

# **A Guide to Surveys of Motor Vehicle Fleets**

**November 1996**

**Energy Information Administration  
Office of Energy Markets and End Use  
U.S. Department of Energy  
Washington, D.C. 20585**

# Contacts

General information about Energy Information Administration (EIA) data on energy consumption can be obtained from **Lynda T. Carlson**, Director of the Energy End Use and Integrated Statistics Division (EEUISD) (202-586-1112). General information regarding the contents of this publication can be obtained from **Dwight K. French**, Chief of the Transportation and Industrial Branch (202-586-1126, Internet [dfrench@eia.doe.gov](mailto:dfrench@eia.doe.gov)). Specific questions regarding this publication can be directed to the author, **Jennifer Reichert** (202-586-5736, [jreicher@eia.doe.gov](mailto:jreicher@eia.doe.gov)). **Leigh Carleton** (202-586-1132, [lcarleto@eia.doe.gov](mailto:lcarleto@eia.doe.gov)), the survey manager, can be contacted regarding information on the Clean Cities Fleet Surveys.

EIA maintains a Home Page on the Internet at <http://www.eia.doe.gov>.

If you have any data or report-related requirements or suggestions for any of the other EIA consumption surveys, please contact the appropriate survey manager directly or use the address below.

Commercial Buildings Energy Consumption Survey (CBECS): Martha Johnson, Survey Manager, at 202-586-1135 (Internet: [mjohnson@eia.doe.gov](mailto:mjohnson@eia.doe.gov)).

Manufacturing Energy Consumption Survey (MECS): Mark A. Schipper, Survey Manager, at 202-586-1136 (Internet: [mschippe@eia.doe.gov](mailto:mschippe@eia.doe.gov)).

Residential Energy Consumption Survey (RECS): Robert Latta, Survey Manager, at 202-586-1385 (Internet: [rlatta@eia.doe.gov](mailto:rlatta@eia.doe.gov)).

Residential Transportation Energy Consumption Survey (RTECS): Ron Lambrecht, Survey Manager, at 202-586-4962 (Internet: [rlambrec@eia.doe.gov](mailto:rlambrec@eia.doe.gov)).

You are encouraged to provide your comments to the survey managers. Your feedback is important to us.

EI-63, Mail Stop 2G-090  
1000 Independence Avenue, SW  
Washington, DC 20585  
FAX: 202-586-0018

# Contents

	<b>Page</b>
<b>Preface</b> .....	1
<b>Defining the Target Population</b> .....	5
<b>Constructing the Survey Frame</b> .....	9
Fleet List Frame Development .....	10
Supplementing the Fleet List Frame .....	11
<b>Implementing the Fleet Survey</b> .....	15
<b>Conclusion</b> .....	19
<b>Appendices</b>	
A. Changes in Survey Procedures Between Atlanta and Denver .....	23
B. Estimation Procedures .....	27
<b>Tables</b>	
A1. Percent of In-Scope Respondent Fleets Identified on TRINC Data File by Survey City .....	23
<b>Figures</b>	
1. Survey Frame Construction .....	9
2. Fleet Survey Sample Design .....	17



# Preface

In response to directives in Section 407 of the Energy Policy Act of 1992 (EPACT), the Energy Information Administration (EIA) developed a data collection program designed to provide information useful to persons interested in the alternative fuels market. The target audience includes those seeking to manufacture, convert, sell, own, or operate alternative-fuel vehicles (AFVs)<sup>1</sup> or alternative fueling facilities.

Among the various projects EIA conducted as part of this data collection program were two fleet surveys conducted in Department of Energy-designated “Clean Cities.” The Clean Cities program is a locally-based government/industry partnership coordinated by the Department of Energy to expand the use of alternative transportation fuels. These surveys were designed to collect a broad range of information regarding the fleets and fleet vehicles in operation in the Atlanta, Georgia and Denver, Colorado areas. One of the main objectives of these surveys was to attempt to identify and describe the market for AFVs. Due to inherent limitations associated with AFVs and limited alternative-fuel infrastructure, it is believed that the first practical applications for AFVs will be within private and government fleets.

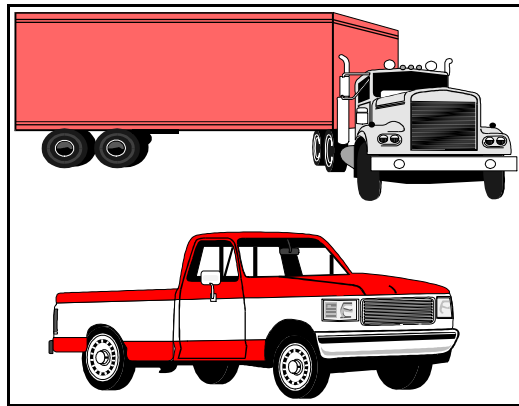
Another objective in conducting the “Clean Cities Fleet Surveys” was to develop a useful methodology for accessing and surveying private and municipal fleets that would aid other interested parties in conducting similar surveys. This report is intended to provide a description of how EIA gathered information on private and municipal fleets, but the basic survey design could be used to design surveys of other difficult-to-access populations. There are three basic steps to any survey: defining the target population, constructing the survey frame, and implementing the survey. This guide presents these three steps as they were conducted for the Clean Cities Fleet Surveys.

The procedures outlined in this report are, for the most part, the procedures used for the fleet survey conducted in Denver. There were several enhancements in the survey process instituted between the survey in Atlanta and the one in Denver. These enhancements resulted in a more efficient and reliable survey. The major changes between the two surveys are described in Appendix A.

---

<sup>1</sup> Alternative-fuel vehicles (AFVs) are defined as vehicles that have the capability of being fueled by an alternative fuel. This category of vehicles includes dual-fuel, bi-fuel, and flex-fuel, as well as dedicated vehicles. Alternative fuels include methanol, ethanol, natural gas, liquefied petroleum gas, hydrogen, coal-derived liquid fuels, fuels derived from biological materials, electricity, and any other fuel that is substantially not petroleum.





## Defining the Target Population

# Defining the Target Population

Although it was clear that the target population for the Clean Cities surveys was private and municipal fleets, the term “fleets” is an ambiguous term and needed to be defined for the purposes of these surveys. In the broadest sense, a fleet can be defined as any collection of vehicles operated by a single business or government entity. However, the business or government that operates a fleet of vehicles can separate its own fleet into smaller sub-fleets in a variety of ways. For example, a business with several locations could be considered to operate one fleet out of several locations or several fleets from several locations, or a single collection of vehicles operated by a single business could be broken down into several fleets based on vehicles size classes (i.e., light-duty, medium-duty, and heavy-duty fleets) or use (delivery fleet, executive fleet, sales representatives’ fleet, etc.). Because of these various ways a business or government can define its own fleet, it is crucial to clearly define the term “fleet” for the purposes of the current study. Four criteria were used to define what would be considered a “fleet” for the purposes of these surveys: geographic area, fleet size, vehicle ownership, and business type. The resulting definition was used to determine which fleets were in-scope for these surveys.

## Geographic Area

The areas designated as nonattainment areas by the Environmental Protection Agency (EPA) were defined as the geographic boundaries for the two surveys. These boundaries were chosen because they were consistent with other Clean City initiatives, and were previously defined by the EPA’s Clean Air Act Amendments (CAAA) to include those areas of the metropolitan regions most affected in terms of ozone levels. The nonattainment region for the Atlanta metropolitan area consisted of 13 counties, and 6 counties made up the Denver nonattainment region. For a fleet to qualify as a potential respondent for the surveys, it must be associated with one or more addresses in the designated nonattainment area and have vehicles that operate from one or more locations in that area.

## Fleet Size

The requirements of the EPACT affect fleets with more than 20 light-duty vehicles (vehicles 8,500 pounds Gross Vehicle Weight (GVW) or less) at a single location and 50 light-duty vehicles nationwide. Other Federal regulations, namely the CAAA, affect fleets with more than 10 vehicles and include in that count vehicles up to 26,000 GVW. Since both legislative mandates create marketing opportunities for marketers of alternative-fuel vehicles (AFVs), alternative fuels, and related equipment, and since the purpose of the Clean Cities Fleet Surveys was to provide useful market-based information to this group, it was important to include fleets of 10 or more vehicles. Therefore, the qualifying fleet size for the Clean Cities Fleet Surveys was set at 10 vehicles or more. However, in order to cover fleets of 10 or more that could be misclassified as fleets with fewer vehicles, fleets identified as having 6 or more vehicles were included in the sampling operations for these studies. Fleets that were identified as having fewer than 10 vehicles at the time of survey fieldwork were eliminated from the studies.

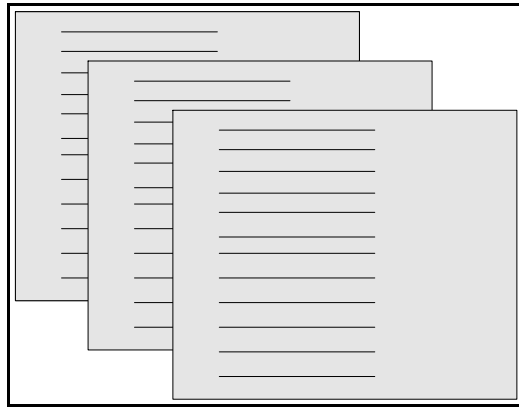
## Vehicle Ownership

Both “purchased” and “leased” vehicles were included in the survey population. Vehicles owned by employees were not within the scope of the surveys. Vehicles of interest were those that the firm operated, managed, and controlled. Because the EPACT requirements cover vehicles that are leased for terms greater than 120 days, fleets consisting of long-term leased vehicles were considered in-scope for the Clean Cities Fleet Surveys and were required to meet the same criteria as owned fleets. Shorter-term rental vehicles were not within scope of the surveys.



## **Business Type**

All businesses were included in the study with the exception of Federal and State government fleets. Federal and State fleets were excluded because it was thought that data regarding those fleets were more directly available from other sources, and State and Federal fleets could not necessarily be clearly associated with the particular areas being surveyed. Vehicle leasing companies had to be handled carefully to avoid the possibility of double counting vehicles which had been leased to businesses that were also respondents to the survey (i.e., vehicles being reported by both the lessor and the lessee). These types of businesses were instructed to respond for only those vehicles which the company maintained for its own use (e.g., shuttles operated by car rental businesses).



## **Constructing the Survey Frame**

# Constructing the Survey Frame

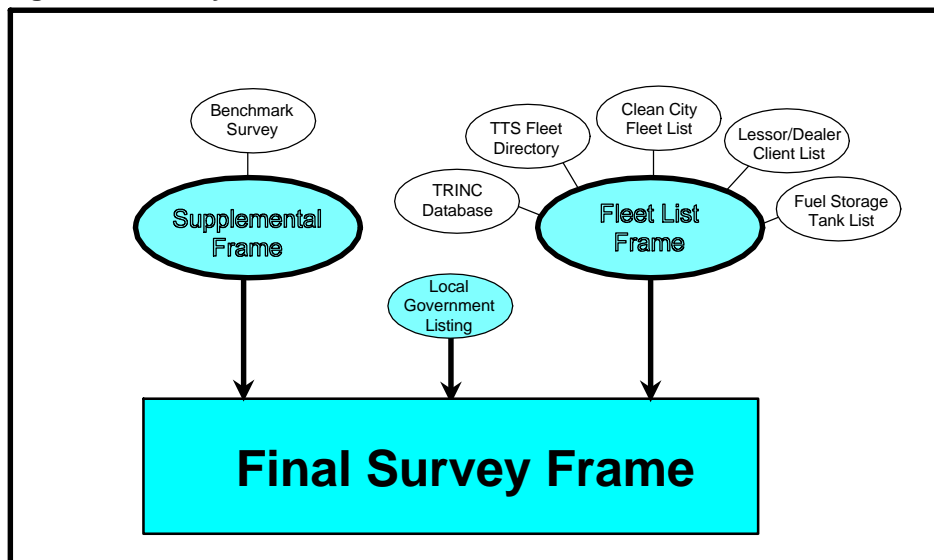
The major obstacle to overcome in the planning stages for a fleet survey is creating a well-defined listing of population units from which to select a probability-based sample for the main study. Even though an acceptable operational definition of the population can be constructed, the elements of the population are difficult to identify because:

- The population is diverse. All kinds of highway vehicles, used by a variety of businesses for a variety of purposes, under a variety of ownership/leasing/operating arrangements, are included in the population. While some listings may do a good job of covering certain homogeneous parts of the fleet population, there are not totally reliable lists of the whole.
- The population is not one that is commonly of interest to researchers. Most interest in vehicle travel and uses is focused at the vehicle level, not at the fleet (owner/operator) level. Because fleet operation ranges from the primary activity of some entities to a very peripheral activity for others, there is not clear economic motivation for comprehensive coverage of fleets. And because vehicles are subject to significant government regulation while fleets generally are not, there is no political reason to identify the fleet population specifically.
- The existence and organization of fleets are easily changed, and are thus continuously changing. Fleets come into and go out of existence as companies are established, merged, split, and dismantled. Even stable companies may decide to expand, centralize, decentralize, or outsource their fleet activities. Thus, even the most conscientiously compiled fleet list must be updated often, or it will quickly become very much out of date.

## A Dual-Frame Approach

To deal with the complexities of creating a suitable frame, EIA decided upon a “dual-frame” approach. Fleet frames, constructed from industry fleet list sources, served as the base frames for both the Atlanta and Denver surveys. However, because the coverage of these frames was uncertain, for the reasons noted above, a supplemental frame was constructed in each area--starting with a more general frame of businesses. A sample of businesses was selected and contacted to identify fleets that were not on the fleet frames. These fleets were then incorporated into the sample design for the main fleet surveys (Figure 1).

**Figure 1. Survey Frame Construction**



TRINC = The identifier used by Dun and Bradstreet to indicate their truck list database. The acronym TRINC does not stand for anything.  
TTS = Transportation Technical Service.

A common conception of a sampling frame is that it lists all the elements of the target population. However, in this case, as with other multistage, large-population surveys, it is necessary to indirectly access the target population, at least to some extent. In such situations it is commonly impractical to identify and list the total population. As long as the sampling that leads to the listing is probability-based, such an “incomplete” frame is suitable, in a probabilistic sense, for the subsequent selection process that provides the final sample for the study of interest. Frames of this type are designed to cover a substantial proportion or nearly all of the population. Efforts can be made to identify the remaining population units, but, in most cases, this process is a costly one. The results of such efforts often are not sufficient to justify the expense.

The next two sections document the procedures for developing the base Fleet List Frame and the supplement.

## **Fleet List Frame Development**

### **Fleet List Compiling**

An exhaustive search was undertaken in order to obtain as many sources as possible to construct the base fleet frame. Several sources were identified and researched. Those sources that appeared to be the most comprehensive and able to provide all needed information were purchased. All of these sources were able to provide lists of private businesses and fleets specific to areas or regions of the country.

In addition to the fleet lists purchased from industry list sources, EIA attempted to get additional lists of fleets local to the survey areas. EIA contacted numerous vehicle manufacturing and leasing companies to request lists of their client fleets in the survey areas. A small number of these companies responded and provided customer lists. These lists were provided contingent on confidentiality of the information. The Clean City Commissions within the two survey cities were also able to provide EIA with a list of known fleets. For the Denver survey, EIA obtained a list of companies that operated private fuel-storage facilities. Many of these companies were obviously public service stations, but some were selected as potentially representing private companies providing fuel for their own fleet of vehicles.

The most useful list obtained from industry sources was Dun and Bradstreet’s TRINC database. The TRINC file served as the primary list and was supplemented with Transportation Technical Service’s (TTS) Private Fleet Directory. Each of these list sources provided basic identification and contact information in addition to a fleet-size indicator. The fleet-size indicator is a number or range which indicates the number of vehicles in the fleet operated by the company. The size indicator on the TTS list was a national fleet size, whereas the TRINC file size indicators represented the size of the fleet in the survey area.

Although the industry list sources did include local and municipal government fleets, it was thought that a more reliable list of these entities could be compiled from other direct sources, such as local municipal leagues and associations, local departments of education and transportation, and directories of local/municipal officials. Therefore a list of local/municipal government agencies was created separately and added to the sampling frame. Because of this separate effort, EIA requested that all government fleets be deleted from the lists purchased from industry sources. However, there were some instances of government fleets that were inadvertently included in the purchased lists. Federal and State fleets along with duplicate municipal fleets were removed during list processing.

### **Fleet List Processing**

**Standardization.** As would be expected, all of the lists obtained from the various sources were in different formats. To create a single format for all of the records, it was necessary to standardize the records taken from each source. Only the necessary information was retained for each record from the individual sources. The desired data fields to be retained were (not all data fields were available from all list sources):

Company Name  
Street Address  
Post Office Box (if applicable)  
City  
State  
Zip Code  
Telephone Number  
Contact Name  
Fleet List Source (if available)  
Fleet-Size Indicator (if available)  
Standard Industrial Classification (SIC) Code (if available).

The records were then standardized by performing the following:

- 1) All punctuation and multiple blank spaces between words were removed.
- 2) The representation and placement of post office box numbers were standardized.
- 3) The spelling and abbreviations of individual words in the company name, street address, and city fields were changed to common forms.

**Match/Merge.** After all of the records had been standardized, the sources were matched and merged to create a single list. A specialized software designed to do matching and merging was used to accomplish this.

The TRINC database was used as the main database to which all others were matched because it contained the largest number of records. The matching software was designed to identify the seven closest matches on the main database to each company on the database being matched by using a weighting technique based on the total number of individual word matches for various data fields, such as: company name, address, contact name, and telephone number.

The software did not automatically designate matches and nonmatches but produced a report listing the possible matches. These reports were manually reviewed and each case was designated a match or nonmatch after such fields as the following were reviewed: name, address, telephone, contact, etc. If the match status of a record could not be determined from the information available, follow-up contacts were made as necessary. All records from the TRINC database and all records that did not match with the TRINC database were added to the fleet list frame. A final screening process was undertaken to eliminate any internal duplicates and companies without telephone numbers which were determined to be out of business.

**Roll-up of Frame Records.** Results from the benchmark survey in Atlanta indicated that frame coverage for medium and larger fleets improved when the entire company as a business unit was used. To facilitate a company-level (as opposed to a location-level) survey, any units on the fleet list frame which represented an individual location of a multi-location company were “rolled up” to the company level by using a company identifier which was available for most frame listings. Where no company identifier was available, the roll-up was done by manually reviewing the records. The vehicle counts were summed for all locations linked to a single company and this sum was used as the fleet-size indicator for each company. This roll-up took place prior to sampling, so an individual location that operated fewer than 10 vehicles could roll up into a company with a resulting fleet size of 10 or more. One possible drawback to performing the roll-up on the sampling frames is that the assumption must be made that the rolled up fleets behave the same as large centrally located fleets. Further analysis would be required in order to validate or invalidate that assumption.

## Supplementing the Fleet List Frame

Because the coverage on the fleet list frame was uncertain, it was supplemented with a business list frame that was used to cover businesses that had fleets but were not identified on the fleet list frame. The business list frame had no indication of whether or not a vehicle fleet was associated with any individual listing. Therefore, in order to identify fleets using the business list frame, a screening survey was conducted on a sample of cases selected from the business list frame. The screening survey was designed to determine whether or not the sampled business operated a fleet meeting

the scoping criteria for the Clean Cities Fleet Surveys. Any in-scope businesses identified that were not previously identified on the fleet list frame were incorporated into the final survey frame.

## Benchmark Survey

**Survey Frame.** The sample for the benchmark survey was drawn from the Dun and Bradstreet (D&B) comprehensive database of all businesses within the United States. D&B was able to provide a list of businesses (excluding State and Federal government entities) in just the counties being used in each of the fleet surveys. Each record on the database carries a D&B number that allows the company to be accurately identified in other lists produced by D & B. This proved to be particularly useful because the D&B TRINC database served as the core list for the fleet list frames constructed in the first stage of the dual-frame design.

Prior to performing the interviews for the benchmark survey, the frame was “rolled-up” to the company level as was done to the fleet list frame.

**Sample Design.** A stratified sampling scheme was used to select the companies to be surveyed on the benchmark survey. The sample was first stratified into two primary strata: businesses with less than 40 employees and businesses with 40 or more employees. Within each of those two strata, three substrata were defined according to the presence-of-multiple-locations categories: businesses operating from a single location, businesses operating from more than one headquarters location, and businesses operating from more than one branch location.

Although the frame used for the benchmark survey consisted of businesses in the survey area, the goal was to identify businesses that operated fleets. Because only a relatively small proportion of businesses operate fleets, it was important to ensure that the sample drawn from the business list was large enough to capture an adequate number of businesses with fleets. For both surveys, the sample for the benchmark survey amounted to approximately 5 percent of the total businesses in the survey area. It would be desirable to evaluate the sample drawn to verify that the fleets captured were representative of the population with regard to distribution across certain fleet characteristics. However, in the case of the Clean Cities Fleet Surveys, this was not possible because there was no existing information on fleets in the survey areas to use for comparison and evaluation.

**Screening Questions.** The benchmark survey was conducted by using the Computer Assisted Telephone Interview (CATI) data collection procedure. The questionnaire for the benchmark survey was designed to collect these simple data:

- Whether the company owns or leases any vehicles for use by the business within the geographic area of interest.
- How many locations does the company have in the geographic area of interest?
- How many vehicles are managed from this location?
- The name, title, and mailing address of the person within the business most responsible for decisions regarding the business’ fleet vehicles in the survey area.

However, if during the benchmark survey interview it was determined that the respondent operated a fleet of 10 or more vehicles, the fleet survey was conducted at that time. This was done in order to avoid the possibility of inconsistencies occurring between the benchmark and fleet surveys. Companies listed on the fleet list frame that had been sampled for the benchmark screening survey and had completed a fleet survey at that time were later removed from the final survey frame prior to sample selection. This was done so that when the fleet survey sample was drawn, those respondents would not be contacted again.



## **Implementing the Fleet Survey**

# Implementing the Fleet Survey

The final step in the survey process is to implement the survey instrument. For the Clean Cities Fleet Surveys, this step consisted of developing and administering a detailed questionnaire to collect a reasonable amount of information on the fleets and fleet vehicles operated by private and municipal fleets in the survey areas. The process is described in detail below.

## Questionnaire Development

A critical step in any data collection project is designing a survey instrument that asks questions in a manner that ensures the desired data will be collected. The questionnaire used for the Clean Cities Fleet Surveys went through an exhaustive, detailed design process.

A variety of AFV-related questionnaires directed to fleet managers were reviewed for content and question design. These were available from alternative-fuel industry coalitions as well as academic institutions performing work in the area of alternative fuels.

The primary source for the types of information to collect was a Federal Register Notice, which was published prior to the initial design of the questionnaire. This Notice requested suggestions from the public on the types of information that would be of use to users of fleet and fleet vehicle data. Many different types of requests were received.

Another step in the questionnaire design process was to conduct focus groups to gather information to help design the final questionnaire and develop procedures for administering this type of survey. Participants in the focus groups included fleet managers from various types of businesses that operate fleets. Focus groups were conducted before each of the Clean Cities Fleet Surveys. The focus groups were designed to:

- Determine what company representative would be the most appropriate respondent for the survey
- Refine questionnaire language
- Determine the types of data that are readily available
- Identify those questions for which the respondent might require advance notice.

In addition to the focus groups conducted with fleet managers in both Atlanta and Denver, an additional focus group was held prior to the survey in Denver which solicited comments and suggestions from vehicle marketers. This focus group was designed to collect information from the marketers regarding the types of data they would find useful and would like to see collected on a survey of this type.

## Questionnaire Design

After all comments from the Federal Register Notice and focus groups had been compiled and analyzed, the final questionnaire was developed. The questionnaire was designed to conform to the interviewing technique used for the survey. Both the benchmark survey and the fleet survey were conducted using CATI. The final questionnaire followed the following format:

### 1. Introduction

This first section of the questionnaire concentrates on reaching the appropriate person and understanding the nature of the organization. It includes several questions designed to locate the individual within the responding company who is most knowledgeable about the fleet vehicles and their usage and to ensure that the company actually operates fleets in the designated area.



## 2. Awareness of Alternative Fuel Laws/AFV Acquisitions

This section of the questionnaire explores the respondent's familiarity with laws and regulations relating to AFVs and any plans a company may have to acquire AFVs in the future.

## 3. Vehicle Stock Information

This section determines the number and types of conventional-fuel (gasoline and diesel) vehicles, in terms of vehicle body type and fuel used, the respondent has in stock. Various characteristics of those vehicles are also collected in this section. (Data on the number of AFVs by body type are collected here, but characteristics data are collected in a special AFV section.):

- Average annual miles traveled
- Average fuel economy in miles per gallon
- Age of vehicles in stock
- Average miles before replacement/retirement.

## 4. Fleet wide Vehicle Operating Characteristics

This section asks questions about various operating characteristics of the company's fleet:

- Vehicle operating cycles (days per week or year)
- Vehicle miles traveled and typical driving range
- Percent of vehicles owned and leased
- Vehicle retirement/disposal procedures
- Garaging practices
- Refueling practices.

## 5. AFV Questions

This section contains questions which collect information on the number and types of AFVs, in terms of vehicle body type and fuel used, operated by the respondent. Characteristics similar to those collected in Section 3 are also collected:

- Type of fuel delivery system (dedicated, bi-fuel, flex-fuel, dual-fuel, or hybrid)
- Secondary fuel used (if applicable)
- Average annual miles traveled
- Average fuel economy
- Maintenance cost comparison to conventional-fuel vehicles
- Driving range per fuel fill-up.

All of the above vehicle operating characteristics were collected for the following vehicle size classes:

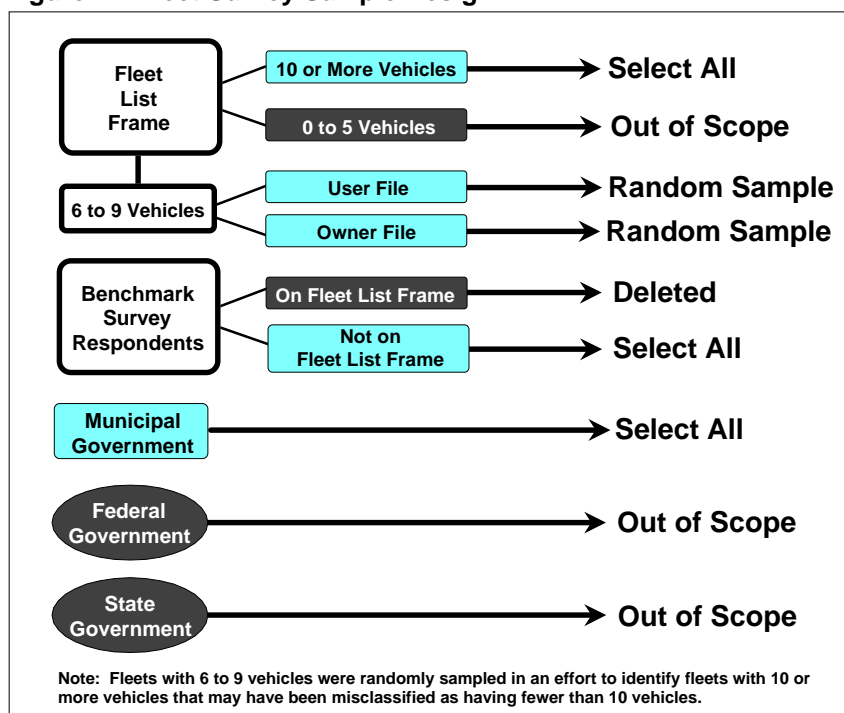
- Passenger cars
- Mini- and full-size vans
- Small and large pickup trucks
- Sport/utility vehicles
- Small, medium, and heavy trucks
- Buses.

The questionnaire for the fleet survey was designed to collect as much detail as possible; however consideration was also given to the length and complexity of the questionnaire. Although more detailed questions tend to result in more accurate data, it was important to avoid overburdening the respondents. For this reason, many questions were simplified in order to make it easier for the respondents to provide responses. It was understood that the records kept by respondents did not contain all of the data being requested on the questionnaire. In the cases where the respondents did not have recorded data to report, estimates or best guesses were permitted.

## Survey Sample Design

The sample design for the fleet survey was stratified sampling with four strata: fleets with 10 or more vehicles from the fleet list frame, local/municipal government fleets, fleets identified on the benchmark survey that were not included on the fleet list frame, and fleets with 6 to 9 vehicles. All strata but that containing fleets with 6 to 9 vehicles were sampled with certainty (Figure 2).

**Figure 2. Fleet Survey Sample Design**



For the sampling of the fleets with 6 to 9 vehicles, an additional stratification scheme was implemented. The source for the sample of small fleets was the D&B TRINC database. Within this database, there are two files. The first is a user file and the other is an owner file. The TRINC database (limited to just fleets with 6 to 9 vehicles) was stratified into two strata determined by these two files. The sample was then drawn randomly with approximately an equal number of fleets was drawn from each stratum in order to obtain approximately equivalent precision within each stratum.

## Survey Administration

Prior to fielding of the fleet survey, a pretest of the questionnaire was conducted. The telephone survey was first administered to a small group of respondents drawn from the fleet survey sample. This pretest checked the questionnaire for any errors in its logical flow and verified the clarity of the interview script. The CATI system was programmed, for both the pretest and the full survey, so that there were built-in data checks to decrease the possibility of illogical or out-of-range responses.

The fleet survey was administered to a total of 2,003 respondents in Atlanta and 1,434 respondents in Denver. The response rates achieved were 71 percent in Atlanta and 65 percent in Denver. Eliminating known ineligibles decreases these response rates to 65 percent for Atlanta and 55 percent for Denver.



# Conclusion

Useful data were obtained in Atlanta and Denver, but the nonresidential vehicles/fleet population remains a very difficult population to access and survey. Fleet lists are not complete even for the very large fleets, and they have unreliable descriptive information. If thorough coverage is required, a dual-frame approach, such as the one used for the Clean Cities Fleet Survey, is highly desirable.

In any survey of company-operated fleet vehicles, the survey design and questionnaire must take into consideration the potential diversity and complexity of fleets:

- Some companies have multiple locations in and out of the survey area.
- Fleet vehicles can be operated from one or more company locations.
- Fleets can be managed by one or more fleet managers.

If these issues are not carefully considered, respondents may be providing data for certain units not intended by the interviewer, a situation that could lead to double counting or undercoverage.

It is important to design the questionnaire to collect an appropriate level of detail from respondents. Questions that are too detailed end up with many categories, but little data, and yet valuable interview time has been spent collecting those data. Questions should also be designed to collect data that conform as closely as possible to the way fleet managers record their data. Because fleets are diverse and have numerous types of vehicles, designing for the correct level of detail is challenging and requires as much prior knowledge of the respondents as possible.

Despite the difficulties, useful data on fleets and their vehicles can be obtained from a carefully designed and monitored survey. This guide is designed to provide a starting point for developing such a survey. It should be expected that certain aspects of the process will require customization in order to conform the survey to the user's requirements and resources.

Results from the two Clean Cities Fleet Surveys in Atlanta and Denver are available as follows:

- The data from the survey in Atlanta were published in the EIA report "Profile of Motor-Vehicle Fleets in Atlanta 1994" (DOE/EIA-0601).
- The data from the survey in Denver are presented in an electronic report that is available on EIA's homepage on the Internet. This report is titled "Profile of Motor-Vehicle Fleets in Denver 1995."



# **Appendices**



# Appendix A:

## Changes in Procedures Between Atlanta and Denver

After the initial survey in Atlanta was completed, several discoveries were made of ways in which the survey processes for the Clean Cities Surveys could be improved. Based on these discoveries, changes were made to the survey processes for the survey in Denver in order either to improve the reliability of the data collected or to better facilitate the data collection process.

### Privacy States

The lists used to develop the fleet list frame are, for the most part, created using motor vehicle files available from most State Departments of Motor Vehicles (DMV). However, some States are classified as “privacy States,” which means that the DMV records are not available to the public and, therefore, cannot be used to construct fleet lists for that State. For those States where the DMV records are not available, the fleet lists are compiled from other sources and contain modeled fleet size data. The modeled fleet size information is based on the observed relationship between certain business characteristics and fleet sizes for businesses in non-privacy states. For nonprivacy States, the lists are primarily compiled from DMV vehicle records but also use modeled fleet size data for those records where the fleet size cannot be determined.

During the process of conducting the two Clean Cities Fleet Surveys, both types of fleet lists were used. For example, Georgia is a privacy State and Colorado is not. Therefore, the fleet list frame constructed for the survey in Atlanta contained modeled fleet size information and the fleet size data on the list frame for Denver was based on actual listings from DMV files. The natural assumption to be made here is that the lists used for the Denver frame construction should be more complete and reliable than those used in Atlanta. However, analysis of the final disposition codes from the two surveys reveals that, for those respondents taken from the primary list source (TRINC file), the percent of qualifying fleets identified did not improve substantially between Atlanta and Denver (Table A1).

**Table A1. Percent of In-Scope Respondent Fleets Identified on TRINC Data File by Survey City**

Survey City	Percent Qualifying	Percent Non-Qualifying
Atlanta, Georgia .....	39	61
Denver, Colorado .....	44	56

Source: Energy Information Administration, Office of Energy Markets and End Use, 1994 Atlanta Vehicles Fleet Survey and 1995 Denver Vehicles Fleet Survey.

### Fleet-Size Indicators/Fleet Survey Stratification

By the time the fleet survey in Atlanta had been completed, it was apparent that the fleet-size indicators provided on the fleet lists did not accurately reflect the true fleet sizes. These discrepancies could be due to insufficient research in the production of the fleet lists or a lack of timely updates to the fleet-size indicators on the files. Because these indicators are one of the characteristics used in stratification of the sample for the fleet survey, it is important to have reliable indicators. Otherwise, the efficiency of the sample design can deteriorate.

This problem is greatly ameliorated by separating the sample into only two strata, large and small fleets. When stratification is performed in this manner, the discrepancies between the frame fleet sizes and the actual fleet sizes will have less effect on the sample, and hence there will be less uncertainty associated with the sample design. This solution was implemented for the survey in Denver. The stratification procedures for the fleets of primary interest (fleets with 10 or more vehicles) changed significantly between the two surveys, resulting in a more simplified and efficient sample



design. For the survey in Atlanta, the fleet-size indicators from the fleet list frame were used to stratify the fleets with 10 or more vehicles into three strata. So that the unreliable size indicators would have minimal impact on the sample for the survey in Denver, all fleets of 10 or more vehicles were placed into a single stratum to be sampled with certainty.

## **Frame Coverage of Large Fleets**

One of the most basic assumptions made in the process of constructing the sampling frames for the two Clean Cities Fleet Surveys was that, at the very least, the fleet lists would contain reliable coverage of the very large fleets. It was expected that there would be coverage problems related to the smaller fleets. For both surveys, however, it turns out that there were several very large fleets that were identified on the benchmark survey but were not on the constructed fleet list frame.

The sample designs for the two fleet surveys were such that all companies identified on the benchmark survey that were not included in the fleet list frame were included in the sample for the fleet survey. The sampling weights of these companies' data were based on the original probability of selection from the benchmark survey. These weights tended to be much larger than the weights in the other strata because the population of the benchmark survey (all businesses in the survey area) was quite large relative to the number of companies selected (see Appendix B for weighting procedures).

Most of the fleets identified on the benchmark survey were smaller fleets, but the few very large companies identified through the benchmark survey reported much higher fleet vehicle counts than the majority of the sample cases. This variability within the benchmark stratum, along with the high sampling weights, created large error measures in the tabulation cells in which these large fleets were classified.

These problems with estimates with large sampling errors were a direct result of insufficient coverage on the fleet lists used to construct the fleet list frames. The ideal solution to the problem would be to somehow identify all large fleets in the survey area to be included in the fleet list frames such that the likelihood of previously unidentified large fleets being included in the benchmark stratum is low. However, it is highly improbable that this solution could be implemented because exhaustive searches of list sources were conducted prior to the Atlanta and Denver surveys, and several large fleets still were identified on the benchmark survey.

The problem could be mitigated somewhat by lowering the weights being applied to the data from the benchmark survey. Because the sampling weights are defined as the reciprocal of the probability of selection, one way to lower the weights is to increase the size of the sample being drawn. Due to budget restrictions, this was not an option for the Clean Cities Fleet Survey. To achieve lower sampling weight without increasing the sample size, a stratified sampling scheme was implemented for the benchmark survey in Denver (see below). The stratification was designed so that the strata most likely to contain the large fleets were sampled in such a way that the sampling weights were kept relatively low. This stratification of the benchmark survey sample caused the overall sampling weights from the benchmark survey in Denver to be significantly lower than those from Atlanta, while the sample size actually decreased between Atlanta and Denver.

## **Benchmark Sample Stratification**

The benchmark sampling methodology for the survey in Atlanta called for a random sample of all businesses in the target area to produce estimates of the number of fleets and identify additional fleets not on the fleet list frame. However, results from the Atlanta benchmark survey indicated that the use of other variables for stratification of the business population and drawing of the benchmark sample would improve the precision of the survey design. Analysis of the benchmark survey responses from the benchmark survey in Atlanta identified two variables, which were related to the presence or absence of fleets at a particular business location: number of employees and presence of multiple locations.

Therefore, for the benchmark survey in Denver, the sample was stratified into a total of six strata, resulting from two-level stratification, representing the breakdown of the two variables chosen. The number of employees variable was split into two categories and created two primary strata: (1) businesses with 1-39 employees, and (2) businesses with 40 or more employees. The presence-of-multiple-locations substrata within each number-of-employees stratum were defined as (1) businesses operating from a single location, (2) businesses operating from more than one location - headquarters

locations, and (3) businesses operating from more than one location - branch locations. This stratification of the benchmark survey sample was designed to improve the efficiency of this part of the sample design. However, this stratification process was unrelated to the problem of fleet coverage on the fleet lists. Therefore, in Denver as in Atlanta, large fleets identified in the benchmark sample that were not on the fleet list frame induced many large sampling errors.

## The Questionnaire

Although much effort was put into researching the target population prior to implementing the survey in Atlanta, there was still a great deal of uncertainty as to how much and what types of information the survey respondents would be able to provide. For that reason, the questionnaire was designed with extensive branching of the questions to allow for numerous possible responses. This branching resulted in a data matrix that was too large for the CATI system to manage. Due to this complexity of the questionnaire, a separate survey instrument was used to collect information on AFVs. If a respondent was determined to operate AFVs during the fleet survey, that respondent was marked for a callback on the AFV survey.

As it turned out, many of the questions asked (both on the fleet survey and the AFV survey) applied only to a very small proportion of the fleets interviewed. For the survey in Denver, the complexity of the questionnaire was reduced by disintegrating many of the data categories. The complexity was reduced and it was possible to integrate the questions regarding AFVs and AFV use into the main questionnaire.

## Survey Operations

**Benchmark Survey.** The benchmark survey and the fleet survey were two completely separate survey efforts in Atlanta. If an in-scope business was identified on the benchmark survey, that business was marked for a callback on the fleet survey. The separate fielding of these two surveys resulted in inconsistencies between responses on the benchmark and fleet surveys. For example, respondents who during the benchmark survey indicated they had a fleet of vehicles later indicated they had no vehicles when called back to administer the fleet survey.

In order to avoid these inconsistencies, the survey processes for the effort in Denver were streamlined so that there was a single survey administered to all respondents. If an in-scope business was identified on the benchmark survey, the fleet survey was conducted at that time.

**Governments and Vehicle Lessors.** Although the Clean Cities Fleet Surveys were designed to collect information from local/municipal governments and vehicle leasing companies along with other businesses, it was understood that the fleets operated by these entities were different from the traditional commercial fleets. It was suspected that government fleets might operate differently from private company fleets, and vehicle-leasing companies may operate their own fleets as well as holding vehicles for lease for outside clients. For these reasons, the decision was made for the survey in Atlanta to survey local governments separately and to conduct an exploratory survey with vehicle-leasing companies. This latter instrument was used to determine what problems, if any, would arise if lessors were surveyed for data on vehicles they operate for their own use.

These separate surveying efforts turned out to be inefficient and added to the time and effort required to complete the Atlanta survey. The survey design in Denver was simplified so that local government entities and the vehicle-leasing companies were surveyed along with all other respondents to the survey. Government entities made up their own stratum for sampling but were handled in the same manner as all other respondents once the sample had been drawn. The questions that collected information on the stock of vehicles operated by a respondent were worded so that it was clear that only those vehicles that were held for the respondent's own use were to be reported. The survey was not intended to collect information on those vehicles the leasing companies were leasing to outside clients, as double counting could occur if other respondents reported information on those vehicles.



# Appendix B:

## Estimation Procedures

Once the data collection process was completed, there were several steps involved in preparing the data for presentation. The appropriate weights had to be determined and applied to the data and there were instances of unit nonresponse that had to be accounted for (there was no adjustment for item nonresponse). After the weighted estimates had been computed, measures of error were estimated to indicate the reliability of the data.

### Estimate Weighting Procedures

Because the Clean Cities Fleet Surveys were conducted as sample surveys, it was necessary to adjust the collected raw data to represent the entire population from which the respondents were sampled. This is achieved by weighting the data with weights based on the probability of selection in each stratum. All companies in the stratum containing fleets of 10 or more vehicles and the stratum of municipal government fleets were selected with certainty, there are no sampling weights applied to the data collected from the those companies.

For the two strata where the respondents were sampled (benchmark stratum and small fleets stratum), the weights,  $W$ , are computed as the inverse of the probability of selection as follows:

$$W = \frac{P}{S}$$

where,

$P$  = number of units in the target population and

$S$  = number of units in the sample.

The companies that were identified on the benchmark survey but not on the fleet list frame were selected with certainty in the fleet survey. The weights for these companies are based on the probability of selection within the six substrata when they were originally sampled for the benchmark survey. For the stratum of small fleets (6 to 9 vehicles), the weights are calculated on the basis of the probability of selection within the two substrata.

Once the sampling weights were calculated, the estimates were computed by multiplying the data reported by the respondent by the appropriate sampling weight.

### Data Imputation

For both the benchmark survey and the fleet survey, unit nonresponse was imputed for using a weighting adjustment. An imputation factor was used to adjust the base sampling weights upwards, so that the respondent fleets would represent not only unsampled fleets, but also nonrespondent fleets. Because the fleet list frame contained many companies that did not actually have a fleet or had a fleet with fewer than six vehicles, nonrespondent fleets which would be out of scope for the surveys were excluded from the computation of the imputation factors. The imputation factors were computed as follows:

$$\frac{S_T - NR_{os}}{R}$$

where,

$S_T$  = Total number of companies sampled,

$NR_{os}$  = Number of out-of-scope nonrespondents, and

$R$  = Number of responding companies.

These imputation factors were applied to the sampling weights within each of the individual fleet survey stratum. As with the calculation of the sampling weights, the imputation factors are computed at the stratum or substratum level. There was no imputation performed for item nonresponse.

## Sampling Error Estimation

The random differences between the survey estimate and the true population value that occur because of the particular sample that was selected are known as sampling errors. The average sampling error, averaged over all possible samples, should be zero. Although the sampling error is nonzero and unknown for the particular sample chosen, the sample design permits sampling errors to be estimated. The typical magnitude of the sampling error is measured by the “standard error” of the estimate. Standard errors for the data collected on the Clean Cities Fleet Surveys were computed as percents of their estimated values, that is, as relative standard errors (RSE’s).

For a given survey statistic,  $y$ , the  $RSE(y)$  is computed as follows:

$$RSE(y) = \frac{\sqrt{V(y)}}{y}$$

$$V(y) = \sum_h [n_h W_h (W_h - 1) S_h^2]$$

$$S_h^2 = \frac{\sum_i (y_i - \bar{y}_i)^2}{n_h - 1}$$

where,

$n_h$  = sample size in stratum  $h$  and

$W_h$  = sampling weight for stratum  $h$ .