Technical Report Documentation Page

1. Report No.	*2. Government Accession No.*	*3. Recipient's Catalog No.*
4. Title and Subtitle Improved Vehicle Occupancy Data Collection Methods		*5. Report Date* April 1997 *6. Performing Organization Code*
7. Author(s) K. Heidtman, B. Skarpness, and C. Tornow		*8. Performing Organization Report No.*
9. Performing Organization Name and Address Battelle Memorial Institute 505 King Avenue Columbus, OH 43201-2693		*10. Work Unit No. (TRAIS)* *11. Contract or Grant No.* DTFH61-93-C-00055
12. Sponsoring Agency Name and Address Office of Highway Information Management Federal Highway Administration 400 7th Street, SW Washington, DC 20590		*13. Type of Report and Period Covered* Final Report 1994-1997 *14. Sponsoring Agency Code*
15. Supplementary Notes Project COTR: Ralph Gillmann (202) 366-5042		

16. Abstract

This report evaluates current and emerging vehicle occupancy data collection methodologies. Five primary methods for collecting vehicle occupancy data were identified: the traditional roadside/windshield observation method, a recently developed carousel observational method, photographic surveillance, mail-out or telephone surveys, and accident database extraction method. The findings show that other methods besides the traditional windshield method may be advantageous for collecting vehicle occupancy information. The key factors in selecting a collection method are the conditions under which vehicle occupancy is to be estimated. For example, the accident method and mail-out surveys are well suited for developing regional vehicle occupancy estimates, while the windshield method and carousel method are well suited for corridor-specific estimates. Another key finding, which impact the selection of a method and sampling period, is the fact that averae vehicle occupancy estimates vary significantly by time-of-day, day-of-week, and month-of-year. The implications of this finding is that agencies cannot continue to infer yearly AVO estimates from data collected during a limited time period such as the morning rush hour period. Recommendations for selecting and implementing a vehicle occupancy data collection method are also provided.

17. Key Words vehicle occupancy, data collection, accident databases, biasing factors		*18. Distribution Statement*	
19. Security Classif. (of this report) Unclassified	*20. Security Classif. (of this Page)* Unclassified	*21. No. of Pages* 97	*22. Price*

Form DOT F 1700.7 (8-72)

ACKNOWLEDGMENTS

The authors thank the U.S. Federal Highway Administration, particularly Ralph Gillman of the Office of Highway Information Management and Monica Francois of the Office of Environment and Planning, for direction and support of this project.

This project would not have been possible without extensive cooperation from the personnel at each participating metropolitan planning organization. These include the Baltimore Metropolitan Council in Baltimore, Maryland, the Capital District Transportation Council in Albany, New York, the Chicago Area Transportation Study in Chicago, Illinois, the Spokane Regional Transportation Committee in Spokane, Washington, and Kentuckiana Regional Planning and Development Agency in Louisville, Kentucky.

The assistance of Cambridge Systematics, Inc. in the early phases of this project, particularly the review of the state of the practice in vehicle occupancy data collection, is appreciated. In addition, the authors would like to acknowledge R.D. Mingo & Associates for their assistance in evaluating the Fatal Accident Reporting System database.

Final Report

on

Improved Vehicle Occupancy Data Collection Methods

Work Order Number BAT-94-011

Contract DTFH61-93-C-00055

April 1997

Prepared by

Battelle
. . . Putting Technology To Work

EXECUTIVE SUMMARY

Transportation planners and policy makers are increasingly concerned with improving the efficiency and effectiveness of transportation investments, and with developing performance measures to evaluate changes in transportation systems. One key transportation performance factor which reflects total person mobility is the average number of occupants traveling in automobiles. Transportation officials would benefit from systematic data on average vehicle occupancy (AVO) because vehicle occupancy data can be used for a variety of site-specific, sub-regional, area-wide, and even national purposes.

This project, jointly sponsored by the Federal Highway Administration's Office of Highway Information Management and the Office of Environment and Planning, focuses on evaluating the current state of the art and practice for collecting vehicle occupancy information; identifying and reviewing potentially valuable new methods; recommending improvements to existing data collection methodologies; and assessing the cost-effectiveness of different vehicle occupancy data collection methods.

One key finding is that AVO varies significantly by time-of-day, day-of-week, and month-of-year. This finding has profound implications for transportation professionals who often may think of AVO as a "single" value for a given region or as an input into transportation demand modeling applications. Generally, AVO is lowest during the morning peak-hour period, higher during the middle of the day, and highest during evening periods. Vehicle occupancy is generally higher on weekends than during the weekdays. In addition, summer months tend to have the highest AVOs, and the winter months generally have the lowest.

Current State of the Practice

Five primary methods for collecting vehicle occupancy data were identified: the traditional roadside windshield method, a recently developed carousel method, photographic surveillance, mail-out or telephone surveys, and the relatively new accident data extraction method. A summary of each method follows this executive summary. The overwhelming majority of vehicle occupancy data is collected by the windshield observational method, in which an observer is stationed along the side of the road and counts occupants in vehicles as they pass. A smaller number of metropolitan areas collect occupancy data through periodic surveys, in which occupancy is typically just one data element of a large regional personal travel survey. Only a few areas have experimented with relatively new photographic, accident database, or carousel collection methods.

The windshield method is generally considered the most accurate and reliable method. However, no previous studies were located that actually compared two or more vehicle occupancy data collection methods. In order to compare data collection methods, five different areas participated in this study: Albany, New York; Baltimore, Maryland; Chicago, Illinois; Louisville, Kentucky; and Spokane, Washington. Metropolitan planning organizations (MPOs) from these areas were selected based on a number of factors, including an interest in participating and previously planned data collection activities. At least two methods were evaluated in each area.

New Vehicle Occupancy Data Collection Methods

Two new vehicle occupancy data collection methods (carousel and accident databases) showed potential for providing more accurate, lower cost data, but innovative photographic methods were found to be unfeasible given current technology. The carousel method, where observers travel at slower than prevailing speeds on multi-lane highways and collect occupancy data from passing vehicles, consistently reported higher AVOs than windshield observations on heavily traveled urban highways. This finding seems logical given the close proximity between carousel observers and passing cars, typically only a few feet, whereas roadside windshield observers are usually 30 feet or more from travel lanes. In addition, roadside windshield observers frequently report difficulty observing occupants in the front passenger and rear seats of vehicles, particularly for vans and automobiles with tinted windows. Carousel observers, on the other hand, reported fewer visual problems because they could observe the passing vehicle from several angles which allowed better detection of multiple occupants in vehicles. However, the carousel method only works on multi-lane freeways, and is not appropriate for all vehicle occupancy data collection needs. For example, to estimate AVO for an entire region, data must be collected from randomly selected locations from all functional highway classes, not only Interstate highways and other multi-lane freeways.

Accident databases offer the greatest promise of readily available, low-cost vehicle occupancy data. The study found accident data to be a valid method for estimating AVO, although agencies need to exercise care in developing the estimates. The accident method relies on police reports about vehicles involved in traffic accidents, and is only available when police reports ask for the *total number of occupants* in each vehicle. As of 1992, accident forms in 16 states had specific fields for vehicle occupancy data. If this information is collected and entered into an accident database, occupancy can be extracted at relatively little cost through a database query of historical accidents. Accident data was available in four of the five metropolitan areas included in this study.

The most frequent criticism in the transportation literature against using accident databases has been that accident records are potentially biased. Potential biases include over-representation of younger drivers, and alcohol- and drug-related accidents. This study found that while some biases were present in accident databases (e.g., younger drivers are involved in proportionally higher numbers of accidents), these biases did not result in statistically significant differences in AVOs for the Chicago region in 1994 and the last six months of 1993. The Chicago data was throughly analyzed because it provided extensive demographic information that supported comparisons to the general population. National accident statistics confirm that younger drivers, who tend to have higher AVOs, are involved in more accidents. However, this may not significantly bias AVO estimates because younger drivers represent a very small portion of the entire population and are at least partially offset by older drivers who tend to have lower AVOs.

A detailed analysis should be conducted before occupancy data is used for the first time in an area. This analysis should compare the population distribution of the accident database to that of the general population. By stratifying the data by age group, weights can be applied to each group. New AVO estimates can be constructed from the weight-adjusted data, which can be evaluated for statistical significance. Once data is successfully extracted, additional analyses can be completed at a small cost.

Video and still photographic vehicle occupancy data collection methods are not recommended at this time. Although cameras can be positioned to observe vehicles, additional labor is required to set up and secure cameras, to replace film or videotapes as needed in the field, and to remove the equipment at the end of the data collection. In addition, automated methods for extracting occupancy data from the photographic record are not practical, which requires staff to review the videotapes to collect and enter the vehicle occupancy data.

Improvements to Traditional Vehicle Occupancy Data Collection Methods

Several improvements can be made to traditional methods of vehicle occupancy data collection. In many situations, roadside windshield data is collected at predetermined locations rather than at randomly selected locations. While there are legitimate reasons such as observer safety and specific study objectives that might exclude certain locations, subjective sampling prevents any statistical estimation of the accuracy of collected data. Statistically acceptable methods are available to transportation officials who would like to sample more heavily traveled highways.

Given temporal variations in AVO, data collection times also should be randomly selected, including off-peak and weekend hours (unless a study's objective is just weekday, rush-hour periods). However, many studies that were reviewed collected data on Tuesdays, Wednesdays, and Thursdays. Another finding from this research was that data collected on mid-week days only are likely to underestimate actual weekly AVO.

To improve the accuracy of the collected data, data collection staff need to be properly trained. For windshield methods, it is important for observers to accurately report data which may mean that not all cars are captured. It is not necessary for observers to record an occupancy value for every vehicle. In more than one occasion, observers were found recording an occupancy value of "one" when it was not clear how many occupants were present. Instructing observers that it is more important to record accurate values, rather than data for every vehicle, can help to lessen observer anxiety. In addition, observer fatigue has been found to significantly impact data collection. To minimize these effects, the sampling plan should included regular breaks (e.g., 10 minutes every hour) to help maintain sufficient observer attention.

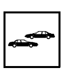
Cost-effectiveness of Different Vehicle Occupancy Data Collection Methods

Overall, no single vehicle occupancy data collection method is appropriate for all data needs. The least costly data collection method can only be determined after specific needs for the data are articulated and study objectives are agreed upon. For example, if the objective of a vehicle occupancy data collection study is to evaluate the success of an individual company's ridesharing program, a region-wide survey would not be cost-effective. However, if basic AVO estimates are desired as inputs to an area-wide travel demand forecasting model, a region-wide survey may be the most cost effective. The relevant question is: what method will result in the necessary amount of data, with the necessary amount of accuracy, at the lowest possible cost? To answer this question, a sampling plan is needed. Such statistical analyses may illustrate ways to reduce data collection activities below what was originally contemplated, thereby reducing collection costs, while still maintaining an acceptable level of accuracy. Specific methods for estimating these costs are discussed in this report. Frequently, the best approach may utilize more than one method, such as using the carousel method for Interstate highways and windshield method for lower functional class roads. However, the lowest cost method of collecting occupancy data is extracting values from existing accident databases (once the initial data set is obtained and internal expertise is developed). No other method can provide systematic data as inexpensively as accident databases.

The Five Data Collection Methods

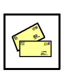
Summaries of the five data collection methods that were field tested are provided on the following pages. The five methods discussed are:

 Roadside Windshield

 Carousel Observation

 Photographic and Video Surveillance

 Accident Data Extraction

 Mail-Out Survey

Each method description is broken down into several sections:

- *Methodology* provides a general overview on the how the method attempts to collect occupancy information.
- *Field Operations* gives detailed information on the specifics of collecting occupancy data.
- *Data Processing* highlights general issues on the storage, retrieval, and filtering of the collected data.
- *Equipment Requirements* describes some of the materials needed during data collection.
- *Sampling Issues* addresses issues related to sampling of the driving population.
- *Operational Features* highlights major activities.
- *Strengths* describes the advantages and strong points of the method.
- *Limitations* describes the weaknesses of the method.
- *Recommendations* is an overall rating on the usefulness of the method.
- *References* provide several sources of further information on the method.

Methodology

The most widely used method for collecting vehicle occupancy data is the windshield method. This method involves stationing one or more observers along the roadside to count vehicles and occupants. Record keeping may be by paper and pencil, by an electronic counter or by a computer. For safety reasons, the observers must stand in a protected area, such as behind a guardrail, on an overpass, or at a reasonable distance from the traffic stream. This may limit the observer's view.

In implementing the windshield data collection method, the planning process has to consider conditions that can bias the results. The following groups of conditions have to be considered:

- *Traffic Characteristics* - type of location, spacing and time of sampling along with temporal variability, speed of traffic, traffic density, traffic weave, average occupancy, type of vehicle
- *Temporal Conditions* - time of day, day of week, season of year
- *Environmental Conditions* - weather conditions, amount of light
- *Human Factors* - observer fatigue, observer comfort, observation distance.

Ulberg and McCormack (1988) have evaluated the impact of some of these conditions. Their findings have contradicted some of the commonly held hypotheses about the influence of these conditions on counting. For instance, it is commonly thought that recording accuracy would deteriorate with traffic density. Their findings contradict this; instead, it appears that low traffic activity may result in boredom and inattention in the observer, which leads to more errors than when the traffic activity is high. Another finding is that weather conditions can influence accuracy. For example, clear skies may actually cause problems because they increase windshield glare and sun in the eyes.

Field Operations

After a region, corridor, or site has been selected, several observation sites and survey time periods are randomly chosen for collecting data. At each observation site, one or more observers are stationed to monitor the lanes of traffic. Each observer is assigned one or more distinct lanes to monitor depending on the traffic volume. Observers are instructed to collect data only on certain vehicles types. Usually, the targeted vehicle types are passenger cars, pick-ups, vans, and motorcycles. For each targeted vehicle passing the observer, the observer records the number of occupants within the vehicle and the vehicle type. Given the quantity of data that is collected with this method, it is more important for observers to record occupancy data for vehicles that they clearly see rather than recording a “value” for every vehicle. Selective sampling of the traffic, such as collecting data on every other vehicle or for certain time periods, could help improve the observer's accuracy while obtaining a representative sample of the traffic.

Data Processing

- Computer/electronic counter data is downloaded to a spreadsheet.
- Paper and pencil data is typed into a computer spreadsheet.

Roadside Windshield Method

Equipment Requirements

- Transportation to sites.
- Counters/computers.
- Paper and pencil survey forms.

Sampling Issues

- For low volume traffic, 90% or more of the vehicles can be captured from a single lane.
- For high volume traffic, an average of 75% of the vehicles can be captured from a single lane.
- For a multi-lane road segment, all lanes are typically monitored (usually requiring more than one observer).
- Collecting data for 20 minutes with a 10 minute break (or some other time allotment) still produces valid sample sets and may reduce observer fatigue and improve the accuracy of collected data.

Operational Features

- Hiring personnel.
- Training observers.
- Transportation to and from sites.

Strengths

- Requires little equipment.
- Easy to implement.
- Standard method which produces results comparable to historical data.

Limitations

- Labor intensive.
- Degradation in observer's performance over time.
- Limited time to view vehicles.
- Visual perspective may hinder ability to see children or backseat occupant.
- Collection possible during daylight hours only.
- Concerns for observer's safety.
- Inability to use method on highways with inadequate shoulder areas.

Recommendations

- Use random sampling techniques.
- Randomly select collection sites and collection times.
- Use selective sampling to improve accuracy of data and/or reduce costs.

References

- Ferlis, R.A. (March 1981). *Guide for Estimating Urban Vehicle Classification and Occupancy*. FHWA Report No. HS-032-518.

Carousel Observation Method

Methodology

The carousel method, an alternative to the roadside windshield method, positions observers in vehicles traveling with traffic in order to collect vehicle occupancy data on neighboring vehicles. On multi-lane highways, the observation vehicle is in the right most lane and drives slightly slower than the prevailing traffic speed, resulting in the traffic continuously passing by the observation vehicle. When the survey vehicle reaches the defined limits of the survey area, it then turns around and drives the same route in the other direction. Thus, bi-directional observations are recorded for the selected road segments.

Field Operations

During the planning process, road segments are selected and turn-around areas are marked. Prior to the collection phase, the average traffic volumes and speeds traveled along the selected highway segment can be reviewed to determine safe traveling speeds for the survey vehicles and to determine the number of survey vehicles that will be needed to collect an adequate sample size.

For each survey vehicle, a designated driver and an observer are assigned. The project coordinator is usually the driver in the lead vehicle. While the survey vehicle drivers focus on the task of maintaining safe speeds and proper distance from the other survey vehicles, the observers monitor all lanes and record occupancy counts with an electronic counter or a computer. Collection is conducted for a specified period, usually a two or three-hour interval during morning or evening peak hours.

The role of the project coordinator during the collection session is to coordinate and properly phase the survey vehicles to avoid double counting. The coordinator develops a consensus of the minimum safe speed for all survey vehicles. The coordinator tries to maintain a speed about 10 to 15 miles per hour slower than the average traffic speed to achieve the highest sampling rate possible under safe conditions. Since traffic conditions can fluctuate, communication by a portable telephone (or some other means) helps to resolve any problems encountered.

Data Processing

Data collected with the electronic counters or laptop computers must be downloaded to a central computer, and additional information on the road segment, such as number of lanes or weather, may have to be attached to these records.

Equipment Requirements

- Survey vehicles for observers.
- Portable telephones/other communication devices (optional).
- Electronic counters/computers.

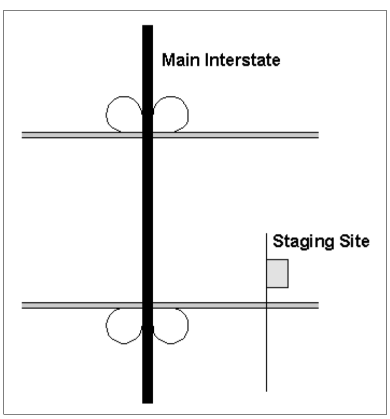
Carousel Observation Method

Sampling Issues

- All traffic that flows by the survey vehicle is captured.
- Sampling rate is dependent upon prevailing speeds and congestion.
- Adequate sampling rate occurs when survey vehicle is 10 to 15 mph slower than prevailing traffic.
- Approximately 25% of the vehicles are captured, compared to 75-90% for the windshield method (when three survey vehicles used).

Operational Features

- Transportation to site.
- Staging area near site (see map).
- Training observers and drivers.
- Spacing between observation vehicles.



Strengths

- Observer's perspective and viewing time is better than roadside observation.
- More accurate data, particularly for heavy flows.
- Safety of observers can be enhanced.

Limitations

- Need continuously moving traffic stream.
- Can only be performed on a multi-lane highway.
- Can obstruct traffic because the vehicle is slower than the surrounding traffic.
- Counts of vehicles observed is less than stationery roadside counts.
- Best used during daylight hours with potential for limited night-time use.
- Proper coordination to avoid double counting.

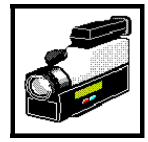
Recommendations

- May be best used in conjunction with other methods. (e.g., carousel method on Interstates and windshield method on other principal arterials.)
- For a multi-lane corridor, can be more cost effective than the windshield method.

References

- Battelle. (February 1997). *Improved Vehicle Occupancy Data Collection Procedures.*
- Baltimore Metropolitan Council. (1997). *FHWA Comparison Survey of Data Collection Methodologies*. Final Report.

Photographic and Video Surveillance Method



Methodology

Like the windshield method, the photographic and video collection methods observe traffic from a stationery position along the roadside; however, instead of human observers, cameras collect occupancy data. The use of video and photographic surveillance systems to capture occupancy information has been limited because of the costs. However, interest in this method has been growing. MPOs in Florida are using video to collect travel time data because of a state requirement to implement congestion management systems, which may have to rely more heavily on real-time traffic monitoring devices (Pietrzyk, et al, 1996). The biggest drawback is that collection of vehicle occupancy counts requires manual viewing of the videotapes. Automated methods for accurately extracting occupancy data from video are not available and appear to require further years of development.

Field Operations

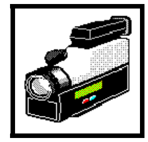
Prior to the data collection day, the selected survey sites should be visited in order to determine camera placement and camera exposure settings. A professional photographer or video technician should be consulted during this phase because proper positioning and operation of cameras will reduce the loss of data due to windshield glare. Cameras are usually positioned near underpasses to reduce or eliminate windshield glare. One camera often will be set to record vehicles traveling on two lanes of a uni-directional traffic stream. Test film should be collected and reviewed to determine image quality. For some conditions, special film may have to be purchased that can capture and enhance the visibility of images.

During collection of the traffic data, methods for monitoring and operating the equipment may vary depending on the type of equipment and manpower available. Cameras may be left unattended at a site or persons may stay with the equipment until the end of the collection session. If the cameras record everything to videotape or film, then periodical replacements of the film need to be scheduled. After the collection session, equipment may have to be retrieved from the collection sites, and all videotape should be gathered together and appropriately marked with route, road segment, date, and time period.

Data Processing

Vehicle occupancy data is extracted by observing the videotapes. Currently, image processing software that can effectively analyze the film is not available; therefore, occupancy data has to be manually extracted by viewing the footage. One way to minimize labor cost is for an observer to view random segments (e.g. 15 minute blocks) of videotape, rather than recording data for all vehicles. Special film enhancement equipment may be required to conduct manual reduction of film. During the extraction process, the person viewing the film should directly record all information into a computer.

Photographic and Video Surveillance Method



Equipment Requirements

- Cameras.
- Monitors and VCRs.
- Film.
- Operation manuals.
- Transportation to site.

Sampling Issues

- Vehicles in two uni-directional lanes captured per camera.
- Specific time segments can be obtained from videotape records.

Operational Features

- Professional photographer to train field operatives.
- Video equipment needs to be protected from theft.
- Periodic film replacement required.

Strengths

- Video provides a permanent record.
- No problems with observer fatigue.
- Variety of data collected besides occupancy (e.g., weather).
- Variety of sampling is available from complete video record.

Limitations

- Equipment can be expensive.
- Data loss due to windshield glare.

- Hiring a professional photographer in order to reduce data loss.
- Training time on using equipment.
- Need to develop photographic film.
- Manually observing videotapes to record vehicle occupancy counts is time consuming unless samples are taken.

Recommendations

- Not a recommended method. Requires further development.
- Limit use to specific sites, such as monitoring HOV lanes.

References

- Billheimer, J., Kaylor, K., and Shade, C. (March 1990). *Use of Videotape in HOV Lane Enforcement: Final Report*. California Department of Transportation.
- Battelle. (February 1997). *Improved Vehicle Occupancy Data Collection Procedures.*
- Miller, K., Harvey, T., Shuldner, P., and Ho, C. (1993). "Using Video Technology to Conduct 1991 Boston Region External Cordon Survey." *Transportation Research Record 1412*, pp. 46-56.
- Pietrzyk, M. et al. (January 1996). *Demonstration of Video-Based Technology for Automation of Data Collection*. Center for Urban Transportation Research, University of South Florida.

Accident Data Extraction Method

Methodology

The number of passengers in a vehicle involved in an accident is routinely recorded in many states by law enforcement officers at the site of the accident. Because of the accessibility and low cost associated with obtaining vehicle occupancy data from existing accident databases, several states have recently shown some interest in this new approach (Ahuja and Hanscom, 1996; Asante, Adams, Shufon, and McClean, 1996; Gaulin, 1991). Accident reports typically record information about the location and time of the accident, and in certain states they also report the number of occupants in each vehicle involved. From a collection of these accident records, it may be possible to derive an estimate of vehicle occupancy for a particular road segment, corridor, metropolitan area, or route for specific time periods. The accuracy of this method to estimate AVO is still being explored, along with alternative methods for calculating AVO that are based on different weighting schemes to adjust for possible oversampling of certain subpopulations (e.g., younger drivers who are involved in more accidents).

Field Operations

No field operations are conducted for this method. However, it is important to understand how the field data is collected by the law enforcement agencies.

Data Processing

After obtaining the accident data from the state DMV or other appropriate agencies, the data usually has to be filtered before it can be analyzed. Filtering may exclude records for vehicle types outside the set of interest, e.g., non-passenger vehicles. Other filtering criteria may be necessary upon further examination of the data. For instance, some states assign an empty parked car an occupancy of one, and data for certain time periods (e.g., midnight to 6AM) may not be of interest.

Since accident data may not provide a representative sample of the driving population, the data will have to be examined and adjusted to avoid biasing effects in the calculation of an AVO. Different methods for adjusting the sample to make it more reflective of the population are being developed based on census data.

Sampling Issues

- Depends on the size of the metropolitan area that an AVO estimate is required for.
- For example, a corridor may require several months of accident data to get a reliable AVO estimate.
- A specific time period must be selected since AVO differs by time of day, day of week, and season.

Accident Data Extraction Method

Strengths

- Requires no collection effort.
- Provides good regional samples.
- Records may contain additional information such as driver's age.
- Availability of new data that is regularly updated.
- Even if certain biases are present in the data, accident data seems to reflect seasonal and other long term trends and can "fill-in" between other data collection methods.

Limitations

- Potentially, small sample sizes for specific sites and time periods.
- Not necessarily a representative sample of the driving population.
- AVO may be biased. Analysis of the population represented is recommend for appropriate weighting.
- High occupancy vehicle (HOV) lanes underrepresented because fewer accidents are reported for those lanes.

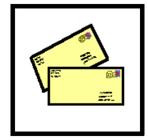
Recommendations

- Select a specific time period because AVO varies across temporal conditions.
- Evaluate possible biasing factors in data and adjust for their impact on AVO estimates.
- Increase the sample size by selecting a wider corridor for analysis.

References

- Ahuja, V., and Hanscom, E. (1996). "Feasibility Study to Estimate the Average Vehicle Occupancy by Utilizing a Traffic Accident Database." Presented at the 1996 Annual TRB Meeting.
- Asante, S., Adams, L., Shufon, J., and McClean, J. (1996). "Estimating Average Automobile Occupancy from Accident Data in New York." Presented at the 1996 Annual TRB Meeting.
- Battelle. (April 1997). *Improved Vehicle Occupancy Data Collection Procedures.*
- Gaulin, R. (April 1991). *A Procedure to Calculate Vehicle Occupancy Rates From Traffic Accident Data*. Connecticut Department of Transportation, Bureau of Planning.

Mail-Out Survey Method



Methodology

Many types of surveys have been developed to collect information directly from travelers. They include telephone interviews, travel diaries, parking lot surveys, employer surveys, household surveys, and mail-out surveys. This section describes data collection using the mail-out survey method, which obtains occupancy information by contacting residents via postcards or letter questionnaires.

Field Operations

The mail-out survey method collects vehicle occupancy information from a selected population of residents living in specific geographic regions such as mailing zip-code areas. A random sample of residents living in each area are sent questionnaires on their general travel patterns. For maximum efficiency, the number of surveys mailed should be proportional to the population in each area or zip code. A common practice is to send out surveys to 10 percent of the population in hopes that 10 to 15 percent of that sample population will respond.

The number and type of questions asked on the postcard questionnaires is limited. A simple postcard survey may want to address the following six issues on vehicle occupancy: use of specific routes, home/work zip codes, time of travel to/from work, number of persons in the vehicle when traveling to/from work, and frequency of use of public transportation or ridesharing.

Data Processing

Once the surveys are returned, the data has to be manually entered into a computer. In the data analysis stage, adjustments may have to be made to insure representative samples.

Sampling Issues

- Approximately 10% of the population is contacted.
- Usually 10 to 15% of the residents contacted respond, resulting in a total response rate of approximately 1% of the total population.

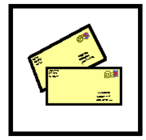
Operational Features

- Survey design and sampling plan.
- Data entry.
- Promotion of survey study through local newspaper to increase response rate.

Strengths

- Logistically easier than training observers, although care must be exercised in collecting a “representative sample”.
- Additional information obtained.
- No safety concerns.

Mail-Out Survey Method



Limitations

- Low response rate.
- Obtains data relative to work trips only.
- Only obtain typical occupancy rates, which may not provide insight on what really happens.
- Large surveys can be expensive.

Recommendations

- Sample subgroups, such as zip code areas, within the area of interest.
- To be cost effective, estimate sample size (i.e., number of questionnaires to be mailed) for each subgroup.

References

- Beaton, W.P., Carragher, F.J., and Meghdir, H. (1993). “Impact of Nonresponse Bias On Forecasts of Average Passenger Occupancy.” *Transportation Research Record 1390*, pp. 43-49.
- Grush, W., and Gross, B. (September 1995). *Vehicle Occupancy: A Comparison of Household Survey Data and Roadside Observation Data*. FHWA Office of Highway Information Management.
- Smith, M. (1977). “Design of Small Sample Home Interview Travel Survey.” *Transportation Research Record 701*, pp. 29-35.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AADT	Average Annual Daily Traffic
AVO	Average Vehicle Occupancy
BMC	Baltimore Metropolitan Council
CAAA	Clean Air Amendment Act
CATS	Chicago Area Transportation Study
CBD	Central Business District
CDTC	Capital District Transportation Council (in Albany, New York)
CMS	Congestion Management System
CTTP	Census Transportation Planning Package
DMV	Department of Motor Vehicles
DOT	Department of Transportation
FARS	Fatal Accident Reporting System
FHWA	Federal Highway Administration
GIS	Geographic Information System
GPS	Geographic Positioning System
HOV	High Occupancy Vehicle
ISTEA	Intermodal Surface Transportation Efficiency Act
KIPDA	Kentuckiana Regional Planning and Development Agency
MPO	Metropolitan Planning Organization
NASS	National Accident Sampling System
NHTSA	National Highway Traffic Safety Administration
NPTS	Nationwide Personal Transportation Survey
O-D	Origin-Destination
PSU	Primary Sampling Unit
PTMS	Public Transportation Management Systems
SRTC	Spokane Regional Transportation Committee
TMSFH	Traffic Monitoring System for Highways
VMT	Vehicle Miles Travelled

1.0 INTRODUCTION

Metropolitan planning organizations (MPOs) and transportation planning agencies are facing greater challenges to increasing the efficiency and level of service of highway transportation systems. With limited highway capacity and growing demands on the system, transportation planners are searching for innovative methods to relieve congestion on their roadway systems. In many areas, road structures have reached their limit and land is not available for expanding the roadways; therefore, the goal of trying to fit more vehicles on roadways is impractical.

The focus is turning to more efficient methods of moving people. MPOs have to review transportation management decisions from a multimodal perspective, where all modes are considered simultaneously and where "person" movements rather than vehicle movements are evaluated. For example, increasing automobile occupancy during peak hours through carpooling, vanpooling, and transit use has been identified as a potential low-cost solution to many transportation problems, including traffic congestion, energy consumption, and air pollution. This growing shift toward a multimodal orientation is an acknowledgment that the purpose of transportation systems is to provide for the overall mobility of people, not simply to maximize the flow of a particular class of vehicles.

To measure the impact that new policies or decisions have upon the transportation structure, more attention has been directed to methods of measuring vehicle occupancy and to the availability and statistical accuracy of these data. The passage of the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA) and recent revisions to the joint Federal Highway Administration (FHWA) and Federal Transportation Administration (FTA) urban transportation planning regulations have strengthened the need for transportation planners to develop multimodal transportation improvement plans. Vehicle occupancy data are critical for determining the mobility of persons and for supporting the ISTEA requirements and the Traffic Monitoring System for Highways (TMSFH), as well as the Clean Air Act Amendments (CAAA) of 1990.

Although no direct mandates in the ISTEA and CAAA regulations require collecting vehicle occupancy data on a certain schedule, it is implicit that current area-wide vehicle occupancy data are essential to achieving the objectives of public transportation management systems (PTMS). Most state, regional, and local transportation agencies rely upon existing data and data collection efforts to meet their data needs. Limited resources and reduced agency budgets and personnel make it nearly impossible to fund new data-monitoring initiatives. This report identifies vehicle occupancy practices that are cost effective, multimodal, and provide area-wide or corridor coverage to facilitate ISTEA- and CAAA- related management and monitoring systems and long-range transportation planning.

1.1 Vehicle Occupancy Data Collection Objective

State and local MPOs collect vehicle occupancy data to address a diverse range of purposes. The following are examples of purposes for collecting vehicle occupancy data:

- To evaluate the effectiveness of a transportation system management action, such as a new congestion management program.
- To verify compliance with state regulation.
- To assess changes in air quality indices based on person consumption and to monitor the energy efficiency of travel.
- To validate urban transportation planning models, which require occupancy information.
- To identify emerging transportations trends that may occur due to increases in fuel prices, restrictions on fuel consumption, weight restrictions on commercial vehicles, changes in speed limits, and motorist response to reduce unnecessary travel.
- To predict trends that would allow the policy formation and planning process to address emerging issues.

In designing and implementing a data collection program, the intended use of the data has to be clearly defined before selecting a collection methodology. Selecting a methodology to collect occupancy data should consider that often the data will be used by different agencies for multiple purposes. It is common for data collected by a state or local government agency to be later aggregated with other MPO data to estimate regional occupancy rates. Because agencies find that examining previously collected data is more economical than collecting new data, any effort to collect occupancy data should always consider its later uses.

1.2 Background on Vehicle Occupancy Data Collection Methods

The traditional procedure for collecting vehicle occupancy data primarily used the roadside windshield method or the household survey method. These traditionally were used to focus on occupancy rates for either an area-wide or facility-specific location. With the growing interest in intermediate geographic coverage areas, new cost-effective methods are being explored by agencies. One new method, which is being used by some states, collects occupancy information from secondary sources, such as accident reports. Other new observational methodologies based on video, optical, and electronic technologies are still currently either under development or in the experimental stage. The following descriptions provide a brief summary about the various collection methods.

Roadside Windshield Surveys

The most widely used method for collecting vehicle occupancy data is the roadside windshield method. This method involves stationing one or more observers along the roadside to perform physical counts of vehicles and occupants. This method is best suited for collecting data at the site-specific or corridor levels, but it can be used to get regional estimates.

Carousel Observational Method

The carousel method, an alternative to the roadside windshield method, positions observers in vehicles traveling with the traffic to collect vehicle occupancy data on neighboring vehicles. On multi-lane highways, the observation vehicle drives slightly slower than the prevailing traffic speed, resulting in the continuous flow of traffic by the observation vehicle. When the survey vehicle reaches the defined limits of the survey area, it turns around and drives the same route in the other direction. Thus, bi-directional observations are recorded for the selected road segments. This method is well suited for corridor or regional studies.

Photographic and Video Surveillance Methods

Like the windshield method, the photographic and video collection methods observe traffic from a stationery position along the roadside; however, instead of human observers, cameras collect occupancy data. The use of video and photographic surveillance systems to capture occupancy information has been limited (Law, 1981) because of the costs. However, interest has been growing in this method. MPOs in Florida are investigating the use of video to collect travel time data because of a state requirement to implement congestion management systems, which may have to rely more heavily on real-time traffic monitoring devices (Pietrzyk, et al, 1996). The biggest drawback of photographic methods is that the collection of vehicle occupancy information from the photographic footage requires a human observer to view the film. Currently, this process has not been effectively automated. For this reason, its usefulness is limited to site-specific or possibly corridor coverage.

Machine Vision Methods

In high occupancy vehicle (HOV) lane enforcement practices, machine vision technologies that utilize electro-optical sensors and pattern recognition image processors are being investigated for the purpose of identifying the number of vehicle occupants (TTI, 1994). This method would use highly sensitive infrared radiation sensors to detect hot spots caused by people in a vehicle. Because radiation sensors can detect other heat sources, like the engine, the pattern recognition image processor is needed to distinguish the occupants from warm vehicle parts.

Currently, an established electronic system that can be used to collect vehicle occupancy information automatically is not available. However, since machine vision technologies have been successfully implemented in military applications, this technology is being explored for potential uses in non-military applications.

In-Vehicle Surveying Devices with Global Positioning System

A new method of collecting personal travel surveys through in-vehicle electronic devices has been proposed by FHWA and is currently being investigated by Battelle (Wagner, Murakami, and Neumeister, 1996). The in-vehicle surveying method involves equipping residents with small computers (e.g., Sony Magic Link), which can be easily set up in their vehicles. Upon starting a trip, the driver would enter the number of occupants and the trip purpose into the device. The global positioning system (GPS) technology in the device will use the global satellite system to track the location of the vehicle throughout the entire trip, and the computer will record the latitude and longitude of the vehicle periodically throughout the trip. For each trip record, the minicomputer will store vehicle occupancy, trip purpose, date, trip start and end time, and extensive information on the actual route driven. From this information, there is the ability to plot onto a city map the entire routes traveled by residents, to evaluate route selection, and to derive travel speeds for different road segments.

The following field operation procedures were developed and implemented by Battelle in a feasibility study of the technology. First, a population to be surveyed was defined. The leading criterion for selecting a population is whether electronic regional maps have been collected by the local MPO. This is very important because a geographic information system (GIS) is used to link the global position of the vehicle onto a road map. Next, a sample of residents who satisfy the demographic and sample size requirements would have to be selected. Once the zip-codes from which residents will be selected have been decided upon, telephone calls are made to the residents in each zip-code area to get their consent to participate in the collection effort. Upon agreement, the residents are mailed one of the minicomputers. Prior to sending the unit, the collection agency checks the operation of the unit and personalizes each unit to contain all of the household members' names. Upon receiving the unit, the resident only has to place the unit in their vehicle, connect the power cord to the cigarette lighter, and place the GPS receiver on their dashboard or car roof. Upon getting into the vehicle for a trip, the driver activates the unit and uses the touch screen interface to enter in occupant data and trip purpose. The device will automatically record date, time, and global position. Each household uses the device over a 5 to 6 day period. After that time, the device is recollected by telephoning the resident. Upon receiving the device back, the data is downloaded and the device is reconfigured and mailed to the next resident.

This method hopes to provide an alternative method to the current self-reporting questionnaire methods (i.e., a written travel diary, interviews, or mail-out surveys), which have problems with trip omissions and rounding of travel times. In addition, the level of participation for this method was extremely high compared to the other household survey methods.

Accident Data Extraction Method

The number of passengers in a vehicle involved in an accident is routinely recorded in many states by officers at the site of the accident. Because of the accessibility and low cost associated with obtaining accident data, several states have recently shown some interest in extracting vehicle occupancy estimates from police crash reports (Ahuja and Hanscom, 1996; Asante, Adams, Shufon, and McClean, 1996; Gaulin, 1991).

Crash reports typically record information on the location and time of the accident, and in certain states crash reports are also required to record the number of occupants in each vehicle involved. From a collection of these accident records, it may be possible to derive an estimate of vehicle occupancy for a particular road segment, corridor, or metropolitan area for specified time periods. However, the accuracy of this method to estimate AVO is still being explored along with alternative methods to calculate AVO based on different weighting schemes to adjust for possible oversampling of certain subpopulations, such as teenage drivers. In addition, further work on the ability of accident data to accurately capture occupancy rates on corridors with HOV lanes has to be investigated. In a conversation with the Connecticut DOT, their experiences show that accident data underrepresents HOV lanes; therefore, they have to periodically collect occupancy data on HOV lanes by the windshield method to supplement the accident data.

Origin-Destination Surveys

Another source of secondary data on occupancy is origin-destination (O-D) studies. The objective of O-D studies is to determine the amount and character of travel along a road segment. Information that can be gathered is trip origin and destination, type of vehicle, and number of occupants. Occupancy data is not the primary data of interest in an O-D study; however, since it is collected, it may be useful for other studies depending on the sampling method.

Several methods can be used to conduct an O-D study. They are roadside pullover interviews, postcard surveys, license plate “trace” surveys, license plate “mail-out” surveys, and tag-on-vehicle surveys. The roadside pullover interview is a commonly used method for collecting O-D data, but it greatly inconvenient to the motorist. Handing out O-D surveys to motorists as they pass through a toll booth is not as intrusive to the motorist as pullover surveys, but the return rate is significantly lower and the sampled population may not be representative of the larger area (URS Consultants, 1996).

Household Surveys

In general, household surveys use three methods for collecting data: questionnaires, telephone interviews, and personal interviews. Primary data sources that can use household surveys are travel time surveys, home interview surveys, origin-destination surveys, and polls. Individual travelers are contacted or sent questionnaires at their place of residence. Generally, origin-destination information is requested along with travel activities for a typical day.

National Surveys

Like accident databases, various national databases exist that capture vehicle occupancy information as part of a larger survey of transportation characteristics. These databases contain a rich source of data on travel characteristics not normally surveyed or capable of being captured by the other collection methods. One drawback is that these databases may be outdated for an agency’s purpose, since the collection surveys are not administered on a frequent basis. However, they do provide historical data on vehicle occupancy, which could be useful in modeling.

The *Nationwide Personal Transportation Survey* (NPTS) is a national household survey conducted by the Bureau of Census to characterize personal travel. The NPTS is a random telephone survey designed to collect household information on the daily trips made by household members. The data include information on household size, family income, availability and proximity of public transportation, number of household vehicles, family life cycle, race, age, income category, and the primary mode of transportation used by individuals to commute to work are recorded. Travel data most relevant to vehicle occupancy studies include the trip purpose (e.g., to-work), the primary mode of transportation, trip mileage, number of household members on a trip, and a wide variety of other related statistics. NPTS provides information that other sources cannot provide regarding vehicle occupancy by trip purpose, household characteristics, and numerous other factors. It cannot, however, provide information on vehicle occupancy by type of highway or time of day.

The *Census Transportation Planning Package* (CTPP) is a set of transportation-related data extracted from the 1990 decennial census. The 1990 Census of Population conducted a journey-to-work survey that could be used to obtain vehicle occupancy estimates, trip origin-destination data, and trip purpose information for each census tract in the United States. With the 1990 census surveying about 18 million households, a significant number of persons are represented in this database. Because the census data is decennially collected, it may not serve as a good recent source for estimating AVO. However, it may serve as a good source of data for examining the relationship between vehicle occupancy and demographic characteristics such as the number of hours worked a week, earnings, age, sex, mobility limitations, and vehicles available. Also, the influence of the more traditionally evaluated factors (i.e., mode to work, workers per vehicle, origin-destination, travel-time, time of day to/from work) can be studied to determine their effect on vehicle occupancy. When comparing census-derived AVO estimates with AVO estimates derived from the other collection methods, caution must be used since the census vehicle occupancy data only focuses on work-related trips, while the other methods may include vehicles on a variety of trip purposes.

1.3 Project History

Over the last several years, many MPOs have conducted vehicle occupancy surveys. However, no state has done a direct comparison of the various methods of collecting vehicle occupancy data. A few studies have, at most, compared two of the methods. For example, some states, like Connecticut, Maine, and New York, have conducted comparisons between the traditional roadside windshield method and the accident data extraction method. Their findings have shown a fair correlation between the two methods, although further work needs to be conducted to account for potential sampling problems (Ahuja and Hanscom, 1996; Asante, Adams, Shufon, and McClean, 1996; Gaulin, 1991). In addition, the Hillsborough County MPO conducted a feasibility study in 1996 on video-based traffic data collection that contained some comparisons with the traditional roadside windshield method (Pietrzyk et al., 1996). Grush and Gross (1995) compared the Nationwide Personal Transportation Survey (NPTS) to a roadside windshield study and concluded that the overall vehicle occupancy estimates were similar, but differed at more stratified levels. Based on the findings of previous comparison studies, it appears that the alternative methods capture the temporal trends in vehicle occupancy rates; however, the numeric estimates of average vehicle occupancy do not always agree across the methods.

This study differs from previous research because it evaluates several prominent collection methods. This project was initiated and funded by the Federal Highway Administration's (FHWA) Offices of Highway Information Management and Environment and Planning to assist states and metropolitan areas in selecting a vehicle occupancy methodology. The first two tasks of this project, which were to review the state of the practice in vehicle occupancy collection methods and to develop and initiate a test plan for comparing alternative methodologies, have been previously reported (Battelle et al, 1995; Heidtman and Skarpness, 1996). This report discusses the final task of summarizing the findings of the field studies and assessing the costs associated with the alternative vehicle occupancy methodologies.

To compare vehicle occupancy methods, five MPOs were selected to participate in the collection of occupancy data: the Capital District Transportation Council (CDTC) in Albany, New York; the Baltimore Metropolitan Council (BMC) in Baltimore, Maryland; the Chicago Area Transportation Study (CATS) in Chicago, Illinois; the Kentuckiana Regional Planning and Development Agency (KIPDA) in Louisville, Kentucky; and the Spokane Regional Transportation Committee (SRTC) in Spokane, Washington. These MPOs were selected primarily because they were already planning to conduct a vehicle occupancy study using one of the current methods and expressed a willingness to collect more data by an additional method.

Based on a review of vehicle occupancy data collection methods, only five were selected for comparison: the roadside windshield method, the carousel observation method, video surveillance, mail-out survey, and accident data extraction. Each MPO collected vehicle occupancy data by at least two of the selected methods (see Table 1.1), and the findings of these methods were compared. The traditional windshield method, in which observers along the roadside count the number of occupants in passing vehicles, was chosen by all of the MPOs.

Table 1.1 Participating MPOs and Vehicle Occupancy Methods Evaluated

Collection Method	Albany CDTC	Baltimore BMC	Chicago CATS	Louisville KIPDA	Spokane SRTC
Roadside Windshield	X	X	X[1]	X	X
Carousel Observation		X			
Video and Still Photography	X				
Mail-Out Survey				X	
Accident Data	X	X	X		X

[1]For the Chicago area, occupancy information was obtained from the US DOT's NHTSA survey.

In addition to comparing these five data collection methods, the general procedures involved in implementing a vehicle occupancy survey study are summarized with highlights on issues that are often ignored in practice. Particularly, the issues focused upon were sampling procedures, measurements of the accuracy and reliability of AVO estimates, methods for adjusting biased data, and cost analyses. These issues surfaced in the review of the literature and in the field test experiences.

2.0 STATE-OF-THE-PRACTICE FOR VEHICLE OCCUPANCY DATA COLLECTION

A review of the vehicle occupancy data collection methods identified two primary sources of information: published reports and state DOT studies. Published reports on data collection practices and procedures were obtained through a literature search. State DOT reports on vehicle occupancy collection practices were requested by FHWA's Office of Highway Information in 1992. Several papers that were most relevant to this research were published reports written by state DOTs, thus the review of the literature complements the review of these state reports.

The review of the literature on vehicle occupancy collection methods has highlighted the lack of reports discussing the details of the collection methodology and those reports that do exist are not being referenced by practitioners. If practitioners are not applying the basic principles for conducting a statistically sound study, such as random sampling and collection of an adequate sample of data, then their findings will be biased regardless of the collection method selected.

2.1 Review of Literature

Literature searches were performed on several databases, with the Transportation Research Information Service (TRIS) database producing the most useful references. The literature search found many reports (about 400) that collected vehicle occupancy data; however, few of these reports discussed data collection methodology issues. In general, most reports published prior to 1988 described the results of vehicle occupancy studies without any discussion of the collection methodologies. A few reports published after 1988 discussed data collection methodology issues, but these reports were initiated and funded by FHWA.

A summary of relevant publications is presented in Appendix A. Robert A. Ferlis's paper, a *Guide for Estimating Urban Vehicle Classification and Occupancy* (1981), has been highly referenced by many researchers. This manual provides sampling and data collection procedures for conducting roadside surveys to estimate vehicle classification and occupancy rates. Ferlis's research showed that a reliable occupancy estimate using statistical survey techniques of stratification and sampling could be obtained at a lower cost than traditional approaches that used less frequent, but longer data collection periods. Another report that complements Ferlis's work is the *Guide to Urban Traffic Volume Counting* (FHWA, 1981), which presents methods for traffic data counting programs. Field tests of these methods in Atlanta (Fisher, Williams, and Boyd, 1980) indicated that the sampling techniques were very useful in obtaining reliable data and that the methods were cost effective when compared to previous surveys.

A detailed report for persons involved in the design and implementation of a vehicle occupancy survey was written by the California Department of Transportation (Levine and Wachs, 1994). The report presents details on using the roadside windshield survey method for collecting occupancy data at a regional and corridor level. Procedures are discussed for conducting the field operations, recruiting and training field personnel, assembling the necessary equipment, implementing the collection procedures, recording and compiling the collected data, and analyzing the data. The report provides many of the minute details necessary to carry out a successful windshield survey.

In 1992, the FHWA's Office of Highway Information Management requested FHWA Regional and Divisional offices to collect reports describing the procedures for collecting vehicle occupancy data in their jurisdictions. A total of 24 reports were collected. The review of the 24 state reports identified several methods for vehicle occupancy data collection, which are presented in Table 2-1. None of these reports discussed the carousel method (which is described in Chapter 3).

Table 2-1 Vehicle Occupancy Data Collection Methods Cited in Literature

Traditional Approaches	
	Windshield Surveys
	Overpass Surveys
	HOV Facility Enforcement
Advanced Technologies	
	Still Photography
	Video Surveillance
	Infrared Radiation
Surveys	
	Mail-out Questionnaires
	Home Interview Surveys/Trip Diaries
	Telephone Surveys
	Employer/Employee Surveys
	Origin-Destination Surveys
Intercept Surveys	
	Parking Lot/Destination Surveys
	Roadside Pullover Interviews
Accident Data	
	Crash (Accident) Databases

A closer assessment of these identified methods indicated that only three primary survey techniques are being used: direct observation, household surveys, and intercept surveys. Of the 24 reports submitted to FHWA, only 12 discussed survey methodology to any great degree. Table 2-2 summarizes these 12 reports, and Appendix A provides a more detailed (one-page) summary for each of the reports.

Table 2-2 Summary of Reviewed Vehicle Occupancy (VOC) Data Collection Reports

Study Title	Report Date	Sponsoring MPO/Agency	City/State	Study Purpose	Methodology Used	Spatial Coverage	Temporal Coverage	Accuracy/ Precision
Auto Occupancy Monitoring	May/June 1987	Washington State DOT	Puget Sound Area, WA	To test methodology for long-range VOC program planning	-Literature review, test development and evaluation -Direct observation field counts	Subarea: -Activity center -Employment center -Highway facility	-AM and PM peak -Evenings -Weekdays and weekends	97%
Auto Occupancy Monitoring Program	October 1989	Washington State DOT for Puget Sound Council of Governments	Puget Sound Area, WA	To update travel demand model for VOC program planning	Direct observation field counts	Subarea: -Random selection of 54 cities	AM and PM peak randomly selected throughout one year	
Comprehensive Travel Telephone Survey	Spring 1989	New York State Metropolitan Transportation Authority	New York City and surrounding areas	To support transit agency planning	Telephone survey of random, stratified sample on travel patterns	Regional	Over 24-hour day	
Metropolitan Providers Transportation Improvement Study Design-Destination Survey	December 1991	Rhode Island DOT Division of Planning	Providence, RI	To study traffic patterns to alleviate congestion on I-95 corridor	-License plate sample -Mail-back postcard survey of above sample on O-D, trip purpose. length, purpose, duration, and VO	Corridor: -Two high volume facilities bisecting the corridor	AM peak, weekdays	
National Safety Belt Survey Sample Design	April 1994	US DOT National Highway Traffic Safety Administration	Nationwide	To propose method to measure level of seatbelt use nationwide	-Clustering and stratification, select sampling cities -Direct observation of front seat occupants	Nationwide: -50 primary sampling units -500 tracts -4,000 cities	To be probability determined	Focus is on how to get a valid sampling on a large scale.
Origin-Destination Survey	June 1991	Executive Transportation Committee for Chemung County	Elmira, NY	To support general planning and development strategies	-Origin- destination survey -Random roadside pullover/interview	Urban area: -High volume point-of-entry sites	AM peak, six week-days over two week period	
Photo System for Auto Occupancy Counting and Vehicle ID	November 1981	FHWA Office of Research and Development	Washington, D.C.	To test methodology to measure VO by prototype camera system	-Study of technology -Photographing field tests with camera equipment	Corridor: -Two HOV lanes, one ramp	Nine days different weather conditions, days and evenings	Good: In most cases, number of occupants could be determined by pictures

Study Title	Report Date	Sponsoring MPO/Agency	City/State	Study Purpose	Methodology Used	Spatial Coverage	Temporal Coverage	Accuracy/ Precision
A Procedure to Calculate VOC Rates from Traffic Accident Data	April 1991	Connecticut DOT	Connecticut	To investigate capacity of traffic accident data base to predict VO rates	Calculate VOC from accident data base, compare to field surveys from previous years	Statewide: -Accidents on interstates	Can be by hour of the day, week, season, or year	Good: Results similar to field survey, allows more frequent data collection
Proposal to Provide Home Interview Survey for Average VOC	April 1992	Delaware Valley Regional Planning Commission	Selected counties in New Jersey and Pennsylvania	To design statistically valid telephone survey method to measure VO	Telephone survey, random sampling of phone numbers	Regional: -4,800 interviews across 13 counties		Fair: There were budgeting and data constraints
Route 9A Reconstruction Project Origin-Destination Corridor Survey Data Collection Plan	November 1-17, 1968	New York State DOT	New York City, NY	To determine impacts of proposal development	Cordon line origin-destination mail-back questionnaire and associated direct observation counts	Urban subarea: -around Route 9A reconstruction area -Primary and secondary impact zones	AM peak, Monday, Tuesday and Thursday	Potential non-response and sample selection bias
Vehicle Occupancy Determinators	Spring 1988	Arizona DOT	Phoenix, Arizona	To determine factors influencing VO rates	-Direct observation field survey -Mail-back questionnaire of selected vehicles arriving at parking lots	-Direct observations stratified by geographic area and highway type -Survey: Metropolitan Phoenix parking garages and parking lots		Good: Driver behavior was measured statistically; problems getting enough surveys back
VOC Program CFVS (1991)	April, May, September 1990, and May 1991	Association of Central Oklahoma Governments	Oklahoma, Canadian, and Cleveland Counties, OK	To update survey methodology to better meet planning needs including CAAA requirements	Direct observation field survey-Clustering for site selection, four types of roadways	Regional: -Twenty stations in three counties, four roadway types	A.M. and P.M. peak midday, Monday, Tuesday, Thursday, two seasons, two years	-Two consecutive years compared -95% coverage -95% confidence rate

2.2 Assessing Vehicle Occupancy Data Collection Practices

The following conclusions emerged from the literature reviewed and from an overall assessment of the state reports:

1. Study Coverage: Most vehicle occupancy data collection studies were one-time special efforts undertaken as part of a larger project, typically the development of a travel demand model system. Regular monitoring of vehicle occupancies on a predefined periodic basis is much less common. The notable exception to this is the Puget Sound region of Washington state, which has considered methods for collecting vehicle occupancy data on a continuous basis.

2. Regional Studies: Traditionally, multiple data collection techniques are often used in the development of regional travel demand models. For example, vehicle occupancy data by trip purpose may have been collected by both a trip diary and an external cordon roadside pullover survey. Emerging purposes are oriented more to the support of congestion and travel demand management programs. This type of data collection was undertaken in Elmira, New York; in Phoenix, Arizona; and in Providence, New Jersey.

3. Traditional Methods: Collection of vehicle occupancy data has been primarily conducted by using the traditional collection methods of windshield surveys and mail-out trip diaries or postcards. Some state DOTs (Connecticut, Maine, and New York) have recently started to use secondary sources like accident data to estimate vehicle occupancy. Newer methodologies such as video, optical, and electronic technologies are still in the developmental stage.

4. Geographic Coverage: The traditional interest has been to estimate vehicle occupancy at either a facility-specific or area-wide level. The current focus is more oriented toward the intermediate levels of geographic coverage, which include transportation corridors and employment centers. Overall, an adequate sampling of vehicle occupancy data at a variety of geographic levels (site, corridor, regional) is required for many new programs.

5. Temporal Variations: Interest in understanding the variation in vehicle occupancy by different time periods has been growing; however, the literature is lacking and many studies do not consider this factor's effect on occupancy measures. The temporal conditions that may impact occupancy rates are time of day, day of week, and season of year. As urban area traffic congestion increases, there is a corresponding interest in examining the effects of peak traffic on various aspects of traveler behavior, including effects on vehicle occupancy. These variations have been examined in the Washington State Puget Sound region.

6. Mode Orientation: Often vehicle occupancy data has been limited to the private passenger vehicles (car, van, or light truck). In planning major transportation investments and in operationally managing existing transportation systems, a multimodal orientation is required that accounts for buses that operate over the street system in addition to passenger vehicles. No reports were identified that discussed these issues.

7. Calculating AVO: Even though Ferlis (1981) provided procedures and formulas for obtaining reliable and accurate vehicle occupancy estimates over 20 years ago, none of the states appeared to be utilizing the recommended stratified sampling techniques and weighted average procedures. Many states appear to use the simple mean calculation procedure which can greatly bias their estimates and result in unreliable and inaccurate findings.

8. Using Existing Data: Many state, regional, and local transportation agencies have a strong preference to rely upon existing data and data collection efforts to meet any new needs. This reaction is in response to limitations on, if not outright reductions of, both agency budgets and personnel and are aggravated by requests for new data-monitoring initiatives. Consequently, staged guidelines for collecting vehicle occupancy information may be useful where recommended procedures can be easily adapted to fit different purposes and resources.

3.0 FIELD STUDY TO COMPARE METHODS

One of the primary objectives of this study was to compare different methods of measuring vehicle occupancy. Therefore, the focus of the following comparisons is on the differences between the methods rather than a detailed description of vehicle occupancy rates by metropolitan area. The factors that influence AVO are discussed, since the relative performance of the data collection methods depends on these factors.

Since variations in AVO are related to numerous factors, as part of the comparison, it was investigated whether the various methods were able to capture the known trends in AVO. When possible, the field studies were evaluated to see if their data captured the following trends:

(1) AVO for weekday AM peak hours is usually lower than AVOs for weekday midday and PM peak hours.
(2) Weekend AVO is usually higher than weekday.
(3) Off-peak AVO is usually higher than peak hour AVO.
(4) AVO is lower in the winter and higher in the summer (at least for the northern cities).

Since AVO varies by region, the methods could only be compared if an MPO used both to collect occupancy data. For example, the mail-out survey can only be compared with the windshield method, since only these two methods were used in the KIPDA area. Table 3-1 summarizes the comparisons that were conducted.

Table 3-1 Method Comparisons by Region

Windshield vs. Carousel	**Accident vs. Carousel**
Baltimore BMC	Baltimore BMC
Windshield vs. Accident	**Windshield vs. Mail-out**
Albany CDTC	Louisville, KIPDA
Spokane SRTC	**Windshield vs. Video**
Baltimore BMC	Albany CDTC
Chicago CATS	

TROUBLE ! ISSUE

When comparing methods, what should serve as the baseline AVO estimate? Windshield AVO has been traditionally used, but is it truly reflective of the population?

3.1 Windshield vs. Carousel

Baltimore BMC

The Baltimore BMC is one of the few agencies who have been investigating the use of the carousel method to collect occupancy data. Because of the design of their highways, which often have sound barriers along the roadsides, the BMC had previously used the carousel method under conditions that the windshield method was not an option because of safety reasons.

The BMC participation in this study provided the only comparison between the carousel method and the windshield method. In addition, the design of BMC data collection effort allowed for rigorous comparisons between the two methods. The collection design was coordinated such that both methods were used simultaneously to measure the same traffic under the same conditions, specifically the same AM peak traffic during the Spring, Summer, and Fall months of 1995.

Prior to this study, it had been conjectured that the carousel observational method provides a more precise method for measuring the number of persons traveling in vehicles. Although fewer vehicles are actually surveyed (about one-fourth of those observed by the windshield method), the ability to discern the number of persons in each passing vehicle is greater because of the observer’s perspective and relative speed to the passing vehicles. The analysis of the data does support this conjecture.

Since the carousel method collects data on vehicles traveling in both directions on a road segment, the analysis of the collected data had to group vehicles by their direction of travel. For most of the road segments observed, vehicles were grouped into peak or off-peak travel directions. The peak directions had higher volumes of traffic than the off-peak directions. In general, the types of travelers in the peak directions mostly consisted of commuters going to work.

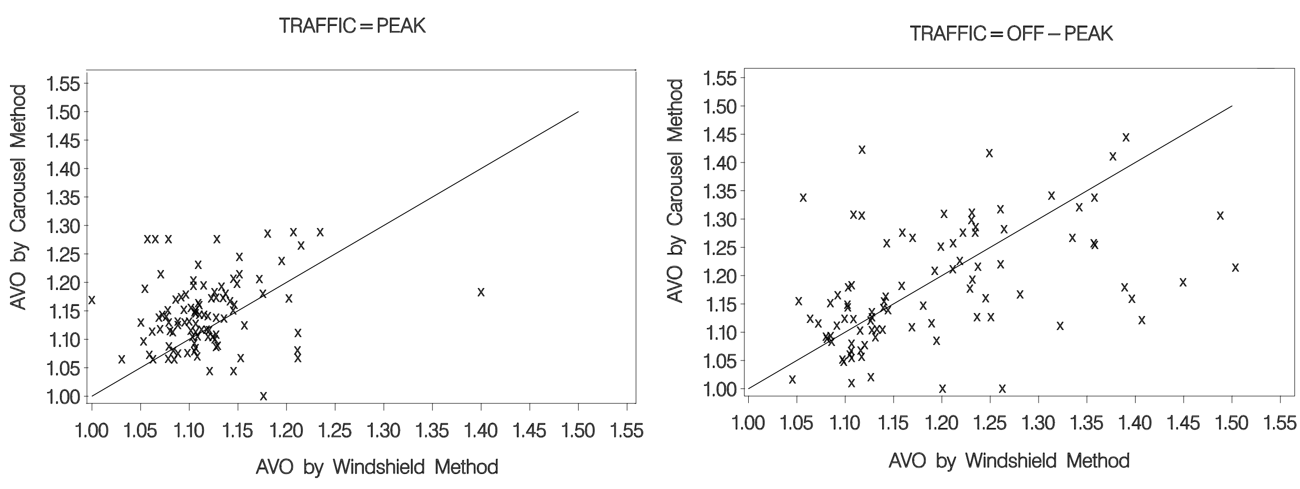
The first comparison of the methods evaluated the overall AVO estimates obtained by each method for the two travel directions. The estimates derived from this analysis are presented in Table 3-2. The findings showed that the carousel AVO estimates were higher than the windshield estimates for the peak direction traffic; however, no differences in occupancy rates were found between the methods in the off-peak traffic data.

Table 3-2 BMC Morning Peak AVOs by Method and by Direction of Traffic

Traffic	METHOD: Windshield AVO	METHOD: Carousel AVO	Carousel to Windshield Ratio
Off-peak	1.167 ± .050	1.163 ± .024	0.991 ± 0.018
Peak	1.112 ± .022	1.141 ± .012	1.026 ± 0.013

This consistent pattern of carousel AVOs being higher than windshield AVOs appeared across most links and is graphically depicted in Figure 3-3. In this figure, each point represents the AVO for an individual link during a half-hour observation period between 6 and 9 AM. The point's x-axis coordinate reflects the link's windshield AVO, while its y-axis coordinate reflects the link's carousel AVO. If the windshield and carousel AVOs were fairly similar, then the points should group evenly around the angled reference line representing the perfect match between carousel and windshield AVOs. In the peak traffic graph, the points are more frequently above the reference line. This indicates that peak travel period AVO estimates based on carousel data are generally higher than AVO estimates based on windshield data. For non-peak directions, the points are more evenly distributed around the line, which indicates no differences between the AVO estimates calculated from carousel data and windshield data.

Figure 3-3 Correlation between Windshield and Carousel AVOs for Peak and Off-Peak Directions

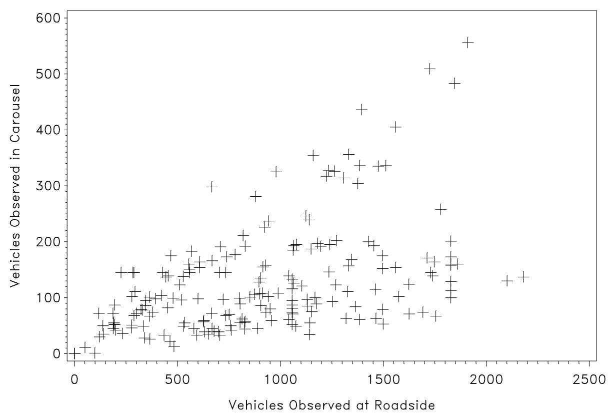


The last column in Table 3-2, which presents the carousel to windshield ratio, statistically justifies this conclusion. For the peak travel period, the ratios range from 1.013 to 1.039 with 95% confidence, which implies that on average the carousel method yields higher AVOs than the windshield method. The off-peak travel ratio was around 1.0, which implies that the two methods yield similar estimates.

To gain a better understanding of the number of vehicles observed by the carousel method, the carousel data was compared to the windshield data, which served as baseline since the windshield method was able to record of all the vehicles traveling through the link. Figure 3-4 displays the observed relationship between the number of vehicles observed by the windshield and by the carousel method. Each plotted point represents the number of vehicles observed during a specific hour on a specific link.

Two relationships appear in the plotted data. The first finding is that the number of vehicles observed by the carousel method increases proportionally with the number of vehicles observed by the windshield method. The second finding was that the carousel method has an upper bound on the number of vehicles that can be observed regardless of increased traffic volumes. In the figure, this upper bound on the carousel method starts appearing around the 200 vehicle count level. Even though counts above 200 were recorded, they were few.

Figure 3-4 Correlation between Windshield and Carousel Counts



It should be pointed out that the number of observer vehicles employed differed across the links. By increasing the number of observer vehicles, the proportion of the traffic stream observed by the carousel method can be increased. However, if the number of observer vehicles is increased, careful planning and coordination are required to avoid double counting of vehicles. Thus, more observation vehicles can capture more of the traffic, but to avoid double-counting the carousel method can never quite capture 100 percent of the traffic flow.

Since a strongly defined relationship between the counts obtained by the two methods does not appear to exist, some corrections to the carousel data may have to be considered. In calculating carousel AVO estimates, an approach analogous to the windshield method was applied that divided the total number of occupants observed by the total number of vehicles observed. To use this procedure, it was assumed that the carousel data was collected proportionally to the traffic flow. However, the carousel method may violate this assumption because the number of vehicles encountered when traveling near the prevailing speed is limited and the vehicles observed may not be directly proportional to general traffic flow on the link.

A weighted mean procedure for estimating carousel AVOs was evaluated since it was recognized that the carousel method only captures a portion of the actual traffic stream. This procedure weighted the carousel counts in proportion to the flow on each of the links, which in this case was inferred from the windshield results. The findings of using the weighted AVO approach, as shown in Table 3-5, revealed no significant differences from the unweighted method. Nonetheless, weighting the regional AVO estimates by traffic flow can be very important, especially if there are large differences between AVOs on busy roads and non-busy roads.

Table 3-5 Carousel AVO Estimates Using a Weighted Approach

	Direction[1]	
Method	**Off-Peak AVO**	**Peak AVO**
Windshield	1.154 ± .040	1.107 ± .022
Carousel (Unweighted)	1.171 ± .037	1.142 ± .014
Carousel (Weighted)	1.160 ± .037	1.140 ± .015

[1] Off-peak data from 15 links and peak data from 17 links.

Conclusions from Baltimore Carousel:Windshield Test

- During peak travel times, AVO estimates derived from the carousel observation method are, on average, higher than AVOs based on the windshield method. This systematic bias was not observed in AVOS for links with off-peak traffic levels.

- From this comparison, it is conjectured that the windshield method may underestimate AVOs, since the observer's perspective may hinder his/her ability to accurately observe the true number of occupants in a vehicle, especially during peak travel periods. This observation challenges the use of windshield AVOs as the baseline method for evaluating the accuracy of other collection methods.

- The proportion of vehicles observed by the carousel method is limited, and it is not consistent. This heightens the possible need to weight AVO estimates by link traffic flows, especially in regional AVO calculations. In the typical applications where parallel roadside sampling is not performed, this may require additional data collection to estimate total traffic flow or the use of prior VMT or average annual daily traffic (AADT) estimates.

The quality of observations made by windshield observers is influenced by a number of factors such as fatigue and traffic volume. The sampling procedure needs to accommodate for these issues by scheduling periodic breaks and possibly sampling only a portion of the traffic stream.

3.2 Windshield vs. Accident

Albany CDTC

With the assistance of the New York State Department of Transportation (NYSDOT) planning staff, the Albany CDTC compiled vehicle occupancy information from the 1993 and 1994 accident records on state and county roads serving the Wolf Road area, a corridor in Albany. After filtering the collected accident data, 16 sites within the Capital District were selected and summarized. Windshield data from 13 of the 16 selected sites was collected in 1995[1]. During the comparison of the two methods, only data from the 13 sites were analyzed.

To extract occupancy information from the accident data, NYSDOT modified a procedure first developed by the Connecticut Department of Transportation (Gaulin, 1991). The CDTC elected to evaluate occupancy data on passenger vehicles only and to use only the first four vehicles listed at the site of an accident. In estimating the regional AVO from the accident data, two procedures for calculating AVO, the simple mean and AADT weighted mean procedures, were used to determine which provided estimates most comparable to the windshield method. Table 3-6 presents the results of the comparisons.

Table 3-6 CDTC Comparison of Windshield AVO to Accident AVO Estimates

	Windshield	Accident Data	
Time Period	**AVO**	**AVO**	**Weighted[2] AVO**
Morning Peak	1.15	1.23	1.26
Midday Peak	1.31	1.36	1.35
Evening Peak	1.27	1.48	1.42

[2] The weighted means were calculated using AADT as relative weights.

The findings revealed that the weighted and unweighted accident methods gave similar results. Neither of these methods produced AVO estimates that were similar to the windshield estimates. However, the accident data and the windshield data did capture the same temporal trends in AVO. A more detailed comparison of AVO estimates at particular sites was not possible since the accident database did not contain a sufficient number of records at this geographic level.

---

[1]The three excluded sites were sites along ramps. A direct comparison of occupancy data from accident and windshield methods was not possible due to delayed availability of accident records.

Conclusions for Albany Accident:Windshield Test

- The accident simple mean and weighted mean AVOs were more similar than when compared to the windshield AVO means.
- NYSDOT has concluded that its procedure for calculating vehicle occupancy rates from traffic crash reports correlates well with data generated from direct observational methods under current travel conditions.
- Comparisons of the windshield data to accident data for each collection site were very problematic and not recommended because of the small number of accidents observed at many of the sites.

Spokane SRTC

Spokane collected windshield data on arterial roads in the summer and fall of 1995. The objective of the collection effort was to obtain AVO data for off-peak and weekend time periods to get a better perspective on personal hours of travel. Comparing the accident data directly with the windshield data was very difficult, since only one roadside location had a sufficient amount of data. The AVO estimates for this one location were very different such that the windshield method reported an AVO of 1.58 and the accident method reported an AVO of 1.36.

Conclusions from Spokane Accident:Windshield Test

- Because of the lack of sufficient windshield data at the various collection sites, it was impossible to adequately compare the two methods.
- Only one location had enough data to compare the two methods. The results of this comparison showed that the windshield estimate was substantially higher than the accident estimate.

Baltimore BMC

The Baltimore BMC compared field data collected by the windshield method to accident data collected by the Department of Motor Vehicles. Only a casual comparison between the accident AVOs and the windshield AVOs was possible, since the AVOs for the two methods were calculated for different temporal conditions. The accident AVOs were based upon accidents occurring during all hours and days of the week; therefore, they represent daily AVOs. The windshield AVOs were based on AM peak-rush hour traffic; therefore, they represent AM peak AVOs. The comparison of accident AVOs to windshield AVOs for each road segment revealed that accident AVOs were higher overall than the windshield AVOs. This agrees with the known trend that morning AVOs are lower than the daily AVOs. This comparison did not allow a determination of whether one method is better than another, since different time periods were sampled.

Conclusions from Baltimore Accident:Windshield Test

- The accident (daily) AVOs were higher overall than the windshield (AM peak) AVOs.
- Temporal trends in AVO may be captured by data collected from different methods.

TROUBLE ! ISSUE

> It is better and less misleading to compare AVOs for the same time period than for different time periods.

TROUBLE ! ISSUE

> Be aware of data differences between methods. For example, accident data collects information on all occupants, whereas the windshield method may exclude small children. Also, the windshield method was just used only on multilane highways, while accident data included lower functional class roads.

Chicago CATS

In 1994, National Highway Traffic Safety Administration (NHTSA) conducted a national survey to collect occupancy counts at controlled intersections in the Chicago MSA during October and November 1994. A controlled intersection was defined as an intersection where the flow of traffic is controlled by a traffic signal or stop sign. The primary sampling unit, Chicago MSA, consisted of Cook, Du Page, Grundy, Kane, Lake, and Will counties in Illinois. The vehicle types, which were monitored in the collection effort, were passenger cars, light trucks, large trucks, vans, and motorcycles. Occupancy information for buses was not collected.

In an effort to make the most direct comparison between the accident data and the windshield data, the accident data were restricted to those accidents occurring during daylight hours between October and November of 1994 and at the same sites as in the NHTSA study. The comparison results are presented in Table 3-7, which presents AVO estimates for all vehicle types and for passenger cars only. This table shows that AVO estimates based on accident data are substantially lower than AVO estimates based on the NHTSA windshield data. These differences were shown to be statistically significant.

Table 3-7 Comparison of AVO Estimates

Vehicle Group	Chicago Accident (Census Weighted)	NHTSA
All Vehicles[1]	1.29 ± .02	1.48
Passenger Cars	1.29 ± .02	1.44

[1]Chicago data included passenger cars, pickup, van/minivan, truck-single unit, motorcycle, and sport utility; NHTSA included passenger cars, light truck, large truck, van, and motorcycles.

Further comparisons between the two data sources were conducted to examine where the differences between the two data sets might exist. The occupancy distributions from both sources were compared. It can be seen in Table 3-8 that the NHTSA windshield data recorded fewer single occupant vehicles and more double occupant vehicles than the accident data.

Table 3-8 Comparisons of Occupancy Distributions for Passenger Cars Only

	Number of Occupants			
Data Source	**1**	**2**	**3**	**AVO**
NHTSA	61.71%	28.80%	9.49%	1.48
Accident Data	80.19%	14.06%	5.75%	1.29

To make the Chicago traffic accident data more comparable to the NHTSA survey data in the Chicago MSA, only accidents occurring between 7:00 AM and 4:30 PM on Monday through Friday were used and the accident locations were restricted to intersections. The results of an analysis of these data were similar to the previous analysis, and thus could not further explain why the overall accident estimates are significantly different from the NHTSA estimates.

Conclusions from Chicago Accident:Windshield Test

- Accident AVOs were substantially lower than the NHTSA windshield AVOs for the Chicago area.

- The NHTSA data reported fewer single occupant vehicles and more double occupant vehicles than the accident data.

3.3 Accident vs. Carousel

Baltimore BMC

The Baltimore BMC compared field data collected by the carousel method to accident data collected by the Department of Motor Vehicles. Only a casual comparison between the accident AVOs and the carousel AVOs was possible, since the AVOs for the two methods were calculated for different temporal conditions. The accident AVOs were based upon accidents occurring during all hours and days of the week; therefore, they represent daily AVOs. The carousel AVOs were based on AM peak-rush hour traffic; therefore, they represent AM peak AVOs. The comparison of accident AVO to carousel AVO for each road segment revealed that accident AVOs were higher overall than the windshield AVOs. This agrees with the known trend that morning AVO is lower than the daily AVO. This comparison does not allow determination of whether one method is better than another, since different time periods were sampled.

Conclusions from Baltimore Accident:Carousel Test

- The accident (daily) AVOs were higher overall than the carousel (AM peak) AVOs.
- Temporal trends in AVO may be captured by data collected from different methods.

TROUBLE ! ISSUE

It is better and less misleading to compare AVOs for the same time period than for different time periods.

TROUBLE ! ISSUE

Be aware of data differences between methods. For example, accident data collects information on all occupants, whereas the windshield method may exclude small children. Also, the windshield method was used only on multilane highways, while accident data included lower functional class roads.

3.4 Windshield vs. Mail-out Survey

Louisville KIPDA

The Louisville KIPDA was the only agency that collected occupancy data by the windshield method and a mail-out survey. The windshield method collected data on occupants 16 years of age or older along several highway entrance and exit ramps in seven areas surrounding the Louisville, Kentucky area during the summer and fall of 1995. Morning and afternoon observations were recorded at several locations within an area. These data served as a source of information for characterizing day-to-day variability at specific sites.

In 1995, 42,000 questionnaires were mailed to residents in seven geographic areas surrounding Louisville. About 5,596 or 13 percent of the surveys were returned and determined to be useable. Each respondent provided the number of persons 16 years of age or older traveling in the respondent's vehicle on trips *to and from work* and indicated the routes that they traveled to-work on a regular basis. The number of occupants was averaged across drivers for specific links and also across all respondents to estimate AVO. On average, when combining summer and fall data, the windshield method yielded AVO estimates approximately 5 percent higher than estimates based on the questionnaire data with buses included and only 3 percent higher when buses were excluded. However, the windshield method was not restricted to vehicles traveling to-work; all vehicles passing observers were counted.

Depending on the intended use of occupancy data collected by mail-out survey, an MPO may want to take into account distances traveled or the time respondents spent on the road. A weighted AVO procedure using travel time was evaluated. This method of calculating mail-out AVO resulted in a value of 1.092 compared to the unweighted mail-out AVO of 1.083. This value was closer to the windshield value of 1.12, but still somewhat lower.

Conclusions from Louisville Windshield:Survey Test

- On average, the windshield AVOs were higher than mail-out AVOs.
- Adjusting mail-out AVO estimates by travel time did not greatly impact AVO estimation.

TROUBLE ! ISSUE

How can occupancy data on "to-work" trips be compared with occupancy data on all trips?

TROUBLE ! ISSUE

In the observational methods, observers should not have to subjectively determine the age of occupants for inclusion in the data set. For example, only including occupants 16 years of age or older.

3.5 Windshield vs. Video

Albany CDTC

The Albany CDTC was the only MPO that used the photographic and video methods to collect occupancy data. The CDTC staff had limited success at obtaining useable footage because windshield glare was a common problem that made it impossible to accurately count vehicle occupants. This problem may have been minimized if a professional photographer had been consulted. After trying this method at three sites, the test was discontinued because it was more labor intensive than the windshield method. As a result of the problems with collecting reliable and consistent data by the photographic and video methods, no comparisons to the windshield data were conducted.

Conclusions from Albany Video:Windshield Test

- No comparison was conducted because of problems with collecting useable footage.

3.6 Insights from Field Tests

This section compared the AVO estimates derived by the various collection methods. This was challenging since the collection conditions did not always perfectly coincide across the methods compared, or not enough information was available to perform the analysis. Only in the Baltimore BMC comparison of carousel method to windshield method were the collection conditions identical, i.e., the same traffic on the same days at the same sites was measured using two different methods. When these ideal conditions are not met, no definite conclusions can be made, since any variability in AVO estimates may have resulted from the differences in collection conditions. Furthermore, it is not possible to state which method is best, since the "true" vehicle occupancy rate is not known, and all methods have been found to have weaknesses. Nevertheless, the following insights were developed from the field tests:

- Windshield observation, typically considered the most accurate method, appears to underestimate vehicle occupancy during peak-travel/high-vehicle-density situations. Under these situations the carousel method consistently gave higher AVOs, which raises a valid concern, since the carousel observers were in a significantly better position to observe multiple occupants in a vehicle. For the windshield method, observer fatigue may contribute to poorer windshield data collection under heavy traffic conditions.

- The carousel method seems to work well on multi-lane, limited-access highways. However, this method alone would be insufficient for a regional AVO estimate.

- In two metropolitan areas (Albany and Baltimore), AVOs from accident data were higher than windshield observations. However, this is not totally unexpected given the specific conditions. The accident data for Albany was one to two years older than the windshield data. Baltimore's accident data included all accidents in a 24-hour period, whereas the windshield observations were only for the AM peak period, which is usually the lowest AVO period in the day.

- Chicago's accident database was very detailed and supported extensive analysis. In Chicago, AVO's from accident data were lower than observed AVOs. However, the temporal trends present in the accident data tracked well with the observational method.

- A mail-out survey in Louisville reported lower AVOs than were observed by the windshield method, but the mail-out survey was limited to "work trips," which are likely to have lower AVOs than non-work trips.

- The video and photographic methods were found to be impractical given "readily-available" equipment and technology.

4.0 EXTRACTING OCCUPANCY DATA FROM ACCIDENT DATA

Of all the vehicle occupancy data collection methods identified in the literature review or tested in the field study, the accident method is the most unique and controversial. With limited funding to conduct vehicle occupancy studies, many agencies are considering extracting vehicle occupancy data from accident databases instead of conducting a field study to collect new data. In addition, since accident data is updated frequently, it is being viewed as a potentially ideal source for annual vehicle occupancy data. This chapter explores the potential utility of various accident databases as sources of ongoing information regarding vehicle occupancy. In addition, it investigates and refines the method for extracting occupancy data from the accident databases.

Both national and regional accident databases were evaluated. In general, the national databases do not provide sufficient coverage on the most important road systems, so accurate estimates of vehicle occupancy cannot be derived from them. The most useful databases appear to be those produced by the individual states who receive information from local law enforcement agencies. Accident data from the Chicago and Spokane areas were used to demonstrate a method for analyzing accident data, which adjusts for biasing influences present in the data. This method for adjusting accident data differs from other methods that perform corrections on the AVO estimates obtained from accident data by using windshield data to adjust the accident AVO estimates. Since the previous chapter mentioned that the windshield method may underestimate occupancy rates, using windshield estimates as baseline estimates for adjusting accident AVO rates may not be the best procedure for correcting the inherent biasing influences in accident data. The chapter concludes with a summary of biasing factors in accident data that an MPO should account for if they plan on using the accident method.

4.1 Sources of Accident Data

Fatal Accident Reporting System (FARS)

The Fatal Accident Reporting System (FARS), which is produced by National Highway Transport Safety Administration, is a compilation of all motor vehicle events (collisions, rollovers, etc.) that resulted in one or more fatalities on United States highways. Currently, annual FARS data is available for 1975 through 1994. Databases for more recent years will require about 2 to 3 years of processing by NHTSA before they can be released to the public. These yearly databases do vary in content and format, so caution must be taken when using these data in a trend analysis.

An analysis of FARS data from the most recent years was conducted for this study. From the data, it appears that some 60,000 vehicles have been involved annually in 40,000 fatal accidents. Under the assumption that these 60,000 vehicles fairly represent vehicle occupancy patterns in general, several alternative analyses of the FARS data were undertaken to evaluate the accuracy and reliability of vehicle occupancy estimates derived from the FARS data.

In these analyses, only data from light vehicles (autos, pickups, vans, and recreational vehicles) were included in the calculation of vehicle occupancy rates, since passenger vehicles are the primary type of vehicle observed by the direct observational methods, e.g., windshield. The various types of trucks and buses were excluded, except for large vans operating as buses which could not be easily excluded.

The first method of evaluating occupancy data in the FARS involved investigating occupancy estimates by highway type. All passenger vehicles involved in fatal accidents were divided into four highway categories: rural interstate, rural other, urban freeway, and urban other. Vehicles in these highway groups were further divided by the time of the accident's occurrence with particular interest on the nominal peak periods of 7-9 AM and 4-6 PM.

An example of the type of occupancy data collected from urban freeways in 1993 is presented in Table 4-1. This table shows the distribution of vehicles by time period and by number of occupants in the vehicle. For each time period, the total number of vehicles observed was summarized. The AVO estimates for each time period were then derived by the total number of occupants in these vehicles divided by the total number of vehicles.

Table 4-1 AVO Estimates for National Urban Freeways

	NUMBER OF OCCUPANTS IN LIGHT VEHICLES (LV)											
TIME PERIOD Time & Day	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10+**	**Total Vehicles**	**L.V. AVO**
0:00-6:59 (All)	707	302	114	46	19	9	4	4	2	1	1208	1.72
7:00-8:59 (M-F)	108	23	6	4	1	1	1	0	0	0	144	1.43
7:00-8:59 (S/S)	42	12	6	1	2	0	1	0	0	0	64	1.64
9:00-15:59 (All)	610	300	72	61	19	11	4	0	0	2	1079	1.74
16:00-17:59 (M-F)	148	59	20	9	2	2	0	2	1	0	243	1.68
16:00-17:59 (S/S)	61	20	12	5	5	1	0	1	0	0	105	1.87
18:00-24:00 (All)	728	321	106	68	26	11	7	5	0	0	1272	1.76
All Hours	2404	1037	336	194	74	35	17	12	3	3	4115	1.73

Source: Fatal Accident Reporting System data for light-duty vehicles accidents on urban freeways in 1993.

This table shows that the 1993 FARS contained information from 4,115 passenger vehicles involved in fatal accidents on urban freeways throughout the US. The AVO estimate derived using all records was 1.73. For fatal accidents occurring on urban highways during the morning peak hours from 7 to 9 AM, only 144 vehicles were reported and their AVO was estimated to be 1.43. Only 243 vehicles were involved in evening peak hour (4 to 6 PM) accidents with an AVO of 1.68. Since the morning and evening peak estimates are probably the most important estimates for transportation planning, the extremely small sample of vehicles poses a significant problem. For a national estimate of AVO, samples of 144 or 243 vehicles cannot capture the regional/state/local variability in occupancy rates. In addition, national estimates are of little interest to most planners, and the sample sizes make it problematic to conduct a regional analysis.

Another issue with the FARS database, which Table 4-1 highlights, is the inclusion of all light vehicles into the calculation of occupancy estimates. Using all reported vehicles may introduce a systematic bias toward overcounting occupancies. Safety experts have long theorized that in light vehicle collisions, passengers die more frequently than drivers. If this is true, then automobile collisions will more likely result in a fatality when passengers are onboard than when only a driver is present. If this theory is true, multi-occupant vehicles would appear with greater frequency in the FARS database than on the actual highway.

An attempt to correct for this bias was investigated by only including light vehicles involved in incidents where at least one light vehicle *driver* died. Table 4-2 shows the results of using this method to correct the bias. However, the method cannot correct for any biases relating to the presence of occupants and its effect on the driver's behavior (e.g., car full of teenage males joyriding).

Table 4-2 Adjusted AVO Estimates for National Urban Freeways

	NUMBER OF OCCUPANTS IN LIGHT VEHICLES (LV)											
TIME PERIOD	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10+**	**Total Vehicles**	**L.V. AVO**
0:00-6:59 (All)	445	121	43	16	4	1	1	2	1	0	634	1.48
7:00-8:59 (M-F)	76	11	3	1	0	0	0	0	0	0	91	1.22
7:00-8:59 (S/S)	34	10	1	1	0	0	0	0	0	0	46	1.33
9:00-15:59 (All)	385	136	31	16	8	2	1	0	0	1	580	152
16:00-17:59 (M-F)	88	24	9	3	1	1	0	1	0	0	127	1.53
16:00-17:59 (S/S)	43	10	5	2	2	0	0	0	0	0	62	1.55
18:00-24:00 (All)	377	117	28	15	5	4	0	3	0	0	549	1.51
All Hours	1448	429	120	54	20	8	2	6	1	1	2089	1.49

Source: Fatal Accident Reporting System data for light-duty vehicles accidents on urban freeways in 1993.

The effect of using this correction method to adjust for the overcounting of higher occupancy vehicles can be seen in the reduced AVO estimates presented in Table 4-2. For example, the AVO estimate for all hours decreased from 1.73 to 1.49. To show that this was more than an anomaly, unadjusted and adjusted AVO estimates were derived from the 1977, 1979, 1991, 1992, and 1993 FARS databases. Table 4-3 confirms that failing to exclude accidents where only non-drivers died will produce a significant bias in the AVO estimates obtained from FARS.

Table 4-3 Comparison of Unadjusted and Adjusted AVO Estimates for Urban Freeways

	YEAR				
	1977	1979	1991	1992	1993
Unadjusted Data Sets					
# of Light Vehicles	2181	2377	4290	4051	4115
AVO	1.75	1.66	1.74	1.75	1.73
Adjusted Data Sets					
# of Light Vehicles	1175	1380	2262	2148	2089
AVO	1.53	1.45	1.50	1.52	1.49

Overall, the national analysis of five years of FARS data on all highway systems indicated relatively small numbers of light vehicles are contained in the FARS database in urban areas. In addition to the small sample size problem, differences in format and content among the yearly databases make it inherently difficult to compare yearly occupancy rates for anything more than the most cursory trend analysis. In conclusion, the national vehicle occupancy rates provided by FARS are of limited interest, since occupancy rates vary widely by location.

National Accident Sampling System (NASS)

NHTSA's National Accident Sampling System (NASS) contains accident data from selected locations in the United States. Beginning in 1979, NASS collected data from ten locations, by 1987 it had grown to cover 50 locations, and in 1988 it decreased to 36 locations. Because the focus of NASS changed so much between 1987 and 1988, NHTSA does not recommend comparing data from the earlier and later periods in NASS history.

The NASS sampling technique for selecting a manageable number of accidents to investigate is a complex procedure. First, the U.S. is divided into 1195 primary sampling units (PSUs), which are stratified by geographic region and type (12 strata). Every year, NHTSA selects at least two PSUs from each stratum, with more than two selected from those strata with more accidents. From each of geographic location, an unspecified number of police jurisdictions are selected. For the selected police jurisdictions, all reported accidents are classified into eight groups, based on severity and type of vehicle involved, and accidents are selected at varying sampling rates within each stratum. By using varying sampling rates, NHTSA allows more investigation of accidents involved in serious injuries and with later model automobiles.

Because of the way NHTSA selects accidents for NASS, the NASS has little utility for estimating occupancy rates. The limited geographic coverage practically rules out the utility of the data source for estimating national occupancy rates, while the inconsistency in choice of geographic areas from year to year lessens the utility of the data as a source for trend analysis in any particular urban area, even for those selected as PSUs.

The bias introduced by the sampling frame's selection of police jurisdictions (suburban, urban, municipal, county, or state), precludes using a PSU as a source of occupancy data in a given year. In addition, highways under any one of these jurisdictions may have different occupancy characteristics than the population of all highways within an urbanized area.

The greater sampling of later model automobiles is another potential biasing factor, since the occupancy characteristics of older and newer automobiles may differ. For example, many drivers in large urban areas tend to choose a commuter vehicle by different criteria than a general-purpose automobile. Trip characteristics of older automobiles differ from those of later-model automobiles, which further supports the argument that occupancy characteristics may vary with automobile age.

In addition, the non-driver fatality bias, which was described in the FARS analysis, is also present in the NASS data. Severe non-fatal injuries to non-drivers can similarly potentially bias occupancy estimates by overcounting the number of vehicles with high occupancy rates. While it was possible to correct for this biasing phenomenon in the FARS data by excluding accidents where no driver died, a similar correction for the NASS data that is based upon a differentially sampled multiple strata subpopulation would at best be cumbersome. It is not clear that sufficient information is available in NASS to make such a correction. For all of these reasons, NASS has little or no potential for estimating vehicle occupancy.

State Accident Databases

State accident databases appear to provide a rich source of vehicle occupancy information at a minimal cost to MPOs. Also, since state accident databases include information on many more types of accidents than just fatal accidents, they have the potential to provide more accurate vehicle occupancy estimates than a national accident database. Many contain enough records to be able to evaluate occupancy rates at the state, local, and corridor levels. However, their usefulness in estimating national occupancy rates is questionable. Even regional estimates, which would aggregate data from various sources, may be challenging to obtain because of differences in database formats and contents among jurisdictions.

State accident databases cover a large and diverse number of accidents, which make them a viable source for estimating occupancy rates. However, not all states and jurisdictions require police departments to record occupancy information in their accident reports. For example, in 1992 only 16 state accident forms asked for the number of "total" occupants[1] (NHTSA, 1992). For those states that require occupancy information, these databases can provide detailed information that can be used in a vehicle occupancy study. Typically, accident reports always include time and date of the accident, the name of the road or street where it occurred, and the direction of travel for each vehicle involved. They usually include route number and milepost; however, they generally do not include functional classification of the roadway, which may be of interest to transportation planners. The roadway functional class information can probably be incorporated later into the accident databases by special software maintained by the states.

[1] The 16 states were: Alabama, Alaska, Illinois, Indiana, Kansas, Louisiana, Maine, Michigan, Nebraska, New Jersey, New Mexico, North Carolina, Oklahoma, South Carolina, South Dakota, and Wyoming.

Although accident forms from 16 states request vehicle occupancy data, the general availability of state accident databases is not presently known. Of the five metropolitan areas selected for this study, four had some form of accident-based vehicle occupancy data available. Presumably, states and MPOs could obtain this information for their own locations. Also, the proprietary formats of many databases may require a special compilation run by the agency possessing the data to make it available to state and/or local transportation personnel.

A Connecticut DOT report, *A Procedure to Calculate Vehicle Occupancy Rates from Accident Data*, (Gaulin, 1991) has served as the basic method for analyzing accident data. In Gaulin's report estimates from roadside windshield surveys, which were conducted in 1982 and 1984, were compared to estimates obtained from accident data. The results indicated a relative difference in estimates of less than 2.4% over comparable time periods. Several factors that can influence accident estimates were identified in the report as part of the verification process for using this new accident data procedure. For example, the impact of male versus female drivers, good versus bad weather, and alcohol/drug-related accidents were considered. In addition, sample size requirements for different tolerance levels were discussed.

Since this report, several concerns have emerged that question the accuracy of accident data to estimate AVO rates. For example, young drivers are involved in more accidents than older drivers, and younger drivers also rideshare more frequently. Drivers with passengers may report marginal accidents more frequently than drivers without passengers. The non-driver fatality bias, which was discussed in the FARS section earlier, is also present in the state accident databases. In addition, drivers with passengers may drive more or less aggressively than drivers alone.

To assess these and other concerns, two large accident databases were obtained from the Chicago CATS and Spokane SRTC. The following section discusses the method used to analyze these databases and the analytical results obtained, which includes a description of possible biasing factors in accident databases.

4.2 Analysis Method for Accident Data

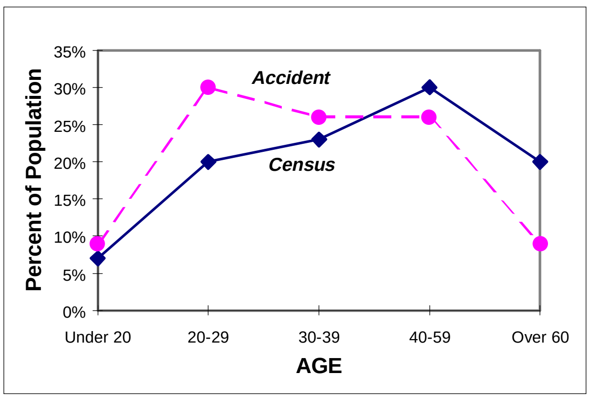
To demonstrate a method for analyzing accident data, the Illinois DOT and the Chicago Area Transit Authority provided this study with extensive accident data for the Chicago area. In addition to the Chicago accident data, some accident data for the Spokane area was obtained from the Spokane SRTC and Washington DOT and compared to Chicago data findings. The goal of analyzing these data was to determine the best way to estimate average vehicle occupancy from accident databases, including identification and correction of any biasing factors in the data.

The vehicle occupancy analyses were conducted on Chicago area accidents that occurred during August to December of 1993 and January to December of 1994 and on Spokane area accidents that occurred during 1992 through 1994. In the analyses, vehicles were stratified by time of day, type of road, condition of driver, and other factors. As a means of correcting observed biases in the data sets, census data and data from other sources were considered. A discussion of other potential biases follows.

Age Bias and the Census Weighting Correction Method

The first analysis of the Chicago accident database was to determine if the age distribution of drivers in the accident database was different from the age distribution for the entire driving population in the Chicago area. To generate a picture of the age of the Chicago driving population, the 1990 census data for the Chicago area was analyzed for persons 16 years of age or older. A comparison of the distribution of ages obtained from the accident database and the 1990 census was conducted and is presented in Figure 4-4. This figure reflects the significant differences in the ages of drivers involved in accidents from the ages of the general driving population. Specifically, younger drivers were more likely to be involved in accidents than older drivers.

Figure 4-4
Distribution of Age for the Chicago Accident Data and the Chicago Census



Even though differences in these age distributions were observed, these differences would only bias AVO estimates if occupancy rates differed by age of the driver. However, Figure 4-5 illustrates that occupancy rates vary by age of the driver for the Chicago area. AVO is clearly highest for the youngest drivers, dropping sharply by the age of 20, with slight increases during childbearing years and retirement. Even for different time periods, which are represented by the three curves plotted in Figure 4-5, this trend across age groups is very similar with only slight differences in AVO rates. An analysis of accident data from the Spokane area also revealed similar age trends, as seen in Figure 4-6.

Figure 4-5 Average Vehicle Occupancy by Age for Chicago Area

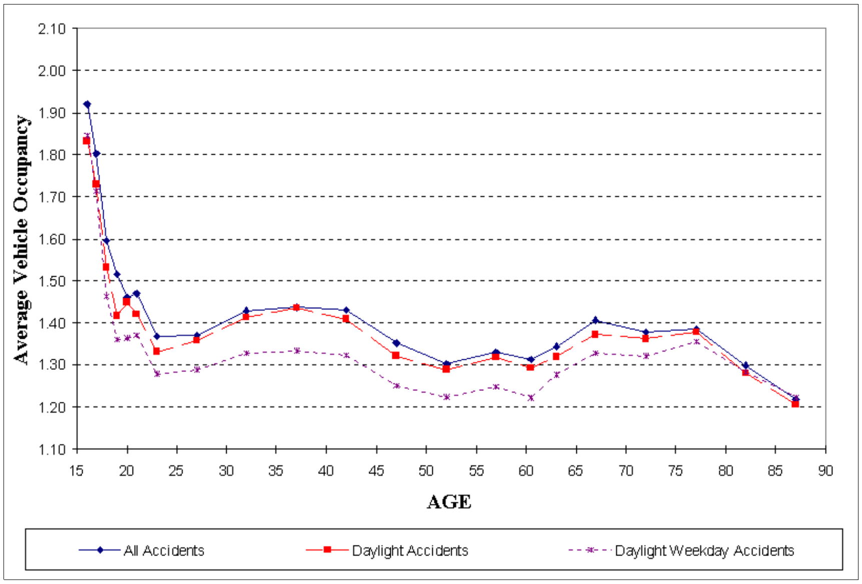
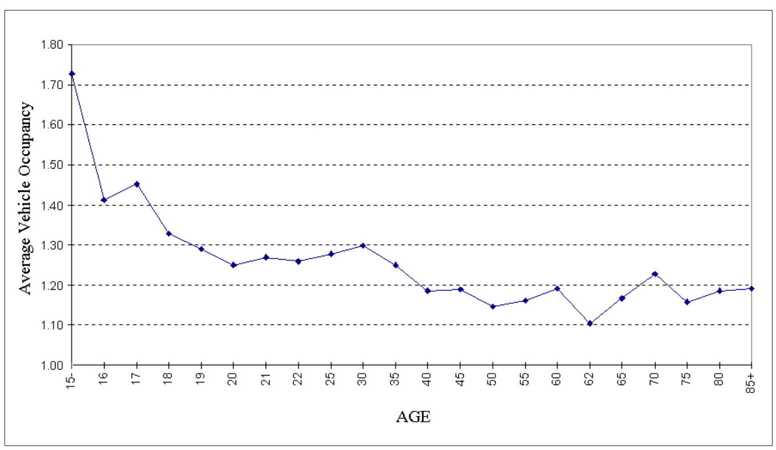


Figure 4-6 Average Vehicle Occupancy by Age for Spokane Area for Weekday Daylight Hours



Because of this driver's age bias in the accident data, any estimates of AVO based upon accident data should be adjusted to minimize the effect of this bias on occupancy calculations. To adjust for this bias, a weighting procedure using the census data was used in this evaluation. To obtain an overall AVO estimate, the weighted mean procedure would weight each age group's AVO in proportion to its age group's contribution to the overall driving population, as defined by the census data. Differences between the weighted mean and unweighted mean procedures are highlighted in Table 4-7 for the Chicago area.

Table 4-7 Adjusted and Unadjusted AVOs by Time Period for Chicago Area

	All	Daylight	Daylight Weekend	Daylight Weekday	Weekday AM-Peak	Midday	Weekday PM-Peak
Census Weighted AVO	1.407	1.384	1.642	1.318	1.182	1.357	1.358
Unadjusted AVO	1.417	1.390	1.664	1.319	1.172	1.359	1.366

Although there were substantial differences in the AVO estimates for the various times of the week, the Chicago accident data set did not show significant differences between the unadjusted and census-weighted estimates. This lack of difference may be an anomaly peculiar to this one data set. Overrepresentation of vehicles driven by younger drivers may have been offset by underrepresentation of vehicles driven by older drivers. Because it is known that occupancy rates do differ by age, the following estimates on the Chicago accident database used the weighting procedure in order to remove the potential age bias.

Vehicle At-Fault Bias

Another factor that has been used to criticize accident databases is that unsafe drivers, such as those that are fatigued or under the influence of drugs or alcohol, are overrepresented in the accident data. Thus, the other vehicles or the struck vehicles are more likely to represent a random sample of the population. In Chicago, the vehicle at-fault is usually reported as the first vehicle. The vehicle at-fault was subjectively judged by the reporting officer at the scene of the accident. The analysis of the vehicles "more likely at-fault" showed that these vehicles did have significantly lower AVOs than the additional vehicles involved in the accident. This phenomenon persisted across apparent driver physical condition (as noted elsewhere on the accident report) whether the driver was sleeping, drunk, medicated, normal, or other.

Examples of AVO estimates segregated by whether the driver was perceived to be at-fault are presented in Table 4-8. For example during daylight hours, the vehicles judged to be more "at-fault" had an AVO of 1.34, while the other vehicles involved in the accident had an AVO of 1.42. During weekday daylight hours, the estimates were 1.29 and 1.35, respectively. This table clearly shows that the more at-fault vehicles do impact AVO estimates for different time periods, except the weekday morning rush hours. However, when interpreting this data, it should be remembered that the reporting officer was using personal judgment in initially determining the vehicle at-fault for the accident and that this information does not reflect the court's opinion of the vehicle at-fault or responsible for the accident.

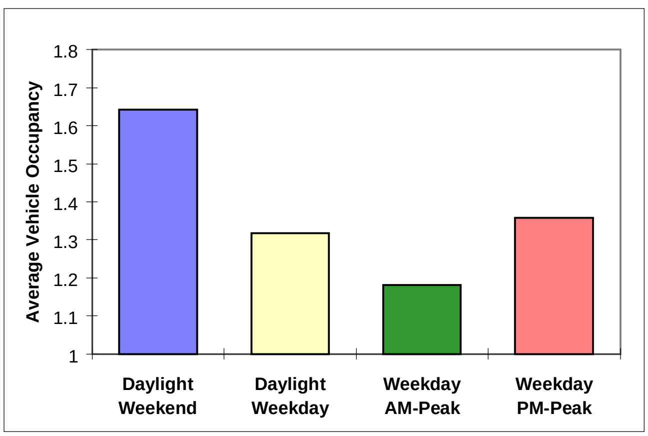
Table 4-8 At-Fault Vehicles Compared to Additional Vehicles

Time/Day of Week	"At-Fault" Vehicle	Additional Vehicles
Weekday	1.29 ± .01	1.35 ± .01
Weekend	1.56 ± .03	1.71 ± .03
Weekday Morning Rush	1.18 ± .02	1.185 ± .025
Weekday Non-Morning Rush	1.32 ± .02	1.39 ± .02
Daylight	1.34 ± .01	1.42 ± .01

Temporal Variability in AVO

Temporal variability in AVO, which is a common issue across all data collection methods, was investigated using the accident data. Three types of temporal factors were shown to influence occupancy estimates. They are the time of day, the day of week, and the season of year. Occupancy rates for the time periods of most interest to transportation planners are highlighted in Figure 4-9. Differences between morning peak period and evening peak period are clearly shown in addition to differences between weekend and weekday rates.

Figure 4-9 Temporal Variability in AVO



Source: Chicago accident data 1994-1995.

The month-to-month variability in vehicle occupancy was analyzed and is illustrated for the Chicago area in Figure 4-10. The AVO estimates plotted in this figure are within 0.03 of the true AVO with 95 percent confidence, so it is safe to say that in the Chicago area vehicle occupancy is higher in the summer months than in the winter months.

Figure 4-10 Monthly Variability in Chicago Area AVOs For Weekday-Daylight Hours in 1994

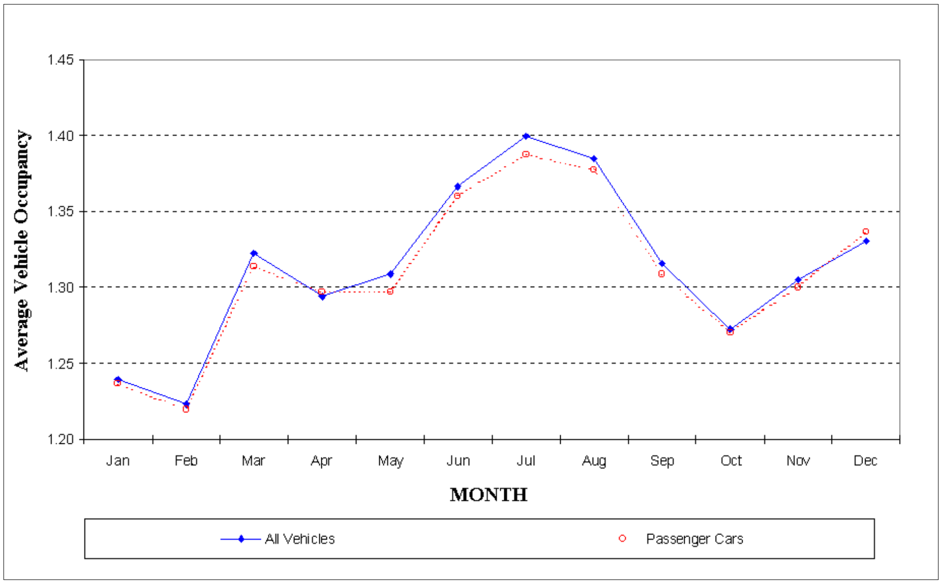


Figure 4-11 for the Spokane area showed very similar trends with the summer months having the highest occupancy rates and the winter months having lower rates. These data support the theory that occupancy rates for the Spokane region are influenced by seasonal variability. The Chicago and Spokane analysis indicate that any study on vehicle occupancy has to consider the monthly variability of occupancy rates, regardless of the data collection method used.

Figure 4-11 Monthly Variability in Spokane Area AVOs For Weekday-Daylight Hours in 1992-1994

4.3 Potential of Accident Databases for Vehicle Occupancy Data

Unlike the other vehicle occupancy collection methods, the accident method does not concentrate its efforts on collecting data because it uses existing databases. Instead, its focus is on the analysis of the data. Since this is a relative new method, the method of analyzing accident data still needs further refinement, especially to adjust for potential biasing influences in the data that can impact AVO estimates. However, accident databases seem to be useful and offer significant potential to provide systematic data for vehicle occupancy performance measurement.

TROUBLE ! ISSUE

> Accident data does not accurately represent the driving population, and it contains inherent biasing influences.

TIP

> AVO estimates for the weekday morning peak period may be less influenced by the inherent biases in accident data.

The most useful accident databases for testing refinements to the accident data analysis procedures are state databases. State databases contain more records at the local and regional levels than any national database. In this study, a demonstration of the procedure to analyze accident data was performed on data from the Chicago and Spokane areas. With these databases, the following biasing factors were investigated:

(1) Driver At-Fault: It has been hypothesized that a number of biases could contribute to producing an accident. Thus, the driver at-fault may have an AVO different from the AVO of the other vehicles involved in an accident. In this study, AVO for drivers 'more likely at-fault' was significantly lower than the additional vehicles involved in the accident. This relationship was tested for weekdays, weekends, weekday morning rush, weekday non-morning rush, and daylight hours. Only the weekday morning rush hour period revealed no significant differences between the AVOs for drivers at-fault versus the other drivers.

(2) Age of Driver: Since young drivers are more likely to be involved in an accident and since they tend to carpool more often, it was hypothesized that an unadjusted AVO estimate will be higher than a weighted AVO estimate that adjusts the accident data to be more reflective of the general population age distribution. An analysis of AVO by age group revealed that AVO is the highest for young drivers and that the accident data overrepresented teenagers; therefore, estimates of AVO were adjusted by using census data. For the Chicago area, a comparison of unweighted accident AVOs to census-weighted accident AVOs for different time periods revealed no statistically significant differences. However, this may be an

anomaly, and the weighted AVO procedure is still recommended because of the known influence of age on AVO. Further testing of this hypothesis using data from other states should be conducted to compare the two estimation procedures.

(3) Fatality Bias: It has been observed by safety experts that in light vehicle collisions, passengers die more frequently than drivers. This was hypothesized to have an effect on AVO estimates. The analysis of the FARS data revealed that non-driver fatalities do impact AVO estimates produced from fatal accident databases.

(4) Temporal Variability: Temporal variability, which can affect occupancy rates obtained by any of the collection methods, was also explored in the accident databases. This analysis showed that AVO rates vary by time of day, day of week, and can vary significantly by month of year.

The first two biases, driver-at-fault and driver's age, did not result in a significantly different AVO results in Chicago, but these biases should continue to be analyzed in future applications. It is possible that the driver's age bias may not greatly impact AVO estimates, since the annual mileage driven by young drivers and older drivers was reported at the national level by Massie and Campbell (1993) to be much lower than the average driver. In addition, the fatality bias affects all accident databases built with only fatal accident records. Finally, temporal variability impacts AVO estimates, and this bias applies to all collection methods, not just the accident method.

All geographic areas should investigate whether their region collects the total number of occupants on their police accident forms. If so, then a potentially valuable, but relatively inexpensive data source may be used for vehicle occupancy data collection. However, it should be recognized that there is a delay between the time of the accident and when occupancy data can be extracted from an accident database. For some areas in our study, data was delayed by one year.

TIP

State and local transportation agencies should investigate whether their accident database collect occupancy information. If so, agencies should begin initiating contact with the organizations that maintain these data because it will take some time to collect the data.

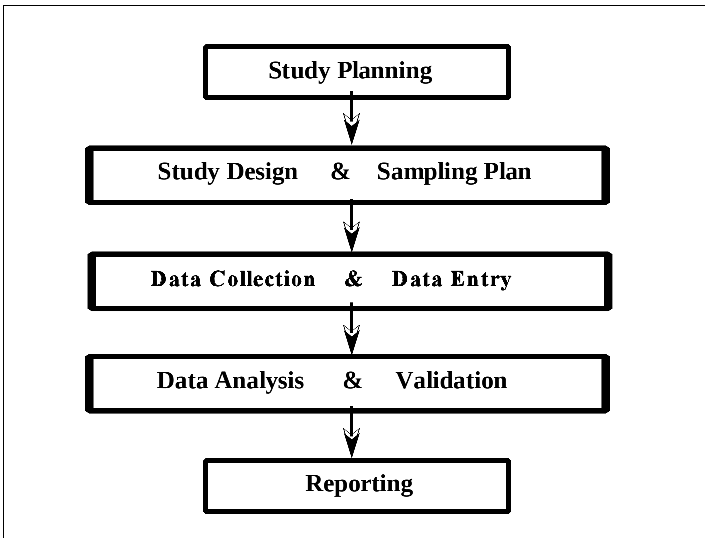
TIP

Sixteen states have been identified that collect number of occupants on their state accident report forms (NHTSA, 1992). They are: Alabama, Alaska, Illinois, Indiana, Kansas, Louisiana, Maine, Michigan, Nebraska, New Jersey, New Mexico, North Carolina, Oklahoma, South Carolina, South Dakota, and Wyoming.

5.0 GATHERING DATA THROUGH STANDARD SURVEY PROCEDURES

Regardless of the collection method selected, the guidelines for conducting a vehicle occupancy study are similar. Several activities, which are defined in Figure 5-1, are involved in the process of collecting vehicle occupancy data. The initial steps involve clearly defining the collection objectives and the type of conditions under which vehicle occupancy will be monitored. Based on this information, the actual collection method can be selected and the collection plan formalized. After receiving the field data, the data has to be processed, filtered, and analyzed. From the statistical analysis of the data, vehicle occupancy trends can be interpreted from the data and summarized.

Figure 5-1 Stages of a Vehicle Occupancy Study



5.1 Study Planning

The primary purpose of the survey planning stage is to determine the intended usages of the collected occupancy information and to clearly define the collection objectives. The collection objectives should establish the scope of the collection effort and the conditions for monitoring traffic. Since estimates of AVO are influenced by the conditions under which occupancy data is collected, the scope will largely determine the types of conclusions that can be inferred from the data.

ACTIVITIES

- Define the collection objective(s).
- Select a geographic area to be targeted:
  - – site-specific (e.g., employer, public facility)
  - – corridor (e.g., interstate)
  - – local, regional, or national (e.g., MPO, state)
- Select the time period:
  - – season of year (e.g., summer vs. winter)
  - – day of week (e.g., weekdays vs. weekends)
  - – time of day (e.g., rush hours vs. off-peak hours)
- Select collection method based on collection objectives, targeted geographic area, and time period to be covered.
- Identify equipment needs and start acquiring equipment.
- Identify the personnel requirements and start hiring or reassigning personnel.
- Determine budgetary restrictions.

TROUBLE ! ISSUE

Using temporary help was found to be less expensive on an hourly basis, but they required more training and supervision than regular personnel.

TIP

Equipment procurement (if needed) can be a lengthy endeavor, so an agency should start obtaining equipment early in the study process.

Defining the Collection Objectives

Seven fundamental questions have to be addressed to determine the vehicle occupancy data requirements for any transportation planning program. These seven points outline the basic design of a field study to estimate vehicle occupancy. With new planning and decision-making processes, which have been initiated in part by TMSFH, PTMS, and CAAA rideshare monitoring programs, vehicle occupancy data needs will grow and these seven issues will have to be adequately handled.

- **What is the purpose(s) for collecting vehicle occupancy data?** Is the purpose to update a regional travel demand model or to establish a baseline estimate for an employee trip reduction program? Is information on trip purpose and travel time required in addition to vehicle occupancy?

- **What is the geographic and temporal scope of the data collection effort?** For the planning activity, does the data collection need to cover a metropolitan area, urban area, corridor, activity centers, or cordon lines? Is information required during the morning, midday, or afternoon peak hours, on certain day(s) of the week, for a specific season, or during special events?

- **How often do data need to be collected?** Will it be necessary to collect vehicle occupancy data on a regular/periodic basis or is a one-time data collection effort sufficient to update estimates? If the effectiveness of a program is being measured, then the occupancy rates before and after the program implementation will have to be collected and compared.

- **What type of vehicle should be included in the survey?** Vehicle occupancy surveys generally only count passengers in private passenger automobiles, light trucks, vans, and motorcycles. When bus and rail are considered, counts are usually obtained from transit authorities.

- **What type of passenger should be included in the survey?** What constitutes a passenger varies across vehicle occupancy surveys. Some studies only count vehicle occupants if they are old enough to drive themselves (i.e., 16 years or older). Others acknowledge some children as occupants but may define children as being a certain age (e.g., 5 years or older in the NPTS). Other efforts may include all children (e.g., infants in child seats).

- **Have sound statistical procedures been used to design the data collection effort?** Have statistical sampling techniques been used to determine collection sites and times? Has the appropriate sample size been predetermined to insure that enough data are collected? If not, the accuracy of the estimates will be questionable and the findings inconclusive.

- **What are the costs for each survey method?** The costs associated with the various vehicle occupancy collection methods are strongly influenced by the design of the study and by the desired precision of accuracy wanted in the estimates. Each method is best suited to a specific temporal and geographic scope. Within that scope they can be very cost effective, outside of that scope they may not effectively obtain AVO estimates.

The purposes for collecting vehicle occupancy data in past studies has been primarily to support the development of regional travel demand models. Other potential areas where vehicle occupancy data can support transportation planning agencies are:

- ISTEA congestion management systems
- Development of travel demand models
- Preferential treatment of high occupancy vehicles
- Major highway and transit investment studies
- Employee trip reduction and commute option programs
- Area-wide ridesharing/travel demand management programs
- Development impact assessment
- General transit and/or growth planning
- Traffic monitoring system for highways
- Air quality modeling
- Mass transit feasibility study, policy options and evaluation

Once the purpose and needs of a transportation planning program are determined, the agency should translate these needs into specific objectives that can be measured. For example, the most common occupancy measure is the estimate of average vehicle occupancy (AVO), which indicates the average number of occupants in a vehicle traveling in a specific geographic area during a specific time period. Depending on the study objectives, AVO by vehicle type or trip purpose may be of interest.

Explicitly stating the geographic scope and temporal conditions for data collection will influence the selection of a particular collection method. For example, suppose the success of a new rideshare program in an air quality non-attainment area in southwest Chicago is to be measured by average vehicle occupancy. Using the household questionnaire method to survey residents in the southwest Chicago region both before and after the program may not be an appropriate method, since most of the population that travels through that corridor resides outside of this coverage area. In this case, a direct observational method, like the roadside windshield method, may be a more appropriate method for monitoring traffic through this geographic area.

Traditionally estimates of vehicle occupancy were geographically focused at the area-wide or site-specific level. Because of changes in needs of MPO for vehicle occupancy data, the current orientation has moved to intermediate levels of geographic coverage, including transportation corridors and employment centers. Table 5-1 lists the different conditions for evaluating vehicle occupancy.

Table 5-1 Desired Spatial and Temporal Coverage for Vehicle Occupancy Information

Geographic	Temporal
◦ Statewide	◦ AM Peak Hour
◦ Air Quality Non-attainment Area	◦ AM Peak Period
◦ Urban Area/Regional	◦ PM Peak Hour
◦ Subarea	◦ PM Peak Period
◦ Corridor	◦ Hour of Day
◦ Activity Center	◦ Daily
◦ Employment Site	◦ Variation by Day of the Week
◦ Highway Facility	◦ Seasonal
◦ Highway Functional Class	◦ Special Events
◦ External Cordon Line	

Because vehicle occupancy varies by the temporal and geographic factors, it is important to select the right conditions for collecting occupancy data. Vehicle occupancy estimates can be potentially influenced by the time of day, day of week, or season of year. In addition, within the geographic area of interest, the selected collection sites can greatly influence occupancy estimates. To account for the impact that temporal and geographic factors have upon the accuracy of AVO estimates, proper sampling procedures should be followed with respect to a given method.

5.2 Study Design & Sampling Plan

Once the general details about the conditions for collecting occupancy data have been established, the actual survey procedures and the sampling plan have to be formalized. Within the selected geographic area, candidate locations for observing traffic have to be determined. For observational collection methods, the exact location of the observers is established by visiting the sites of interest and evaluating good observation points that are a safe distance from the traffic but provide a clear view of the occupants.

The development of the sampling plan will define the number of sites to be observed, the number of collection days at each site, and the times of day traffic is monitored. Because most agencies are constrained by a budget, the proposed sampling plan may have to be modified to stay within an allotted budget. This can result in a tradeoff between the precision of the AVO estimates derived from the collected data and the cost associated with gathering the required sample data.

ACTIVITIES

- Formulate a sampling plan which will include number of links to be surveyed to obtain an accurate and reliable estimate of vehicle occupancy.
- Define the sampled population:
  - Target vehicle types (e.g., all vehicles vs. passenger vehicles)
  - Define occupant (e.g., all persons vs. adult drivers excluding children)
- Randomly select the specific times and locations/sites within the targeted geographic area where the observers will be placed to monitor traffic or where residents will be surveyed.
- Coordinate and schedule collection activities
  - Personnel assignments
  - Equipment assignments
  - Site checks/preparation
  - Dry-run and refinement of methods
- Train personnel on collection procedures
  - Discuss safety issues
  - Familiarize them with equipment and collection procedures
  - Define data elements to be collected (e.g., counts by vehicle type) and conditions for collection (e.g., lanes to be monitored and time periods)
- For mail-out surveys, design survey collection forms, cover letter, and advertisements.

TIP

If a roadway does not have good locations for positioning observers, do NOT place observers along rampways because this can bias results. Instead, consider changing data collection method.

TROUBLE ! ISSUE

What is an occupant? MPOs vary in their opinion of what constitutes an occupant. Some studies only consider occupants who are old enough to drive (16+ years). Other studies count children over 5 years as occupants because they generate trips, while others consider all persons of every age to be an occupant.

TIP

In accident databases that record the number of occupants, all occupants are reported regardless of age. This consistency in the definition of occupant makes it easier to conduct trend analyses.

Summary of Sampling Plan Procedure

During the Study Planning phase, the transportation agency identifies their specific needs for collecting vehicle occupancy data and translates these needs into specific survey objectives and a survey approach. From this, a sampling plan is formalized that identifies the target population, defines the method of sampling the population, and defines the data sample size requirements needed to make any statistically sound judgements from the collected data.

This section discusses probability sampling methods for obtaining vehicle occupancy information appropriate for generating useful inferences. This will allow an MPO to obtain occupancy estimates that should be unbiased and whose precision can be quantified to assess their level of uncertainty. However, in developing the survey design, the MPO will have to use their personal judgement to weigh statistical and cost tradeoffs. Generally, sampling designs incorporate clustering to be cost effective and use stratification to increase the precision of the AVO estimates. For example, data are clustered by geographical regions, such as counties or municipalities, and stratified by road type, such as interstate versus minor arterial.

Step 1. Target Population and Subgroups

As part of the planning processes, the target population of interest has to be defined. A number of factors have to be considered when defining the population of interest since these factors will influence the design of the study. Some of these factors are the geographic area of coverage, the temporal period, vehicle types, types of occupants, trip types, and other factors. These factors are used to define the population of interest. For example, the population could be all passenger vehicles on a major interstate traveling to work during the morning rush hour period. Occupancy data collected under these conditions and any conclusions based on these data will only apply to this defined population.

Step 2. Sampling Procedure

Because logistically and financially it would be impossible to collect data from every vehicle in the entire population, a subset of vehicles will be observed or chosen to represent the population. This subset is assumed to reflect the same characteristics as the target population. The best way to insure that a representative sample of the population will be chosen is to use a random selection process.

To reduce the number of vehicles to be sampled, the population may be stratified into homogeneous subgroups upon which sampling will occur. This stratification process divides the population into smaller subgroups under the assumption that the members within a subgroup have similar attributes. For example, a mail-out questionnaire method may stratify residents in the region of interest by grouping them into subsets based on their hometown or zip code.

Random selection is important because it insures that every member in the population has a chance at being selected. Random sampling should also be used to determine the collection sites and days. The simplest approach for randomly sampling is to give each member an *equal* chance of being selected. In selecting roadways/links to monitor, it has been argued that the lower volume roads should have a lower probability of being selected than the higher volume roads because roads with higher traffic volumes provide more data. In the current practice, the geographic locations or collection sites where occupancy data are collected are generally not selected at random. Higher volume roads are often deterministically selected to the exclusion of lower volume roads, which can result in biased AVO estimates and limited extrapolation of the findings outside the geographic locations sampled.

A compromise to this problem is to use a weighted random sampling procedure that takes into account traffic volumes when selecting road segments. Unlike the simple case where every member has an equal chance at being selected, in the weighted random selection method the higher volume roads have a greater chance at being selected over the lower volume roads. Data collection methods based on *weighted* random selection processes are less likely to be biased, and they permit assessment of precision in the AVO estimates.

Step 3. Estimating Minimum Sample Size

Before starting the collection effort, it is possible to estimate the degree of sampling required to obtain a certain level of precision in the AVO estimate. By knowing the minimum sample sized needed, the financial costs of collecting data can be predicted and kept to a minimum.

Since the estimate of sample size is related to the precision required in the measures of occupancy (e.g., AVO), the first step is to define the degree of error that is considered acceptable in the estimates. In other words, the desired level of precision in the AVO measure has to be specified. For example, a common approach is to set the tolerance level for an estimate to be within +/-.02 (or +/- .05) of the true AVO with 95 percent confidence, i.e., there is a 95 percent likelihood that the estimated AVO will fall within a range of +/-.02 to +/- .05 from the true population average. The tolerance level represents the acceptable difference between the estimated measure and the true value. The confidence interval is an estimated range of values with a specified (95%) probability of covering the true population mean. The tolerance level helps to limit confidence interval range (e.g., 1.22 to 1.31) at the cost of having to sample more vehicles. A more precise estimate will require more sampling. Therefore, the sample size and the survey costs will be directly affected by the desired level of precision in the estimated AVO.

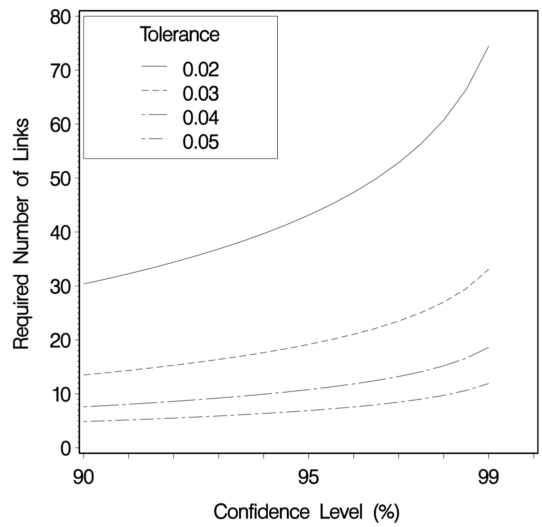
Since each of the data collection methods is fundamentally different, the formulas for calculating the minimum sample size required will vary by method. Appendix B provides more detail on the formulas for each of the methods. Table 5-2 lists the sampling issues associated with the five methodologies evaluated and field tested in this study. This table provides a description of the sampling decision to be faced by an MPO. It should be noticed that the sampling unit varies by method. Typically for the observational collection methods, the sampling unit is the number of link-days observed, which is a randomly selected link sampled on a given randomly selected day for a specified time period. It is assumed that the number of vehicles observed on the link-day is sufficient (i.e., 100-200 vehicles). Other issues of concern are possible sources of variability or error in the estimate. The two types of variability that will be shown to significantly impact sample size requirements will be temporal (day-to-day) variability and geographic (link-to-link) variability. Usually, situations with higher variability will require more sampling.

Table 5-2 Sampling Issues Associated with Each Collection Methods

Method	Sampling Decision	Sampling Unit
Roadside Windshield	How many links to sample and how often.	The link-day. Each link has a certain duration associated with it during which vehicles are observed.
Carousel	How many links to sample and how often.	The link-sample. *Same sampling units as windshield surveys.*
Video Surveillance	How many links to sample and how often.	The link-sample. *Same sampling units as windshield surveys.*
Mail-out surveys	How many questionnaires to mail out and where.	The questionnaire. The observation is the number of persons that travel with the respondent during the day. (It may be restricted to to-work trips.)
Accident data analysis	How many accidents have to be observed in an area or a corridor to obtain a desired precision.	The vehicle. The observation is the number of vehicles involved in an accident along a specific link during a specified time.

An example of the various factors that influence the sample size estimates is presented in Figure 5-3. The link-to-link variability, which is not highlighted by this figure, will be dependent upon the geographic area being sampled by the MPO. For the calculations used to generate this figure, a link-day variability estimate was obtained from a previous vehicle occupancy study done by Ferlis (1979). Ferlis reported that the composite standard deviation estimate of 0.067 is a reasonable estimate of the link-to-link and day-to-day variability.

Figure 5-3 Required Number of Links for Estimating AVO is Dependent on Confidence Level and Tolerance



Note: Assumes composite standard deviation of 0.067.

Figure 5-3 demonstrates the effect that the defined tolerance level and confidence interval will have upon sample size requirements for a particular population, specifically for a study that would use one of the observational collection methods. In this figure, the horizontal-axis represents the typical range of confidence levels that may be selected by an MPO. The four curves each correspond to a different tolerance level that may be selected by an MPO. The value of most interest to MPOs will be the sample size estimate that is displayed on the vertical axis, which in this example is the number of links to be sampled.

Using this type of graph to determine sample size requirements is fairly easy. For example, a regional AVO estimate with a 95 percent confidence at the 0.03 tolerance level can be obtained by randomly sampling a minimum of 10 sites. This is assuming that the study is only interested in a regional estimate and does not want to obtain estimates for subpopulations within the region, such as weekday vs. weekend or interstate vs. other highway types. If the sample size estimate is based only on the entire population of interest and not on each of the various subgroups comprising the population, then only an AVO estimate for the entire population can be derived. AVO estimates for the subgroups can be made, but they will not have the same precision as the population estimate because of insufficient data on the individual subgroups.

If AVO estimates are desired for each subgroup, then in this example for each group 10 links would have to be sampled. As the number of subgroups increases, the total number of links sampled in the population will increase. For example, if an agency wants to derive AVO estimates by type of day (weekends and weekdays), time of day (rush and non-rush), and type of road (interstates and other), sample size estimates for the eight (2x2x2) different subgroups would have to be calculated. In this example, to obtain estimates with the specified precision for each group and for the region, it would require sampling 80 links.

5.3 Data Collection Procedures

For most of the collection methods, the field data collection stage is the most important and most demanding phase. In this stage, occupancy data is collected by field observation or through mail-out surveys, which are distributed to residents. After collecting the data, this information has to be uploaded or recorded into an electronic database for later retrieval and analysis.

ACTIVITIES

- Equipment/site set-up procedures
- Survey in Session
  - Transportation to site
  - Data entry using paper and pencil, electronic counters, or laptop computers.
- End of Survey Day
  - Breakdown
  - Equipment care
  - Battery recharging/replacement
  - Equipment replacement
  - Collection of survey data and data entry
  - Transportation back to office

5.4 Data Analysis

After collecting the occupancy data, a number of measures can be derived from the data. The most common measure, the average vehicle occupancy (AVO), estimates the average number of persons per vehicle on the road. Most often the AVO calculation is simply derived from two other measures, the total number of occupants and vehicles counted. However, for the accident method, some adjustments and filtering of the data are required to remove biasing influences inherent in these types of data.

ACTIVITIES

- Processing and filtering data
- Calculation of AVOs by geographical area, time of day, etc.
- Validation of results - Since the actual data sample size collected will differ from the sample plan, MPOs may want to derive the actual level of precision in the AVO estimate based on the actual sample size. This is especially important if the sample size is smaller than planned.
- Comparison to previous findings
- Refinement of estimation procedures

Summary of Calculating AVO

Calculating AVO for Observational Methods

The following calculations assume that *the number of vehicles observed on each link should be proportional to the actual traffic flow on the link.* Table 5-4 describes some conditions where this assumption is held and where it is violated.

Table 5-4 Conditions Where Data Collection Assumption is Satisfied and Not Satisfied

Satisfied	Not Satisfied
• If every vehicle was observed. • If every *other* vehicle was observed. • If the observers included in their counts all vehicles passing each hour except during a fifteen minute break.	• If observers focused on only one lane of traffic (e.g., the curb lane). • If the number of lanes varied across links and each lane was not monitored. • If observers non-randomly counted as many vehicles as they could accurately examine. • If there was a deviation in the collection schedule on one or more of the segments. For example, if a crew either got a late start, or finished early.

AVO for a link is simply estimated as the number of persons observed divided by the number of vehicles observed (i.e., P/V). An unweighted estimate of AVO for a collection of links can be derived by dividing the total number of persons counted by the total number of vehicles counted:

$$AVO = \frac{\sum_i P_i}{\sum_i V_i},$$

where P_i equals the number of persons observed on link i, V_i equals the number of vehicles observed on link i.

This standard error estimate for the unweighted estimator is[1]:

$$\text{SO} = \left(\left(\frac{N}{N-1} \right) \frac{\sum (P_i - \text{AVO} \cdot V_i)^2}{(\sum V_i)^2} \right)^{1/2}$$

where N is the number of links sampled.

The tolerance (±) for a 95 percent confidence interval can be computed by:

$$SO \cdot t_{.025}(N-1)$$

where $t_{.025}(N-1)$ represents the 97.5th percentile of the t-distribution with N-1 degrees of freedom.

When the assumption of proportional sampling across links is not satisfied, a weighting procedure may have to be applied in the calculation of AVO. For example, when using the carousel method, the assumption is more likely to be violated than when using the windshield survey. This is because the carousel method is limited by the number of vehicles that can be encountered when traveling near the prevailing traffic speeds.

If estimates of the total traffic flow on each link sampled (M_i) are available, they can be used to improve the area-wide estimate of AVO_R as follows:

$$AVO_R = \frac{\sum \frac{P_i}{V_i} M_i}{\sum M_i}$$

---

[1]The standard error estimate is similar to the formula presented by Ferlis (1981); however, Ferlis did not include the correction factor of *N/(N-1)* when the sample size is small.

Combining across road types. When calculating regional estimates, if data was collected on roads of different types (e.g., roadway functional classes), it is best to use a stratified estimator of AVO. In other words, a separate AVO estimate should be obtained for the different road types, and then these AVOs should be combined in proportion to the vehicle miles of travel or number of vehicles on those road types. For example, if AVO_A, AVO_B, and AVO_C represent the AVOs on interstates, principal arterials, and minor arterials, respectively, (as estimated by the method described above), and VMT_A, VMT_B, and VMT_C, represent the vehicle miles of travel on these roads in the specified region, then an appropriate aggregate estimate is as follows:

$$AVO_{aggregate} = \frac{VMT_A\ AVO_A + VMT_B\ AVO_B + VMT_C\ AVO_C}{VMT_A + VMT_B + VMT_C}$$

The standard error estimate for $AVO_{aggregate}$ is:

$$SO_{aggregate} = \left(\frac{VMT_A^2\ SO_A^2 + VMT_B^2\ SO_B^2 + VMT_C^2\ SO_C^2}{VMT_A + VMT_B + VMT_C} \right)^{1/2}$$

where SO_A, SO_B, and SO_C are the standard error estimates of AVO for the three types of roads.

TIP

DO NOT fall into the trap of calculating regional AVO by taking the simple mean of the link-day AVOs. Link-days need to be weighted to reflect traffic flows.

Calculating AVO for Mail-out Questionnaires

Mail-out questionnaires are most often used to collect regional information. The sampling of the region often involves stratifying the population into subgroups or strata (e.g., zip code area). For each stratum *i*, an AVO estimate will be derived by dividing the number of occupants reported by the number of questionnaires returned. Mathematically, the equation would be:

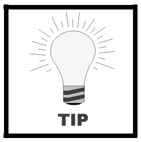
$$AVO_i = \sum_{j=1}^{n_i} \frac{VO_{ij}}{n_i}$$

where VO_{ij} is the number of occupants recorded on questionnaire j in stratum i, n_i is the number of questionnaires mailed to stratum i, L is the number of strata.

To estimate the regional AVO, the weighted mean of the strata AVOs would be calculated, as shown in the following equation:

$$AVO = \frac{1}{N} \sum_{i=1}^{L} N_i \left(AVO_i \right)$$

where N_i is the driver population in stratum *i*, and N is the driver population in the region of interest. A common mistake in deriving the regional AVO estimate is to estimate the simple mean of the stratum AVOs. The problem with using the simple mean procedure is that it gives all strata equal weighting in the estimation of regional AVO, such that an area with a relatively low population can be overrepresented and an area with a higher population can be underrepresented.



DO NOT calculate the cumulative (regional) AVO by the taking the mean of the subgroup AVOs.

Calculating AVO for Accident Data

The most common method of deriving an AVO from accident data is to use a method similar to the observational method's way of deriving AVO. This method simply divides the number of occupants by the number of vehicles for a specified road segment and the time period. This method is not recommended since accident data may be a biased and unrepresentative sample of the population. To accommodate these biases, the data has to be filtered and a weighted means procedure used. More detail on these issues and how to correct for them were presented in Chapter 4.

5.5 Reporting

It is a good practice to summarize the procedures and findings of a vehicle occupancy collection effort in a report. The report should reiterate the collection objective, the design of the study, the collection conditions and criteria, the procedures for estimating AVO, and the findings. The report may serve as a reference to agencies who may later use the collected data or try to follow the reported procedures in a new occupancy study.

ACTIVITIES

- Document field collection effort.
- Report findings.
- Defend any new methods for estimating AVO.
- Provide a cost analysis of the study.

6.0 COST COMPARISONS

In selecting a data collection method, a leading factor for many MPOs will be cost considerations. This section provides a means of calculating and comparing costs across collection methods. However, it does not place a fixed price tag on each method, since a number of factors can influence the total cost of using a particular method. The most influential factors on collection costs are the study design (e.g., site-specific vs. regional study) and regional labor wages. Therefore, this cost analysis presents examples of cost analyses for different survey designs.

Since collection methods are so diverse, developing a means for comparing costs between methods can be difficult. The cost factors that significantly affect the cost of a direct observational study (e.g., observers' wages, transportation to collection sites, computer equipment, number of observers per site, number of collection days) are very different from the factors affecting a mail-out questionnaire survey (e.g., cost of developing a questionnaire, acquiring a target list, printing, postage, and data entry). The commonality among collection methods is that standard survey procedures for conducting a vehicle occupancy study are similar regardless of the collection method. Guidelines for comparing and disaggregating method costs were developed around these survey procedures.

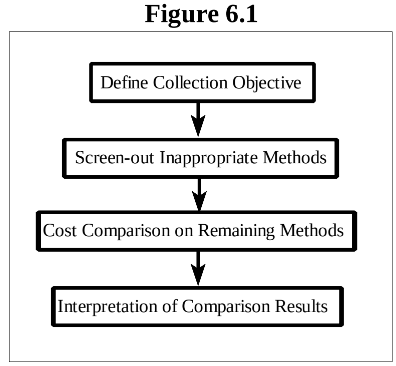
TROUBLE ! ISSUE

It is impossible to put a fixed price tag on each method, since the study design and regional labor wages greatly impact the overall cost.

6.1 Cost Comparison Procedure

The following 4-step cost comparison procedure, which is summarized in Figure 6.1, was developed to compare costs associated with the different collection methods. In addition, it can be used to compare costs between different survey designs using the same method. The first two steps of the procedure involve clearly defining the collection objective, and then determining which collection methods are best suited to accomplish this objective. The last two steps involve calculating method costs by the five survey stages and comparing the overall costs associated with the various methods.

Figure 6.1



Step 1. Define collection objective.

Before calculating costs, the collection objective should be clearly defined, so an agency can make an informed decision on the most suited collection methods. Therefore, the first activities to be completed are to define (1) the study objective, (2) the geographic area, (3) the time period, and (4) the sampling plan. Defining the study objective will help determine the geographic area and time periods to be covered in the collection effort. Geographic area can be a site, corridor, region or nation. The time period defines the time of day, day of week, and the season of year over which collection will occur. Once the geographic area and time periods are determined, a sampling plan has to be constructed, which considers the desired level of precision required in the AVO estimate and the sample size need to obtain this precision. The precision of the estimate will indicate how closely the AVO estimate matches the true AVO value for the population of interest. Too often in practice though, precision is sacrificed to reduce costs.

TROUBLE ! ISSUE

Cost/Precision Trade-off: How many link-days have to be sampled to maintain a reasonable level of precision in the AVO estimate?

Step 2. Screen-out inappropriate methods.

Once the criteria for collecting data have been established, these criteria can be used to screen-out collection methods that are not suited for the intended survey design. Not all collection methods can be used under every condition, so calculating costs for an inappropriate method should be avoided. For example, the carousel method and the accident method are not well suited for collecting occupancy data at a specific site such as a factory.

Step 3. Cost comparison on remaining methods.

After screening out the inappropriate methods, costs are estimated for the remaining methods. The costs associated with each method are broken down by the five stages involved in collecting occupancy data: survey planning, survey design, data collection, data analysis, and reporting. This cost comparison method follows the historical practice of disaggregating costs by survey stage. This will allow an MPO to determine where most costs are incurred.

$$
\begin{gathered}
\textit{Method Cost} = \\
\textit{Planning Costs} + \textit{Survey Design Costs} + \textit{Data Collection Costs} + \textit{Analysis Costs} + \textit{Reporting Costs}
\end{gathered}
$$

For many of the methods, the costs associated with particular stages will be similar. The one stage that is most influenced by the method selected and the study design is the Data Collection & Data Entry Stage. This stage’s costs include field personnel wages over the total collection period, data entry personnel wages, and equipment costs.

Step 4. Interpretation of comparison results.

Completing the analysis of the total costs associated with each method, provides a foundation for comparing between the methods. The overall method costs will be most important to an agency; however, they should review the breakdown of costs for each method. This will allow an agency to determine which phase of the project is most costly and provide guidance on ways to reduce costs. However, any effort to reduce costs by modifying the study design and sample plan should recognize that the level of precision in the AVO estimate may be affected.

6.2 Cost Scenarios

Cost scenarios are provided in this section to demonstrate the use of this cost comparison methodology and to highlight the impact of various study designs. Three hypothetical examples are presented, which range in geographic coverage from site-specific to corridor to regional. The costs associated with each stage and method are subjective estimates of the real costs and are not based on actual cost data because few MPOs provide a cost analysis for the vehicle occupancy studies they have conducted. In addition, because of regional variability in labor wages, each MPO will want to substitute their regional costs into the model. Therefore, the following examples should serve as illustrative examples of the relative differences in method costs.

Hypothetical Corridor — Example #1

Step 1. Define collection objective.

Study Objective:	To determine the morning rush hour AVO for Interstate traffic into the central business district (CBD) of a major metropolitan area in order to support policy planning.
Geographic Area:	10-mile Interstate corridor extending beyond the CBD limits. Only peak traffic directions monitored.
Time Period:	Morning (6am-9am) rush hours during the week (Monday-Friday) for the Fall season only.
Sampling Plan:	For the observational methods, four observation sites (Links A-D) are selected along the interstate corridor to monitor all 3 lanes of peak direction traffic. Traffic will be observed at each of these sites on 5 different days.

Step 2. Screen-out inappropriate methods.

Windshield Method: Possible.

Carousel Method: Possible. However, non-peak traffic direction is not of interest.

Video Method: Possible. However, video reduction may be labor intensive.

Survey Method: Possible, but a mail-out survey may not be appropriate if there are too many alternative routes into CBD or if there are many long-distance commuters. However, a license plate survey may be used to identify travelers along the corridor and address surveys to these residents.

Accident Method: Depends upon the number of records available in the database for this corridor during the specified time period.

Step 3. Cost comparison of remaining methods.

The costs associated with each method can be broken down into the various activities involved in conducting a vehicle occupancy study. The individual costs can be estimated using cost-breakdowns from previous studies and/or derived on estimates of the amount of labor involved. The costs in Table 6-1 were subjectively derived from previous knowledge of the relative differences in costs across methods and not from real data.

Table 6-1 Breakdown of Method Costs by Activity for Corridor Example #1

Method	***Planning & Design***	***Collection & Data Entry***	***Analysis & Reporting***	***Total***
Windshield	$1,500	$5,940	$500	$7,940
Carousel	$1,500	$3,960	$500	$5,960
Video	$3,000	$7,500	$500	$10,500
Questionnaire	$3,500	$5,500	$500	$9,500
Accident	$550 (Note: Potential sample size problem)			$550

Calculation Assumptions:

- Accident method costs assume MPO has prior experience with their accident database, which includes having filtered the data, established programs for analyzing the database, and some initial testing of biasing factors present in the database.
- The windshield and carousel methods will use electronic counters or laptop computers, not pencil and paper, to record observations and to save money by reducing data entry costs.

The following tables give more detailed information on how the collection and data entry costs can be calculated. Since collection and data entry costs contribute most to the overall costs, this breakdown will be useful in the comparison of methods. The following formula was used to estimate collection costs, and Table 6-2 gives examples of the inputs required:

$$
\begin{aligned}
&Collection\ \ Costs(Links) \\
&= [Wages\ x\ (Hrs.\ of\ Collection)\ x\ (No.\ of\ workers\ )\ +\ (Misc.\ Costs)\ x\ (No.\ of\ workers)]_{link1} \\
&+\ [Wages\ x\ (Hrs.\ of\ Collection)\ x\ (No.\ of\ workers\ )\ +\ (Misc.\ Costs)\ x\ (No.\ of\ workers)]_{link2} \\
&+\ ... \\
&+\ [Wages\ x\ (Hrs.\ of\ Collection)\ x\ (No.\ of\ workers\ )\ +\ (Misc.\ Costs)\ x\ (No.\ of\ workers)]_{link_n}
\end{aligned}
$$

Table 6-2 Collection Costs Associated with Each Link and Method for the Corridor Example #1

Observation Site	Method	Wages /Hr.	Hours /Link	# of Workers /Link	Misc. Cost/ Worker	Cost/ link	# of Days	Collection Costs
Link A	Windshield	$22	4	3	$11	*$297*	5	*$1485*
	Carousel	$22	4	6	$11	*$594*	5	*$2970*
	Video	$22	4	1	$11	*$99*	5	*$495*
Link B	Windshield	$22	4	3	$11	*$297*	5	*$1485*
	Carousel	$22	4	0	$11	*$66*	5	*$330*
	Video	$22	4	1	$11	*$99*	5	*$495*
Link C	Windshield	$22	4	3	$11	*$297*	5	*$1485*
	Carousel	$22	4	0	$11	*$66*	5	*$330*
	Video	$22	4	1	$11	*$99*	5	*$495*
Link D	Windshield	$22	4	3	$11	*$297*	5	*$1485*
	Carousel	$22	4	0	$11	*$66*	5	*$330*
	Video	$22	4	1	$11	*$99*	5	*$495*
Total for Link A-D	Windshield	-	-	-	-	-	-	*$5940*
	Carousel	-	-	-	-	-	-	*$3960*
	Video	-	-	-	-	-	-	*$1980*
**Note: Carousel Method uses same personnel on consecutive links.*								

The costs associated with entering the data collected from field observations or from mail-out surveys can be predicted by estimating the number of records to be processed and dividing that number by the average number of records processed in an hour. However, for some methods this will not have to be calculated. For example, it was assumed that the windshield and carousel methods used electronic counters, so no separate data entry costs are associated with these methods. Data entry costs are a concern mostly for the video surveillance and mail-out questionnaire methods.

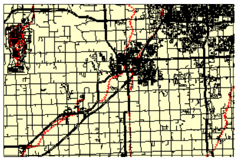
Step 4. Interpretation of comparison results.

In this corridor scenario, a comparison of the overall costs shows that the accident method had the lowest costs; however, the accident costs did not take into account the initial costs that an MPO has to invest when acquiring the data and becoming acquainted with it. Even though the accident method had the lowest cost, an MPO may not be able to use this method if there are not a sufficient number of accident records on this corridor for the specified period. Only an investigative assessment of the accident data can determine if enough records exist for the desired level of precision required in the AVO estimate. This assessment of the database should not be too costly.

If the accident method is not an option, the carousel method may be a good alternative and may be safer for the observers than monitoring traffic from the roadside of an interstate. For this scenario, the carousel is more cost effective than the windshield method; however, this can change depending on the study design. If the study design was to evaluate only one link instead of four links, the windshield method would be more cost effective than the carousel method. For a study of consecutive links (at least 3) on multi-lane highways, the carousel method in general can be more cost effective than the windshield method. Table 6-2 on collection costs highlights the differences in costs per link for each method. For the first link A, the costs for the windshield method are less than the carousel ($297 to $594). However, for the following links B through D, the carousel costs drop considerably since the same survey team is used on all four links ($66 to $297).

The mail-out survey method can be costly, but it is the only method that provides a means of capturing additional data such as travel time. The primary contributor to the cost of this method is in the printing and postage associated with the survey instrument and data entry costs. Since only 10-15 percent of the surveys are generally returned, a large number of surveys have to be sent out in hopes of obtaining an adequate sample of returned surveys. This lack of responsiveness is a major contributor to the costs, but it is generally beyond the control of the agency even when advertising is used. One of the few costs that an MPO can try to minimize is data entry costs. In this example, it was assumed that personnel would have to enter the data into a computer, but new methods such as computer scanners with text editing software can reduce the need for persons to perform this task.

The video method can have low field collection costs but very high data entry costs. Until there is a better way to extract occupancy data from videotape instead of using human observers, this method will be too costly for most agencies.



Hypothetical Region-Wide — Example #2

<u>Step 1. Define collection objective.</u>

Study Objective:	To determine AVO values for an area-wide transportation modeling program.
Geographic Area:	All roads in the metropolitan area.
Time Period:	Weekdays and weekends. Rush hours and non-rush hours.
Sampling Plan:	Data will be needed to estimate AVO separately by time of day (rush and non-rush), day of week (weekdays and weekends), and type of road (interstate and other), which amounts to 8 (2x2x2) subgroups being sampled. For observational methods, the sample size calculation presented in Chapter 5 indicates a minimum of 80 link-days (10 per subgroup) would need to be sampled to produce AVO estimates with 95 percent confidence at the 0.03 tolerance level. (Note: Greater precision would require more data collection. If less precision is acceptable then fewer links would be required. See Chapter 5.)

<u>Step 2. Screen-out inappropriate methods.</u>

Windshield Method:	Possibly inappropriate and to labor intensive because it may require a minimum of 80 days of data collection along various roadways.
Carousel Method:	Inappropriate because only higher functional highway classes are captured.
Video Method:	Inappropriate and too labor intensive.
Survey Method:	Well-suited.
Accident Method:	Well-suited (if available).

Step 3. Cost comparison for remaining methods.

Table 6-3 Breakdown of Method Costs by Activity for Regional Example #2

Method	***Planning & Design***	***Collection & Data Entry***	***Analysis & Reporting***	***Total***
Windshield	$2,500	$23,760	$1500	$27,760
Survey	$3,500	$10,500	$500	$14,500
Accident	$550			$550

Table 6-4 Collection Costs Associated with Each Link and Method for the Regional Example #2

Observation Site	**Method**	**Wages /Hr.**	**Hours /Link**	**# of Workers /Link**	**Misc. Cost/ Worker**	**Cost/ link**	**# of Days**	**Collection Costs**
One Link (i.e., subgroup)	Windshield	$22	4	3	$11	*$297*	10	*$2,970*
Total for All Links	Windshield	$22	4	3	$11	*$297*	80	*$23,760*

Step 4. Interpretation of comparison analysis.

Comparing the overall costs, the accident method has the lowest costs and is relatively the quickest method if the data is already in-hand. However, if additional travel data such as travel time is needed, the survey method may be better suited. An alternative source for obtaining travel data that wasn't examined is to use existing databases such as the census or NPTS. Since these sources are at best a few years old, their usefulness is in providing supplemental information rather than recent changes in occupancy rates.

The windshield method would be very expensive to use in this type of situation. An agency may be able to reduce some costs and the total number of links sampled, if they are willing to make some simplifying assumptions and identifying a model. For example, in this scenario, the agency may have already assumed that the weekdays, Monday through Friday, have similar vehicle occupancy characteristic. Thus, using this assumption, they will not need to collect exorbitant amounts of data for each day of the week, and in this scenario only 70 links would have to be sampled. Another assumption might be that the differences between AVO estimates for interstates and for other highways is not dependent on the day of week (weekend or weekday). These types of assumptions may help to reduce the required number of links to be sampled; however, they must be based on sound reasoning and tested on data where possible.

Hypothetical Site-Specific — Example #3

<u>Step 1. Define collection objective.</u>

Study Objective: To evaluate an employer's rideshare program.
Geographic Area: Entrance to employer complex.
Time Period: Peak times.
Sampling Plan: Morning and evening rush hours for 3 days.

<u>Step 2. Screen-out inappropriate methods.</u>

Windshield Method: Possible.
Carousel Method: Inappropriate.
Video Method: Possible.
Survey Method: Possible with employer survey.
Accident Method: Inappropriate.

TROUBLE ! ISSUE

> For an employer survey of AVO, if employees carpool with employees from another company or with their spouses and/or children, the observational methods may not capture this type of information. However, the survey method may be able to capture this information and other items of possible interest.

<u>Step 3.</u>
<u>Cost comparison for remaining methods.</u>

Table 6-5 Breakdown of Method Costs by Activity for Site-Specific Example #3

Method	***Planning & Design***	***Collection & Data Entry***	***Analysis & Reporting***	***Total***
Windshield	$1,500	$594	$500	$2,594
Video	$3,000	$2,500	$500	$6,000
Survey	$1,500	$2,000	$500	$4,000

Calculation Assumptions:

- Employer survey method would require the employer's cooperation in distributing and collecting the surveys. Some costs may be incurred by the employer.
- Video costs are assuming the need to purchase and to setup equipment. If the employer has CCTV or security cameras, it may be possible to use this equipment if the viewing angle is correct or adjustable.

Table 6-6 Collection Costs Associated with Each Link and Method for the Site-Specific Example #3

Observation Site	Method	Wages /Hr.	Hours /Link	# of Workers /Link	Misc. Cost/ Worker	Cost/ link	# of Days	Total Costs
Entrance (AM)	Windshield	$22	4	1	$11	$99	3	$297
	Video	$22	4	1	$11	$99	3	$297
Entrance (PM)	Windshield	$22	4	1	$11	$99	3	$297
	Video	$22	4	1	$11	$99	3	$297
Total Collection Costs	Windshield							$594
	Video							$594

<u>Step 4. Interpretation of comparison results.</u>

For a site specific case, the windshield method may be best suited since the cost to implement the method is relatively low compared to the other methods. One potential problem with the windshield method may be its inability to capture occupants who are dropped-off/picked-up at another location. This is not a problem for the survey method. Since the survey method costs are largely incurred in the data entry tasks, alternative methods for reducing these costs may be investigated. For example, many employers have internal computer networks, which could be used to electronically distribute the occupancy questionnaire and electronically record the replies. This method could reduce the MPOs data entry costs.

6.3 Insights on Data Collection Costs

When comparing survey costs for the various occupancy data collection methods, care should be taken to ensure that the same reasoning and structure are used. For example, the same geographic and temporal scope must be used across methods for a fair comparison to occur. This is also true when comparing costs associated with using the various methods. A mail-out survey should not be compared to a windshield survey if one focuses on a region and the other on a corridor.

Since there are several factors that influence costs, a procedure for establishing relative costs was presented. This procedure broke down a method’s cost by the 5 survey stages involved in a collection effort. This allows an MPO to pinpoint the stage and task that contribute most to the overall costs. Most often the actual activity of collecting the data and entering the data into a computer is the most costly stage.

To summarize the relative costs associated with the various methods, Table 6-7 provides a cost breakdown by the five survey methods and five study activities. The cost of each activity is ranked as low, medium, or high based on the relative level of costs associated with the given collection method. These assessments attempt to illustrate the general cost and complexity of a given method, rather than to precisely quantify costs.

TIP

Don't forget to include the costs to plan, analyze, and report occupancy study findings. These can add significantly to the collection costs.

Exact or final cost estimates to collect vehicle occupancy data will depend on a number of decisions made during the planning and design processes. All methods, except video, may be the most cost-effective depending on the specific study objectives.

TIP

If a more innovative method is selected, such as accident data, additional costs may be incurred to educate interested parties about the collection method and to defend the validity of the method.

Table 6-7 Comparison of Occupancy Data Collection Costs by Method and Activity

	METHOD				
Types of Costs (activities)	***Windshield***	***Surveys***	***Carousel***	***Accident***	***Video/Photo***
Planning	M	H	M	L	M-H
Mostly staff costs. (Define objective, area, and time period; select method; hire staff)	• Identify major areas of study	• Define area • Select survey type	• Identify major corridors (ensure multi-lane)	• Request data	• Develop equipment plan
Design	L	H	M	L	H
Staff and equipment costs. (Site selection, sample sizes, time segments, test duration)	• Pick sites and times • Assign staff • Acquire counting equipment.	• Obtain target list. • Design survey instrument • Sample as necessary • Prepare logistics (Letters, Ads, PO)	• Pick sites and times • Assign staff • Acquire counting equipment.	• Pick sites and times	• Pick sites and times • Assign staff • Acquire counting equipment.
Collection	*study design and regional labor costs greatly impact collection costs*				
Staff and equipment costs. (Perform steps required to obtain data)	• Deploy staff • Enter data	• Mail-out or phone respondents • Receive survey data • Screen responses • Enter data	• Ensure spacing between vehicles • Deploy staff • Enter data	• Download accident files • Make sure records are readable	• Restock film • Develop film • Review tapes • Enter data • Deploy staff
Analysis/ Validation	L	L	M	H	M
Mostly staff costs. (Compile and check data, interpret results)	• Simple calculation	• Simple calculation	• Adjust/weight responses	• Adjust/ weight responses	• Simple calculation • Potential missing data problem
Reporting	L	L-M	L	M	L
Mostly staff costs. (Prepare reports,		• Defend method		• Defend method	
H = relatively higher cost, M = relatively medium cost, and L = relatively lower cost.					

7.0 FINDINGS, EXPERIENCES, AND RECOMMENDATIONS

MPOs are facing greater challenges to increase the efficiency of highway transportation systems, which have reached their capacity limit. Increasing the average vehicle occupancy on freeways and arterials has been identified as a potential low-cost solution to improving the efficiency of the existing transportation system. To evaluate the effectiveness of programs to increase vehicle occupancy, such as carpooling, vanpooling, and transit use, measuring changes in vehicle occupancy rates is important to MPOs.

This project studied ways of utilizing vehicle occupancy data collection methods to obtain reliable and accurate estimates while keeping the costs to a minimum. Overall, it was found that no method was suited to monitoring traffic under all geographic and temporal conditions. When selecting a collection method, an MPO has to be familiar with the specific geographic and temporal conditions that each collection method is suited for. By defining these and other study objectives, different methods were found to be more appropriate for different situations.

7.1 New Methods

This review of vehicle occupancy data collection methods has investigated newer methods for collecting occupancy data. The most promising methods are the carousel method and the accident method. The high technology methods require further development.

The carousel method, which has been overlooked by many MPOs, seems to provide better and less costly vehicle occupancy data for multi-lane, heavy traffic highways than the windshield method.

- The carousel method does not have the windshield problem of misinterpreting the number of occupants in a vehicle, since the observer is in a vehicle moving with the traffic. In comparing windshield and carousel field collection data, the carousel AVO estimates were consistently higher than the windshield AVOs during peak traffic periods.

- On multi-lane, heavy traffic corridors, the carousel method can be more cost effective than the windshield method because fewer observers are needed to monitor the corridors.

The accident method, which is becoming more popular because it is a low-cost, temporally-rich data source, is a viable method. Whenever possible, the data should be adjusted to remove potentially biasing influences.

- Trends in AVO rates obtained from the accident data compared well with other sources.

- Younger drivers, who tend to have higher AVOs, were over-represented in the accident data. However, for the Chicago region, this factor did not result in a statistically significant difference between AVO estimates that were corrected for this bias versus AVO estimates that did not correct for this bias. This may not be true for other areas or other years.

- Similarly, there was not a significant difference between AVOs, which included all types of accidents, and AVOs, which excluded drug/alcohol-related accidents.

The use of high technology devices for measuring vehicle occupancy has primarily focused on automating the collection stage and has largely ignored developing an automated process for extracting occupancy data from the film. These methods can sometime reduce collection costs; however, the data extraction costs are exorbitant. In general, it is quicker and less costly to use the windshield method.

- High technology methods (e.g., infrared, automated recognition) are not practicable at this time.

- Modern technologies for replacing the human observer with cameras or radar are still in the developmental stages and currently are not effective methods of collection. The methods require human observers to view the film and record occupancy rates.

7.2 Improvements to Existing Methods

Based on observations of the state-of-the practice in conducting a vehicle occupancy study, improvements can be made to survey procedures that can increase the accuracy of AVO estimates and/or lower costs:

- Prior to data collection, planning decisions regarding the expected level of precision in the AVO estimates should be clarified to ensure that sufficient data are collected and to avoid the collection of unnecessary data, which can reduce collection time and costs.

- The precision of AVO estimates obtained from field data should be reported with the estimates. This information will clarify the accuracy of the data and the strength of any conclusions based on the data.

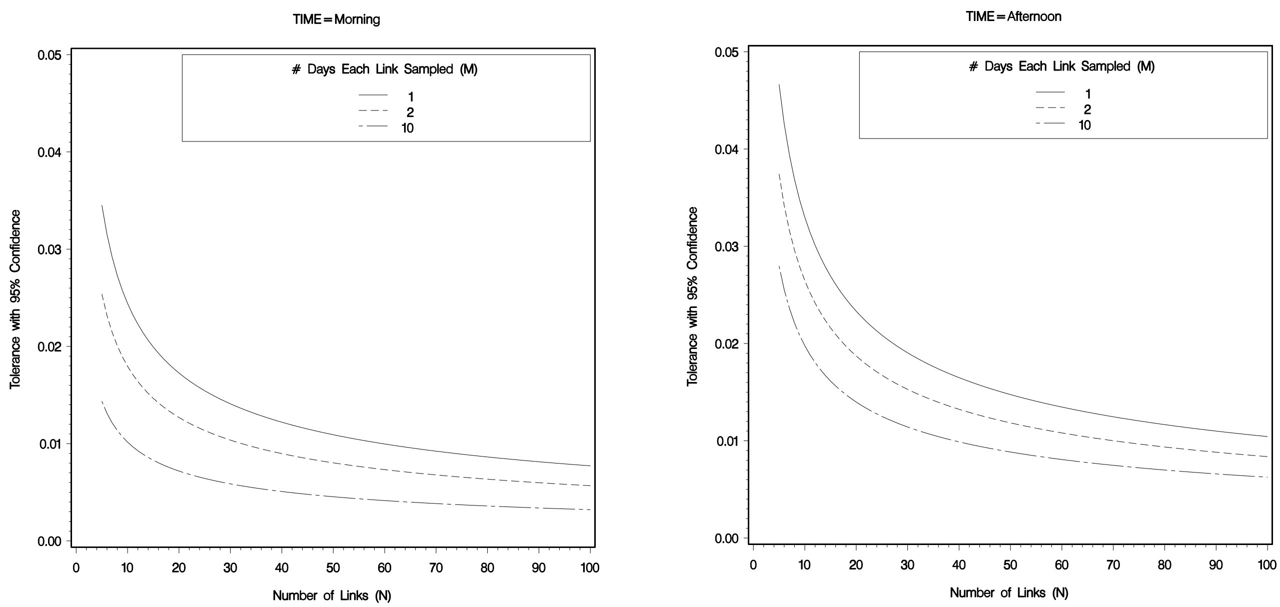
- Only random sampling and selection of collection sites and days will allow for the evaluation of statistical precision in the AVO estimates.

- The sampling procedures used in occupancy studies need to randomly select the geographic sites/links. In the current practice, MPOs often will deterministically select higher volume roads to the exclusion of lower volume roads. This practice can result in biased estimates, and any conclusions generated cannot be extrapolated outside the locations sampled. To account for traffic volumes when selecting road segments, a weighted random sampling procedure should be used.

- The sampling procedures used in occupancy studies need to randomly select the collection days. For example, it has been observed that MPOs sometimes intentionally exclude collecting data on Mondays and Fridays, but will report their AVO estimates as weekday estimates when in truth the estimates only represent midweek traffic.

- When conducting a regional or corridor study using one of the observational collection methods, it is better to collect data from more sites than to collect data multiple times from the same sites/links (see Figure 7-1). This is because increasing collection sites has a much greater impact on the precision of the AVO estimate.

Figure 7-1 Tradeoff Between Number of Links Sampled and Number of Days Sampled Per Link
(Based on Variability Estimates from KIPDA)



- MPOs need to become more aware of the sensitivity of AVO estimates to variability in temporal conditions. It has been shown that AVO rates do differ across time of day, day of week, and season of year. Therefore, data collected for one specific period cannot be used to make inferences about another different time period. For example, yearly AVO rates cannot be based solely on winter AVO rates.

- In a windshield study, all cars do not need to be observed to have a statistically accurate AVO estimate. Better data on fewer cars is more desirable than an observer guessing the occupancy of a vehicle. However, some procedure for randomly sampling part of the traffic stream does have to be followed, such as only recording occupancy on every other vehicle.

- Further evaluations of possible biasing factors in accident data need to be conducted on accident reports collected from different regions over different time periods.

- It should be investigated for various regions whether their regional accident data contains occupancy information. If the regional accident database does not contain occupancy information, an agency should coordinate with the law enforcement department to see if an occupancy field can be added to the next revision of the police report form.

7.3 Relative Costs

In comparing the costs associated with the various method, it was highlighted that the design of the study will greatly determine which method is most cost effective. In this study when comparing method costs, relative costs were used since exact costs will be greatly impacted by regional labor wages. When estimating the cost to implement a particular collection method, each MPO will have to adjust the costs to reflect their local rates.

- Where available, the accident data method was found to have the lowest costs. However, this cost did not consider the initial expenses of setting up the database and getting familiar with any peculiarities. It has been mentioned that accident costs should not be compared against the costs accrued by the other methods because the accident method is not a true collection method since it relies on existing databases that were primarily created to record accident information, not to reflect the occupancy patterns of the driving population.

- In comparing the costs associated with the various methods, it was discovered that the design of the proposed occupancy study will greatly impact the overall costs and findings. For instance, windshield is less expensive than carousel when only one or two links are evaluated; however, carousel is less expensive than windshield when a number of sequential multi-lane links are selected for evaluation.

- The survey activities that impact total costs the most are the activities of collecting the data and recording data. Based on this information, a survey planner may want to focus on new methods for performing these tasks that will lower their costs. For example, video surveillance could possibly reduce collection costs but its data entry cost are so large that the collection cost savings have little impact on the overall cost.

- Combinations of methods may result in better results or lower costs. For example, ConnDOT uses the windshield method to supplement accident data for roadways that contain HOV lanes since the accident method is known to underrepresent HOV lanes.

7.4 Conclusions

Federal regulations require the implementation of a Congestion Management System (CMS) to measure transportation system performance and for procedures to assess alternative congestion mitigation strategies. Vehicle occupancy is one way to measure system efficiency and effectiveness, to validate modal split in transportation models, to assess total person trips with vehicle trips, and to evaluate rideshare programs.

Traditional methods of collecting vehicle occupancy data, such as roadside windshield and intercept methods, require the use of field and office personnel to manually record and download vehicle occupancy information. The costs associated with roadside surveys are generally high due to extensive use of labor and to personnel costs. Careful planning is required to ensure the safety of the observers and to minimize the slowing of traffic because of "rubber-necking". It has been demonstrated that on multilane major arterial roads the "carousel" method can be more cost effective and still provide accurate estimates of vehicle occupancy without compromising safety or slowing the traffic. Alternative methodologies, such as video camcorders and mechanical vision vehicle occupancy readers, are more expensive than the observational roadside windshield method. However, these methods may in the future overcome some of the shortcomings associated with the manual methods.

The availability of existing vehicle occupancy data has recently prompted many MPOs to investigate the possibility of using accident databases to estimate average vehicle occupancy in their region. This relatively inexpensive method of using accident data has been shown to give reasonable estimates for routes and locations that have frequent crashes, although the data may not be timely. Also, the accident data extraction method has corroborated findings from other methods that average vehicle occupancy rates do vary in some locations by time of day, day of week, season of the year, and geographical location.

An MPO's selection of an appropriate methodology to estimate average vehicle occupancy is highly dependent on the objectives of their study. For example, a roadside windshield study is the most straightforward method to implement and generally used to provide a one-time, location specific average vehicle occupancy estimate. The carousel method appears to be the most effective method for estimating vehicle occupancy on multilane major arterial highways. The use of accident data provides temporal and geographical trends of vehicle occupancy. Video cameras might be best suited to monitor vehicle occupancy at specific locations, intersections, or facilities. If a study requires additional information, such as trip purpose or trip frequency, as well as the number of occupants, then the mail-out or handed-out survey approach should be used. Some combination of methods may provide the best solution to meeting of the objectives and requirements of a particular MPO study.

Each of the methods evaluated in this study can provide accurate and precise estimates of vehicle occupancy if the collected data is aggregated properly. Developing a sampling design will ensure that an effective (reduced time and cost) plan will be implemented that will provide the data required to meet the goals of the study.

REFERENCES

Agyemang-Duah, K., and Hall, F.L. (July 24-27, 1991). "Some Issues Regarding the Numerical Value of Freeway Capacity: Higway Capacity and Level of Service." *Proceedings of the International Symposium on Higway Capacity*.

Ahuja, V., and Hanscom, E. (1996). "Feasibility Study to Estimate the Average Vehicle Occupancy by Utilizing a Traffic Accident Database." Presented at the 1996 Annual TRB Meeting.

Arnold, E.D., Jr. *Review of Travel Data Issues in VDOT's Transportation Planning Division. Final Report*. Virginia Transportation Research Council, Report No. VRTC 92-R11.

Asante, S., Adams, L., Shufon, J., and McClean, J. (1996). "Estimating Average Automobile Occupancy from Accident Data in New York." Presented at the 1996 Annual TRB Meeting.

Barton-Aschman Associates, Arizona Department of Transportation, and Federal Highway Administration. (December 1986). *Phoenix Metropolitan Area External Trip Study.Volume I: Final Report.* Report No: FHWA/AZ 86/226-1.

Battelle Memorial Institute, Cambridge Systematics, Inc., and R.D. Mingo & Associates. (March 30, 1995). *Improved Vehicle Occupancy Data Collection: Task 2 -- Test Plan*. Submitted to the FHWA Office of Highway Information Management.

Beaton, W.P., Carragher, F.J., and Meghdir, H. (1993). "Impact of Nonresponse Bias On Forecasts of Average Passenger Occupancy." *Transportation Research Record 1390*, pp. 43-49.

Ferlis, R.A. (1979). "Field Data Collection and Sampling Procedures for Measuring Regional Vehicle Classification and Occupancy." *Transportation Research Record 701*.

Ferlis, R.A. (March 1981). *Guide for Estimating Urban Vehicle Classification and Occupancy*. FHWA Report No. HS-032-518.

FHWA Office of Urban Planning. (September 1981). *Guide to Urban Traffic Volume Counting*. Office of the Secretary of Transportation.

Fisher, R.F., Williams, G.J., and Boyd, J.P. (1980). "Atlanta Vehicle Occupancy Monitoring." *Transportation Research Record 779*, pp. 27-32.

Gaulin, R. (April, 1991). *A Procedure to Calculate Vehicle Occupancy Rates From Traffic Accident Data*. Connecticut Department of Transportation, Bureau of Planning.

Grush, W., and Gross, B. (September 1995). *Vehicle Occupancy: A Comparison of Household*

Survey Data and Roadside Observation Data. FHWA Office of Highway Information Management.

Heidtman, K., and Skarpness, B. (1996). *Improved Vehicle Occupancy Data Collection: Task 2 -- Test Plan*. Prepared for the National Traffic Data Acquistion Conference, May 8, 1996.

Hupp, R.C., and Palombo, C. (1980). "Evaluation of the FHWA Vehicle Classification and Automobile Occupancy Sampling Manual." *Transportation Research Record 779*, pp. 21-26.

Law, C. (November 1981). *A Photographic System for Auto Occupancy Counting and Vehicle Identification: Vol. I, Principal Finding and Conclusions*. FHWA Report No. FHWA-RD-82-000.

Levine, N. and Wachs, M. (October 1994). *A Methodology to Measure Vehicle Occupancy Rates at Corridor and Regional Levels, Part I: Operations Manual -- Instructions & Procedures.* Report prepared for The California Air Resources Board and The Office of Travel Forecasting and Analysis, California Department of Transportation.

Massie, D., and Campbell, K. (1993). *Analysis of Accident Rates by Age, Gender, and Time of Day Based on the 1990 Nationwide Personal Transportation Survey*. The University of Michigan Transportation Research Institute, Report No.:UMTRI-93-7.

Metropolitan Washington Council of Governments. (1992). *1992 Beltway Cordon Count of Vehicular and Passenger Volumes.*

Metropolitan Washington Council of Governments. (1989). *1989 Beltway Cordon Count of Vehicular and Passenger Volumes.*

NHTSA. (1992). *State Accident Report Forms Catalog 1992 Update*. National Highway Traffic Administration, US DOT Report No. DOT-HS-807-847.

Pietrzyk, M. et al. (January 1996). *Demostration of Video-Based Technology for Automation of Data Collection*. Center for Urban Transportation Research, University of South Florida.

Rutherford, G.S., Kinchen, R.K., and Jacobson, L.N. (1990). "Agency Practice for Monitoring Violations of High Occupancy-Vehicle Facilities." *Transportation Research Record 1280*, pp. 141-147.

Texas Transportation Institute. (February 1994). *Use of Advanced Technology in HOV Lane Enforcement*. Intermin Report.

Ulberg, C. (1994). *HOV Usage and Evaluation*. Washington State Department of Transportation.

Ulberg, C. and McCormack, E. (1988). “Accuracy and Other Factors Affecting a Continuous Vehicle Occupancy Monitoring Program.” *Transportation Research Record 1206*, pp. 35-47.

Ulberg, C. and McCormack, E. (June 1988). *Auto Occupancy Monitoring Study. Final Report*. Washington State Transportation Center, Report No. WA-RD-157-1.

URS Consultants, Inc. (August 1996). *Vehicle Occupancy Study: Alternative Data Sources and Methodologies*. Prepared for Florida Department of Transportation, Systems Planning Office. Draft Report.

Wagner,D., Murakami, E., and Neumeister, D. (1996). “Global Positioning Systems for Personal Travel Surveys .” Presented at the National Traffic Data Acquisition Conference, May 5-9, 1996.

Young, C.P. (1988). “A Relationship Between Vehicle Detector Occupancy and Delay at Signal-Controlled Junctions.” *Traffic Engineering and Control* 29(3), pp. 131-134.

ADDITIONAL LITERATURE

Battelle Memorial Institute, Cambridge Systematics, Inc., and R.D. Mingo & Associates. (March 30, 1995). *Improved Vehicle Occupancy Data Collection: Task 2 -- Test Plan*. Submitted to the FHWA Office of Highway Information Management.

Benke, R., and Sjoberg, R. (March 1977). “Auto Occupancy Parameter Variations.” *Traffic Engineering* 17(3), pp. 21-27.

Benuska, M., Hansen, M. and Ulberg, C. (February 1994.). “HOV Monitoring and Evaluation Tool.” Washington State Department of Transportation Report No: WA-RD-318.1.

Billheimer, J., Kaylor, K., and Shade, C. *(*March 1990). *Use of Videotape in HOV Lane Enforcement : Final Report*. California Department of Transportation.

Borg, W., and Meyburg, A. (1981). “Nonresponse Problem in Travel Survey: An Empirical Investigation.” *Transportation Research Record 775*, pp. 34-38.

Dunphy, R. (1979). “Workplace Interviews as an Efficient Source of Travel Survey Data.” *Transportation Research Record 701*, pp. 26-35.

Eash, R. (1987). “Management of a Small Home Interview Travel Survey Using a Microcomputer.” *Transportation Research Record 1134*, pp. 49-56.

Ettema, D., Timmermans, H., and van Veghel, L. (1996). *Effects of Data Collection Methods in Travel and Activity Research*. European Institute of Retailing and Services Studies.

Fehon, K. (March 1982). "A Technique for Estimating Car Occupancy in a Parking Study." *Australian Road Research* 12(1), p. 41.

Ferlis, R.A. (1979). "Field Data Collection and Sampling Procedures for Measuring Regional Vehicle Classification and Occupancy." *Transportation Research Record 701,* pp. 1-6.

Lester, M.S., Dare, J.W. and Roach, W.T. (1979). "Techniques for Monitoring Automobile Occupancy: Research in the Seattle Area." *Transportation Research Record 701*, pp 7-15.

Miller, K., Harvey, T., Shuldner, P., and Ho, C. (1993). "Using Video Technology to Conduct 1991 Boston Region External Cordon Survey." *Transportation Research Record 1412*, pp. 46-56.

National Roads Board. (1973). "Vehicle Occupancy, Trip purpose, Traffic Composition and Vehicle Operating Speeds." *RRU Bulletin 17*, pp. 1-48.

Ohstrom, E.G., Ohstrom, J.B., and Stopher, P. (1984). "Successful Administration of a Mailed 24-Hour Travel Diary." *Transportation Research Record 987*, pp. 14-20.

Rutherford, G.S., Kinchen, R.K., and Jacobson, L.N. (1990). "Agency Practice for Monitoring Violations of High Occupancy-Vehicle Facilities." *Transportation Research Record 1280,* pp. 141-147.

Smith, M. (1977). "Design of Small Sample Home Interview Travel Survey." *Transportation Research Record 701*, pp. 29-35.

Steuart, G.N., Mori, B., and Noehammer, P.J. (1987). "Direct Data Entry Using Microcomputers--A Travel Survey Pilot Project." *Transportation Research Record 1134*, pp. 40-48.

Texas Transportation Institute. (February 1994). *Use of Advanced Technology in HOV Lane Enforcement*. Intermin Report.

Young, C.P. (1988). "A Relationship Between Vehicle Detector Occupancy and Delay at Signal-Controlled Junctions." *Traffic Engineering and Control* 29(3), pp. 131-134.

Zakaria, T. (1991). "Cordon Line Highway Survey for the Delaware Valley Region." *Transportation Research Record 1305*, pp. 93-102.

APPENDIX A

SUMMARY REVIEW OF VEHICLE OCCUPANCY DATA COLLECTION REPORTS

MPO/Agency:	New York State Department of Transportation
City/State:	New York City, New York
Survey Date:	November 17, 1988
Study Purpose:	To determine current and future traffic and air quality impacts of proposed development scenarios for Route 9A reconstruction.
Methodology Used:	Of primary interest, a cordon line origin-destination mailback questionnaire and a parking lot survey. Direct observation vehicle occupancy counts of 10 percent of vehicles to be taken at time of survey distribution.
Geographical Coverage:	Primary and secondary impact zones surrounding Route 9A Reconstruction Project area, New York City. Survey stations at selected intersections and ramps.
Temporal Coverage:	Surveys distributed 6:00 to 8:00 A.M., Monday, Tuesday, and Thursdays in November 1988.
Accuracy/ Precision:	Author notes potential non-response and sample selection bias.
Cost Effectiveness:	N/A
Notes:	The survey includes one question about how many people in the vehicle. Determining vehicle occupancy is not the primary purpose of this plan

MPO/Agency: Association of Central Oklahoma Governments

City/State: Oklahoma, Canadian and Cleveland Counties

Survey Date: April, May, September 1990 and May 1991

Study Purpose: Update VOC survey methodology to better meet Clean Air Act and transportation planning requirements to update computer model with new/more accurate data.

Methodology Used: Field survey, two years in a row, two seasons, three one-hour periods per day, Monday, Tuesday, or Thursday. Thirty stations; clustering for selection, four types of roadways.

Geographical Coverage: Twenty stations throughout Oklahoma County, in Canadian and Cleveland County.

Temporal Coverage: Three one-hour sessions -- 7:00 to 8:00 A.M., 9:30 to 10:30 A.M., 4:30 to 5:30 P.M. -- April, May, and September.

Accuracy/ Precision: * 1989 data was reviewed, erroneous counts were detected, 1.96 standard development was used to calculate 95 percent coverage, 95 percent confidence interval.

* Cluster sampling was used to avoid bias.

* Survey conducted both in 1990 and 1991 to validate results (found 1.24 vs. 1.20 VOC). An average was used for the model.

Cost Effectiveness: N/A

Notes: It was recommended to consolidate subareas of study for purposes of air quality monitoring and transportation planning work (not clear why).

A Procedure to Calculate VOC Rates from Traffic Accident Data

MPO/Agency: Connecticut DOT

City/State: Connecticut

Survey Date: April 1991 (1989 last year accident data available)

Study Purpose: To document a procedure for calculating VOC rates from traffic accident database to be easily implemented to investigate bias entering into calculations.

Methodology Used: Using calculated minimum sample size, extract relevant vehicle and occupancy data from accident data base, to be customized to individual user, computer system and accident database.

Geographical Coverage: Statewide/accidents on interstates/selected vehicle types (i.e., private auto).

Temporal Coverage: Can be as specific as by hour of the day, week, season, or year.

Accuracy/ Precision: Work vs. non-work trips must be estimated making assumptions about time-of-day.

Can calculate results on more frequent basis given ease -- helps with trend analysis. Compared to 1982 and 1984 field survey data, results were similar, varying by as much as eight percent.

Cost Effectiveness: Very, if system for entering proper data exists. Huge time savings. $250.00 vs. $12,800 for field survey.

Notes: Bias problems:

* Weather
* Male vs. Female
* Drug Involvement

MPO/Agency: U.S. DOT National Highway Traffic Safety Administration

City/State: Nationwide

Survey Date: April 1994

Study Purpose: A proposal of how to measure level of seat belt use of front seat occupants of passenger vehicles that are covered by the state safety belt laws -- autos, vans, and light trucks included.

Methodology Used: Using clustering and stratification, representative sampling sites were selected around the country using census data and state seat belt usage rates. Direct observation field survey counting of persons using seat belts and helmets.

Geographical Coverage: Nationwide.

Fifty primary sampling units, 500 tracts, 4,000 sites.

Temporal Coverage: To be probability -- determined.

Accuracy/ Precision:

Cost Effectiveness: See notes.

Notes: Outlines a basic design for measuring seat belt use nationwide to balance cost effectiveness with statistical accuracy. Focus is how to get a valid sampling in such a large scale survey.

MPO/Agency: Arizona Department of Transportation

City/State: Phoenix, Arizona

Survey Date: Spring 1988

Study Purpose: To determine what factors have the greatest influence on vehicle occupancy rates, and which conclusions are transferrable to other metropolitan areas.

Research/Application Objectives: 1) to calibrate the shared ride component their mode choice model to effect carpooling characteristics; 2) to distinguish between determinators that can be affected by public policy; and 3) to determine what elements are transferable to other metropolitan areas.

Methodology Used: Two surveys: 1) direct observation at destination by time-of-day, stratified by geographic area and highway facility type (for validation and calibration of model); and 2) mail-back intercept survey of selected arriving vehicles at parking lots to find out more detailed information (to confirm or identify characteristics of travelers/journey/destination which influence VOC).

Geographical Coverage: Direct observation: locations throughout metropolitan Phoenix stratified by geographic area and highway facility type. Questionnaire distributed to vehicles arriving at stratified sample of parking garages and lots throughout metropolitan Phoenix.

Temporal Coverage: Direct observation: 7:00 A.M. to 5:30 P.M., with breaks. Covers two-hour A.M. peak, four-hour P.M. peak, eight hours all together. Questionnaire: handed out until all cars had entered lots for the day.

Accuracy/ Precision: * Driver behavior measured by statistics about characteristics, rather than a survey, proved more reliable and cost effective.
* Given small portion of carpoolers, survey was designed to generate sufficient valid responses. Still, problem getting enough surveys back.

Cost Effectiveness: N/A

Notes: Much effort in creating a technique which would provide information relevant to specific goals of the project. Trip purpose and, therefore, also time-of-day, were by far the most significant determinators of VOC.

MPO/Agency: Delaware Regional Planning Commission

City/State: Selected Counties in New Jersey and Pennsylvania

Survey Date: April 1992

Study Purpose: To design a statistically valid procedure for a telephone survey to determine average vehicle occupancy in 13 counties.

Methodology Used: Telephone survey of randomly selected telephone numbers to determine vehicle occupancy rates.

Geographical Coverage: Total 4,800 completed interviews, 13 counties.

Temporal Coverage:

Accuracy/ Precision: Given budgeting constraints and lacking county commuting data, not possible to be precise.

Cost Effectiveness: N/A

Notes: * Author notes non-response bias as an issue.
*Related to ozone attainment program.

MPO/Agency: Washington State Department of Transportation

City/State: N/A

Survey Date: May/June 1987

Study Purpose: To investigate various methods to measure AVO in order to determine feasibility and costs of a continuous, ongoing data collection program. The data would be used to help evaluate the effectiveness of programs to increase vehicle occupancy.

Methodology Used: A literature review, design of data collection method, testing of the field collection method, and development of a proposal for assessing vehicle occupancy in the Puget Sound region. Field tests focused on the influence of different conditions on counting error.

Geographical Coverage: Sixteen observations at 10 locations across the region, at employment sites, arterials and freeways.

Temporal Coverage: From 6:00 to 9:00 A.M., 3:00 to 6:00 P.M., 6:00 to 9:00 P.M., for 14 days in a row.

Accuracy/ Precision: Three sets of measurements were made independently to measure accuracy of test. Conclusion that vehicle observers were correct 97 percent of the time.

Cost Effectiveness: Quarterly counts at 26 sites that are accurate within 1.5 percent were able to be conducted for about $50,000 per year.

Notes: Comprehensive discussion of factors influencing accuracy, amount of data required and types of survey design. The authors conclude that data collection by direct observation offers the most cost-effective and accurate means of assessing vehicle occupancy rates.

MPO/Agency: Washington State Department of Transportation for Puget Sound Council of Governments

City/State: Puget Sound Area

Survey Date: From October 1989

Study Purpose: To evaluate data collection methodology proposed in the previously completed "Auto Occupancy Monitoring Study," initiate a time series database, and incorporate the resulting data into regional travel demand modeling.

Methodology Used: Direct observation field counts at randomly-selected sites.

Geographical Coverage: Fifty-four sampling sites in the Puget Sound area.

Temporal Coverage: AM and PM peak hours randomly selected throughout one year period.

Accuracy/ Precision:

Cost Effectiveness: N/A

Notes: Report focuses on administration of the data collection effort and reducing cost and little on sampling methodology.

A Photo System for Auto Occupancy Counting and Vehicle Identification: Principal Findings and Conclusions

MPO/Agency: FHWA Offices of Research and Development

City/State: Washington, D.C.

Survey Date: November 1981

Study Purpose: To redesign and test a prototype camera system to meet desired physical specifications and performance requirements for documenting vehicle occupancy and vehicle type.

Methodology Used: The feasibility of using a pulse/time lapse camera (equipped with an electronic flash and an optical vehicle sensor), or two synchronized cameras, to determine vehicle occupancy and identification, is discussed and tested. Study examines camera's ability to meet different performance requirements, including time-of-day, weather, and distance.

Geographical Coverage: Two HOV lanes and a ramp on Interstate 395 near Washington, D.C.

Temporal Coverage: Nine tests were conducted on different days at different times of day.

Accuracy/ Precision:

Cost Effectiveness: N/A

Notes: Traffic was not impacted. In most cases the number of occupants could be determined. Some types of car windows could not be penetrated by the flash.

MPO/Agency: Executive Transportation Committee for Chemung County

City/State: Elmira, New York

Survey Date: June 1991, six-day period

Study Purpose: To gather travel patterns and demographic data to help market and develop declining areas of downtown Elmira and to support planning strategies to attract new development.

Methodology Used: 1. Origin-destination survey. Traffic counts and vehicle classification information was obtained for all vehicles crossing designated cordon line entry points.
2. Pullover roadside interview to a random sample of motorists. Data collected on trip timing, origin, destination, socioeconomic characteristics, and number of vehicle occupants.

Geographical Coverage: Thirteen point-of-entry sites selected at high volume roads.

Temporal Coverage: A.M. peak, six weekdays over a two-week period.

Accuracy/ Precision:

Cost Effectiveness: N/A

Notes:

MPO/Agency: New York Metropolitan Transportation Authority

City/State: New York City and surrounding areas

Survey Date: Spring 1989

Study Purpose: To collect travel-related information to support New York area public transit agency planning.

Methodology Used: Telephone survey of a random sample of 21,000 respondents (drawn by stratifying the sample by county and conducting a simple random sample within each county) on travel patterns within a 24-hour period, including vehicle occupancy.

Geographical Coverage: New York City and surroundings areas.

Temporal Coverage: One full 24-hour day in the Spring of 1989

Accuracy/ Precision:

Cost Effectiveness: N/A

Notes:

MPO/Agency: Rhode Island Department of Transportation, Division of Planning

City/State: Providence, RI

Survey Date: December 1991

Study Purpose: To determine the magnitude and the orientation of vehicle traffic movements within the study area to study ways of alleviating congestion on the Interstate 95 corridor.

Methodology Used: A two-stage survey:
1. License plate samples of vehicles crossing the two facilities.
2. Postcard survey sent to addresses matching license plates on trip purpose, length, and duration, along with the question: "How many people are usually in the vehicle (including driver)?"

Geographical Coverage: Two high volume facilities which bisected the study area.

Temporal Coverage: A.M. peak, weekday

Accuracy/ Precision:

Cost Effectiveness: N/A

Notes:

APPENDIX B

SAMPLE SIZE FORMULAS

Calculating Sample Size for Each Collection Method

Direct observational methods are probably the most widely used vehicle occupancy collection method in transportation planning. They typically involve counting people through a windshield at randomly selected geographical locations within a defined region. The purpose of most direct observational methods is to quantify vehicle occupancy, vehicle type, traffic volume, or to evaluate characteristics of people or vehicle movements at specific locations.

Windshield Sampling Plan

In the windshield method, occupancy data can be collected on a number of road segments (links) in the geographic area of interest during a number of different time periods. Before conducting the collection effort, the number of sampling units to be observed in order to get a reasonable estimate has to be determined. The sampling unit for the windshield method is the link-day which is defined as a particular sampling period on a particular link where a link is an uninterrupted section of road. The sampling period is a specified time of day on a specific day of the week and optionally during a specific season. For each link-day, the windshield method is used to collect data on the number of vehicles and number of vehicle occupants observed. Determining the number of samples required to obtain a reasonable AVO estimate depends on whether the estimate required is site-specific, corridor, or area-wide. Separate calculations are presented for these three levels.

Site-Specific Study

For a site-specific study, since only one site/link is of interest, the sampling plan only has to consider the appropriate number of days to collect windshield data. This number is dependent on the level of tolerance and confidence required in the AVO estimate and on the temporal (day-to-day) variability, σ_d, of vehicle occupancy at this site/link. If no day-to-day variability in vehicle occupancy patterns existed, then collection would only have to occur on one day. The formula for calculating the number of collection days, M, for a given tolerance level, B, with 1-α confidence is:

$$M = \left(\frac{z_{1-\alpha/2}\,\sigma_d}{B} \right)^2$$

where $z_{1-\alpha/2}$ is the upper 1-α/2 percentile of the standard normal distribution.

For example, the minimum number of collection periods required to obtain an estimate within 0.05 of the true AVO with 90% confidence (α set to 0.1) would be written as:

$$M = \left(\frac{1.645\sigma_d}{0.05} \right)^2 = 1082\sigma_d^2$$

If day-to-day variability, σ_d, in occupancy rates were known to be low, say 0.07, then M=6. If σ_d was high, say 0.13, then the number of collection days would be M=19.

Regional Study

To estimate the area-wide AVO within a desired tolerance level, a sufficient number of links (road segments) have to be sampled for an appropriate number of time periods (e.g., mornings or afternoons). Factors that will influence the number of links and the number of time periods to be sampled include the geographic variability in vehicle occupancy across the links (road segments), and the temporal (day-to-day) variability of vehicle occupancy within links. For simplicity, it is assumed that variability is the same across links.

The standard deviation in vehicle occupancy across links is assumed to equal some unknown constant, $\sigma_{L.}$ The standard deviation in vehicle occupancy across days is assumed to equal some unknown constant, σ_d. These are referred to as measures of the geographic and temporal variability, respectively. Estimates of these standard deviations can be obtained from prior research.

To determine the number of links, *L*, and the number of days per link, *M*, to be sampled to get a regional AVO estimate within a tolerance level of *B*, with 1-α confidence, the following formulas would be used:

$$L = \frac{z_{1-\alpha/2}}{B^2}\sigma_L\left(\sigma_L + \sqrt{\frac{C_d}{C_L}}\sigma_d\right)$$

$$M = \sqrt{\frac{C_L\sigma_d^2}{C_d\sigma_L^2}}$$

where $z_{1-\alpha/2}$ is the upper 1-α/2 percentile of the standard normal distribution, C_L is the cost of identifying a site for sampling the chosen link and C_d is the cost of sampling an additional day.

To insure that the number of days, M, is a whole number some adjustments may have to be made with the derived M obtained in the previous calculation. If M is less than 1 then M is set to 1 day. If M is a non-integer greater than 1, then the number of days that each link is sampled should be varied across links between the two integers surrounding M such that they average to M. For example, if M=1.4, then 60 percent of the links should be sampled once and 40 percent of the links should be sampled twice.

In calculating the number of days that one link should be sampled this formula takes into consideration the average costs associated with each link. Two cost components were considered: the start-up costs of assessing each location for safety and adequacy, C_L, and the daily collection cost for a day, C_d. It was assumed that the costs associated with a link are the same across links, i.e., even though links may differ in characteristics such as number of lanes, the same procedure will be used on each link. If this is not the case in practice, the average costs (C_L and C_L) across links should be estimated for these computations. Table B-1 presents a breakdown of these costs for the windshield method.

Table B-1 Windshield Link Costs

New Site Cost (C_L)	Sampling Cost of Each Data Collection Period (C_d)
• Ensuring the safety of data collectors • Examining the adequacy of the site	• Transportation to and from site • Labor of observation • Downloading data • Interpreting or filtering data. • Hand-held counter, laptop etc.

For the windshield method, the new site costs, C_L, may be insubstantial compared to the daily collection costs, C_d; however, for other methods like video surveillance they can be a bigger factor. If C_L is very small in comparison to C_d and the variability in AVO between links is large relative to day-to-day variability on the same link, the most efficient design is to sample each selected link for only one day.

Examples of estimates for σ_L and σ_d are presented in Table B-2. These estimates were obtained from the KIPDA field test of the windshield method that was conducted for this study.

Table B-2 Geographical and Temporal AVO Standard Deviation Estimates (KIPDA)

Source of Variability		Morning	Afternoon
σ_L - (Link-to-link)		0.011	0.028
σ_d - (Day-to-day)	Highest	0.059	0.074
	Average	0.037	0.044

Corridor Study

If an AVO estimate is required for an entire corridor, which consists of more than one link, then a compromise between the site-specific and the area-wide sampling calculations is appropriate. If there is a strong belief that the link AVOs within the corridor are strongly correlated, i.e., occupancy characteristics are very similar, then it is probably sufficient to use the site-specific method. This means that one link may be able to represent the corridor, thus the only calculations to be made are the number of days to sample. However, it is advised that the number of sites sampled should be slightly increased to accommodate for any unknown differences between the links in the corridor.

Carousel Surveys

The sampling procedure for the carousel method is very similar to the windshield sampling method. As in the windshield procedure, the sampling unit is the link-day, which is a particular sampling period (usually a morning or an afternoon) on a particular link. Theoretically, the link-to-link variability and the day-to-day variability for the carousel observations should be the same as those of the windshield observations. The only differences will be the actual value of the cost parameters, C_L and C_L. The factors that contribute to the carousel collection costs are described in Table B-3.

Table B-3 Carousel Link Costs

New Site Cost (C_L)	Sampling Cost of Each Data Collection Period (C_L)
• Identification of the route • Locating a staging area	• Transportation to and from site • Cost of Observer Vehicle (Driver + Vehicle costs per mile) • Labor of Observation • Downloading data • Instruments cost - computers, hand-held counters, etc.

Mail-out Questionnaires

When designing sampling approaches for residential surveys, it is important to achieve a maximum response rate from the residents who receive the questionnaire. The most troubling problem in survey research is non-response bias. Non-response bias usually refers to a bias introduced into the data because not all survey respondents return the surveys and the persons who do respond may differ from those who do not respond. The statistical validity of the survey results will be directly related to the sample design, sampling plans, and procedures that compensate for such biases.

In mail-out surveys for vehicle occupancy, questionnaires are mailed to a particular subpopulation to inquire about the residents' travel behavior. A simple approach is to address the questionnaire to a resident and ask the resident to answer questions regarding their typical trips. A primary question used on many surveys asks the respondent to indicate the number of occupants traveling with them to or from work. The sampling unit is the returned questionnaire. The observation is the number of persons that travel with the respondent during the day on the trips described.

One approach in sampling is to stratify the population of interest by zip code (stratum), and to send out questionnaires to a percentage of the driver population in each stratum. The sampling design for this procedure (based on zip codes) is described next.

Regional Study

To estimate regional AVO within a desired tolerance, one needs to collect a sufficient number of questionnaires. Because sampling will be done in proportion to the population in each stratum, it is only necessary to determine the total number of returned questionnaires needed. This number depends on the variability of vehicle occupancy within areas. With prior estimates of the standard deviation in vehicle occupancy for each area, it is possible to estimate the required number of questionnaires to be sent and to be returned.

When the region is broken into multiple strata, the number of returned questionnaires, n, needed to obtain an AVO estimate within a tolerance level of B with 1-α confidence can be calculated using the following formula:

$$n = \frac{z^2_{1-\alpha/2} N \sum_{i=1}^{L} N_i \sigma_i^2}{N^2 B^2 + z^2_{1-\alpha/2} \sum_{i=1}^{L} N_i \sigma_i^2}$$

where $z_{1-\alpha/2}$ is the upper 1-α/2 percentile of the standard normal distribution, N_i is the driver population in stratum i, N is the driver population in the region (the sum of the N_i's), and σ_i is the standard deviation for stratum i (the variability in occupancy in the stratum).

When only one stratum is of interest or when the standard deviations for the various strata i are known to be equal, the following equation can be used. In this case the standard deviation for the stratum, σ_i, is the same as the standard deviation for the regional, σ_r.

$$n = \frac{z^2_{1-\alpha/2} N \sigma_r^2}{N B^2 + z^2_{1-\alpha/2} \sigma_r^2}$$

The standard deviation estimates of vehicle occupancy for different strata, σ_i, can be obtained from prior research. Examples of these estimates are provided in Table B-4. These estimates are based on the data collected by KIPDA using the mail-out questionnaire method. The strata in this example are the different zip code areas sampled.

Because not all questionnaires will be returned, the number of surveys distributed, N_d, has to be calculated after the number of returned questionnaires, n, is derived. A formula for determining the number of questionnaires necessary to distribute is as follows:

$$N_d = \frac{n}{r}$$

where r is the assumed response rate. Depending on the population and the length and complexity of the questionnaire, it is reasonable to assume a 10-15% response rate.

Using the KIPDA example, Table B-4 describes the minimum number of surveys to be distributed and the minimum number of surveys that need to be returned. It should be observed that the sample size estimates have a strong dependence on standard deviation. For example, area code 40059, which has a σ_i of 0.163, only requires a minimum of 110 questionnaires to be distributed to that area, while area code 47106, which has a higher σ_i of 0.475, requires 900 questionnaires to be distributed.

Table B-4 Example of Minimum Samples Sizes Required by Stratum

KIPDA Stratum	σ_i - Standard Deviation	*n* - Returned Questionnaires	N_d - Distributed Questionnaires
40014	0.324	42	420
40047	0.341	47	470
40059	0.163	11	110
40204	0.276	31	310
40207	0.225	21	210
47106	0.475	90	900
47111	0.396	63	630
47119	0.233	22	220

Note: Tolerance level, B, was set to 0.1 with 95% confidence.

Accident Data

Unlike the previous procedures, the sampling plan for the accident method is not used to determine the number of sample points to be collected; instead it is used to determine if the database contains a sufficient number of records to make the required estimations. Because the accident data method is based on existing databases, an MPO can only determine whether the accident data will allow estimation of AVO within a specified level of precision. If not, the MPO will have to use another method for collecting vehicle occupancy data.

The use of accident data to estimate AVO was discussed in more detail in the report. To use the data, the MPO has to clearly define the population of interest in terms of geographic coverage and temporal conditions. Within each selected subpopulation, the variability in occupancy rates will dictate the number of accidents that are required to obtain estimates of a specified precision level. Using prior estimates of the standard deviation for each subpopulation of accidents, the formula for calculating the number of vehicles in subpopulation i that are required to estimate AVO within a tolerance level of B with 1-α confidence is:

$$n_i = \left(\frac{z_{1-\alpha/2}\sigma_i}{B} \right)^2$$

where $z_{1-\alpha/2}$ is the upper 1-α/2 percentile of the standard normal distribution, and σ_i is the standard deviation for subpopulation i.

For example, the required number of vehicles needed to derive an AVO estimate within 0.1 with 95% confidence is:

$$n_i = \left(\frac{2\sigma_i}{0.1}\right)^2 = 400\sigma_i^2$$

The following table provides estimates of the standard deviation, σ_i, of vehicle occupancy for several strata, as well as the number of reported vehicles that would be needed to estimate AVO. Note that there are frequently more than one vehicle involved in any one accident, implying that if 200 vehicles are required, then perhaps this could be obtained on the basis of only 100 accident reports.

Table B-5 Chicago Accident Data Vehicle Occupancy Analysis[1]

Time Period (Days and Times)		Actual # of Vehicles in Database	Standard Deviation	Minimum Required # of Vehicles [5]
All Accidents		68,627	0.866	289
Daylight	All	48,923	0.841	272
	Weekend	10,016	1.043	418
	Weekday	38,907	0.765	226
	Weekday-AM Rush[2]	9,033	0.577	128
	Weekday-Midday[3]	15,349	0.786	238
	Weekday-PM Rush[4]	13,264	0.831	266

[1]Restricted to drivers 16 years and older, and vehicles with less than or equal to 18 occupants.
[2]6:00AM-9:00AM.
[3]9:00AM-3:00PM.
[4]3:00PM-7:00PM.
[5]Required number of vehicles involved in the accidents for ± 0.1 precision.