-1996	1998-2000	1963-2002	1988-1994	1997	1996-2001
Population	Private cars, light trucks and vans	Canadian Households who own a vehicle	All registered-vehicle-days	All registered vehicles	Personal use vehicles in the USA	Private and commercial trucks in the US	Vehicles in German households
Sample frame	Vehicle registers	Rotate-outs from Labour Force Survey	Vehicle registers	Vehicle registers	Rotate-outs from RECS	Vehicle registers	Sub-sample from Mobility Panel
Sampling unit	Vehicles	Households	Vehicles	Vehicles	Households	Vehicles	Households
Sample size	160,000 (approx)	39,287	31,404 (in 2000)	16,000 (in 2000)	3,020 (in 1994)	131,000	350 (p.a. in FCS)
Survey technique	Phone/Mailout Fuel Diary	CATI/Mailout Fuel Diary	CATI/Mailout Trip-Log	Mailout, with Pre-advice	CATI, with mail reminders	Mailout recall questionnaire	CATI/Mailout Fuel Diary
Recording period	1 month	1 month	Provinces: 7 days trip log; Territories: quarterly odometer	Quarter	12 months	12 months	7days for mobility panel; 8 weeks for fuel consumption survey
Unit of analysis	1 Vehicle	1 Vehicle	1 vehicle	1 Vehicle	All vehicles in household	1 Vehicle	All vehicles in household
Info collected	Vehicle Info; Start & End Odometer; Fuel Purchases	Demographics; Vehicle Info; Fuel Purchases	Vehicle info; Trip data	Vehicle Info; Est. Usage; Est. FC rate; Usage breakdowns	Demographics; Vehicle Info; VIN; Start and End Odometer	Vehicle Info; Est. Usage; Est. FC rate; Usage breakdowns	Demographics; Full trip details; Vehicle Info Odometer; Fuel purchases
Travel recording	Odometer only	Odometer only	Trip data	Est. kilometres	Odometer only	Est. kilometres	Full details in MP; odometer readings in FCS
Fuel recording	Fuel purchases	Fuel purchases	Fuel purchases within the 7 days	Est. FC rate	EPA FC ratings	Est. FC rate	Start & end odo; Start & end fuel gauge; Fuel purchases
Imputation	When 2 non-fillup purchases	When 2 non-fillup purchases	Major imputation of missing log-days	For INR, esp. FC rate	For INR	For some INR	Unspecified
Response Rate	65%	27% to CATI and fuel diary	40% to CATI and trip-log	79%	100% (?)	84%	75% (for FCS)

5 Objectives of the Proposed Survey

Before embarking on the detailed design of the CIVAFUS options, it is important to clearly define the specific objectives of the survey. Without such specific objectives, it is very difficult to make the necessary decisions about sample design, question content, question design, coding design and a host of other design decisions. In addition, it is almost impossible to make the inevitable trade-offs that will be needed between the various aspects of the survey design.

In general, as stated in the Terms of Reference for this project, the objective of CIVAFUS is to "combine the surveys [CVS, FCS and NaPVUS] in some manner to obtain both trip characteristics and vehicle-specific fuel consumption from each sampled vehicle". However, while this general objective is useful, it is not sufficient to guide the development of the design of CIVAFUS. What is needed is a more specific set of objectives in terms of required outputs from CIVAFUS and an idea of how these outputs might be used.

Transport Canada have previously specified their objectives for the current CVS in terms of a set of output table specifications, as follows:

Physical Characteristics

- Total road vehicle-km in a year
- Vehicle-km by construction class of vehicle
- Vehicle-km by body type of vehicle
- Vehicle-km for trucks by configuration
- Vehicle-km by fuel type
- Vehicle-km by age of vehicle, by construction class

Operational Characteristics

- Vehicle-km by truck carrier type
- Vehicle-km by truck driver type
- Vehicle-km by type of vehicle and month
- Vehicle-km by hour of day
- Vehicle-km by day of week
- Vehicle-km by commodity

Demographic Characteristics

- Vehicle-km by driver age-group
- Vehicle-km by driver sex
- Vehicle-km by trip length
- Vehicle-km by trip purpose

For CIVAFUS, Transport Canada have specified that vehicle-specific fuel consumption together with vehicle use characteristics would be useful for modeling the relationships between vehicle

use and vehicle fuel consumption, as well as for the simpler purpose of estimating average "degradation" of fuel use over the test results. They would therefore like to collect all of the vehicle use data currently collected in CVS, plus the fuel consumption rate of each vehicle for which vehicle use was measured.

In addition to the above objectives from Transport Canada, which are mostly concerned with vehicle use and fuel consumption, there are another set of objectives related to safety issues. Currently, the CVS data is extensively used to examine the exposure to risk incurred by various segments of the driving population. While this has been very useful, it omits the exposure to risk suffered by passengers in vehicles, since few details are obtained about passengers in vehicles in the current CVS. From a road safety perspective, it would be desirable to have comparable data about drivers and passengers.

NRCan has similar objectives in that they would like to obtain vehicle-specific fuel consumption together with vehicle use characteristics which would be useful for modeling the relationships, as well as for the purpose of estimating average "degradation" of on-road fuel use against the lab-tested results. In addition, NRCan have more specific needs for use in the Transportation Energy Demand Model (TEDM). TEDM is a transportation energy-use demand model developed by the Office of Energy Efficiency to assess the energy efficiency trends in Canada's transportation sector. TEDM include road and non-road sectors. The road sector is disaggregated at the provincial level and includes nine categories of vehicles (small car, large car, light truck, medium truck, heavy truck, school bus, intercity bus, urban bus and motorcycle) and five possible fuel types (gasoline, diesel, propane, natural gas and electricity). Using vehicle stock characteristics and usages, the average distance traveled by vehicle age and the vehicle fuel economy, the model estimates the annual fuel use for each province and vehicle category. The annual fuel use estimates, or equivalently, the estimated provincial, road energy demand, are calibrated to the provincial energy demand data for the transportation sector reported by Statistics Canada. For each category of cars and trucks, the model uses annual, new vehicle sales and survival rates to construct vintaged, vehicle stock totals.

The TEDM, as well as some of the OEE Programs, require info on:

- stock of the vehicles (by GVW class and vehicle type),
- VKT (by vehicle type),
- passenger km,
- tonne km,
- fuel consumption rates and ratios (by vehicle type),
- occupancy rate,
- type of fuel (by vehicle type) and
- variables affecting driving patterns.

The details pertaining to these variables are as follows:

Vehicle Stock Data – from vehicle registration files

Car Stock
- By province (the territories included with BC)
- By weight class (0 to 2600 lb., 2600 +)
- By fuel type
- By model year
- By vehicle use sector (Personal use or Commercial use)

Truck stock
- By province (the territories included with BC)
- By weight class (0 to 8500 lb, 8501 to 33000 lb., 33001 +)
- By fuel type
- By model year
- By vehicle use sector (Personal use or Commercial use)

Vehicle Kilometres Traveled
- By province (the territories included with BC)
- By vehicle type (car, truck, bus)
- By fuel type
- By model year
- By city/highway
- By vehicle use sector (Personal use or Commercial use)

Number of passengers per vehicle
- By province (the territories included with BC)
- For passenger cars and light trucks, separately, and by weight class (possibly)
- By vehicle use sector (Personal use or Commercial use)
- By model year

Number of tonnes carried per vehicle
- By province (the territories included with BC)
- By Vehicle type
- By Weight class
- By road use: city/highway

Fuel consumption (and, possibly, on-road fuel efficiency)
- By province (the territories included with BC)
- By weight class
- By fuel type
- By model year

With respect to the weight classes for trucks, it is very important to NRCan that they get the data on truck weight in classes of 0-8500lb, 8500-33000lb, 33000+. In the current CVS, the sample is

stratified by 0-10000lbs, 10000-33000lbs, 33000 lbs or more. Unfortunately this stratification is only useful for safety standards (TC) but NRCan can't use it for TEDM because the GVW classes that are used in TEDM model (and the ones that will be used in NEMS) are 0-8500lbs (the fuel efficiency standards for LDVs), 8500-33000lbs, 33000 lbs or more. It's not important to NRCan to have 8500-10000lbs as a separate set, although this may be one way to satisfy both needs is TC still need the 0–10000lbs classification.

In considering the objectives outlined above, two issues are particularly relevant. Firstly, which variable must be collected in order to fulfill these objectives. Secondly, how can data be collected such that it can be used "for modeling the relationships between vehicle use and vehicle fuel consumption".

Considering the issue of the required variables, it would seem that the following would be a required minimum set of variables (and categories within these variables) that would need to be collected about vehicle characteristics:

- Vehicle type
  - car, truck, school bus, intercity bus, urban bus and motorcycle
- Weight class
  - Cars (0 to 2600 lb., 2600 lb+)
  - Trucks (0 to 8500 lb, 8501 to 10000 lb, 10001 lb to 33000 lb., 33001 lb+)
- Body type
  - Car, light van, light pickup, straight truck, truck tractor, bus
- Configuration (for trucks)
    - Straight truck, straight truck and trailer, tractor semi-trailer, tractor and 1 semi-trailer and 1 full trailer, 2 semi-trailers and a convertible dolly
- Fuel type
  - gasoline, diesel, propane, natural gas and electricity
- Age of vehicle (year of manufacture)
- Truck carrier type
  - Private (ancillary fleet) or For-Hire
- Truck driver type
  - Owner-operator or employee
- Province of registration
  - The territories to be included with BC

The following would be a required minimum set of variables (and categories within these variables) that would need to be collected about vehicle use:

- Date of travel

- Time of travel
- Length of trip (distance)
- Travel time of trip
- Province within which travel takes place
  - The territories to be included with BC
- Environment in which travel takes place
  - Urban driving/highway driving
- Weight of load (for trucks)
- Commodity Carried (for trucks)
  - This variable was identified by Transport Canada as one of their desired output variables from the current CVS, but such data is not currently being collected. If such information is required in CIVAFUS, then a commodity classification such as that used in VIUS could be used as a starting point (i.e agricultural and food products, mining product, building materials, forestry, wood and paper products, chemicals, petroleum and allied products, metals and metal products, other manufactured products, miscellaneous and mixed cargo)
- Driver age and gender
- Number of passengers
- Passenger ages and genders
- Trip purpose
  - This variable was identified by Transport Canada as one of their desired output variables from the current CVS, but defining the trip purpose for a vehicle does not make much sense (unless there is only one person in the vehicle or unless all persons in the vehicle have the same trip purpose). It makes more sense to define trip purpose on a person-basis, which may be possible if more data is to be collected about the passengers on each trip.

The above list of variables would provide a comprehensive description of the vehicle and vehicle-use characteristics. However, they provide no direct information of fuel consumption. Fuel consumption could be obtained via a modification of the methods used in FCS and NaPVUS (with due allowance made for situations where less than two fill-ups were made) by recording the following (based on the DMP fuel purchase survey):

- Odometer readings at start and end of recording period (say 4 weeks)
- Fuel gauge readings at start and end of recording period (this could be done visually on a fuel gauge divided into quarters)
- Date and time of each fuel purchase
- Odometer readings at each fuel purchase
- Litres purchased at each fuel purchase
- Price paid for each fuel purchase

The first two items of data are required to overcome the problems incurred in FCS and NaPVUS when less than two fill-ups were made during the recording period.

While the above data will provide the average fuel consumption between two fill-ups, or over the entire recording period, it does not necessarily provide data that can be used "for modeling the relationships between vehicle use and vehicle fuel consumption". This is because the vehicle-use recorded in the travel diary and the fuel consumption measured in the fuel purchase diary may not relate to the same travel. Indeed, it is most unlikely that it would ever relate to exactly the same travel, unless the fill-ups happened to span exactly the period of the travel diary recording. In all other cases, the two items of data will refer to different aggregations of travel. The travel diary will refer to travel made within the diary period, while the fuel consumption will refer to travel made between two fill-ups or over the entire period of the fuel consumption survey. Given the variability of travel over time (the German MobiDrive survey over 6 weeks (axhausen et al., 2002) showed a Coefficient of Variation in weekly personal VKT of 60% and a Coefficient of Variation in daily personal VKT of 150%), it is quite likely that the travel covered by the travel diary is quite different to the travel covered by the fuel consumption calculation. An extreme example would be a fuel purchase diary over 4 weeks, with a travel diary completed in the first week. It could happen that a fill-up was made just before the start of the recording period, the first week covered urban driving not requiring another fill-up, while the second week contained lots of rural driving requiring two fill-ups. The travel diary would record a moderate amount of urban driving, while the fuel consumption would be based on lots of rural driving in the second week. In this situation, it is clear that any "relationship between vehicle use and vehicle fuel consumption" would be highly spurious.

The only way in which the vehicle use and fuel consumption could be made compatible would be to make them cover exactly the same period. Options for achieving this will be described in Section 6.

Finally, even if vehicle-use and fuel consumption could be measured with respect to the same travel, it is not clear how the "relationships between vehicle use and vehicle fuel consumption" could be established, since the vehicle use would be measured in terms of the multi-dimensional characteristics of a collection of trips while the fuel consumption would be in terms of an average litres/km over that collection of trips. This problem will also be further discussed in Section 6.

6 Design Considerations for the Proposed Survey

Having reviewed the various surveys in Section 4, this section will consider some of the major issues emerging for the design of CIVAFUS, in order to meet the objectives outlined in Section 5. This discussion will outline the alternative ways of addressing each issue, and the advantages and disadvantages of each approach in the context of the objectives outlined in Section 5.

6.1 Sampling Frame

The surveys reviewed in Section 4 used either vehicle registration files (FCS, CVS, SMVU, VIUS) or rotate-outs or sub-samples of other household-based surveys (NaPVUS RTECS, DMP). Note that all the surveys using household-based frames have taken these from another survey. None have compiled the frame just for the purposes of this survey. The choice of a vehicle-based sampling frame or a household-based sampling frame will depend on the population to be surveyed, the needs of the analysis and on the availability of the frame.

Vehicle-Based Sampling Frames

Advantages

- Obtains a complete listing of all registered vehicles, including commercial vehicles.
- Gives detailed information about the vehicle, including VIN.
- Gives contact details for the registered owner.

Disadvantages

- Does not necessarily give contact details for the user of the vehicle.
- Requires significant effort to put the complete registration files together from all provinces and territories.
- Only provides one vehicle per registered owner, unless matching of vehicles from the same owner is undertaken.

Household-Based Sampling Frames from Other Surveys

Advantages

- Uses a frame which is already assembled.
- Enables the use of prior information collected in the previous survey to assist in sample design.
- Provides access to the users of vehicles, not just the registered owners.
- Encourages a higher response rate, since the household has already responded to the previous survey.

- Provides information (from the prior survey) on the characteristics of non-respondents to the latter survey, which assists in identifying non-respondent bias.

Disadvantages

- Does not give access to commercial vehicles.
- Another frame must be used to get access to commercial vehicles.
- The rotate-outs or sub-sample of respondents are a potentially biased sample of the total population, since they have already agreed to undertake the initial survey.

6.2 Survey Period

The various surveys have used different recording periods for their trip diaries, their fuel purchase diaries and their odometer reading surveys. The surveys that used trip diaries (CVS, DMP) both used 7-day diaries, although most other travel surveys using trip diaries have used a shorter recording period (usually 1 day). The surveys using fuel purchase diaries have used either 4 weeks (FCS, NaPVUS) or 8 weeks (DMP). The surveys which simply use odometer readings to record overall vehicle-use have used either a quarter (CVS territories) or 12 months (RTECS). Finally, those surveys which rely predominantly on recall of vehicle-use have used either a quarter (SMVU) or 12 months (VIUS). Since CIVAFUS will concentrate on detailed vehicle-use and fuel consumption, the following discussion will concentrate on these two types of survey, rather than the aggregate vehicle-use surveys based merely on odometer readings or recall.

7-day trip diary surveys

Advantages

- Obtains a larger sample of trips from each respondent.
- Enables distribution of travel over a week to be observed for one vehicle.
- Longer period more closely resembles period used for fuel purchase diary.
- Less likelihood of having a sampled vehicle make no trips during the recording period.

Disadvantages

- Imposes more respondent burden on each respondent.
- Generally increases non-response rate.
- Generally introduces respondent fatigue on later days in the recording period.

1-day trip diary surveys

Advantages

- Less respondent burden.
- Enables more detailed trip reporting to be attempted.

Disadvantages

- Must rely on cross-sectional data to observe differences in travel across days of the week.
- Higher likelihood of having a sampled vehicle make no trips during the recording period.

From the point of view of a vehicle-use survey, this latter disadvantage could be particularly significant (from a survey efficiency perspective). The CVS showed that 26% of all light-vehicle-days, 58% of light-truck-days, 51% of heavy-truck-days and 61% of bus-days had zero km vehicle use. While the figures for trucks and buses seem somewhat high, the figures for light-vehicles are confirmed by results from the Victorian Activity and Travel Survey (VATS) in Melbourne, which showed that 33% of all light-vehicle-days were zero-km days. Further analysis of the VATS data, however, shows some further interesting aspects of zero-km days.

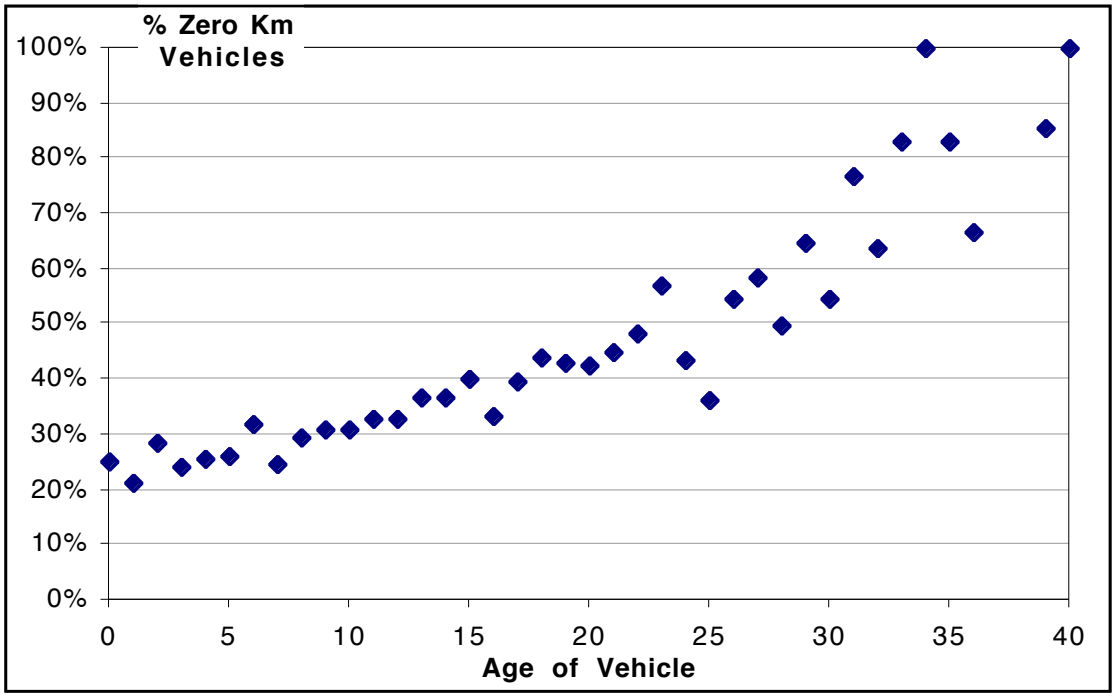
Firstly, zero-km days are more likely to happen on weekends, especially Sunday, where nearly half the cars remain unused.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Total
% Zero Km Vehicles	30%	30%	30%	29%	29%	38%	47%	33%

Secondly, it is the older cars in the household that remain unused. The newest car in the household is unused on only 28% of days, while the oldest cars in the household are unused on 55% of days.

	Newest Car	2nd Newest Car	3rd Newest Car	Older Cars
% Zero Km Vehicles	28%	38%	47%	55%

This relationship between vehicle age and non-use is further illustrated in the figure below, which shows that the probability of a vehicle remaining unused increases with a vehicle's age.



The above discussion assumes that only one vehicle per household is randomly selected for inclusion in the sample. If all vehicles in the household are selected for inclusion, then the probability of no vehicle kilometres being observed in the household decreases (especially for households with more than one vehicle). Overall, 19% of households would have zero kms traveled on any one day. This is highest for single-vehicle households (29%) and much lower for multi-vehicle households (<13%).

	1	2	3	4+	Total
% Zero Km Households	29%	12%	9%	8%	19%

Once again, the probability of having no vehicle-kilometres in the household is the highest on weekends, but the percentage has been reduced to 28% on Sundays (compared to 47% of vehicles having zero kms).

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Total
% Zero Km Households	17%	18%	18%	17%	16%	21%	28%	19%

A similar situation would occur if one extended the survey to seven days on one vehicle, rather than one day. However, being a one day survey, VATS cannot be used to investigate this effect. However, the current CVS data could be used to see how many vehicles were not used for the entire week.

The point of the above discussion is to show that extending the survey from one vehicle on one day to multiple vehicles on one day or one vehicle on multiple days are both effective ways of reducing the probability that no usage data will be obtained from one household. While the proportion of vehicles having zero travel is an essential piece of information to have, the cost efficiency of data collection can be improved by reducing the proportion of vehicles with zero usage in the sample and correcting for this under-representation by the use of weights.

The above discussion has concentrated mainly on light vehicles based in households. The problem is even worse for commercial vehicles (trucks and buses), if the figures from CVS (showing zero travel proportions on one day from 50%-60%) are correct. This proportion could be reduced by increasing the number of days of the survey, or increasing the number of vehicles surveyed within a business. Clearly, it would be infeasible to survey all vehicles in a business, but it would be feasible to survey a given number (e.g. three vehicles) from any sampled business.

It should be realized that collecting data for all days of the week for one vehicle is more respondent burden than collecting data for all vehicles in a household for one day, since the average number of vehicles per vehicle-owning-household is much less than the number of days in the week (e.g. the average vehicles per vehicle-owning-household in VATS was 1.75). Therefore, for a equal respondent burden, the proportion of zero-travel respondents should be compared for all vehicles for one day versus one vehicle for two days (each having approximately 2 vehicle-days of recording for the respondent). Under such circumstances, it would be expected that recording travel for all vehicles for one day would have less zero-travel

respondents that recording all travel for one vehicle for two days, since the correlation between kilometres traveled by vehicles in a household would be expected to be less than the correlation between kilometres traveled by one vehicle on consecutive days. For example, using the VATS data, the correlation coefficient between the kilometres traveled by one vehicle in a multi-vehicle household and all other vehicles in that same household is shown below, for each day of the week. It can be seen that the correlation if small and positive for all weekdays and Saturday, but negative on Sunday, i.e. on weekdays and Saturday if one vehicle in the household is used, then the other vehicles are slightly more likely to be used, while on Sunday if one vehicle in the household is used, then the other vehicles are slightly less likely to be used.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Correlation	0.06	0.02	0.04	0.15	0.08	0.05	-0.08

For multi-day recording of a single vehicle, it would be expected that the correlation between days would be higher, since vehicles which are used a lot on one day are probably also more likely to be used on other days, while vehicles unused on one day (especially older vehicles which are used only rarely) are also more likely to be unused on other days.

It would be useful to test out the assumptions described above about the use of vehicles on consecutive days using one quarter or a full year of CVS data, to determine authoritatively whether more vehicles on one day or more days for one vehicle are the best way to reduce zero-travel respondents while minimizing respondent burden.

The fuel consumption surveys have used recording periods of either 4-weeks (FCS, NaPVUS) or 8-weeks (GMP). The fuel consumption surveys are longer than the travel diary surveys because they are less burdensome and because a longer period is required in order to increase the probability of having more than one fill-up. There appears to be no significant advantage of having the 8-week period, however, since 4 weeks should be sufficient to obtain multiple fuel purchases for all except the most lightly used vehicles.

The fact that the travel diary and the fuel purchase diary cover different periods of time can create problems for matching the two data sets. As was mentioned in Section 5, while the fuel purchase data will provide the average fuel consumption between two fill-ups, or over the entire recording period, it does not necessarily provide data that can be used “for modeling the relationships between vehicle use and vehicle fuel consumption”. This is because the vehicle-use recorded in the travel diary and the fuel consumption measured in the fuel purchase diary may not relate to the same travel. Indeed, it is most unlikely that it would ever relate to exactly the same travel, unless the fill-ups happened to span exactly the period of the travel diary recording. The only way in which the vehicle use and fuel consumption data could be made compatible would be to make them cover exactly the same period. This could be done by starting the travel diary recording period when the next fill-up is performed and continuing until another fill-up is undertaken. These fill-ups could be as they would normally occur for that vehicle, but this could result in very long periods for the travel diary for vehicles with large fuel tank capacities and for vehicles that are not used very much. Alternatively, the fill-ups could be re-timed to coincide

with the travel diary, by asking the respondent to fill-up at the start of the travel diary recording period and to fill-up again at the end of the travel diary recording period. However, this might be expecting too much cooperation from respondents. An alternative is to ask for fuel gauge readings at the start and end of the travel diary recording period (and any fuel purchases during the recording period). This, however, probably places too much reliance on the fuel gauge readings which are very approximate measures of the contents of the fuel tank.

Finally, even if vehicle-use and fuel consumption could be measured with respect to the same travel, it is not clear how the "relationships between vehicle use and vehicle fuel consumption" could be established, since the vehicle use would be measured in terms of the multi-dimensional characteristics of a collection of trips while the fuel consumption would be in terms of an average litres/km over that collection of trips. This problem will be further discussed in Section 6.4, when alternative survey methods are discussed.

6.3 Single or Multiple Respondents

While all the reviewed surveys have used vehicles as the unit of analysis, even if they used a household-based frame to get to the vehicles, some have concentrated on a single vehicle (FCS, NaPVUS, CVS, SMVU and VIUS) while others have considered the usage of all vehicles in the household (RTECS, DMP).

Single-vehicle surveys

Advantages

- Reduces the respondent burden within the household/business.
- If sample frame was the vehicle register, then all details on the vehicle have already been obtained.

Disadvantages

- Does not show interactions between the use of all household vehicles.
- Does not capitalize on having already contacted the household/business.

Multiple-vehicle surveys

Advantages

- Shows the interactions between the use of different vehicles in a household.

Disadvantages

- Could be a substantial burden on multi-vehicle households and, particularly, multi-vehicle businesses. Must impose a limit on number of vehicles surveyed within any one household or business.
- Details of vehicles not sampled from vehicle register must be obtained from respondent.

The above description has assumed that the survey period is the same for single-vehicle and multiple-vehicle surveys. However, as noted in Section 6.2, there is an interaction between the number of vehicles to be surveyed and the number of days on which the vehicles are surveyed, e.g. one vehicle on several days or several vehicles on the one day. In terms of respondent burden, the critical issue is the number of vehicle-days of information asked for from each respondent.

If it was decided to ask about the usage of several vehicles on one day, a question arises as to how one would obtain the information about all the vehicle characteristics. For households, there are several possibilities:

- Use a household-based sampling frame, and then ask the household for the details of all vehicles (including the VIN)
- Use a vehicle-based sampling frame, select one vehicle from this frame to identify the household, and then ask the household for the details of all other vehicles in the household (including the VINs)
- Use a vehicle-based sampling frame, select one vehicle from this frame to identify the household, and then search the sampling frame to find all other vehicles registered at this same address. From these vehicles, select a random sub-set of a maximum of (say) three vehicles, then contact the household to seek usage data on these vehicles.

For the sampling of commercial vehicles from businesses, the first two of the above methods could work in theory, but not in practice. The third of these methods would appear to work much better for businesses. To maintain consistency across all the sample, the third method would therefore be recommended for a multi-vehicle survey.

6.4 Survey Method

All of the reviewed surveys used a repeated cross-sectional survey design, except for the GMP which used a sub-sample of an ongoing panel survey for the fuel consumption survey (even though the fuel consumption survey was not a true panel survey itself). For the survey objectives outlined in Section 5, there appears to be no compelling reasons to opt for a panel survey for CIVAFUS. Indeed the adoption of a panel survey design would simply bring more complications in terms of attrition bias and other biases typically associated with panel surveys. This is not to say that there is no role for panel surveys in such surveys; especially where there is a need to track changes in travel and fuel consumption over time. However, there is no expressed need for such tracking in CIVAFUS, and hence the simpler repeated cross-sectional survey should be adopted.

All of the reviewed surveys have used telephone and/or mailout surveys. Most have used a combination of phone and mailout (FCS, NaPVUS, CVS, RTECS, GMP) while the others have used mailout by itself (SMVU, VIUS). None of the surveys have used in-home or at-business personal interviews.

STC have reported the results of some tests they performed during the conduct of the CVS, comparing CATI and mailout with mailout only. They concluded that the CATI/mailout procedure was superior to the mailout-only procedure, primarily because the CATI/mailout method gave a better response rate to the travel diaries (40%) than the mailout-only method (20%). However, it should be pointed out that the mailout-only method tested was hardly state-of-the-art, with only one reminder and no pre-contact. Mailout surveys with proper pre-contact and reminder regimes have regularly obtained response rates of 40-70% in other parts of the world (e.g. Kunert et al., 2002), and this should also be achievable in Canada. In addition, it is being increasingly recognized that CATI surveys are becoming more difficult to undertake, and response rates are correspondingly reduced, because of the increasing number of ways for respondents to avoid doing a CATI survey (e.g. cell phones, answering machines, caller identification, call barring etc). There it is recommended that the comparison between CATI and mailout surveys be kept continually under review, to ensure that the most appropriate method is selected.

In most state-of-the-art surveys these days, reliance is not placed on a single survey method. Rather, a "mixed-mode" approach is used whereby different methods are used for different sub-groups in the population, or different methods are used to measure different parts of the total information required. In the context of CIVAFUS, for example, it could be that higher response rates might be achieved from businesses (and some households) by offering an Internet version of the survey, whereby information could be provided directly over the web. More importantly, the "relationships between vehicle use and vehicle fuel consumption" might be more reliably identified by instrumenting a sub-sample of sampled vehicles with GPS and fuel consumption monitors, such that the usage and fuel consumption of a sample of vehicles can be recorded over exactly the same set of trips. This would enable fuel consumption models to be developed, whereby fuel consumption rates (litres/km) could be modeled as a function of:

- Vehicle type
- Vehicle load
- Trip length
- Average trip speed
- Engine temperature (as a function of soak length)
- Ambient temperature
- Driver characteristics

Such models could then be applied to the data obtained from all the travel diaries, providing that the diaries contain sufficient detail as described in Section 6.7.

6.5 Information Collected

The information collected in the reviewed surveys can be classified into five categories: demographics/firmographics, vehicle characteristics, trip details, odometer readings and fuel purchases.

Demographics/firmographics

The characteristics of the household within which the vehicle is located and the demographics of the occupants of the vehicle on each trip are required information (for commercial vehicles, the characteristics of the business and the occupants of the vehicle on each trip are the required information).

Vehicle characteristics

Most of the vehicle data can be obtained directly from the VIN, after appropriate decoding. Some other characteristics must be obtained directly from the respondent, such as vehicle body type and configuration, transmission type and brake type.

Trip details

Some of the reviewed surveys have collected relatively detailed information about trips (CVS, GMP), while the rest have collected more aggregate information about vehicle usage. The CVS definition of a trip is relatively coarse, while the GMP definition is more in line with that used in most travel diary surveys.

Odometer readings

Most of the reviewed surveys have collected odometer readings in one form or another (FCS, NaPVUS, CVS, RTECS, GMP). Most have collected odometer readings at each fuel purchase, while some have collected it at the start of every trip (CVS). Others (FCS) have collected odometer readings at the start and end of the fuel purchase recording period.

Fuel Purchases

Four of the reviewed surveys (FCS, NaPVUS, CVS and GMP) have collected data on fuel purchases within the survey period, although the CVS period is only one week (the trip diary period) and hence cannot usually be used to calculate fuel consumption rates. Two of the surveys (SMVU and VIUS) ask the owner for an estimate of the average fuel consumption rate of the vehicle, while the RTECS estimates fuel consumption by applying adjusted EPA fuel consumption ratings to the distance traveled by each vehicle.

The combination of the types of data to be collected in CIVAFUS will depend on the ability of the package of data collected to enable the objectives of the survey to be met.

6.6 Fuel Consumption: Measurement or Modeling

In the reviewed surveys, there been essentially three ways of estimating fuel consumption rates and total fuel consumption. Three of the reviewed surveys (FCS, NaPVUS and GMP) have estimated fuel consumption rates based on fuel purchases within the survey period, two of the surveys (SMVU and VIUS) ask the owner for an estimate of the average fuel consumption rate of the vehicle, while the RTECS estimates fuel consumption by applying adjusted EPA fuel

consumption ratings to the distance traveled by each vehicle. Only the CVS makes no estimate of fuel consumption rates.

For CIVAFUS, where it is desired to estimate "relationships between vehicle use and vehicle fuel consumption", the choice of method for estimating fuel consumption lies between the use of fuel purchase diaries and the use of fuel consumption rate modeling (respondent estimation of fuel consumption rate is deemed to be too aggregate and unreliable).

Already problems have been identified in relating the fuel purchase diaries to the trip diaries because of the different periods used for each survey, and the problem of relating a one-dimensional measure of fuel consumption to a multi-dimensional measure of driving patterns. Although there are possible methods for relating the measured fuel consumption to the observed driving patterns, such methods tend to be somewhat unworkable in practice. For this reason, it is perhaps more desirable to consider the use of fuel consumption models that can convert the observed driving patterns into estimated fuel consumption.

For such fuel consumption modeling to work, two requirements are necessary. Firstly, the fuel consumption models must be detailed enough to be able to predict in-field fuel consumption rates. Secondly, the travel pattern data must be sufficiently detailed to be able to match the detail required by the fuel consumption models.

The fuel consumption models must be able to predict fuel consumption rates (litres/km) as a function of:

- Vehicle type

It is expected that the main vehicle characteristics that will affect fuel consumption rates will be the size of the engine (measured in ccs or by the number of cylinders), the type of transmission (auto or manual) and the vehicle weight.

- Vehicle load

Fuel consumption rates will be affected by the load carried by the vehicle, especially for freight vehicles. While the fuel consumption rate of passenger vehicles will also be affected by the number of passengers carried, it is unlikely that the effect will be as significant as for freight vehicles.

- Trip length

The fuel consumption rate has been shown to vary systematically with the length of the trip, with decreasing rates as the length of the trip increases. However, trip length is often a surrogate for two other factors associated with increasing trip length; average trip speeds, which increase with increasing trip length, and engine temperature, which increases with increasing trip length.

- Average trip speed

It is well-known that fuel consumption rates vary with varying vehicle speed. As speed increases from a very low speed, fuel consumption rates (litres/100km) decrease until an average speed of

approximately 60kph is reached. As speed increases further, the fuel consumption rates increase, at an increasing rate. It is also well-known that average trip speed tends to increase as trip length increases, as vehicles move from local roads near the origin to higher speed roads (arterials and freeways) further from the origin. It would therefore be expected that fuel consumption rate would, ceteris paribus, decrease with increasing trip length until a speed of approximately 60kph was reached, and then increase thereafter.

- Engine temperature (as a function of soak length)

A well-known phenomena associated with fuel consumption rates is that of "cold starts" (Taylor and Stewart, 2001). When a vehicle has been standing for some time (such that the engine temperature drops to ambient air temperature), the fuel consumption rate on start-up and until the engine warms up is higher than when the engine is started warm (after only a short standing time since the last trip). The extent of the "cold start" effect will depend on the length of the "soak" time between the end of the previous trip and the start of the current trip, and the engine temperature at the end of the previous trip. If a sequence of trips and rest periods is known, then a profile of engine temperatures can be estimated via a sequence of warming-up and cooling-down periods. Since engine temperature increases with trip length, it would be expected that fuel consumption rate would, ceteris paribus, decrease with increasing trip length.

- Ambient temperature

The effect of ambient temperature on fuel consumption rate is twofold. As ambient temperature decreases, the "cold start" effect is exaggerated since it takes longer for the engine to warm up to normal operating temperature. Secondly, even after the engine has warmed up, there is a remaining effect of low ambient temperatures in terms of the increased air resistance due to the denser air mass through which the vehicle is moving.

- Driver characteristics

The above discussion on the effect of speed on fuel consumption rate omits one very important variable. It is well-known that fuel consumption rate depends not only on the average speed but also on the extent and severity of acceleration and deceleration maneuvers. Without extremely detailed vehicle trajectory traces, however, it is very difficult to quantify such effects. It may be possible, however, to relate driving styles (incorporating the extent and severity of acceleration and deceleration maneuvers) to certain driver characteristics, such as gender, age and years of driving experience.

Such fuel consumption models can be developed by the instrumentation of a sub-sample of the vehicles selected for the survey. Such instrumentation should consist of the use of GPS vehicle tracking (to obtain speed trajectories of the vehicle over time and space) plus fuel consumption monitoring to provide fuel consumption data which can be related to travel characteristics at various levels of aggregation (e.g. to second-by-second vehicle trajectories, or to trip-based travel patterns).

Given the difficulty of relating fuel purchase data to travel pattern data, and the feasibility of developing fuel consumption models, it is suggested that the use of fuel consumption models in conjunction with detailed trip diary data is the most appropriate method of analysis.

6.7 Travel Recording

If travel patterns are to be measured and related to fuel consumption patterns, it is important that the travel patterns be measured to an appropriate level of accuracy that enables them to be used in the estimation of fuel consumption. The current CVS uses the following definitions of trips in the measurement of travel patterns:

For light vehicles, if any of the following events happened:

- a stop of more than 30 minutes
- a change of driver
- a change in the main trip purpose

For heavy vehicles (trucks) weighing 4.5 tonnes or more if any of the following events happened:

- a stop of more than 30 minutes
- a change of driver
- a change of purpose or use
- a change in the truck configuration
- a change in the status of the load from loaded to unloaded or the reverse

For buses, if any of the following events happened:

- a stop of more than 30 minutes
- a change of driver
- a change in the type of bus service
- all the passengers have been dropped off and another passenger trip begins (does not apply to scheduled urban buses)

These definitions of a trip are somewhat unconventional, compared to normal travel behavior surveys, especially the criteria of "a stop of more than 30 minutes". With this criteria, a trip chain from home to several different shops would be counted as just one trip, so long as no more than 30 minutes was spent at any one of the shops. From a fuel consumption modeling point-of-view (considering the parameters of the fuel consumption model outlined in Section 6.6), such a definition would be satisfactory from a "cold start" perspective, since a new "trip" starting less than 30 minutes after the end of a previous "trip" would still be considered as a "hot start". However, the use of such a definition would result in a lowering of the average trip speed, since the duration of the CVS-defined trip would be increased by the time spent at all the activities along the way. Such a reduction in average speed would result in an over-estimation of fuel consumption because these trips would appear to be very low-speed trips, which have high rates of fuel consumption. Therefore, it would be necessary to use a shorter stop time to denote the end of a trip, in order to not unduly lower the average speed of such trips. Since respondents tend to round trip start and end times to the nearest 5 minutes anyway, a stop of 5 minutes or more could be used to denote the start of a new trip. In the VATS survey, stops of less than 5 minutes made

up 8% of all stops between car trips, while stops of 30 minutes or less made up 33% of all stops between car trips. Thus the number of recorded trips would increase by 37% (=92%/67% -1) by changing the criteria from "a stop of more than 30 minutes" to "a stop of 5 minutes or more".

The "change of driver" criteria is presumably used to identify when an episode of exposure finishes for one driver and starts for another driver. As will be described below in section 6.9, CIVAFUS should be used to extend the measurement of exposure from only covering drivers to also covering passengers in vehicles. Under such circumstances, the "change of driver" criteria should be changed to include any change in the composition of passengers within the vehicle. Thus, picking up or dropping off a passenger would give rise to the start of a new trip.

The "change in the main trip purpose" criteria is presumably used to divide travel into purposes so that VKT by purpose and fuel consumption by purpose can be estimated. However, the current CVS defines "purpose" in terms of the main purpose being served by the vehicle trip. Since "purpose" is more an attribute of the person making the trip (since all people in the vehicle may have different reasons for making the trip), it is more appropriate to measure "purpose" on a personal level, rather than on a "vehicle" level. Since CIVAFUS will be a vehicle-based survey, and since trip purpose has no impact on the measurement of VKT, fuel consumption or exposure, it is possible that this criteria could be dropped from the definition of a trip. Thus the definition of a light vehicle trip could be:

- a stop of 5 minutes or more
- a change of driver
- a change in the number or composition of passengers.

For the definition of trips for heavy vehicles, the same comments as above apply. In addition, the criteria concerning "a change in the status of the load from loaded to unloaded or the reverse" may need to be revised somewhat. Since fuel consumption will vary with the loaded weight of the vehicle, any change in load should theoretically be recorded. However, this could lead to an overly complex diary for commercial vehicles making lots of small pickups and deliveries. On the assumption that the new "a stop of 5 minutes or more" criteria will pick up any major loadings or unloadings, all that might be required is that the load at the start of each new trip be recorded. Thus the definition of a heavy vehicle trip could be:

- a stop of 5 minutes or more
- a change of driver
- a change in the number or composition of passengers
- a change in the truck configuration
- a change in the status of the load from loaded to unloaded or the reverse (with the load at the start of each new trip being recorded).

For buses, the definition of a bus trip could be:

- a stop of 5 minutes or more
- a change of driver
- a change in the type of bus service

- all the passengers have been dropped off and another passenger trip begins (does not apply to scheduled urban buses).

6.8 Emissions

Using the survey methods employed in FCS, NaPVUS or CVS, there would be no way of obtaining information on vehicle emissions (except for carbon dioxide which is directly proportional to fuel consumption). However, if the emphasis was to change from measurement of fuel consumption to the modeling of fuel consumption, then it might also be possible to estimate emissions by means of an analogous modeling procedure. Whether this is feasible depends on the availability and accuracy of emissions models that use the same input as the fuel consumption models, and which could be obtained from the travel diary outputs. Such models have been used in Europe for CO, HC, NOx and CO2 emissions. The feasibility of this approach in Canada requires further investigation.

6.9 Road Safety Exposure

From the current CVS, it is possible to obtain some information about the exposure of drivers to crashes. However, drivers are only one component of the road safety equation. Other user groups include vehicle passengers, motorcyclists, bicyclists and pedestrians. Motorcyclists are required in the NRCan data for the TEDM model and hence should be included in CIVAFUS as a part of the vehicle registers from which the sample is drawn. It should also be possible to calculate the crash exposure of vehicle passengers by obtaining more information about the passengers in vehicles on the trips recorded in the trip diary. The characteristics of these passengers can be obtained in the household survey component of CIVAFUS, but it is recommended that the minimum data required for exposure calculations be obtained in order to minimize respondent burden. Thus age, gender, licence holding and years of driving experience should be required demographic information, but all other demographic variables will need thorough justification before being included in the survey.

6.10 Imputation

All the reviewed surveys used imputation to some extent, although FCS and NaPVUS failed to impute data for vehicles with less than two fill-ups (thereby biasing the respondents towards higher-use vehicles). The CVS probably made the most extensive use of imputation, with nearly half of all vehicle-days of travel diaries being imputed. While imputation is a useful technique for "filling holes" in the data caused by Item Non-Response (INR), one needs to be careful in using data sets where a high proportion of the data is imputed. While imputation is no different in principle to the use of weighting (in that both techniques make up for missing data), it can be more misleading for people working with the imputed data set if they do not fully understand that much of their data is imputed, based on some model of imputation. Analyses based on the imputed data set can be reflecting the imputation model more than they do the underlying data. For example, if missing fuel consumption rates were imputed based on vehicle weight and engine

capacity, and then an analysis was run to determine the factors underlying fuel consumption rates, it would not be surprising (if the full data set was used for analysis) if vehicle weight and engine capacity came out as strong predictors of fuel consumption rate!

Imputation is a useful technique for dealing with INR, but should be used sparingly. It is recommended that imputation should not be used when INR is greater than 10-15%.

6.11 Survey Management

There have been three models of survey management employed in the reviewed studies. In FCS, NaPVUS and CVS, the survey was conducted by a Government agency for another government agency. In SMVU, RTECS and VIUS, the data was collected by a Government agency for themselves or as part of their regulated duty. In the GMP, the data was collected by a private survey company for a Government agency. Given the widespread use of this public-private model in many other areas of government data collection, it is somewhat surprising that only one of the surveys used this model of survey management.

Given the concern of the client agencies with the increasing cost of the CVS, it is possible that alternative models of survey management could be investigated for CIVAFUS. Perhaps the survey could be put out to tender, with STC and other organizations being able to bid for the right to conduct the survey. This, however, may cause some problems with the use of the vehicle registration data files as a sampling frame, since it is STC and not the client agencies that have the permission to use these files.

Another management technique that might be employed is for STC to continue conducting the field work for the survey, but for Transport Canada and NRCan to appoint an independent Project Manager, who has significant experience in survey design and conduct, to act as a coordinator and overseer of the project and to act as a full-time go-between for the client agencies. This would free up the time of people in the client agencies who have many other responsibilities within their own organizations and still ensure that adequate time was put into the management of the project to ensure that the needs of the clients were being fully met. This Project Manager would liaise with STC on a continuing basis, and report back to a Management Committee of the joint clients on a regular basis.

7 Recommended Approach for the Proposed Survey

Having reviewed the various surveys in Section 4, outlined the survey objectives in Section 5 and considered some of the major issues emerging for the design of CIVAFUS in Section 6, this section will outline the approach recommended for CIVAFUS. In doing so, we adopt a staged approach, starting with what is currently done in CVS and gradually transforming into the recommended approach for CIVAFUS over a 2 year period. This allows the current data collection to proceed while various pilot surveys and data assembly activities are undertaken as a prelude to the final CIVAFUS design being implemented in 2004.

It should be obvious from the range of design issues considered in Section 5 that the recommended approach outlined in this section is but one of many possible designs. Some of the major options are outlined below.

Design Issue	**Options**
Sampling Frame	Vehicle Based Household/Company Based
Survey Period	Travel Diaries: 1-day, multi-day, 7-day Fuel Consumption Diaries: 4-week, 8-week
Number of Respondents	One vehicle (person) Multiple vehicles (persons)
Survey Method	CATI Personal Interview Mailout/mailback questionnaire Vehicle monitoring
Information Collected	Demographics/firmographics Vehicle characteristics Trip details Odometer readings Fuel purchases Driver and/or passengers characteristics
Fuel Consumption Estimation	Measurement Modeling
Travel Recording	Vehicle usage Traveler trip patterns

The recommended approach is, however, the one that we feel will best meet the current and future needs of the clients, as described in Section 5.

7.1 The Recommended CIVAFUS Approach

The major features of the recommended approach are as follows:

Name

Canadian Integrated Vehicle and Fuel Use Survey (CIVAFUS, tentatively)

Conducted for

Transport Canada and Natural Resources Canada

Conducted by

Statistics Canada. Although it is an option to put the project out to tender, this would create problems with gaining access to the official vehicle register files from the Provinces. If other options are available for obtaining a sampling frame of vehicles, then the tendering-out of the survey could be considered.

Duration of survey

The main survey would be scheduled to commence in 2004. Until then, the current CVS could be continued while the required pilot surveys and database assembly activities (see Section 8 of this report) are proceeding for CIVAFUS.

Survey Objectives

The survey objectives have been fully specified in Section 5 of this report.

Population

As with the current CVS, the population would be all registered-vehicle-days in the survey year in Canada. One required change (due to an objective of NRCan) would be that motorcycles would be included in the survey population.

Sample frame

The survey population would again be defined by the 13 jurisdiction vehicle registration lists (ten Provincial and three Territorial Governments) created three months before each quarter.

Sampling unit

The sampling unit from the sampling frame would initially be individual vehicles.

To meet the requirements of NRCan, there would be nine vehicle types.

- Motorcycles
- Light vehicles with gross vehicle weights below 2600 lbs;
- Light vehicles with gross vehicle weights of 2601 lbs and above;
- Trucks with gross vehicle weights of less than 8500 lbs;
- Trucks with gross vehicle weights of 8501 lbs or more and less than 33000 lbs;

- Trucks with gross vehicle weights of 33001 lbs or more;
- School buses;
- Intercity buses; and
- Urban buses.

Unit of Analysis

While individual vehicles would initially be selected from the sample frame, the "unit of analysis" about which information is later provided will be a group of vehicles registered at the address of the selected vehicles. The person providing the information about the vehicles and the household/company would be any knowledgeable adult who could answer the questions about vehicle and household/company characteristics. The person providing the information in the travel diary survey would be the driver of the vehicle on each trip.

Sample design

The survey would use a multi-stage, variable-fraction, stratified random sampling design. All vehicles in the sample frame would be stratified into 234 strata. First, the vehicles would be stratified into 9 vehicle types and 13 jurisdictions (ten provinces and three territories). Then, for efficiency of estimates, they would be further divided into two vehicle-age strata of newer and older vehicles.

Next, a sample of vehicles (first stage sample) would be selected from the survey population for each stratum, with the three characters of the postal code used to spread the sample over all regions (this could be done by sorting the sample frame within each stratum by postal code, and then selecting a systematic sample from the sorted list).

The registered address of each of the sampled vehicles would then be found from the vehicle register, and all other vehicles at these addresses would be selected from the sample frame. Since the original selection of household vehicles is biased towards households with more vehicles, the average number of vehicles at the selected households will be higher than the overall average number of vehicles per vehicle-owning household. For example, using the VATS data, the average number of vehicles per vehicle-owning household was 1.75, while the average number of vehicles at the households identified through a random selection of vehicles would be 2.20. For companies owning vehicles, the same process would apply, and a similar upwards bias would be noticed.

Within each of the identified households and companies, up to three vehicles would be randomly selected (second stage sample) from the list of vehicles within each household/company. If there were less than three vehicles in the household/company, all such vehicles would be selected. If there were three or more vehicles, then only three vehicles would be selected.

Finally, a Travel Day within the quarter would be randomly assigned (third stage) to each household/company selected at the second stage. Within each stratum, the Travel Day would be

evenly spread over the quarter to ensure a uniform distribution of the sample over time and for each day of the week.

Sampling stages 2 and 3 would only be applied to vehicles registered in the Provinces. These stages would not be applied to the vehicles registered in the three territories since only odometer readings at the start and end of the quarter are collected from the vehicles selected in Stage 1.

The same household or company would not be selected more than once in any four consecutive quarters, in either the Provinces or the Territories.

Sample size

The final sample size has not yet been determined. However, it is expected that it would consist of approximately 60,000 separate vehicles in the final sample.

Survey technique

The data collection for the vehicles sampled in the ten provinces would be different from the one for the vehicles sampled in the territories.

Provincial collection

Given the growing problems being encountered in obtaining satisfactory response rates to travel surveys using CATI in many countries, including Canada and the United States, it is considered that a more reliable method of conducting CIVAFUS would be via a state-of-the-art mailout/mailback survey, with appropriate reminders, motivational calls and follow-up surveys. While a mailout/mailback survey has previously been tested by StatCan in a pilot survey in the current CVS, and rejected because of low response rates (about 20% response), it is considered that this was not a fair test of the methodology since it did not use techniques that are commonly employed in state-of-the-art mailout/mailback surveys conducted in other parts of the world. Such surveys in Europe and Australia have consistently obtained response rates of 50% and greater. Pending the conduct of a comprehensive pilot survey test of the proposed methodology, it is therefore recommended that CIVAFUS be conducted primarily as a mailout/mailback survey, with appropriate reminders, motivational calls and follow-up surveys being conducted using mail, phone and personal interviews where appropriate.

The registered owners of the sampled vehicles would first receive an introductory letter, outlining the purpose of the survey and informing them that some or all of the vehicles in their household/company had been selected for inclusion in the survey. One week later, a confirmatory letter and a package of questionnaires would be sent to the registered owner, advising them of their Travel Day and providing instructions for how to complete the survey. On the day before their Travel Day, respondents with identifiable phone numbers would be called to remind them of their Travel Day and to answer any questions they may have. Completed surveys would be returned via postage-paid envelopes supplied with the questionnaires. Reminder/thank-you post-cards would be sent to all registered owners on the day after their Travel Day. Subsequently, reminder postcards would be sent to all non-respondents for the next three weeks. In the second

of these weeks, a complete pack of questionnaire material would be re-sent to non-respondents (since by now they may have lost, or perhaps never received, the original questionnaires). After the completion of the reminders, a sample of the non-respondents would be recontacted (by phone or personal interview) by experienced interviewers to ascertain the reasons for their non-response, and to gain some idea of their usage of the sampled vehicles. This will give an indication of the likely direction and magnitude of any non-response bias.

In designing this survey, due attention should be paid to the design features described by Dillman (1978, 2000) which contribute to a Total Design approach. In addition, to maximize response, a thorough understanding of the principles of persuasion developed by Cialdini (1984) and an understanding of the decision to participate in a survey (Groves et al., 1992) should be applied to the design.

Territorial collection

The registered owners of the selected vehicles would be mailed postcards and asked to provide two odometer readings, one at the beginning of the quarter and another at the beginning of the next quarter and information about the vehicle status (owned, sold, scrapped). Appropriate reminders would be used to increase the response rate from the Territories, which have had very low response rates in the current CVS.

Survey recording period

Respondents in the Provinces will be asked to complete a 1-day trip diary for the selected vehicles in their household/company. The reduction in the number of days of trip recording (compared to the current CVS) is balanced to some extent by the increase in the number of vehicles about which trip details are obtained. Since there is less correlation between the travel of different vehicles on the same day than the same vehicle on different days, the amount of new travel information gathered per vehicle-day is higher with more vehicles on the same day than with the same vehicle on more days.

Respondents in the Territories will be asked to provide an odometer reading at the start and end of the quarter for the selected vehicle.

Information collected

The information to be collected in CIVAFUS can be classified into three categories: demographics/firmographics, vehicle characteristics, and trip details.

Demographics/firmographics

The characteristics of the household or the company within which the vehicles are located are required information. However, careful attention needs to be paid to how much of this type of information is collected. As with all questions in the overall survey, only that information that contributes to meeting the specified objectives of the survey needs to be collected. While there are numerous other demographics that are collected in conventional travel surveys (for the

purposes of constructing behavioral models of travel choice), it is not considered that such variables are required from CIVAFUS. Examples of such superfluous variables would be work status, occupation, and income. Considering the objectives as currently specified by Transport Canada and NRCan, the following demographics would be required:

- Number of household members
- Age of household members
- Gender of household members
- Licence-holding status of household members
- Number of vehicles in household (confirmation of vehicle registry data)

For companies, the following firmographics would be collected:

- Type of operation (leasing company, commercial carrier)
- Type of carrier (private or for-hire)

Other demographic and firmographic variables would only be collected where the need for such variables was clearly demonstrated.

Vehicle characteristics

Most of the vehicle data can be obtained directly from the VIN, after appropriate decoding. Some other characteristics must be obtained directly from the respondent. Considering the objectives as currently specified by Transport Canada and NRCan, the following vehicle characteristics would be required (either from the VIN or the respondent):

- Vehicle type
  - car, truck, school bus, intercity bus, urban bus and motorcycle
- Weight class
  - Cars (0 to 2600 lb., 2600 lb+)
  - Trucks (0 to 8500 lb, 8501 to 10000 lb, 10001 lb to 33000 lb., 33001 lb+)
- Body type
  - Car, light van, light pickup, straight truck, truck tractor, bus
- Configuration (for trucks)
  - Straight truck, straight truck and trailer, tractor semi-trailer, tractor and 1 semi-trailer and 1 full trailer, 2 semi-trailers and a convertible dolly
- Fuel type
  - gasoline, diesel, propane, natural gas and electricity
- Age of vehicle (year of manufacture)

Other vehicle variables would only be collected where the need for such variables was clearly demonstrated.

Trip details

Information about the usage of the vehicles will be obtained directly from the trip diaries. Considering the objectives as currently specified by Transport Canada and NRCan, the following trip details would be required (see the next section for the definition of a "trip"):

- Starting odometer reading
- Starting time
- Starting place (city or suburb, and province)
- Ending odometer reading
- Ending time
- Ending place (city or suburb, and province)
- Age, sex and licence-holding status of driver
- Age, sex and licence-holding status of all passengers
- Nature of trip (personal trip or business trip)
- Nature of commercial vehicle driver (owner-operator or employee)
- GVW of truck (for commercial trips)
- Commodity carried (for commercial trips, using VIUS categories)
- Fuel purchased on trip (litres, cost per litre) and fuel gauge reading at time of purchase

Other trip variables would only be collected where the need for such variables was clearly demonstrated.

Travel recording method

If travel patterns are to be measured and related to fuel consumption patterns, it is important that the travel patterns be measured to an appropriate level of accuracy that enables them to be used in the estimation of fuel consumption. Section 6 of this report has discussed the options for various definitions of a trip (based on the current CVS definition) and has identified the following trip definitions as providing the required information at the appropriate level of detail.

For light vehicles, if any of the following events happened:

- a stop of 5 minutes or more
- a change of driver
- a change in the number or composition of passengers.

For heavy vehicles (trucks) weighing 4.5 tonnes or more if any of the following events happened:

- a stop of 5 minutes or more
- a change of driver
- a change in the number or composition of passengers
- a change in the truck configuration
- a change in the status of the load from loaded to unloaded or the reverse (with the load at the start of each new trip being recorded).

For buses, if any of the following events happened:

- a stop of 5 minutes or more
- a change of driver
- a change in the type of bus service
- all the passengers have been dropped off and another passenger trip begins (does not apply to scheduled urban buses).

Travel recording would begin at 4 a.m. on the specified Travel Day and continue until 4 a.m. on the following day. The Travel Day starting time of 4 a.m. (rather than midnight) is used to minimize the problem of trips spanning the start or end of the Travel Day. Time-use surveys have shown that 4 a.m. is the time when most people are at home and asleep, and thus least likely to be undertaking a trip.

Fuel consumption estimation

There are essentially two options for estimating fuel consumption; direct measurement of fuel consumption and modeling of fuel consumption. The direct measurement approach has been used in two previous Canadian studies (FCS and NaPVUS). While there have been some problems in these surveys, especially with respect to respondents who made less than two re-fills during the survey period, such problems could be resolved by using the techniques used in the German Mobility Panel fuel consumption survey (by obtaining odometer readings and fuel gauge readings at the start and end of the recording period). However, the direct measurement approach is not recommended for CIVAFYUS for the following reasons:

- Given that the objective of CIVAFUS is to estimate "relationships between vehicle use and vehicle fuel consumption", vehicle use and vehicle fuel consumption measurements over different periods of time are not able to be uniquely related to each other.
- Even if vehicle use and fuel consumption could be measured with respect to the same travel, it is not clear how the "relationships between vehicle use and vehicle fuel consumption" could be established, since the vehicle use would be measured in terms of the multi-

dimensional characteristics of a collection of trips while the fuel consumption would be in terms of an average litres/km over that collection of trips.

- Collecting fuel consumption estimates by direct measurement would impose an additional burden on respondents, who would have to record trip details and fuel consumption over the same period.

Rather, it is recommended that fuel consumption estimates be obtained by means of fuel consumption modeling based on the trip details recorded by the respondent (supplemented by some parallel studies using instrumented vehicles for a sub-sample of the population). The fuel consumption models will use the following characteristics of the vehicle, the trip and the driver to estimate total daily fuel consumption:

- Vehicle type

The main vehicle characteristics that will affect fuel consumption rates will be the size of the engine (measured in ccs or by the number of cylinders), the type of transmission (auto or manual) and the vehicle weight. The base fuel consumption rate (in litres/100km) will be obtained from the laboratory test results reported in the NRCan Fuel Consumption Guide (adjusted to reflect the fuel consumption rates obtained under the test drive cycle conditions and not the inflated values released to the public).

- Vehicle load

Fuel consumption rates will be affected by the load carried by the vehicle, especially for freight vehicles. While the fuel consumption rate of passenger vehicles will also be affected by the number of passengers carried, it is unlikely that the effect will be as significant as for freight vehicles.

- Average trip speed

It is well-known that fuel consumption rates vary with varying vehicle speed. As speed increases from a very low speed, fuel consumption rates (litres/100km) decrease until an average speed of approximately 60kph is reached. As speed increases further, the fuel consumption rates increase, at an increasing rate.

- Engine temperature

A well-known phenomena associated with fuel consumption rates is that of “cold starts”. When a vehicle has been standing for some time (such that the engine temperature drops to ambient air temperature), the fuel consumption rate on start-up and until the engine warms up is higher than when the engine is started warm (after only a short standing time since the last trip). The extent of the “cold start” effect will depend on the length of the “soak” time between the end of the previous trip and the start of the current trip, and the engine temperature at the end of the previous trip. If a sequence of trips and rest periods is known, then a profile of engine temperatures can be estimated via a sequence of warming-up and cooling-down periods. The

engine temperature at the start of each trip will determine the extent of the cold start effect, while the length of each trip will determine the relative contribution of the cold start to the average fuel consumption rate for the entire trip.

- Ambient temperature

The effect of ambient temperature on fuel consumption rate is twofold. As ambient temperature decreases, the "cold start" effect is exaggerated since it takes longer for the engine to warm up to normal operating temperature. Secondly, even after the engine has warmed up, there is a remaining effect of low ambient temperatures in terms of the increased air resistance due to the denser air mass through which the vehicle is moving.

- Driver characteristics

The above discussion on the effect of average speed on the fuel consumption rate omits one very important variable. It is well-known that the fuel consumption rate depends not only on the average speed but also on the extent and severity of acceleration and deceleration maneuvers. Without extremely detailed vehicle trajectory traces, however, it is very difficult to quantify such effects. It may be possible, however, to relate driving styles (incorporating the extent and severity of acceleration and deceleration maneuvers) to certain driver characteristics, such as gender, age and years of driving experience.

Fuel consumption models that account for the above factors require data about vehicle usage that is detailed enough to provide the information required for use in the fuel consumption models. Therefore, the trip diary records must be able to provide the following information about each trip:

- Type of vehicle undertaking the trip
- Extra load carried by the vehicle
- Characteristics of the driver on that trip
- Location of start of trip
- Ambient temperature at start of trip (obtained from meteorological records for the location of the start of the trip)
- Engine temperature at start of trip (obtained from engine temperature at the end of the previous trip and the soak time since the end of the previous trip)
- Engine temperature at end of trip (obtained from engine temperature at the start of the trip and the length of the current trip)
- Average trip speed (obtained from distance and travel time of trip).

The feasibility of using this approach to estimate daily fuel consumption has been tested using trip data from the VATS survey, and some plausible models of fuel consumption.

In order to make the technique operational in practice, a number of specific models concerning fuel consumption need to be developed in parallel studies using instrumented vehicles. These models would consider:

- The cool down rate when standing at rest

To account for varying soak times between trips, it is necessary to know the rate at which vehicles cool down when standing at rest after a trip. It is expected that this cool down rate will be a function of the initial engine temperature, the ambient air temperature, and the type of engine. It is expected that the cool down rate will be expressed as a function of time.

- The warm up rate when running

To account for the extent of cold-start effects, it is necessary to know how quickly vehicles warm up when running. It is expected that this warm up rate will be a function of the initial engine temperature, the ambient air temperature, the type of engine and the average speed during the warm-up period. It is expected that the warm up rate may be expressed either as a function of time or as a function of distance traveled.

- Fuel consumption penalties with cool engine

The "cold-start" effect is caused by running the engine at less than a "hot engine" temperature. It is therefore important to know what fuel consumption penalty is imposed by running the engine at various engine temperatures. It is expected that this penalty will be a function of the engine temperature, the type of engine, and perhaps the conditions under which the engine is being operated (speed and acceleration rates).

- Instantaneous fuel consumption rates

It is well-known that the fuel consumption rate is a function of the vehicle speed and rate of acceleration. Such "vehicle maps" provide a simple way of describing the instantaneous fuel consumption rate as a function of the operating conditions. These "vehicle maps" should be obtained under fully warmed-up conditions, to avoid confounding the effects of operating conditions and warm-up periods. For example, at the start of a trip, average speeds are usually lower than later in a trip. However, at the start of a trip, engine temperatures are also usually lower than later in a trip. Under these conditions, it is difficult to determine whether it is speed or engine temperature giving rise to higher fuel consumption rates at the start of a trip (or on short trips in general).

- Operating condition matrices

To use the vehicle maps described above, which show instantaneous fuel consumption as a function of vehicle speed and acceleration, it is necessary to have a matrix of operating conditions (speed and acceleration) for a range of drivers, vehicle types and trip types. Laboratory testing of fuel consumption rates uses a limited set of "drive cycles", which can also be expressed as operating condition matrices. However, it is known that the operating condition

matrices used by real drivers on real trips can be substantially different to those obtained from the classic drive cycles used in laboratory testing. The purpose of these real-life operating condition matrices would be to summarize driving patterns for real drivers on real trips.

Emissions estimation

One of the major advantages of estimating fuel consumption by the use of fuel consumption models which use the outputs of the trip diaries is that these very same techniques can be employed to estimate a range of other emissions (such as CO, CO2, NOx and HC) which are useful for a variety of other environmental studies. An example of the use of such modeling can be found in the MEET project conducted for the European Union (1999).

Imputation methods

The current CVS makes extensive use of imputation methods, especially for the generation of missing travel diaries, where nearly half the diary-days are the result of an imputation process. It is recommended that CIVAFUS should use imputation methods in dealing with item non-response (INR), but it should be used sparingly. It is recommended that imputation should not be used when INR is greater than 10-15%. Imputation should not be used to generate entire vehicle-days of trip diary data.

Expansion to population totals

Since the survey population (groups of vehicles owned by the same registered owner) differs from the population of interest (the total population of Canadian registered vehicles), several corrections need to be done to assure that the estimates correspond (as closely as possible) to the population of interest.

Since the registered owners were sampled by having one of their vehicles selected from the vehicle registry lists, registered owners with more vehicles have a higher chance of selection than registered owners with fewer vehicles. Correspondingly, each vehicle on the register has a different chance of selection. Firstly, the owner of a vehicle will have a chance of selection that is proportional to the number of vehicles owned. For example, an owner of ten vehicles will have ten times the chance of selection as the owner of one vehicle. However, since there is a maximum of three vehicles to be selected from any one owner, the vehicles owned by the owner with ten vehicles will only have a 30% chance of selection, given that their owner is selected. Therefore, the "ownership weight" to be applied to any vehicle will be 1.0 if that vehicle is the only vehicle owned by the registered owner, 0.5 if that vehicle is one of two vehicles owned by the registered owner, and 0.33 if that vehicle is one of three or more vehicles owned by the registered owner.

"Sampling weights" will be derived as the inverse of the sampling fractions used to initially select vehicles from within the various strata in the sample design. The sampling weights derived from the sample design will be adjusted and improved using updated registration lists.

"Day-of-week weights" will be calculated to ensure that each day of the quarter is appropriately represented in the final population estimates within each of the sampling strata.

“Non-response weights” will be calculated to allow for differential non-response within each of the sampling strata.

8 Recommended Pilot Studies

In order to proceed to the full implementation of CIVAFUS in 2004, there are a number of pilot studies that need to be completed in 2002 and 2003. Some of these pilot studies are concerned with the conduct of the Trip Diary survey, while others are concerned with the development of the fuel consumption models to be applied to the trip diary data.

8.1 One-day mailout with full reminders for households

This pilot survey is designed to test the implementation of a mailout/mailback questionnaire survey, seeking the completion of household, person and trip diary questions for up to three vehicles in a household.

The registered owners of the sampled vehicles would first receive an introductory letter, outlining the purpose of the survey and informing them that some or all of the vehicles in their household had been selected for inclusion in the survey. One week later, a confirmatory letter and a package of questionnaires would be sent to the registered owner, advising them of their Travel Day and providing instructions for how to complete the survey. On the day before their Travel Day, respondents with identifiable phone numbers would be called to remind them of their Travel Day and to answer any questions they may have. Completed surveys would be returned via postage-paid envelopes supplied with the questionnaires. Reminder/thank-you post-cards would be sent to all registered owners on the day after their Travel Day. Subsequently, reminder postcards would be sent to all non-respondents for the next three weeks. In the second of these weeks, a complete pack of questionnaire material would be re-sent to non-respondents (since by now they may have lost, or perhaps never received, the original questionnaires). After the completion of the reminders, a sample of the non-respondents would be recontacted (by phone or personal interview) by experienced interviewers to ascertain the reasons for their non-response, and to gain some idea of their usage of the sampled vehicles. This will give an indication of the likely direction and magnitude of any non-response bias.

8.2 One-day mailout with full reminders for businesses

This pilot survey is designed to test the implementation of a mailout/mailback questionnaire survey, seeking the completion of company, driver and trip diary questions for up to three vehicles in a company.

The registered owners of the sampled vehicles would first receive an introductory letter, outlining the purpose of the survey and informing them that some or all of the vehicles in their company had been selected for inclusion in the survey. One week later, a confirmatory letter and a package of questionnaires would be sent to the registered owner, advising them of their Travel Day and providing instructions for how to complete the survey. On the day before their Travel Day, respondents with identifiable phone numbers would be called to remind them of their Travel Day

and to answer any questions they may have. Completed surveys would be returned via postage-paid envelopes supplied with the questionnaires. Reminder/thank-you post-cards would be sent to all registered owners on the day after their Travel Day. Subsequently, reminder postcards would be sent to all non-respondents for the next three weeks. In the second of these weeks, a complete pack of questionnaire material would be re-sent to non-respondents (since by now they may have lost, or perhaps never received, the original questionnaires). After the completion of the reminders, a sample of the non-respondents would be recontacted (by phone or personal interview) by experienced interviewers to ascertain the reasons for their non-response, and to gain some idea of their usage of the sampled vehicles. This will give an indication of the likely direction and magnitude of any non-response bias.

8.3 Matching of owner addresses within the vehicle registers

In order to undertake surveys of several vehicles in the one household or company, it is desirable that all vehicles in the household/company have been previously identified on the vehicle register. This ensures that the VIN , and hence the vehicle characteristics, are available for all vehicles. From the list of identified vehicles in the household/company, up to three vehicles would then be randomly sampled for inclusion in the survey. To ensure that this process works satisfactorily, a trail matching needs to be undertaken whereby all vehicles in the register with a common owner are clearly identified. While this should be possible, the purpose of this pilot study is to ensure that no unexpected complications exist when working with the vehicle registers in this way.

8.4 Instrumentation of experimentation vehicles

In order to obtain data for the development of the fuel consumption models, vehicles will need to be instrumented to obtain the following information:

- Ambient air temperature
- Engine coolant temperature
- Transmission oil temperature
- Instantaneous fuel consumption
- Speed
- Acceleration

The above information will be required on a second-by-second basis.

Specialist advice will be needed in the development of the required vehicle instrumentation. GPS monitoring can provide the speed and acceleration information, as well as other information about gradient and the geographic location of the vehicle at all points in time. However, it will also be necessary to have some on-board recording of speed and acceleration as a backup for the

GPS data when then are gaps in the GPS data stream (such as when the vehicle is in tunnels or otherwise out of contact with the GPS satellites).

In the first phase of the instrumented vehicle pilot study, a set of experimental vehicles will be fitted out for the purposes of some controlled experimentation. These vehicles should cover the spectrum of vehicles likely to be encountered in the field, but it will not be necessary to be exhaustive in the selection of vehicle types. The main categories to be covered include various types of passenger car (4,6 and 8 cylinders; automatic and manual transmission; vehicles of different age), pickup trucks, sports utility vehicles, four-wheel-drive vehicles, light trucks, heavy trucks in the different weight categories, and various types of buses. A decision must be made as to whether it is feasible to develop an instrumentation package for motorcycles.

A range of controlled tests would be conducted for each vehicle to measure ambient air temperature, engine coolant temperature, transmission oil temperature, instantaneous fuel consumption, speed and acceleration on a second-by second basis.

These tests would be undertaken according to an experimental design that allows the various effects to be ascertained independently. For example, to identify the independent effects of vehicle speeds and engine temperature on fuel consumption rate, test runs with the same speed/acceleration matrix would need to be run with the engine at various temperatures.

8.5 Instrumentation of household vehicles

While the instrumentation of the experimental vehicles will provide much of the data for the development of the fuel consumption models, it will also be necessary to instrument vehicles belonging to sampled households and companies to see how the vehicles are actually driven under real-world conditions. This is especially important for observing the speed/acceleration matrices employed by real drivers on real trips.

Experience gained during the instrumentation of the experimental vehicles will be useful in developing instrumentation packages that can be used with vehicles belonging to households and companies. It is expected that the instrumentation packages to be used with household and company vehicles will need to be portable, self-powered, easy to fit and remove, and requiring minimum involvement of the driver. It is expected that the package will be able to store information on a second-by-second basis for a period of one week.

8.6 Development of Fuel Consumption models

The models to be developed to enable the estimation of fuel consumption from the information recorded in the trip diaries will include:

- The cool down rate when standing at rest
- The warm up rate when running
- Fuel consumption rates at various operating temperatures

- Fuel consumption rates at various ambient temperatures
- Instantaneous fuel consumption rates as a function of speed and acceleration

Since a speed/acceleration matrix will not be available for every trip recorded in the trip diary, it will be necessary to develop some models that summarize the effects of the speed/acceleration matrix for a trip on the fuel consumption for the trip. Average trip speed might capture the effect of trip speed, while it might be possible to relate the acceleration profile to various characteristics of the vehicle, the driver and the trip.

8.7 Calibration of Fuel Consumption estimations

The fuel consumption models developed above could then be applied to the trip diary data obtained from the pilot studies conducted in households and companies. This would result in an estimate of daily fuel consumption for each vehicle surveyed. These estimates could be checked in two ways, using the information on fuel purchases made on the Travel Day as recorded in the trip diary.

Firstly, at an aggregate level, the total fuel purchased (and assumed to be used) per quarter could be estimated by expanding the fuel purchases on the Travel Day for the sampled vehicles up to the fuel purchases on all days in the quarter for the population of vehicles. An independent estimate of fuel use could be obtained by expanding the estimated daily fuel consumption on the Travel Day for the sampled vehicles (using the trip diary data and the fuel consumption models) up to the fuel consumption on all days in the quarter for the population of vehicles. This will give an estimate of whether the modeled fuel consumption is of the same order of magnitude as the recorded fuel purchases. To the extent possible with the available sample size, this same comparison could be undertaken with various subsets of the data (e.g. by vehicle type).

Secondly, at the level of the individual vehicle, the modeled fuel consumption for each vehicle could be compared with the observed fuel purchasing behavior of each vehicle and a model indicating the probability of a fuel purchase being required could be developed as a function of fuel consumed. This model could also take account of the size of the fuel tank and the fuel gauge level at which drivers tend to make fuel purchases. Analyses using the VATS data have already shown that vehicles using more fuel on the travel day are more likely to make a fuel purchase on that day. While 7% of all vehicles made a fuel purchase on the travel day, only 3% of vehicles using less than 2 litres made a fuel purchase, while 16% of vehicles using 15 litres made a fuel purchase (with a relatively linear transition in between).

9 Conclusions

This project has considered the problem of obtaining estimates of vehicle use and fuel consumption from a single survey of Canadian vehicles. It has reviewed previous Canadian and foreign studies in similar areas, and has highlighted the major design issues to be addressed in the design of the new surveys (tentatively named CIVAFUS – the Canadian Integrated Vehicle and Fuel Use Survey). It has then proposed a recommended survey procedure and a set of associated pilot studies to test the applicability of the method.

CIVAFUS builds on some of the methods already used in the CVS, but also contains some significant differences from the current CVS, as summarized in the table below. The similarities are that it would be conducted by StatCan, would be a survey of vehicles, would use vehicle registers as the sampling frame, and would use a trip diary for provinces and quarterly odometer readings for territories. The differences are that the survey would be managed by an independent project manager on behalf of the two clients (TC and NRCan), would use a mailout/mailback survey with full reminder regime and follow-up surveys, would use a 1-day trip diary, would survey up to three vehicles in selected households and companies, would obtain demographic details of passengers in vehicles, would use imputation more sparingly, would use fuel consumption models in conjunction with the trip diary data in order to obtain estimates of total fuel consumption, and could use the same modeling techniques to estimate a range of emissions.

	CVS	**CIVAFUS**
Conducted by	StatCan	StatCan
Conducted for	Transport Canada	Transport Canada NRCan
Managed by	TC/StatCan	Independent Project Manager
Population	All registered-vehicle-days	All registered-vehicle-days
Sample frame	Vehicle registers	Vehicle registers
Sampling unit	Vehicles	Vehicles
Sample size	31,404 vehicles (in 2000)	60,000 vehicles (approx.)
Survey technique	CATI/Mailout Trip-Log	Mailout with full reminder regime, phone follow-up
Recording period	Provinces: 7 days trip log; Territories: quarterly odometer	Provinces: 1 day trip diary; Territories: quarterly odometer
Unit of analysis	1 vehicle	Up to 3 vehicles in household or firm
Info collected	Vehicle info; Trip data	Demographics; Full trip details; Vehicle Info.
Travel recording	Trip data	Trip data, including details of passengers
Fuel recording	Fuel purchases within the 7 days	Application of FC models to trip data
Emissions	Not considered	Modeled as per fuel consumption
Imputation	Major imputation of missing log-days	Imputation of INR (up to 15% INR)

The recommended method has been proposed after a detailed consideration of the alternative methods. It is considered that the recommended method allows the current objectives of TC and NRCan to be achieved most efficiently, while allowing scope for future developments and broadening of the environmental objectives of the survey.

10 References

Australian Bureau of Statistics (1998). Information Paper: Motor Vehicle Use, Australia. Catalogue 9219.0.

Australian Bureau of Statistics (2001). Survey of Motor Vehicle Use, Australia. Catalogue 9208.0.

Axhausen KW, Zimmermann A, Schönfelder S, Rindsfüser G and Haupt T (2002). 'Observing the rhythms of daily life: A six-week travel diary'. Transportation. 29(2), pp 95-124.

Cialdini RB (1993). Influence: The Psychology of Persuasion. William Morrow & Co.: New York.

Dillman DA (1978). Mail and Telephone Surveys, The Total Design Method. John Wiley & Sons: New York.

Dillman DA (2000). Mail and Internet Surveys, The Tailored Design Method. John Wiley & Sons: New York.

Energy Information Administration (1997). Household Vehicles Energy Consumption 1994. Publication DOE/EIA-0464(94).

Groves RM, Cialdini RB & Couper MP (1992). 'Understanding the Decision to Participate in a Survey'. Public Opinion Quarterly , 56, 475-495.

Kunert U, Kloas J and Kuhfeld H. (2002). 'Design Characteristics of National Travel Surveys; an international comparison for ten countries'. Paper presented at the 81st Annual Meeting of the Transportation Research Board, Washington DC.

Office of Energy Efficiency (2000). National Private Vehicle Use Survey – October 1994 – September 1996 Detailed Statistical Report.

Richardson, A.J. (2000). "Measurement Error Problems in Surveys of Motor Vehicle Usage", Road & Transport Research 9(4), pp. 3-10.

Statistics Canada (1989). Fuel Consumption Survey. Catalogue 53-007. Family Expenditure Surveys Section.

Statistics Canada (1998). Microdata User's Guide: National Private Vehicle Use Survey. Catalogue 53M0003GPE. Special Surveys Division.

Statistics Canada (2001). Canadian Vehicle Survey, Annual Report 2000. Catalogue 53-223-XIE. Transportation Division.

Taylor GWR and Stewart S (2001). Cold Start Impact on Vehicle Energy Use. Paper 2001-01-0221, Society of Automotive Engineers.

U.S. Census Bureau (1997). 1997 Economic Census: Vehicle Inventory and Use Survey. U.S. Department of Commerce, Economics and Statistics Administration.