FINAL REPORT

Contract No. A2-043-32
PES Project No. 608

EVALUATION OF ARB'S IN-USE VEHICLE SURVEILLANCE PROGRAM

Prepared by:

PACIFIC ENVIRONMENTAL SERVICES, INC.

Lowell G. Wayne
Yuji Horie

Prepared for:

CALIFORNIA AIR RESOURCES BOARD
Sacramento, California

October 1983

ABSTRACT

Because the State of California has more stringent emission standards for motor vehicles than those of the United States Environmental Protection Agency (EPA), the Air Resources Board (ARB) conducts a continuing program of measuring emissions from vehicles in customer service. ARB uses data from this in-use vehicle surveillance program not only for assessing emissions but also for developing and guiding its strategies for control of motor vehicle emissions. Main purposes of this study, therefore, were (1) to evaluate sample selection and procurement methods as applied by ARB's procurement contractors and (2) to evaluate the vehicle selection methodology applied by EPA for its relevance to ARB's needs.

To accomplish these purposes, Pacific Environmental Services (PES) reviewed both the selection procedures actually used by procurement contractors and those described in proposals and final reports. Procedures applied in the field were evaluated through interviewing contractor personnel who had worked on procurement contracts and through analyzing data on emissions, procured vehicles, and owners' socioeconomic conditions. Results of the study indicate that a scientific random sampling of in-use vehicles is extremely difficult to achieve and that several changes in currently used selection procedures should be made to obtain a more representative sample of the California vehicle population. The changes needed appear to require procedure modifications mostly by ARB rather than by its contractors. Recommendations are made for improvements in future vehicle surveillance programs, based on the findings of the present study and also on the prospect of the forthcoming California I/M program.

DISCLAIMER

The statements and conclusions in this report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products.

ACKNOWLEDGMENTS

The successful completion of this study depended upon the cooperation of many individuals and groups, especially personnel of the California Air Resources Board and the three procurement contractors: Environmental Resource Management, Systems Control, Inc., and Automotive Environmental Systems, Inc. The Southern California Association of Governments provided useful statistics of population, household, and family income. The following individuals contributed to the completion of this study:

J. Pantalone, ARB Sacramento, Research Contract Manager

M. Hostak, ARB El Monte, Senior Air Pollution Specialist

Y. Horie, PES Santa Monica, Project Manager and Principal Investigator

L. Wayne, PES Santa Monica, Senior Author

T. Leuty, PES Santa Monica, Data Clerk

This report was submitted in fulfillment of ARB Contract No. A2-043-32 entitled "Evaluation of ARB's In-Use Surveillance Program" by Pacific Environmental Services, Inc. under the sponsorship of the California Air Resources Board. Work was completed in April 1983.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION AND SUMMARY

Because the State of California has more stringent emission standards for motor vehicles than those mandated by the United States Environmental Protection Agency (EPA), the Air Resources Board (ARB) conducts a continuing program of measuring emissions from vehicles in customer service. This "in-use vehicle surveillance" program parallels, but is independent of, the "in-use compliance" program conducted by EPA.

ARB uses data from the surveillance program not only for assessing emissions but also for developing and guiding its strategies for control of motor vehicle emissions. It is, therefore, important to ensure that the vehicles selected for the surveillance program adequately represent California's motor vehicle population.

To investigate whether, in the past, vehicles selected for the surveillance program have constituted an adequately representative sample of the California vehicle population, and to recommend improvements in methods of sample selection and procurement, ARB contracted with Pacific Environmental Services, Inc. (PES) for a study entitled "Evaluation of In-Use Vehicle Surveillance Program." The formal objectives of this study were two-fold:

1. Evaluate sample selection and procurement methods as applied by ARB's vehicle procurement contractors and recommend needed measures to ensure that procured samples adequately represent the vehicle populations under consideration.

2. Evaluate the vehicle selection methodology applied by EPA for in-use vehicle compliance testing, and recommend selection methods to ensure that the proportions of tested vehicles failing to meet the emission standards are adequately representative of the vehicle populations under consideration.

In the contract, ARB directed PES to evaluate the sample selection and procurement methods used by ARB's vehicle procurement contractors; to investigate methods used by EPA in its parallel program for in-use vehicle compliance testing, as well as other vehicle selection methods used for vehicle recall programs; and to develop recommendations for the improvement of ARB's selection methodology, based on the findings from these evaluations.

Specific tasks identified were the following:

TASK 1. Evaluation of methodology used for the selection of vehicles for ARB's "In-Use" Vehicle Surveillance Program.

TASK 2. Evaluation of the selection methodology actually practiced by ARB's procurement contractors.

TASK 3. Evaluation of EPA's "In-Use" Vehicle Compliance Program.

TASK 4. Review of Other Sample Selection Methods; Recommendations.

This chapter, Chapter 1.0, presents a summary of findings of the present study and a set of recommendations for improvements in future in-use vehicle surveillance programs. Following chapters of this report provide more complete detail on the considerations involved, the information retrieved, and the findings and conclusions arrived at by PES in carrying out this study.

Chapter 2.0 addresses Task 1, the evaluation of vehicle selection and procurement procedures as documented by ARB and its procurement contractors in proposals, work plans and final contract reports.

Chapter 3.0 addresses Task 2, the evaluation of de facto methodology as reported by representatives of firms having current or completed procurement contracts and as indicated in the reports of completed surveillance test series.

Chapter 4.0 addresses Tasks 3 and 4, evaluation of procurement methodology specified by EPA for its compliance testing and any relevant methodologies found in connection with other vehicle recall programs.

Chapter 5.0 provides an overview and discussion of progress made, technical findings and their implications in regard to various aspects of methodology for vehicle procurement.

References are provided in Chapter 6.0 and a Glossary in Chapter 7.0.

1.1 SUMMARY OF FINDINGS

This study, an evaluation of the In-Use Vehicle Surveillance Program (VSP), was carried out under contract (A2-043-32) to the State of

California, Air Resources Board (ARB). Its main purposes were (1) to evaluate sample selection and procurement methods as applied by ARB's procurement contractors and (2) to evaluate the vehicle selection methodology applied by the U.S. Environmental Protection Agency (EPA) for in-use vehicle compliance testing, for its relevance to ARB's needs.

To accomplish these purposes, Pacific Environmental Services, Inc. (PES) reviewed the selection procedures used by several contractors who have done vehicle procurement for ARB or EPA. PES reviewed documentation which was provided by these contractors and also interviewed in person or by telephone, contractor personnel who had worked on procurement projects.

PES also reviewed extensive documentation provided by ARB and EPA as well as by procurement contractors, to identify significant issues, primarily in relation to the adequacy of vehicle selection procedures to support ARB needs for determination of emission factors. Since it appeared that procedures applied by the contractor were in large part mandated by ARB or significantly constrained by ARB requirements, PES also undertook some exploratory investigation of the statistical status of data derived from the VSP, to determine whether the basic statistical design might merit improvement.

A review of contractors' methodologies indicated that contractors seldom complied effectively with ARB requirements for proportional sampling of vehicles from various socioeconomic levels of ownership. Although reasonable compliance was obtained in at least one instance, emphasis on this aspect seemed to decline in later studies, perhaps in the interest of overall cost savings. A more recently evolved system depends on maintaining separate index files for various types of vehicles, and keeping an adequate backlog of committed candidate vehicles in each file so that a vehicle of any given type can be quickly obtained when called for.

A review of the actual map locations of participating owners in six of the annual procurement programs was done, and the spatial distributors were compared with population distributions based on demographic data from the Southern California Association of Governments. In relation to a hypothetical proportionate sampling scheme covering the entire South Coast

Air Basin, areas near the contractors' facilities have quite regularly been oversampled. Heavily populated areas outside Los Angeles County have usually been oversampled, while those inside Los Angeles County have been undersampled. Thus, although ARB requirements regarding the numbers of vehicles from different counties have been met, the geographical distribution within each county has been quite spotty. The importance of this deficiency is not clear.

Emission data from one of the testing series, LDVSP5, were reviewed to determine some of their statistical parameters. The results support a conclusion that accuracy of ±10 percent in overall emissions, as estimated from the procured sample of vehicles, was probably within a 95% confidence interval for NOx emissions, but not for HC or CO emissions.

A preliminary review of possible sources of bias (i.e., systematic error or error due to unrepresentative sampling) in estimation of emissions based on VSP testing data yielded the following categories (non-exclusive): (1) voluntary participation; (2) policy exclusions; (3) erroneous weighting; (4) overstratification; (5) inappropriate sample quotas; (6) geographic concentration; (7) socioeconomic concentration; (8) inaccurate assumptions in applying test data.

PES elicited the views of procurement contractors about difficulties, shortcomings, and possible improvements in the methods and practices of vehicle procurement. Generally, all contractors felt that ARB funding was insufficient to support important improvements in the procurement system. Specific problems mentioned as having appreciable impact on the system were

(1) problems of security in high-crime areas
(2) problems of communication in minority areas
(3) need for improving quality of communication and contact material
(4) lag in obtaining registration records from public information firms
(5) need of responding quickly to ARB requests for special procurements
(6) insufficiency of available incentives

A statistical treatment of an emissions data base derived from LDVSP5 showed that CO and HC emissions differed significantly for vehicles of different model years, but that effects of other factors were less important. Among the factors not showing significant influence on emissions (within this rather restricted data base) were county of origin, owner income, manufacturer and engine size. There was strong correlation between HC emissions and CO emissions, but little correlation of NOx with either of these.*

Model-year effects differentiated principally between older and newer vehicles, with older vehicles (up to 1973 or 1974) showing higher average HC and CO emissions, significantly wider range of HC and CO emissions, and significantly more outliers in emissions of the same pollutants. PES concludes that these older vehicles contribute, in excess of their proportion in the population, both to high emissions (CO and HC) and to impairment of accuracy in estimating those emissions.

PES reviewed EPA procurement practices for recall testing, and compared them with ARB's practices. The most noteworthy difference in procedure is that EPA provides its procurement contractors with lists of candidate vehicles, which the contractors are required to attempt to obtain, whereas ARB specifies quotas for many types of vehicles but does no preliminary sorting of registration data. The EPA list is said to result from a random selection process applied by EPA.

Other differences are:

---

*Caveat: The statistical testing to illustrate possible bias in sampling, discussed in Sections 3.3 and 3.4, necessarily relied on available data from vehicles used in the surveillance program. However, because of the possibility of substantial bias in the data base due to low response rates in vehicle procurement, the surveillance data base itself may not adequately represent the general vehicle population. Therefore, the results given in these sections apply strictly only to the surveillance data base; due caution should be observed in any attempts to generalize from these results to the characteristics of the vehicle population as a whole.

(1) EPA requires contractors to make specific further attempts to contact non-responding vehicle owners, while ARB does not;

(2) EPA requires sequential action on candidate vehicles, while ARB does not;

(3) EPA appears to have stricter requirements for acceptance of vehicles and, thus, to require a stricter screening of vehicles by their contractors, than ARB.

Considering the information developed in this study, PES has reached certain general conclusions.

First, there is little scope for improvement of the In-Use Vehicle Surveillance Program insofar as contractor initiative is to be involved. The principal constraints on adequacy of sampling appear to lie in overspecification of quotas by ARB and in procedures which constitute logical responses to ARB cost controls.

Second, the goals of random sampling and of random stratified sampling are not currently feasible goals for the VSP. The current system, which is best described as quota sampling, is subject to a number of sources of bias, most of which cannot be readily eliminated or even evaluated. However, there are a number of possible improvements that ARB could institute, which would minimize bias from some of these sources. Further, by analyzing information which is already in the data base, ARB can design its future procurement programs to improve cost-effectiveness and to improve the accuracy of emission factors derived from the test results.

The following principal options for sampling design are discussed in this report: (1) Continuous Random Sampling; (2) Stratified Sampling; (3) Incremental Quota Sampling; (4) Intensive Studies of Specific Subpopulations.

1.2 RECOMMENDATIONS FOR IMPROVEMENTS IN FUTURE VEHICLE SURVEILLANCE PROGRAMS

PES has reviewed documents of past VSP programs, interviewed vehicle procurement experts of both ARB and EPA contractors, analyzed emission test data and background data of tested vehicles and owners, and evaluated both ARB's and EPA's vehicles procurement programs. Based on the study described above, PES has reached the following recommendations for improvements in future ARB VSP programs:

(1) ARB should provide the contractor a written guideline of how the sample selection method proposed by the contractor (or developed by ARB) is to be implemented in the actual vehicle procurement operation. Specifically, ARB should ask the contractor to submit lists of initially selected candidate vehicle owners and the basis of such selections, outcomes of mail solicitation and follow-up telephone contact, causes for non-participation, and a summary table for final capture rates in different strata. An audit of the contractor's operating practices would also be desirable.

(2) ARB should develop a randomized owner list and provide it to the procurement contractor; or, alternatively, arrange access for the contractor to the most recent DMV records. Polk-provided registration records are, at minimum, 6 to 8 months old; this causes loss of access to many candidate vehicle owners who move or trade their cars during that period.

(3) ARB should reduce the extent of sample stratification. The number of strata presently used is about 200 and it has obviously exceeded the point of diminishing return for the benefit of using stratified random sampling. The number of strata for manufacturer can be reduced to 5 or 6 and that for engine displacement to 2 or 3. The number of strata for model years can be reduced to half of the number presently used, or less, by limiting vehicles to those with odd-numbered model years in an odd-numbered calendar year and those with even-numbered model years in an even-numbered calendar year, or by grouping several model years with similar emission rates into a combined category.

(4) ARB should determine the proportion of vehicles in each model year stratum more scientifically. The quota (or sub-sample size) assigned to each stratum should be proportioned to the relative

contribution of vehicles in that stratum to either the total vehicle population or the total vehicle emissions. The currently used quotas for different strata are not proportional either to the vehicle population or to the vehicle emissions. Should the latter criterion be used, considerably greater emphasis must be placed upon older vehicles because of their higher emission rates.

(5) Some adjustment for quotas assigned to different strata should be made annually to improve both the representativeness of the sample and the accuracy of emission parameters to be determined from the sample. To make such adjustment rational, annual statistical review of all existing data bases of emissions, vehicle populations, manufacturing, and use factors is recommended.

(6) Exemption from emission testing in the mandatory I/M program should be used as an additional incentive to encourage vehicle owners to participate in VSP programs after the start of the I/M program in March, 1984. The current flexible-incentive approach should be continued because flexibility can often be more effective than increased dollar value of an incentive.

(7) ARB should develop improved solicitation letters and other contact materials and also urge the contractor to do so because the quality of contact material can greatly affect positive response rates to the solicitation.

2.0 EVALUATION OF METHODOLOGY USED FOR THE SELECTION OF VEHICLES FOR ARB'S "IN-USE" VEHICLE SURVEILLANCE PROGRAM

2.1 VEHICLE SELECTION PROCEDURES

2.1.1 PROCEDURES DESCRIBED BY ARB PROCUREMENT CONTRACTORS

In order to review the selection procedures, PES first reviewed the proposals submitted by various contracting firms in response to solicitations by ARB, then interviewed representatives of those firms for any needed clarifications. PES received, in all, 13 procurement proposals and 8 final reports of vehicle testing programs. Firms contacted included the following:

Automotive Environmental Systems, Inc.
Automotive Control Engineering
EG&G Automotive Research (in Virginia)
Environmental Resources Management
Olson Laboratories, Inc.*
Systems Control, Inc.

ARB required each procurement contractor to endeavor to obtain vehicles corresponding to a specific "fleet composition" which identified vehicles by class (i.e., light duty passenger car, light duty truck, medium duty vehicle, heavy duty truck of a particular weight class, or motorcycle), by model year, by manufacturer, by engine displacement and, sometimes, by additional factors such as control technology.

The geographical area covered was, in most instances, the South Coast Air Basin (SCAB). The SCAB, in early studies, comprised 5 counties: Los Angeles, Orange, Riverside, San Bernardino, and Ventura; in later studies, Ventura County was not included. The northern part of Los Angeles County, within the South East Desert Air Basin, was excluded.

ARB, in most of the relevant solicitations, specified that certain percentages of the procured vehicle fleet should be obtained from each of the counties in the SCAB; for example, for the Fourth Surveillance Series solicitation, dated November 15, 1978, the specified geographical distribution was as follows:

---

*Subsequently a division of Systems Control, Inc.

Los Angeles County	70.2%
Orange County	19.6%
Riverside County	4.1%
San Bernardino County	6.1%

In addition, ARB specified that the procurement procedure should be designed to achieve an appropriate representation of various socioeconomic groups within the SCAB.

Responding firms adopted various strategies to meet ARB's requirements. In 1976, one proposal offered a plan to select 200 1975 vehicles according to a matrix of manufacturer, model, and engine size, using a stratified sample scheme based on California sales data. Statistical data were to be obtained from Automotive News and similar publications and from registration listings supplied by public information service firms such as R.H. Donnelly Corporation. However, it was expected that many or most of the required vehicles would be obtained from owners listed for previous EPA vehicle testing programs.

In a later proposal (1977) for procurement of 400 light duty vehicles, the same firm proposed to solicit vehicles by a mailing list to be obtained from Donnelly Corp. by a random selection process, then use EPA's owner lists and direct solicitation only as needed to fill in the matrix.

A subsequent proposal (for procurement of 300 vehicles for the fourth surveillance series) offered a choice of two approaches. One of these was basically similar to the approach just described. The second provided for a more closely controlled distribution of vehicle owners to be solicited. From automobile registration data, lists of specified vehicles within selected ZIPcode areas were to be generated, from which the required number of owners' addresses would be randomly selected. In order to minimize selection cost and "provide a realistic balance with a rough cross-section of socioeconomic areas," the firm proposed to choose 10 ZIPcode areas from western Los Angeles County, 3 from Orange County, 3 from northern Riverside County and 3 from northwestern San Bernardino County. It appears that the achievement of a realistic balance would depend upon the contractor's judgment in selecting appropriate ZIPcode areas for solicitation.

A somewhat more explicit approach to stratified sampling is described in the final test report of the Medium Duty Vehicle Surveillance Program, Series 3 (MDVSP3; see ARB 1981a), for which ACE was selected to supply 90 vehicles. The technique was based on a two-stage approach. The first level of stratification was according to ZIPcode, where 32 "subpopulations" in the five counties were established. For the second stage, these subpopulations were further stratified in an effort to achieve economic representation based on commercial factors. Specified percentages of the sample were taken from areas with high, medium and low commercial occupancy in Los Angeles County, while the subpopulations in the other counties were not further stratified. After the subpopulations were determined, the total sample was drawn "as evenly as possible" from the 32 ZIPcode sampling regions, but randomly within each.

Stratified sampling was also the basis of contracts carried out by ERM. As described in the final testing report of the Light-Duty Vehicle Surveillance Test Program, 2nd Series (LDVSP2; see ARB 1980b) the total population of vehicles on the source list is divided into strata according to model year, manufacturer and model, county, and socioeconomic level of the area the vehicle is garaged in. Primary Sampling Units (PSU's) are selected, based on ZIPcodes; Secondary Sampling Units (SSU's) are the individual vehicles selected from the PSU's.

Stratification by socioeconomic status (SES) areas was limited to Los Angeles County. All census tracts with median family income below $10,000 per year were identified and related to ZIPcode areas by manual comparison of maps. ZIPcode areas containing a high proportion of the low-SES census tracts were designated as PSU's for the low-SES subsample. The analogous procedure, with a lower cutpoint of $20,000 per year, was used to designate PSU's for the high-SES subsample. Quotas of six PSU's to be selected from each of the two categories were established. ZIPcode areas containing 10,000 or more vehicles were treated as individual PSU's; other ZIPcode areas were combined with nearby ZIPcode areas to create PSU's with 10,000 or more vehicles.

To maintain approximately proportionate sampling, the income cutpoints were adjusted. The PSU list for high SES was extended down to about

$15,000 per year (median family income) to increase the number of vehicles listed to about 20 percent. The upper end of the low SES list was reduced (below $10,000) enough to decrease the number of vehicles listed to about 20 percent of the total. All PSU's not designated as high or low SES areas were retained on the middle SES list, for which a quota of 15 PSU's was assigned.

Based on an estimate that there were about 4 million eligible vehicles in the SCAB, ERM decided to use 40 PSU's throughout the SCAB; thus, selecting 10 vehicles from each selected PSU would provide the desired fleet of 400 vehicles. A randomization procedure was designed for choosing the necessary number of PSU's in each county and SES category from those satisfying the stratification requirements; in this way, PSU's containing about one tenth of the vehicle population were actually selected to yield the SSU's (the delivered vehicles) for the entire sample. The lists of PSU's selected were designated by the ZIPcodes they included and were submitted to the R. L. Polk Company, who provided a printout listing the eligible vehicles in those areas. From these lists, needed vehicles were randomly selected.

In later contracts, ERM modified this system, developing a "computerized vehicle profile" program to simplify vehicle selection. ERM's proposal for procurement of vehicles for the Light Duty Vehicle Surveillance program, Series 5 (LDVSP5; see ARB 1982a), as quoted in the final report of that program, describes the revision as follows: "Census tract data has previously been used as the most realistic indicator of economic level (status) . . . Previously, the ERM data design group has computerized the latest census tract data and appropriately upgraded it based on DOL upgrade grids. Environmental Resource Management will evaluate and update its current 'computerized vehicle profile' program. Since registration records show what the actual distribution of light duty vehicles is in the SCAB counties selected, results from testing this random sample can be used for inferring conclusions about all light duty passenger cars/trucks in the areas . . . The selection criteria will provide a statistical sampling that will be as representative as possible of a realistic cross-section of socioeconomic conditions from the four counties selected..."

This description was much less specific than the one given previously and cited above. It seemed to imply that ERM intended to de-emphasize the role of census tract information in defining PSU's. Subsequent discussion between PES and ERM personnel confirmed this inference.

Principally for the purpose of accelerating access to newer vehicles and thereby improving readiness to respond quickly to ARB requests for delivery of vehicles, ERM instituted a new sampling plan for new cars, based on Polk Company listings taken from annual DMV data. In this system, ERM obtains annual listings of all new vehicles recorded in each of the four couties, sorted according to (successively): vehicle type, manufacturer, engine type, city and ZIPcode. In order to create a "pool" of available vehicles of each type specified by ARB, a page of the corresponding printout is randomly selected. Each page carries about 50 individual vehicles records, and ERM solicits owner participation for about 30 of the corresponding vehicles. Those responding are entered into the appropriate pool, from which vehicles are randomly selected when required for testing.

This system relies on randomization of owner selection (within vehicle subpopulations) to generate appropriate geographic and socioeconomic distributions as desired by ARB.

The methodologies described in these various proposals and reports seem to differ, principally, in two respects: (1) the degree to which they rely on previously generated lists of eligible vehicles, and (2) the degree of emphasis placed on systematic stratification and randomization in selecting vehicles for testing. There was also a marked variation in the attention devoted to ARB's admonition regarding socioeconomic considerations, which, in a 1982 solicitation (ARB 1982a) for procurement of 400 vehicles, was phrased as follows:

"The effect of the economic status of the owner upon vehicle maintenance and therefore emissions is of concern to the ARB. Therefore, the contractor will be required to exercise some method of introducing a socioeconomic factor into the sample design. Geographic sub-areas are acceptable to satisfy this requirement; however, alternate approaches will be evaluated. A description of the methodology for socioeconomic sampling must be included in the proposal."

The question of possible introduction of bias into the estimation of motor vehicle emissions from ARB's test results, due to such variations in procurement procedures, is discussed in Section 2.2.

2.1.2 PROCEDURES SPECIFIED BY EPA FOR SIMILAR PROGRAMS

An indication of procedures applied by EPA in an in-use vehicle surveillance program with objectives similar to those of ARB's In-Use Vehicle Surveillance Program is found in a proposal for "A Study of Emissions from Light-Duty Vehicles in Five Cities (FY 80)", prepared in March, 1980. The proposal presented a plan for obtaining and testing 300 vehicles in the Los Angeles area. A subcontractor would use vehicle registration lists from R. L. Polk Company to provide names and addresses of potential participants. Owners of vehicles meeting the requirements specified by EPA would be contacted by mail and telephone; a "random listing of vehicle owners approximately 50 times larger than the required number of vehicles" would be used. No plan for stratified sampling was offered. The geographical area described was "within 25 kilometers of Los Angeles city limits."

PES also obtained excerpts from an EPA draft "Scope of Work" (EPA 1981) for recall testing programs, which included an attachment describing a procedure for random selection of vehicles from a prepared list of eligible vehicles. This document, however, does not address the problem of formulating the list of eligible vehicles to obtain a representative sample for any particular geographical area. Further discussion of recall programs is offered in Chapter 4.0.

2.2 <u>POSSIBLE SOURCES OF BIAS IN PROCUREMENT PROCEDURES</u>

2.2.1 INTRODUCTION

The purpose of the In-Use Vehicle Surveillance Program as expressed in various solicitations, statements and reports by ARB is, essentially, to accumulate a data base from which accurate estimates of in-use vehicle emissions can be formulated. These emissions data are used for several regulatory activities including formulation of certification standards, preparation of emissions inventories for the State Implementation Plan, development of inspection-and-maintenance programs, and enforcement of emissions warranty requirements. Additionally, the program is expected to provide useful information regarding vehicle maintenance practices and the frequency of tampering with emission control components.

For all these uses it is clearly desirable to have vehicle sample populations which accurately represent the California vehicle population, or appropriate segments of the California vehicle population. However, the uses for which accurate representation is most critically needed are the preparation of emissions inventories and the enforcement of emissions warranty requirements (i.e., vehicle recall programs).

Although optimum sampling strategies might differ slightly depending on which end-use of the data is of primary importance, the development of emissions inventories is the use that requires the largest numbers of vehicles to be provided. Typically, ARB has designed vehicle-procurement programs to utilize the same methodology for both purposes, the only difference being that larger numbers of particular vehicle models would be called for when enforcement testing was to be done, and that more specific criteria for vehicle conditions were applied, modeled after EPA compliance criteria. For example, ARB's solicitation for procurement of 400 vehicles for Light Duty Surveillance Testing, 7th series (ARB 1982a) specifies: "Sixty of the 400 vehicles will be requested for a special enforcement investigation. Three groups of 20 vehicles each, which are suspect problem emitters, will be subjected to special testing. These vehicles will be designated at a later date."

In Chapter 4.0, EPA's procurement methodology for compliance testing is examined in detail and implications for possible improvements in ARB's procurement methodology for recall testing are explored. In the remainder of this section, attention is focused on vehicle sampling methodology as it relates to the general in-use surveillance, excluding vehicles obtained for special enforcement investigations.

2.2.2 APPLICATION OF SURVEILLANCE DATA TO EMISSIONS INVENTORIES

In principle, if an estimate of emissions is needed for an area (like SCAB) with a large vehicle population, an estimate of any desired accuracy can be obtained by testing the emissions from a large enough random sample of vehicles from the population. However, with a population having great diversity of individual emission values, the number of vehicles required for testing, in order to yield emissions estimates of any reasonable accuracy, is prohibitive because of expense and administrative difficulties.

In such cases, stratified sampling is the customary expedient for increasing the efficiency and thereby lowering the cost of obtaining the needed information. This method is appropriate whenever the population to be sampled can be divided into subpopulations whose variability is substantially less than that of the entire population. ARB has chosen to apply the principle of stratified sampling to the estimation of vehicle emissions, using vehicle class, model year, manufacturer, engine displacement, location of owner's residence, and owner's socioeconomic status as parameters to define the subpopulations. Procurement contractors are therefore directed to procure vehicles in accordance with such a stratification scheme.

Sources of error in a general random sampling approach mainly spring from sample variability, which may be very high for samples of manageable size. For example, using statistical estimates based on carbon monoxide emissions found in LDVSP 5, there is a chance of 20 percent or more that, for 339 light duty vehicles, an emission factor based on the whole sample would deviate from the true emission factor for the population by more than 10 percent. To reduce that chance to a statistically acceptable risk of 5 percent (i.e., one chance in twenty) would require a sample of well over 1,000 vehicles.

There are also other possible sources of error in estimating population emissions from a general random sampling approach. It is difficult to achieve a truly representative random sample in practice, because owners of some types of vehicles may be less willing to participate than others; for example, owners of expensive new vehicles may be especially reluctant to furnish them. Restrictions on the acceptability of vehicles for testing also may tend to create discrepancies between the sample composition and that of the entire vehicle population; thus, rental cars, fleet vehicles, and law enforcement vehicles are excluded from the sample but not from the general population. A further complication is that the testing yields not emissions but emission factors for each vehicle in terms of amounts of pollutant emitted per mile, so that estimation of total emissions in any given period requires the application of a usage factor (i.e., vehicle miles travelled, VMT.)

Stratified sampling offers the possibility of reducing errors caused by some of these factors. In applying a system of stratified sampling to the estimation of emissions, the standard approach would be to determine emission factors for various subpopulations by random sampling of the subpopulations, then determine total emissions by summing the emissions attributable to the various subpopulations. However, additional complications are introduced by the need to weight the emission factors by appropriate use factors and to adjust for differences in emission rates between testing conditions and road conditions.

Because of the need to estimate use factors to be applied to various segments of the vehicle population, ARB adopted a weighting scheme (ARB 1980a) promulgated by EPA, which amounted to an assumed fixed schedule of average annual miles travelled for every vehicle, depending only on its age. This reflected the assumption that vehicles are used less as they grow older, and utilized estimates of this effect, which had been published by EPA. For convenience in calculation, ARB further assumes (ARB 1980a, 1981d) that emission factors for all vehicles of a given class and model year are independent of manufacturer and engine displacement and are related in a linear fashion to accumulated mileage. Thus, the stratification on manufacturer, engine displacement, geographical location and socioeconomic status, which ARB requires of the procurement contractor, is not reflected in the calculation of emission inventories. It is intended, essentially, to assure systematic (rather than random) representation of the recognized subpopulations in the overall sample of vehicles to be tested.

It is evident from the foregoing discussion that the stratification system required of procurement contractors differs from the stratification system used in calculations to estimate on-the-road emissions. Any inaccurate assumptions which are embodied in the estimation method may, of course, cause errors in the emissions estimates. However, such biases cannot be corrected by modifying procurement methodology and are not the responsibility of the procurement contractor. Further discussion of such errors will be found in Chapter 5.0.

It is the procurement contractor's responsibility to minimize bias in the systematic representation of the vehicle subpopulations defined by manufacturer, model year, engine displacement. In these categories, ARB designates numbers of vehicles expected; an example is given in Table 2-1, showing numbers of vehicles specified (and those received) in ARB's solicitation for LDVSP 5 (ARB 1981e). Since in most of the categories listed only one or two vehicles are expected, the procurement contractor has little scope for further stratifying the corresponding subpopulations. Efforts to minimize bias have therefore taken the form of assuring that, over the whole sample, vehicle locations are not unduly concentrated in one or a few areas and that participants are not predominantly from one end or segment of the socioeconomic scale. In other words, they seek to achieve, within the whole sample, distributions of participants' locations and socioeconomic status which conform (within reason) to corresponding distributions in the vehicle population at large.

The endeavor to deliver a sample having representative distributions of geography and socioeconomic status is predicated on the concept that the actual vehicle population might exhibit significant non-uniformity in relation to these parameters; in other words, that otherwise equivalent vehicles could be expected to have different emission values if they came from different locations, or from different socioeconomic strata. Little information exists to substantiate this possibility, which is known in statistical terminology as "interaction," but it is a conceivable source of

Table 2-1. LDVSP 5 FLEET COMPOSITION

PASSENGER VEHICLES

1980 Model Year	# of Cars (58)	Vehicles Received	# of Cars (50)	Vehicles Requested**
AMC	1	1 (258 CID)	1	
Chry	6	2 (105 CID), 2 (225 CID) 2 (318 CID)	5	1 (225 CID)
Ford	15	1 (98 CID), 6 (140 CID) 3 (200 CID), 2 (255 CID), 2 (302 CID) 1 (351 CID)	12	2 (140 CID), 3 (302 CID)
G.M.	19	1 (98 CID), 4 (151 CID) 3 (173 CID), 6 (231 CID), 3 (305 CID) 1 (350 CID), 1 (368 CID)	16	1 (98 CID), 3 (173 CID) 3 (231 CID), 1 (260 CID), 7 (305 CID), 1 (368 CID)
Datsun	4	1 (75 CID), 2 (119 CID) 1 (168 CID)	4	2 (85 CID), 2(119 CID)
Toyota	3	1 (108 CID), 1 (134 CID) 1 (156 CID)	2	1 (97 CID)
V.W.	2	2 (97 CID)	1	1 (89 CID) or 1 (97 CID)
Honda	2	2 (107 CID)	2	2 (91/98 CID)
Other Imports	6*		7*	

The vehicles requested were selected from the manufacturer's anticipated sales figures. The actual rates for this model year varied widely from the anticipated figures due to the escalating price of gasoline.

* Various CID randomly selected.

** Where the number of vehicles received is either different or omitted from that requested, the original selection is repeated in the requested column.

Table 2-1. LDVSP 5 FLEET COMPOSITION (CONTINUED)

1979 Model Year	# of Cars (43)	Vehicles Received	# of Car (41)	Vehicles Requested**
AMC	1	1 (258 CID)	1	
Chry	5	1 (105 CID), 1 (225 CID), 2 (318 CID), 1 (360 CID)	4	2 (318/360 CID)
Ford	8	3 (140 CID), 1 (171 CID), 1 (200 CID) 1 (250 CID), 1 (302 CID), 1 (351 CID)	10	3 (351/400 CID)
G.M.	13	1 (98 CID), 2 (231 CID), 1 (250 CID), 1 (260 CID), 2 (305 CID), 3 (350 CID) 2 (403 CID), 1 (425 CID)	13	1 (85/98 CID), 2 (231 CID), 4 (305 CID), 3 (350 CID), 1 (403 CID), 1 (425 CID)
Datsun	2	1 (85 CID), 1 (168 CID)	2	1 (85 CID),1 (119 CID)
Toyota	2	1 (97 CID), 1 (134 CID)	2	
V.W.	3	2 (89 CID), 1 (97 CID	2	
Honda	2	1 (91 CID), 1 (107 CID)	2	2 (91/98 CID)
Other Imports	7*		5*	

* Various CID randomly selected.

Table 2-1. LDVSP 5 FLEET COMPOSITION (CONTINUED)

1978 Model Year	# of Cars (33)	Vehicles Received	# of Cars (34)	Vehicles Requested**
AMC	1	1 (258 CID)	1	
Chry	4	1 (225 CID), 3 (318 CID)	4	
Ford	8	2 (140 CID), 1 (171 CID), 1 (200 CID), 2 (302 CID), 2 (351 CID)	8	
G.M.	9	1 (98 CID), 2 (231 CID), 1 (250 CID), 2 (305 CID), 1 (350 CID), 1 (403 CID), 1 (425 CID)	9	1 (231 CID), 2 (350 CID)
Datsun	2	1 (85 CID), 1 (119 CID)	2	
Toyota	2	1 (97 CID), 1 (134 CID)	2	1 (97 CID), 1 (114 CID)
V.W.	2	1 (89 CID), 1 (97 CID)	2	
Honda	2	1 (91 CID), 1 (98 CID)	2	
Other Imports	3*		4*	

1977 Model Year	# of Cars (33)	Vehicles Received	# of Cars (30)	Vehicles Requested**
AMC	1	1 (258 CID)	1	
Chry	3	2 (225 CID), 1 (360 CID)	2	1 (225 CID), 1 (360 CID)
Ford	7	1 (140 CID), 2 (171 CID) 1 (250 CID) 1 (302 CID), 1 (351 CID), 1 (400 CID)	7	1 (171 CID), 3 (351/400 CID)
G.M.	12	1 (140 CID), 1 (231 CID) 3 (305 CID), 4 (350 CID), 2 (403 CID) 1 (425 CID)	11	1 (140 CID), 1 (231 CID) 1 (250 CID), 2 (305 CID) 4 (350 CID), 1 (403 CID) 1 (425 CID)
Datsun	2	1 (85 CID), 1 (168 CID)	2	
Toyota	2	1 (97 CID), 1 (134 CID)	2	
V.W.	1	1 (97 CID)	1	
Honda	1	1 (98 CID)	1	
Other Imports	4*		3*	

* Various CID Randomly Selected

Table 2-1. LDVSP 5 FLEET COMPOSITION (CONTINUED)

1976 Model Year	# of Cars (25)	Vehicles Received	# of Cars (25)	Vehicles Requested**
AMC	1	1 (258 CID)	1	
Chry	2	1 (318 CID), 1 (360 CID)	2	
Ford	6	2 (140 CID), 1 (171 CID) 1 (250 CID), 1 (302 CID), 1 (460 CID)	6	1 (140 CID), 1 (351/400 CID)
G.M.	8	1 (140 CID), 2 (231 CID) 1 (250 CID), 3 (350 CID), 1 (500 CID)	8	1 (140 CID), 1 (231 CID), 1 (250 CID), 1 (350 CID), 3 (350/400 CID), 1 (500 CID)
Datsun	2	2 (85 CID)	2	
Toyota	1	1 (97 CID)	1	
V.W.	1	1 (97 CID)	1	
Honda	2	2 (91 CID)	2	
Other Imports	2*		3*	

1975 Model Year	# of Cars (20)	Vehicles Received	# of Cars (20)	Vehicles Requested**
AMC	1	1 (318 CID)	1	1 (232/258 CID)
Chry	2	1 (225 CID), 1 (360 CID)	2	
Ford	5	1 (140 CID), 1 (250 CID), 1 (302 CID), 1 (351 CID), 1 (460 CID)	5	
GM	7	1 (140 CID), 1 (250 CID), 4 (350 CID), 1 (500 CID)	6	3 (350/400 CID)
Datsun	1	1 (85 CID)	1	
Toyota			1	1 (97 CID)
V.W.	1	1 (90 CID)	1	
Honda	1	1 (90 CID)	1	
Other Imports	2*		2*	

* Various CID Randomly Selected.

Table 2-1. LDVSP 5 FLEET COMPOSITION (CONTINUED)

LIGHT-DUTY TRUCKS

1980 Model Year	# of Cars (10)	Vehicles Received	# of Cars (10)	Vehicles Requested**
Chry			1	1 (318/360 CID)
Ford	2	1 (300 CID), 1 (302 CID)	1	1 (351 CID)
G.M.	2	2 (250 CID)	2	1 (250 CID), 1 (350 CID)
Datsun	2	2 (119 CID)	2	
Toyota	1	1 (134 CID)	1	
V.W.	1	1 (120 CID)	1	
Other Imports	2*		2*	

1979 Model Year	# of Trucks (11)	Vehicles Received	# of Trucks (10)	Vehicles Requested**
Chry	1	1 (318 CID)	1	
Ford	3	1 (302 CID), 2 (351 CID)	1	1 (350 CID)
G.M.	1	1 (250 CID)	2	1 (350 CID)
Datsun	1	1 (119 CID)	1	
Toyota	1	1 (134 CID)	1	
V.W.	1	1 (120 CID)	1	
Other Imports	3*		3*	

1978 Model Year	# of Trucks (11)	Vehicles Received	# of Trucks (10)	Vehicles Requested**
Ford	1	1 (351 CID)	1	
G.M.	3	2 (250 CID), 1 (350 CID)	2	1 (250 CID)
Datsun	1	1 (119 CID)	1	
Toyota	2	2 (134 CID)	2	
V.W.	1	1 (120 CID)	1	
Other Imports	3*		3*	

* Various CID Randomly Selected.

Table 2-1. LDVSP 5 FLEET COMPOSITION (CONTINUED)

1977 Model Year	# of Trucks (10)	Vehicles Received	# of Trucks (10)	Vehicles Requested**
Chry	1	1 (318 CID)	1	
Ford	1	1 (351 CID)	1	1 (302 CID)
G.M.			1	1 (350 CID)
Datsun	2	2 (119 CID)	2	
Toyota	2	2 (134 CID)	2	
V.W.	1	1 (120 CID)	1	
Other Imports	3*		2*	

1976 Model Year	# of Trucks (13)	Vehicles Received	# of Trucks (10)	Vehicles Requested**
Chry	1	1 (318 CID)	1	
Ford	3	2 (360 CID), 1 (351 CID)	1	1 (360/390 CID)
G.M.	4	1 (250 CID), 3 (350 CID)	3	2 (250 CID), 1 (350 CID)
Datsun	1	1 (119 CID)	1	
Toyota	1	1 (134 CID)	1	
V.W.	1	1 (120 CID)	1	
Other Imports	2*		2*	

* Various CID Randomly Selected.

Table 2-1. LDVSP 5 FLEET COMPOSITION (CONTINUED)

1975 Model Year	# of Trucks (9)	Vehicles Received	# of Trucks (10)	Vehicles Requested**
Chry	1	1 (318 CID)	1	
Ford	1	1 (360 CID)	1	1 (360/390 CID)
G.M.	3	1 (250 CID), 2 (350 CID)	3	
Datsun	1	1 (119 CID)	1	
Toyota	1	1 (134 CID)	1	
V.W.	1	1 (109 CID)	1	
Other Imports	1*		2*	

LIGHT-DUTY VEHICLES
(Passenger Vehicles and Light Duty Trucks)

1974 Model Year	# of Vehicles (14)	Vehicles Received	# of Vehicles (20)	Vehicles Requested**
Chry	1	1 (318 CID)	2	1 (225 CID)
Ford	5	1 (140 CID), 1 (171 CID), 1 (302 CID), 1 (351 CID), 1 (400 CID)	5	
G.M.	6	1 (116 CID), 1 (250 CID), 2 (350 CID), 1 (455 CID), 1 (472 CID)	7	1 (140 CID), 1 (250 CID), 3 (350/400 CID), 1 (455 C[illegible] 1 (500 CID)
Datsun	1	1 (108 CID)	1	
Toyota	1	1 (120 CID)	1	
V.W.			1	1 (Any)
Honda			1	1 (75 CID)
Other Imports			2*	

* Various CID Randomly Selected.

Table 2-1. LDVSP 5 FLEET COMPOSITION (CONTINUED)

1973 Model Year	# of Vehicles (18)	Vehicles Received	# of Vehicles (19)	Vehicles Requested**
AMC	1	1 (232 CID)	1	
Chry	2	1 (318 CID), 1 (400 CID)	2	
Ford	2	1 (302 CID), 1 (429 CID)	4	1 (122 CID), 1 (351 CID)
G.M.	9	2 (140 CID), 1 (250 CID) 4 (350 CID), 1 (455 CID), 1 (472 CID)	8	1 (140 CID)
Toyota	1	1 (120 CID)	1	
V.W.	1	1 (97 CID)	1	
Other Imports	2*		2*	

1972 Model Year	# of Vehicles (17)	Vehicles Received	# of Vehicles (19)	Vehicles Requested**
AMC	1	1 (258 CID)	1	
Chry	2	1 (225 CID), 1 (360 CID)	2	
Ford	3	1 (110 CID), 1 (122 CID), 1 (351 CID)	4	1 (122 CID), 1 (302 CID) 1 (351 CID), 1 (400 CID)
G.M.	7	1 (250 CID), 4 (350 CID) 1 (400 CID), 1 (455 CID)	8	1 (140 CID)
Datsun	1	1 (97 CID)	1	
V.W.	1	1 (97 CID)	1	
Other Imports	2*		2*	

* Various CID Randomly Selected.

Table 2-1. LDVSP 5 FLEET COMPOSITION (CONCLUDED)

1971 Model Year	# of Vehicles (14)	Vehicles Received	# of Vehicles (19)	Vehicles Requested**
Chry	1	1 (318 CID)	2	1 (225 CID)
Ford	4	1 (122 CID), 1 (302 CID), 1 (400 CID), 1 (429 CID)	4	
G.M.	5	1 (250 CID), 1 (307 CID), 3 (350 CID)	8	1 (140 CID), 1 (400 CID), 1 (472 CID)
Datsun	1	1 (97 CID)	1	
Toyota	1	1 (113 CID)	1	1 (120 CID)
V.W.	1	1 (103 CID)	1	
Other Imports	1*		2*	

* Various CID Randomly Selected

bias in sampling and can be appropriately minimized by the measures proposed. At the same time, data already existing in the in-use vehicle surveillance database should enable ARB to evaluate whether the suggested effects are in fact significant enough to deserve continued attention. Further discussion on this subject will be found in Chapter 5.0.

Another potential source of bias lies in the "self-selection" of the sample, which is inherent in the nature of any program where participation is voluntary. Even if the list of candidate vehicles is in all respects accurately representative of the vehicle population, there remains a reasonable possibility that the sample of vehicles actually procured will not accurately reflect the distribution of the vehicles in the candidate list. This occurs because the candidate list covers two distinct subpopulations: one set of vehicles whose owners are amenable to participation, and another set which will not be made available. It is commonly supposed that the differences in the characteristics of these subpopulations are small, if not negligible. However, it is entirely possible that the vehicles obtained may, on the whole, have been better maintained than average comparable vehicles, or be less likely to exhibit signs of alteration, misfueling, or tampering with control systems, or in other subtle ways be less representative of the general population.

It is especially difficult to evaluate this possibility, because the actual sample of vehicles available for testing is ordinarily only a small fraction of those eligible. There is no ready source of information on the condition of the remaining large majority of vehicles and, thus, there is no way of determining whether the discrepancy, if it exists, is small or large. However, since the actual availability of vehicles may depend on incentive offered for their owners' participation, it may be possible to devise experimental approaches to evaluating the importance of the self-selection factor. Further discussion on this subject will be found in Chapter 5.0.

2.2.3 GEOGRAPHIC AND SOCIOECONOMIC DISTRIBUTIONS OF PROCURED VEHICLES

This section examines sample distributions of motor vehicles procured in each of six past surveillance programs (MDVSP series 2 and 3, and LDVSP series 2, 3, 5 and 6). Such examinations reveal some practical problems associated with those vehicle selection methodologies which have been proposed by past and current procurement contractors. In addition to theoretical soundness of each vehicle selection method, real-world usefulness of the method should also be tested in order to correctly evaluate the vehicle selection method. For this purpose, the geographical distribution of vehicles procured in each of the six surveillance programs was first determined from the addresses of the owners of those vehicles, and then was compared with those of the total vehicle population and of household income levels over the SCAB.

Since there was no readily available data summary with convenient geographical resolution for vehicle registration records, a tabulation of household statistics based on 1980 census data was purchased from the Southern California Association of Governments (SCAG) and used as a surrogate for the vehicle registration records. This tabulation covered 55 regional statistical areas (RSA's) constituting the SCAG Planning Region, which covered the SCAB and some adjacent areas.

Figure 2-1 shows the ideal spatial distribution of 400 vehicles if the number of procured vehicles is to be proportional to that of households in each RSA. Figure 2-2 depicts the spatial pattern of high, medium and low income areas over the SCAB. The high income areas are defined by RSAs in which households with incomes less than $20,000 per year make up less than 43.9% of the total, i.e., more than 10 percentage points below the regional average of 53.9%. Similarly, the low income areas are defined by RSAs in which low income households make up no less than 63.9% of the total. The middle income areas include all other RSAs, in which low income households constitute from 43.9% to 63.9% of all households. Figure 2-2 indicates that the low income areas are located in and around the Los Angeles Central Business District (CBD) and in the southeastern part of Riverside County. The high income areas are found in outlying areas of Los Angeles County and in Orange County.

Figure 2-3 shows the spatial distribution of the 339 vehicles which were tested in LDVSP-5. A comparison between Figures 2-1 and 2-3 indicates

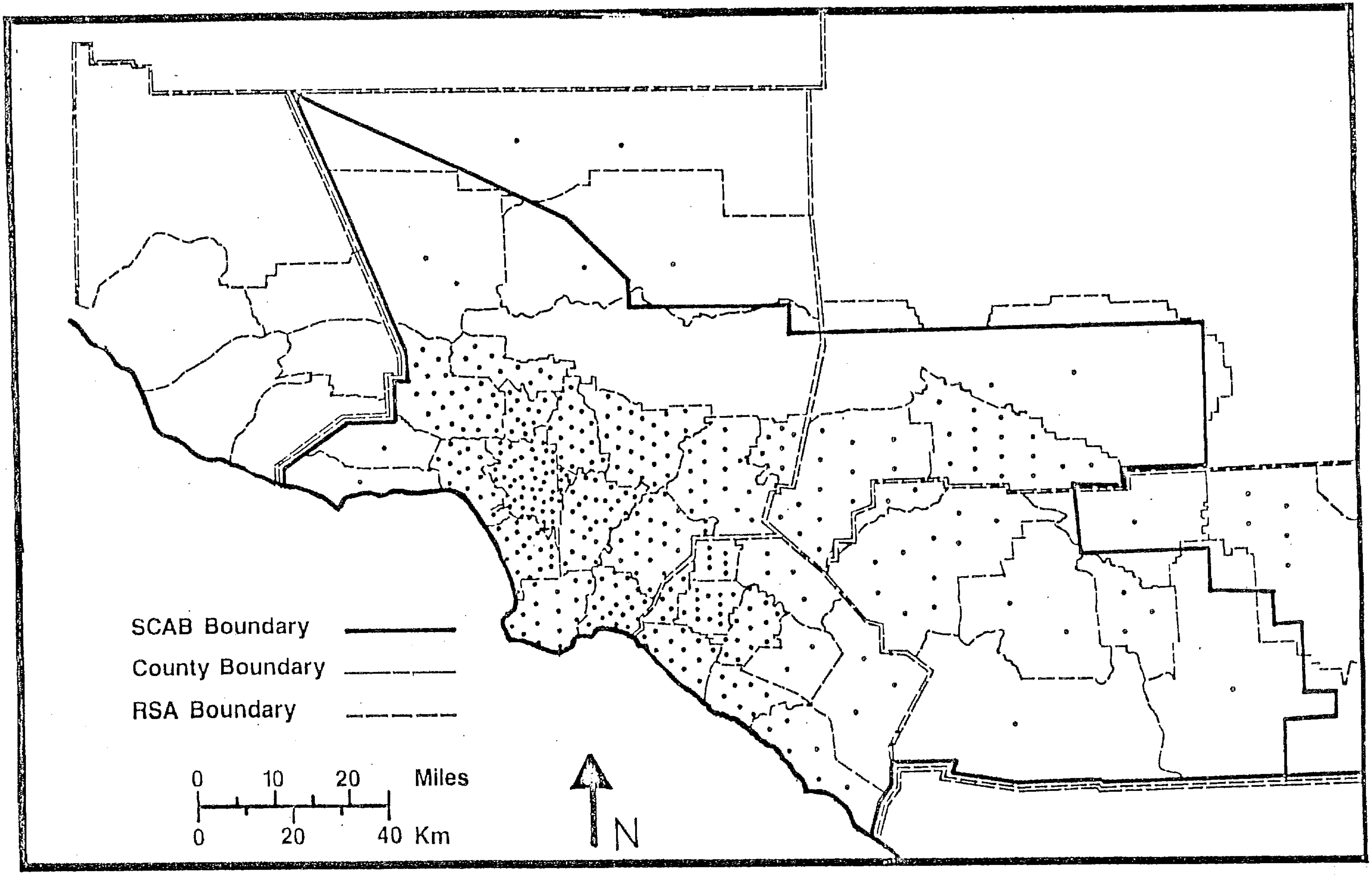


Figure 2-1. Ideal Spatial Distribution for 400 Procurement Vehicles over SCAB

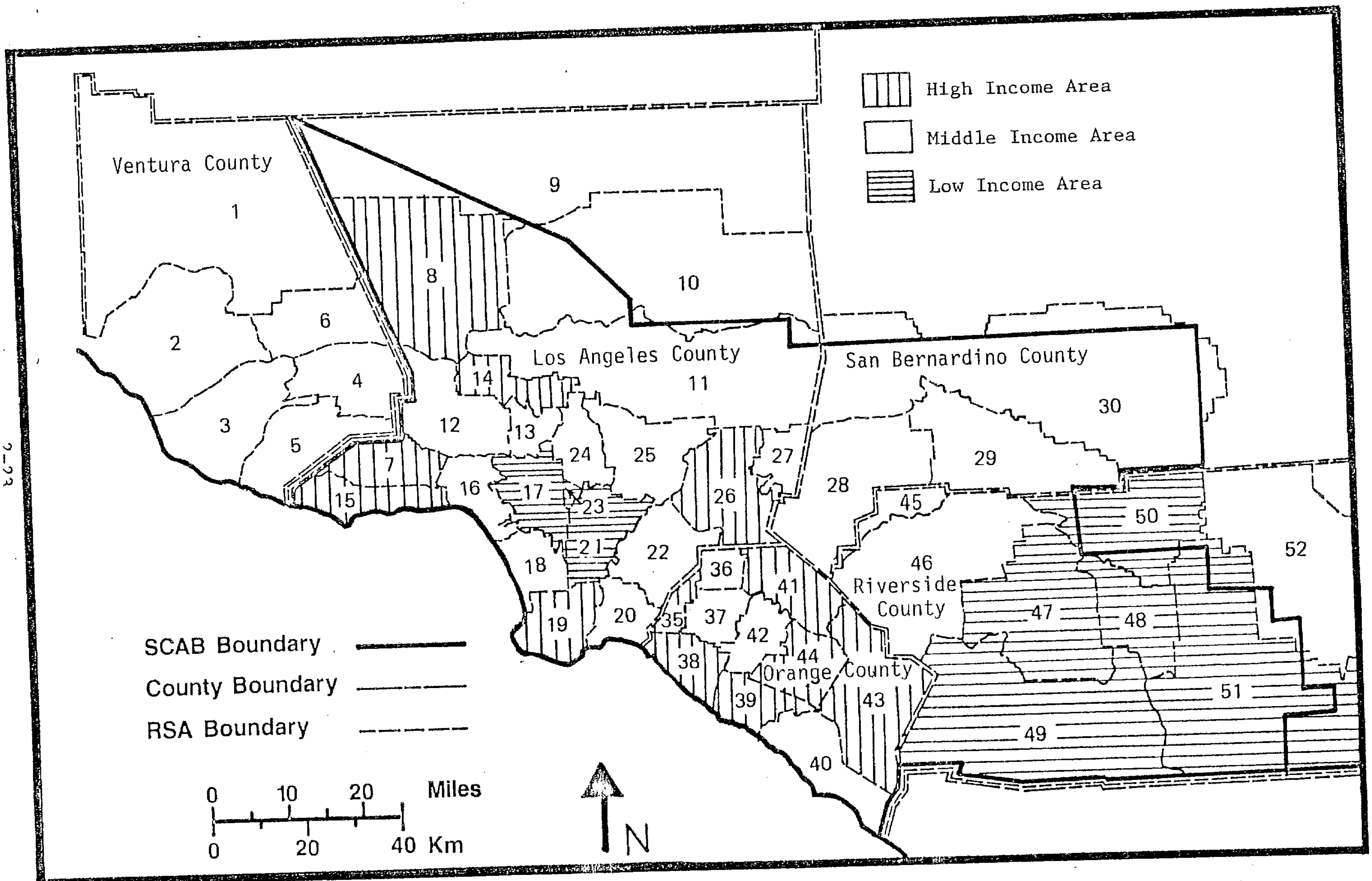


Figure 2-2. Distribution of High, Middle and Low Income Areas over the SCAB (Numerals indicate SCAG's RSA numbers)

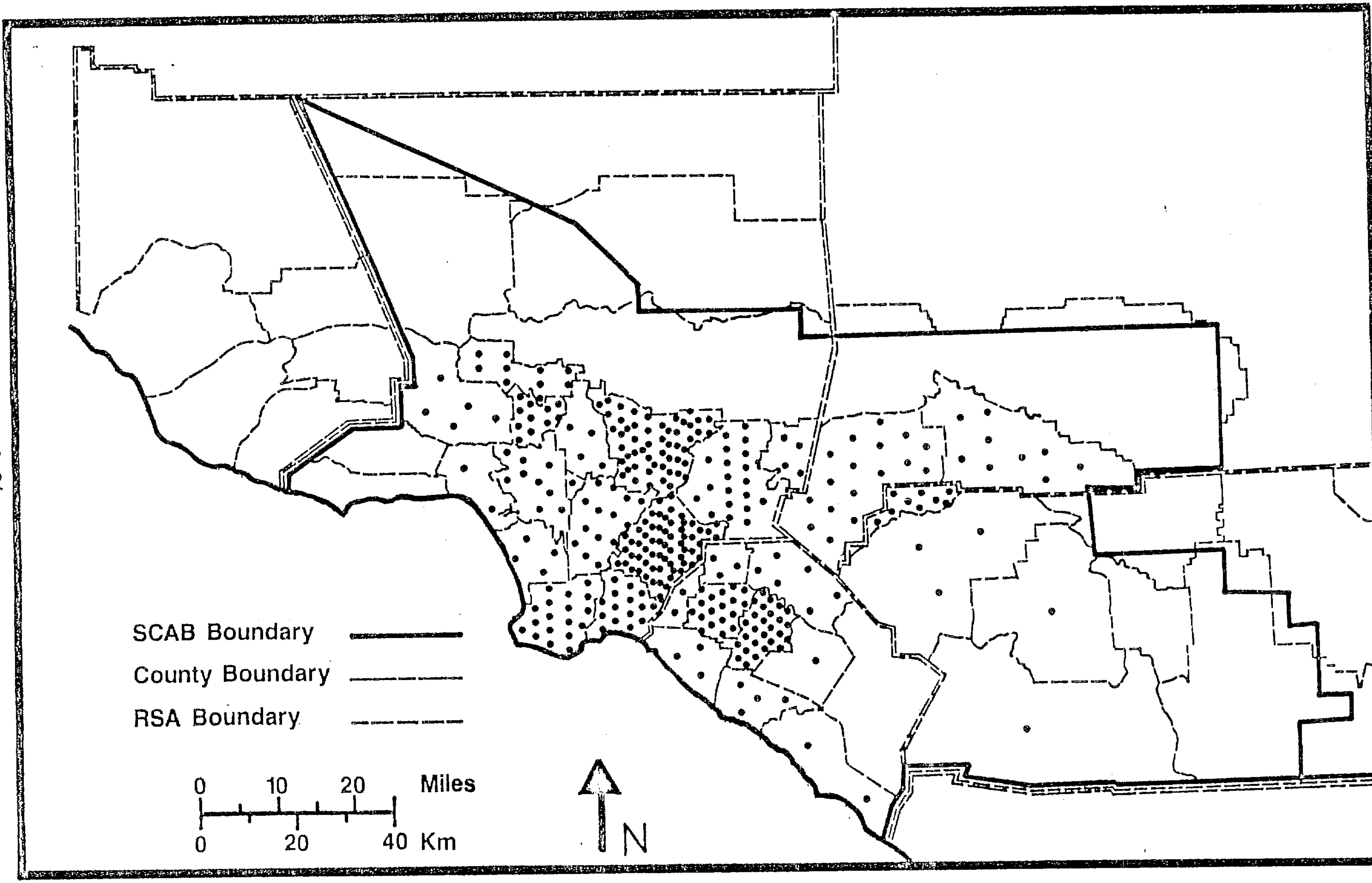


Figure 2-3. Spatial Distribution of 339 Vehicles Procured in LDVSP 5 (Excluding 60 vehicles procured for the Recall Program)

that the sample distribution of LDVSP 5 deviates considerably from the ideal distribution. RSAs near the Los Angeles CBD and those in outlying areas appear to be undersampled while RSAs which are conveniently linked by freeways to the contractor's facility near Anaheim appear to be oversampled. To highlight such differences between sample vehicle distributions and an ideal distribution, Figures 2-4 through 2-9 depict seriously oversampled RSAs, seriously undersampled RSAs, and RSAs from which no sample was taken in the six surveillance programs.

In MDVSP series 3 and 4, numbers of vehicles procured from the SCAB were rather small, 86 and 100 respectively. According to the ideal distribution, the number of samples to be taken from each RSA ranges from 0 for sparsely populated RSAs to 11 for the most populated RSA, No. 17. Therefore, RSAs which were oversampled by three vehicles or more over the ideal sample size are considered to be seriously oversampled. Conversely, RSAs which were undersampled by three vehicles or more are considered to be seriously undersampled. For example, in MDVSP 3 (Figure 2-4), numbers of vehicles procured from RSAs no. 16 (around Santa Monica), 17 (around Culver City) and 21 (around Southgate) are 0, 3 and 0, respectively. The numbers of vehicles procured are less than the ideal sample sizes (3, 9 and 6, respectively) for the three RSAs by three or more. Thus, these RSAs are considered to be seriously undersampled. For RSAs No. 16 and 21, a zero (0) is shown to indicate that no sample was taken from these RSAs.

In LDVSP series 2, 3, 5 and 6, numbers of vehicles procured are close to 400, about four times as large as the sample sizes for MDVSP series 3 and 4. Therefore, RSAs which were oversampled by ten vehicles or more above the ideal sample size are considered to be seriously oversampled. Conversely, RSAs which were undersampled by ten vehicles or more are considered to be seriously undersampled. For example, in LDVSP 6 (Figure 2-9), the actual numbers of vehicles procured from RSAs No. 17 and 21 are, respectively, 8 and 12, which are less than the ideal sample sizes of 40 and 24 for the two RSAs by 32 and 12 vehicles, respectively. Thus, these RSAs are considered to be seriously undersampled. On the other hand, numbers of vehicles procured from RSAs No. 22 (around Whittier), 25 (around Pasadena) and 26 (around Pomona) are, respectively, 50, 62 and 31, which

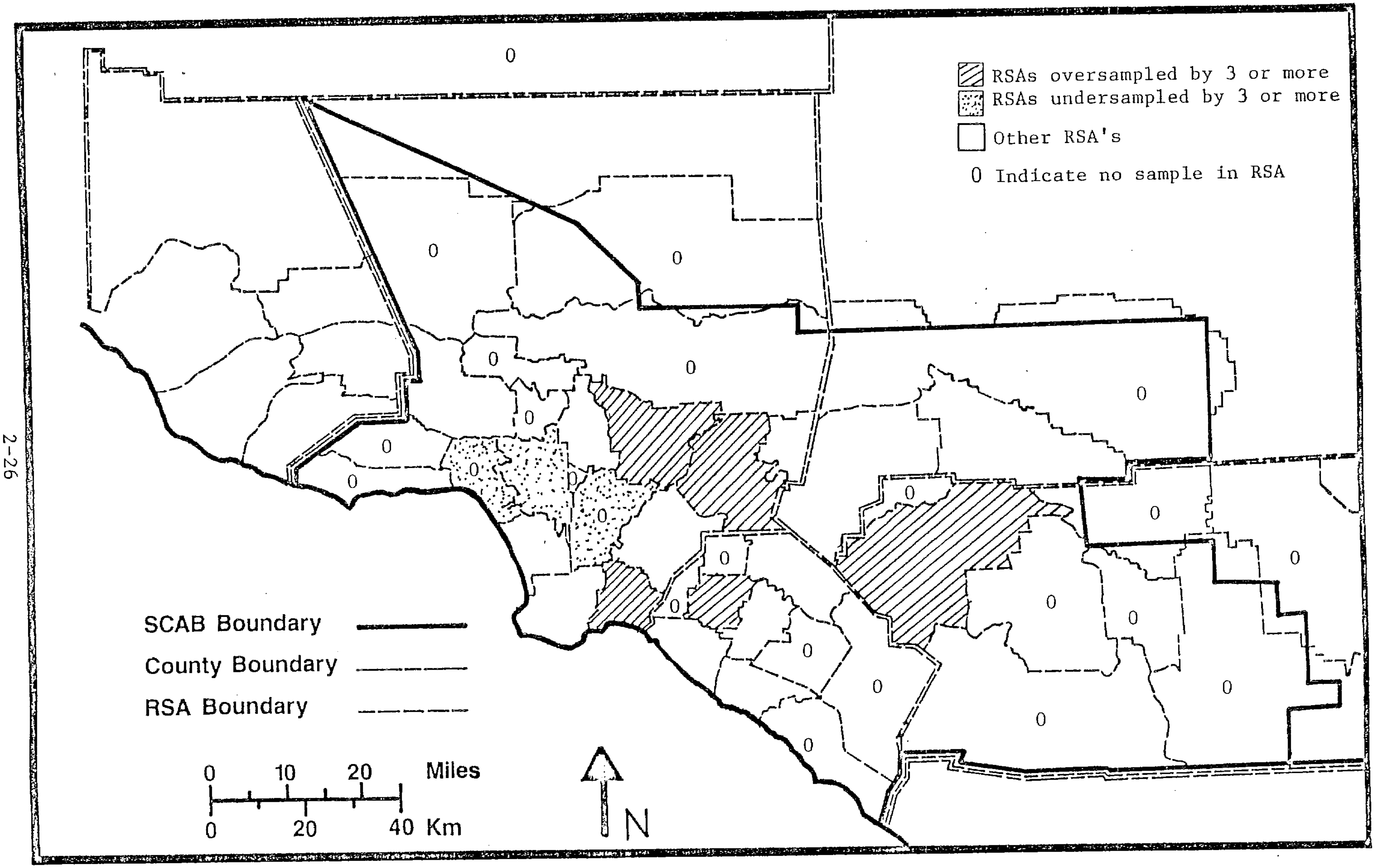


Figure 2-4. Deviation of Sample Distribution from Ideal Spatial Distribution in MDVSP 3 (Number of Vehicles Procured = 86)

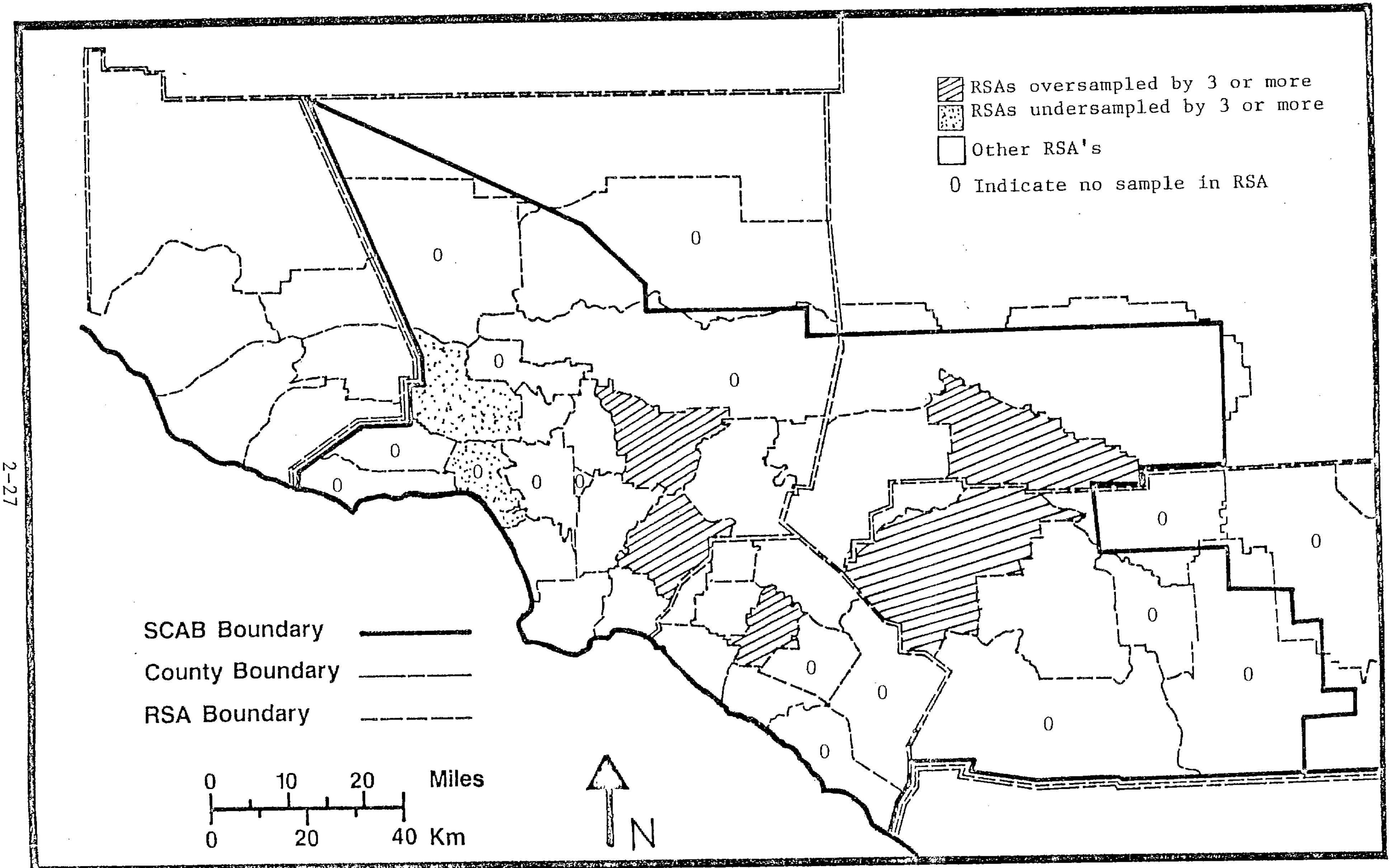


Figure 2-5. Deviation of Sample Distribution from Ideal Spatial Distribution in MDVSP 4 (Number of Vehicles Procured = 100)

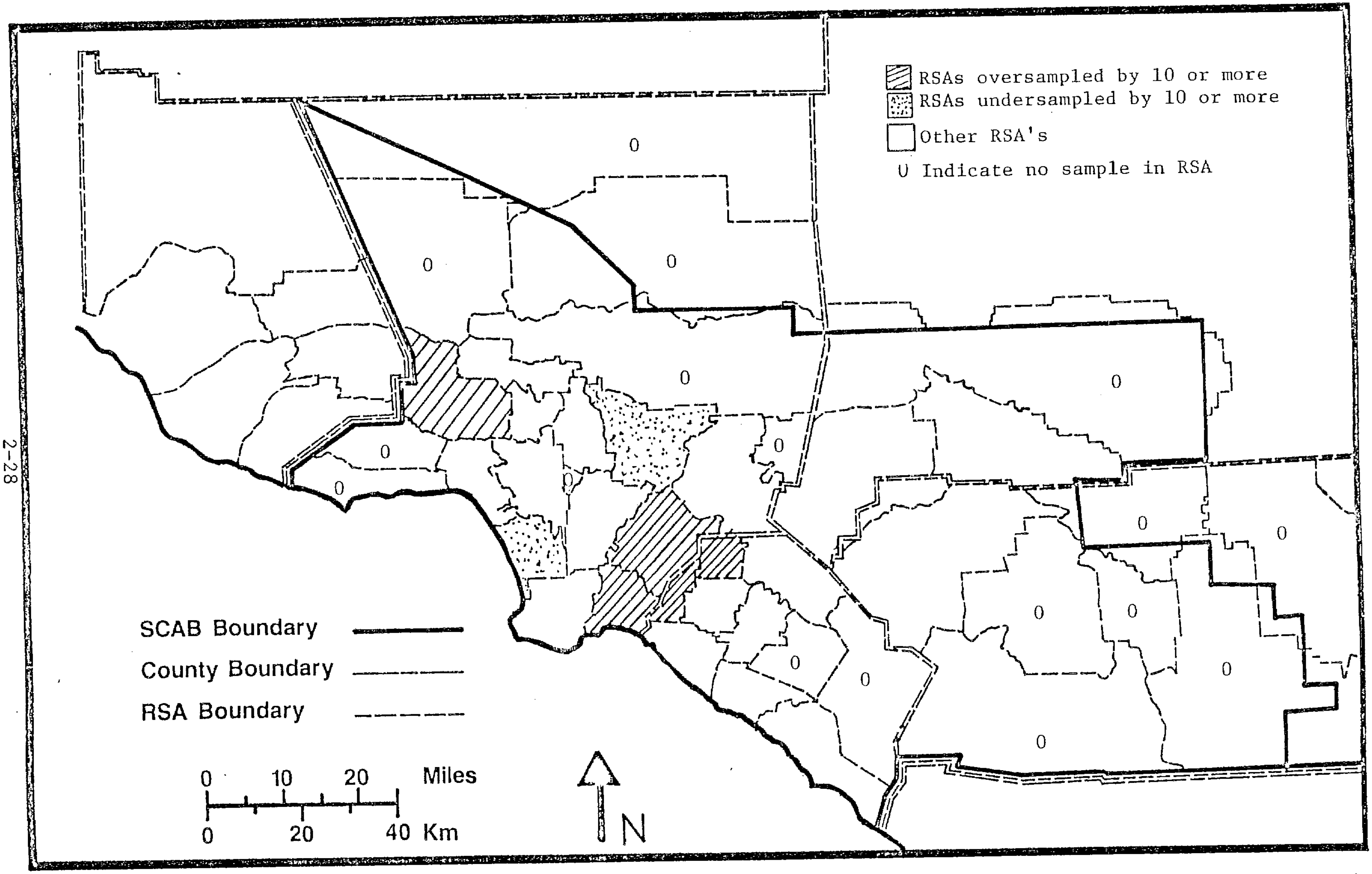


Figure 2-6. Deviation of Sample Distribution from Ideal Distribution in LDVSP 2 (Number of Vehicles Procured = 382)

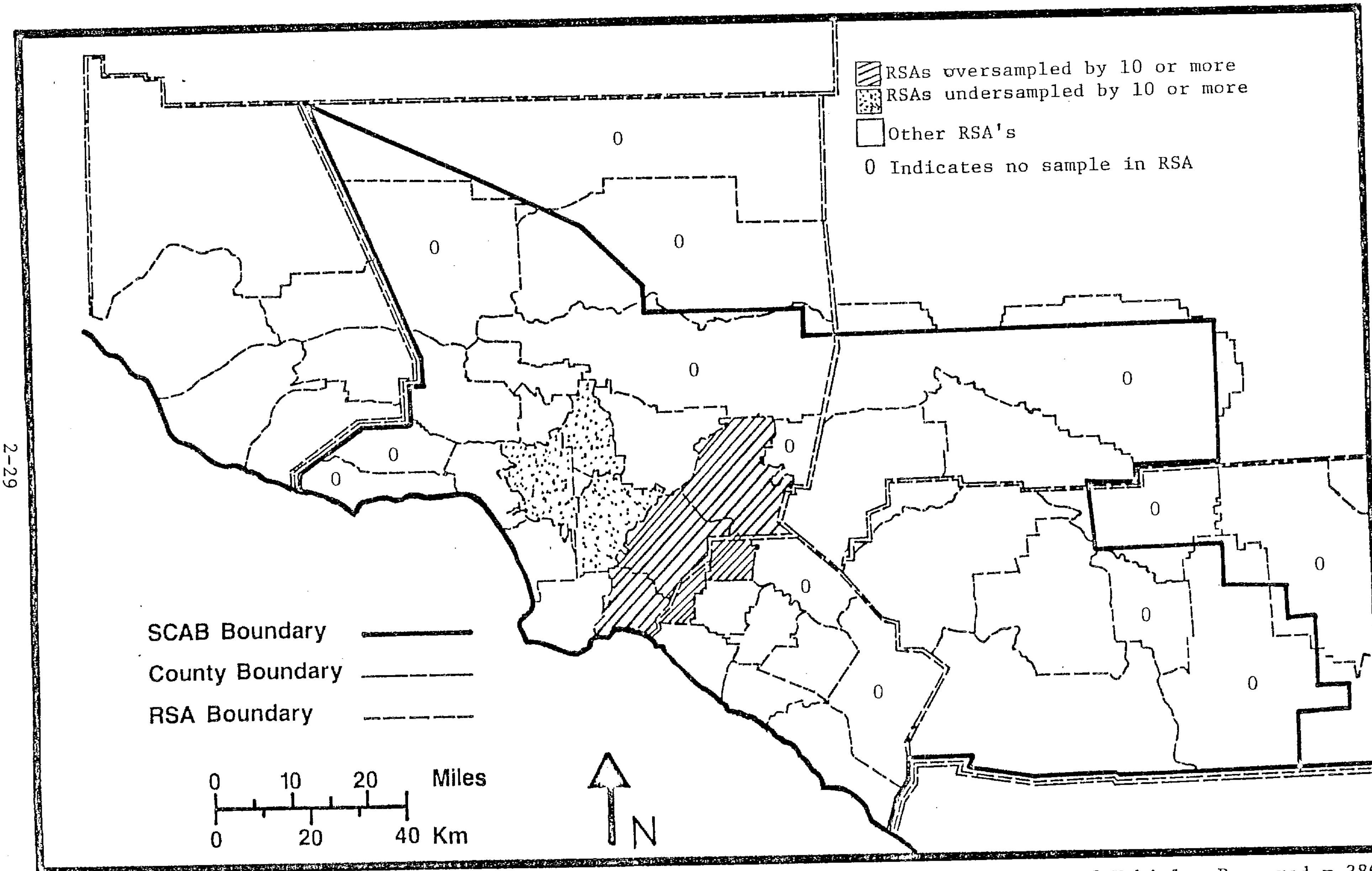


Figure 2-7. Deviation of Sample Distribution from Ideal Distribution in <u>LDVSP 3</u> (Number of Vehicles Procured = 386)

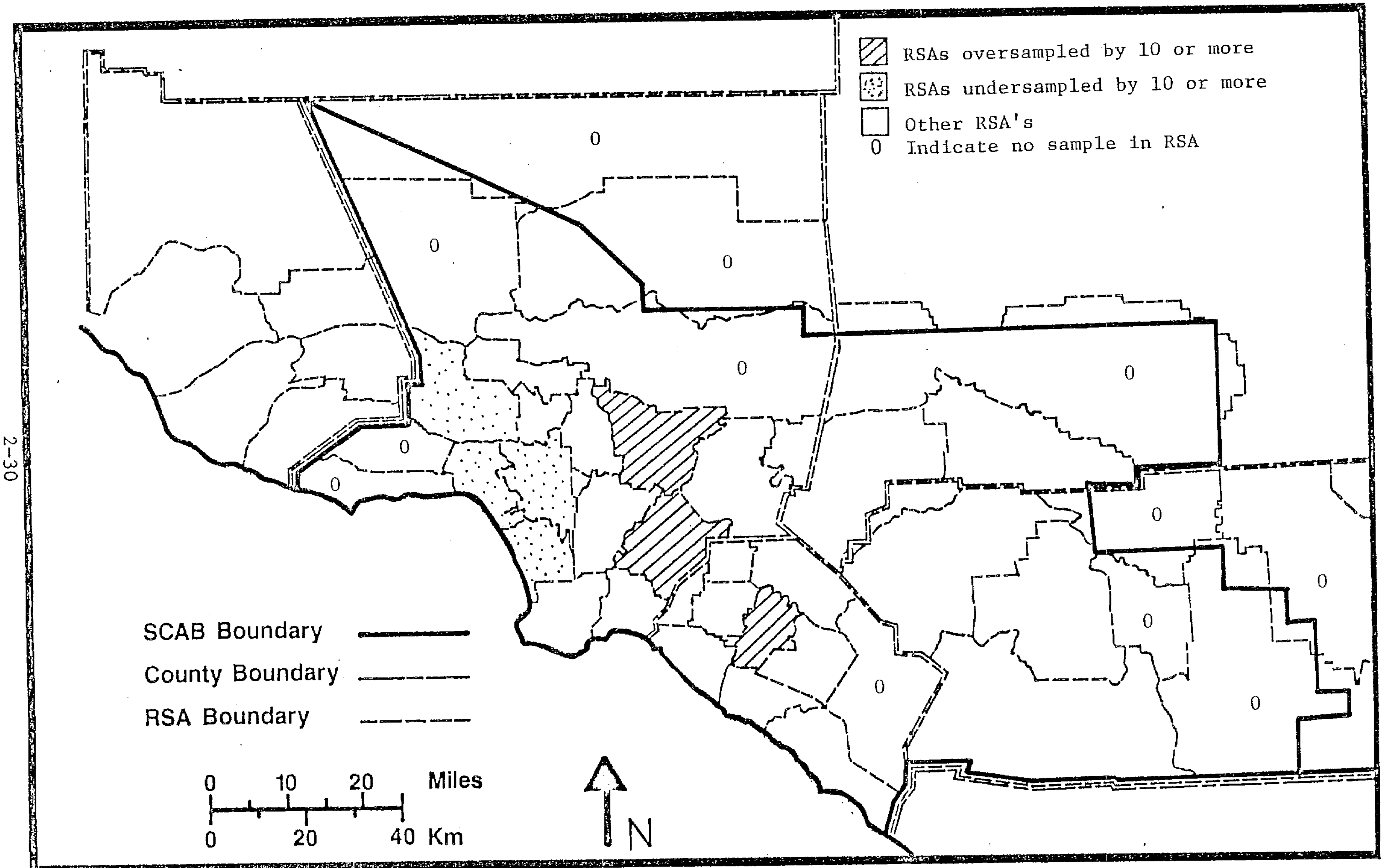


Figure 2-8. Deviation of Sample Distribution from Ideal Distribution in LDVSP 5 (Number of Vehicles Procured = 339)

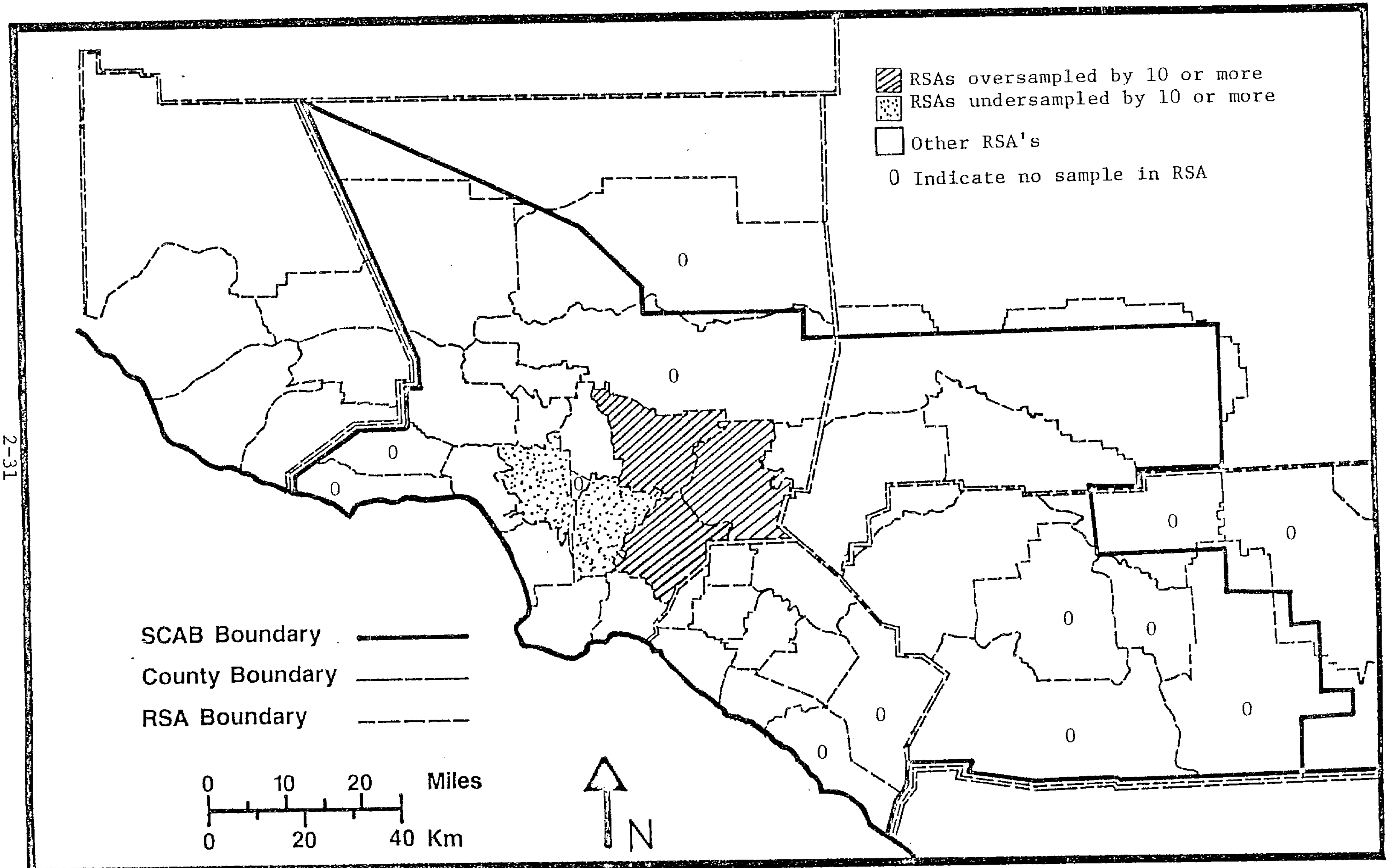


Figure 2-9. Deviation of Sample Distribution from Ideal Distribution in LDVSP 6 (Number of Vehicles Procured = 369)

are considerably greater than the ideal sample sizes of 19, 23 and 15, indicating that these three RSAs were seriously oversampled. It is interesting to note that while outside of Los Angeles County seriously oversampled areas fall in the most populated RSAs (i.e., RSA No. 29 around San Bernardino, RSA No. 46 around Riverside, RSA No. 37 around Anaheim, and RSA No. 42 around Santa Ana), the same is not true in Los Angeles County. In fact, the two most populous RSAs (RSA No. 17 around Culver City and RSA No. 21 around Southgate) are seriously undersampled in three of the four LDVSP programs (see Figures 2-7 through 2-9). These two RSAs were not seriously undersampled, however, in LDVSP series 2, in which vehicle selections were made according to the PSU method described in Section 2.1.1.

It should be noted that the most populous RSAs in Los Angeles County happen to be low income areas while those in the other counties are middle income areas (see Figure 2-2). It can hardly be a mere coincidence that the seriously undersampled RSAs in the three recent LDVSP series all represent low income areas. As procurement contractors learn about the efficiency of the procurement operation, they naturally tend to avoid low income areas where the response rates to their mail and telephone solicitations are low and the security of loaner cars and of vehicle-pick-up personnel is uncertain. It is interesting to note that in all six surveillance programs examined, seriously oversampled areas tended to occur in RSAs near the contractors' facilities, all of which have been located near Anaheim in Orange County. Conversely, RSAs distant from these facilities tended to be unsampled or undersampled in most surveillance programs.

If ARB really requires a geographically representative sample of motor vehicles in the SCAB, this examination clearly indicates a need of more rigorous specifications for the vehicle selection procedure. It is apparent that the currently used quota for each of the four counties does not assure unbiased geographical representation.

2.3 STATISTICAL VALIDITY OF SAMPLES

2.3.1 INTRODUCTION

For ordinary practical purposes, a sample statistic may be said to be reasonably valid as an estimator of a "true" population value if the 95% confidence interval about the estimated value amounts to no more than 10 percent of the estimated value. Put more briefly, this means that there is only one chance in twenty that the estimate is more than 10 percent away from the true value.

The sample size necessary in order for this criterion to be met is determined by the statistical parameters of the population from which the sample is taken. To illustrate the necessary concepts, we will consider data from LDVSP 5, (ARB 1982e) especially the carbon monoxide emissions as determined by constant-volume sampling on 339 vehicles as received by the ARB In-Use Vehicle Surveillance Section, El Monte. PES explored some of the statistical characteristics of this data base and found that CO emissions ranged from 1.03 grams per mile to 229 grams per mile, with a standard deviation (σ) of 27.2 g/mi and an average value of 20.5 g/mi.

For a sample of this size, assuming that the sample is randomly chosen from the population, the standard error of the mean is 1.48 g/mi (equal to $27.2/\sqrt{338}$) and the 95% confidence interval is ± 2.95 g/mi, or $\pm 14\%$ of the mean. Thus, the mean of these test results is not quite as valid an estimator of the population mean as the stated criterion requires, when the sample is viewed solely as a general random sample of the entire population.

However, further investigation of this population shows that stratification of the sample by model years, as mandated by ARB, is likely to improve the validity of the estimate of the mean. Table 2-2 shows how the range of CO emissions varies with model year, in the LDVSP 5 data. It is evident that, for individual years, the range found is less than the range of the entire set. Therefore, for most model years, it can be expected that the standard deviation for the subpopulation will be less than that of the entire population and, when the subgroups are appropriately combined, the residual variance (the sum of squares of

Table 2-2. RANGES OF CO EMISSIONS BY MODEL YEAR, LDVSP 5

MODEL YEAR	MINIMUM (g/mi)	MAXIMUM (g/mi)
1971	19.4	112.2
1972	18.4	102.3
1973	8.2	228.7
1974	15.7	61.3
1975	3.4	87.0
1976	1.8	105.6
1977	1.0	85.5
1978	1.7	52.3
1979	1.5	82.1
1980	1.1	76.7

individual deviations from their appropriate subgroup means) will be substantially reduced. This is the principal reason for the superiority of a stratified sampling approach over a general random sampling approach.

In order to estimate the improvement in accuracy associated with this stratified sampling approach, PES computed means and standard deviations of CO emissions for each model year subgroup and used them to determine the standard error and corresponding confidence limits for the overall mean (which is, of course, the same value as previously cited, 20.5 grams per mile). The results are shown in Table 2-3. Although the residual variance is reduced by about one third by the stratification, the 95% confidence interval is ±2.40 g/mi, or about ±12 percent of the mean. This revised value still does not meet the conventional accuracy allowance of ±10 percent, although it is, indeed, closer to it than the value (±14 percent) obtained without stratification.

Applying the same considerations to hydrocarbon emissions, we find that, viewing the 339 vehicles as a nonstratified collection, the mean emission rate is 1.69 g/mi, with standard deviation 2.37 g/mi; from this the standard error of the mean is ±0.13 and the 95% confidence interval is ±0.26 g/mi, approximately ±15 percent of the mean.

With hydrocarbons, as with carbon monoxide, computation shows that about one third of the overall variance is removed by stratifying the sample according to model year, and this results in a revised 95% confidence interval of about ±13 percent of the mean.

In the case of NOx emissions, the mean emission rate (339 vehicles) is 2.08 g/mi, with standard deviation of 1.25 g/mi; from this, the standard error of the mean is ±0.068 g/mi and the 95% confidence interval ±0.135 g/mi, or about 6.5 percent of the mean. Accuracy for this determination is clearly better than required by the ±10 percent convention. Stratification by model year improves the implied accuracy by reducing the confidence interval about the mean to ±0.11 g/mi, or very slightly more than ±5 percent of the mean.

These results demonstrate that, with or without stratification by model years, the sampling scheme used in LDVSP 5 is less than adequate to

Table 2-3. LDVSP 5: EMISSIONS BY MODEL YEARS
(Means and Standard Deviations)

Model Year	CO, g/mi		HC, g/mi		NO_x, g/mi	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
1971	48.1	27.4	3.90	2.64	3.53	0.87
1972	50.1	25.5	3.84	1.75	3.00	1.06
1973	67.9	62.0	4.90	3.38	2.66	1.07
1974	38.2	15.5	4.68	4.83	3.45	1.17
1975	23.2	20.3	2.02	1.45	2.81	1.38
1976	20.6	24.5	1.41	1.19	2.52	1.05
1977	14.9	16.4	1.64	3.21	2.22	1.44
1978	10.6	10.4	0.82	0.56	1.59	0.77
1979	9.8	12.3	0.69	0.54	1.59	0.87
1980	8.5	13.4	0.59	0.89	1.09	0.70

yield a mean value of emissions of either CO or HC with a conventional ten percent accuracy. However, the mean found for NOx emissions is probably within the suggested ten percent allowance.

2.3.2 VALIDITY OF VEHICLE SAMPLING IN RELATION TO ESTIMATION OF LIGHT DUTY VEHICLE EMISSIONS IN SCAB

The validity criterion defined and discussed in the previous section (viz., 95% confidence interval equal to $\pm$10 percent of the mean for an entire series sample) offers a reasonably simple test in the context of a single year's sample. However, while meeting the criterion would assure that the sample was large enough for practical purposes, it would not assure that emissions estimates based on the testing results would also conform to the desired 10 percent error limits.

Whether or not the estimated mean emissions of a year's sampling of vehicles are accurately determined and adequately representative of the corresponding population, that estimated mean is not utilized in estimating population emissions. In fact, the system devised by ARB's Stationary Source Control Division is much more involved than the mere selection of a representative sample.

As explained in the Stationary Source Control Division publication, "Procedure and Basis for Estimating On-Road Motor Vehicle Emissions," (ARB 1980a), emissions for each of six types of vehicles are calculated as the product of emission factors and use factors. Each emission factor provides an estimate of the rate at which a pollutant enters the atmosphere because of vehicle use, while the use factor provides an index of the extent of the indicated use of such vehicles.

The emission factor (for a given vehicle type, vehicle use and pollutant) is a function of several parameters such as model year, vehicle age, traffic conditions, engine temperature, ambient temperature and so on. In actual on-road use, vehicles cannot be expected to perform with exactly the same emission factors as those obtained by the Mobile Source Control Division under laboratory conditions, which are quite rigorously controlled. SSCD has therefore developed a methodology for estimating vehicle emissions which depends on the prior estimation of composite emission factors (CEF).

Each composite emission factor (CEF_{vpnsm}) pertains to a given vehicle type (v), pollutant (p), calendar year (n), speed (s) and operating mode (m), where the categories are defined as follows:

Vehicle type (v):

light-duty passenger vehicles
light-duty trucks
medium-duty trucks
heavy-duty gasoline-powered trucks
heavy-duty diesel-powered trucks
motorcycles.

Pollutant (p):

exhaust hydrocarbons (HC)
carbon monoxide (CO)
oxides of nitrogen (NOx)
evaporative hydrocarbons
particulate matter
oxides of sulfur (SOx)
lead

Calendar year (n): the year for which emissions are to be estimated; generally not before 1966.

Speed (s): the average speed (km/hr or mi/hr) of a segment of traffic in which the vehicles are assumed to be operated.

Operating mode (m):

cold start: the condition between the time a vehicle is started with a cold engine and the time it is fully warmed up. This time is 505 seconds and the average speed during this time is 25.6 miles per hour (41.2 km/hr).

hot start: A restart condition after the hot stabilized condition and after a ten-minute pause during which time the engine is off. This period from restart to the end of the CVS-75 test is 505 seconds.

hot stabilized: the condition when fully warmed up; that is, 867 seconds after the cold start portion of the test is completed. The average speed is 16 mph.

Each CEF is calculated from the corresponding "FTP mean" emission factors, which are specific for a given vehicle type (v), pollutant (p), calendar year (n) and model year (i), by aggregating all model years weighted by annual miles driven, and applying correction factors to adjust for non-standard conditions of average speed (s) and operating mode (m).

The "FTP mean" emission factor is the emission factor derived from the ARB vehicle surveillance data, using standard testing conditions and procedures, for the in-use vehicle surveillance programs. ("FTP" stands for Federal Test Procedure). The data used are "CVS"--constant volume sampling--emissions as measured on vehicles as delivered by the procurement contractor, before any adjustment or repair is performed by the Mobile Source Control Division.

The FTP mean emission factor, however, is not obtained by taking the mean emissions of all of the corresponding vehicles in any year's surveillance program, or even of all in the MSCD's accumulated data base. Rather, the FTP mean emission factor is calculated from a linear equation,

$$C_{ivnp} = A_{ivp} + B_{ivp}\, Y_{ivn},$$

where C_{ivnp} is the FTP mean emission factor for model year i, vehicle type v, calendar year n and pollutant p, and

A is the exhaust emission rate;

B is the rate of "deterioration" (increase) of exhaust emissions with mileage;

Y is the cumulative mileage attributed to the vehicles in the selected type and age category.

The factors A and B are determined by applying a regression analysis to all data (for vehicles of type v) found in the accumulated data base.

Thus, to recapitulate, the composite exhaust emission factors (CEF) defined by SSCD are composites of FTP mean emission factors (C), corrected for road conditions and weighted by use factors (average annual miles) attributed to each vehicle type. For the convenience of users outside ARB, SSCD has computed and tabulated all necessary estimates of both the CEF's and the corresponding use factors (travel fractions).

Nevertheless, the tabulated values are subject to error from several causes, only one of which is the uncertainty of representative sampling in the procurement of vehicles for the surveillance program. Although a detailed consideration of the effects of these error sources is beyond the scope of this report, some of the readily recognizable sources may be listed, as follows:

1. Nonlinearity of relation of emission factor to accumulated vehicle mileage. (SSCD assumes the relation is linear, as discussed above, but this assumption is probably not strictly correct).

2. Use of a standard table of vehicle mileage as a function of vehicle age. (Any variations in the average annual miles driven, due to differences in location or in the choice of calendar year for an inventory, will not be reflected if the standard table is used.)

3. Use of a standard table of the distribution of vehicle miles traveled (VMT) among vehicle types.

4. Use of a standard table of the distribution of vehicle types and ages among the vehicle population.

5. The necessity of estimating evaporative emissions, not fully evaluated in the vehicle surveillance programs.

6. The omission, by policy, of certain types of vehicles from the population available for testing. (Examples are fleet vehicles, rental vehicles, law enforcement vehicles, vehicles failing to meet requirements as to condition to qualify for testing.)

7. Averaging of speeds before applying speed correction factors (since the correction factors are not proportional to speed, averaging of speeds necessarily introduces a bias.)

8. The necessity of estimating the fractions of trips started with cold and hot engines, respectively.

9. The necessity of estimating numbers of trips per day.

10. The necessity of estimating numbers of vehicle miles traveled per day.

Other sources of uncertainty exist, sources both of possible bias and of statistical variance. Many of these may be of minor importance, but it is probable that, in the aggregate, they outweigh the uncertainties which arise because of possible unrepresentative sampling of the vehicle population for the in-use vehicle surveillance program.

Thus, in considering whether current or proposed procedures for vehicle procurement can be expected to yield an adequately representative sample of the vehicle population, it is reasonable to assume that a correctly stratified sample which gives 95% confidence limits within ± 10 percent of the sample mean will be adequate. This criterion, of course, does not eliminate the problem of assuring that any sample is "correctly stratified."

3.0 EVALUATION OF SELECTION METHODOLOGY PRACTICED BY ARB CONTRACTORS

3.1 INTRODUCTION

A critical evaluation of vehicle selection methodology as practiced by ARB contractors, in the context of completed procurement contracts, has been carried out. The two principal types of information which have been useful in this respect are: (1) information obtained by interviewing the technical staff members who have managed such contract work; (2) statistical tests, on data generated by MSCD's surveillance testing, to determine post hoc whether certain factors considered in ARB's program design are important enough to deserve continued emphasis.

In this chapter, information from interviews with contractor personnel is described and discussed in subsection 3.2; results of statistical tests are presented and interpreted in subsection 3.3.

3.2 PROBLEMS IN SELECTION METHODOLOGY AS SEEN BY CONTRACTORS

PES has interviewed, face-to-face, vehicle procurement staff of Environmental Resource Management, Inc. (ERM), Systems Control, Inc. (SCI), and Automotive Environmental Systems, Inc. (AESI). PES has also interviewed over the telephone technical staff of two other procurement contractors: Automotive Control Engineering (ACE) and EG&G Automotive Research in Virginia. Some of these staff members have worked on both ARB and EPA procurement programs. Vehicle procurement procedures reported by these staff members are all basically the same. Major steps involved in vehicle procurement are:

1. Candidate vehicle identification
    a. Obtain DMV records
    b. Sort DMV records for convenient use
    c. Select owners of desired-type vehicles
2. Owner contact
    a. Prepare a solicitation letter and other contact material
    b. Mail the contact material to vehicle owners
    c. File positive and negative responses

d. Mail secondary contact material

e. Telephone contact for a procurement arrangement

3. Owner screening

   a. Check proficiency of vehicle maintenance

   b. Ascertain completeness of written records

   c. Verify owner's willingness to participate actively in program

4. Vehicle screening

   a. Verify that the car is a desired-type vehicle

   b. Verify that the car is safe for driving and testing

   c. Inspect for evidence of any tampering with the emission control system

The third step is required for a recall program but is often skipped in a non-recall program. Furthermore, EPA recall programs (EPA 1981) appear to place greater importance on this step than ARB recall programs. For instance, EPA requires the following conditions to be met prior to any vehicle procurement for its recall program:

- Demonstrated proficiency of maintenance per manufacturer's recommended maintenance schedule;
- Written and notarized records of maintenance (later used as affidavits); and
- Expressed willingness of the owner to appear in court if necessary.

Almost all of the procurement contractors PES interviewed pointed out the inadequacy of funding in ARB procurement programs. One procurement contractor told PES staff his personal experience of justifying various costs in procurement operation. EPA questioned several cost items of his company invoice to EPA and decided to audit the company. Based on an agreement reached between EPA and the company, an EPA investigator worked for a week with procurement staff of the company to determine whether all cost items were indeed necessary and justifiable. After the week-long study, this investigator found no cause for suspicions about the cost items and approved every one.

Since a car is such an indispensable and personal item in today's life, most people feel that exchanging their car for a loaner car is like trading their favorite clothes for rented garments. When people are content with their cars, they simply do not want to be bothered with any program unless it readily stimulates their interest. In view of this kind of personal attachment of car owners to their vehicles, the quality of the contact material and the personal communication becomes essential for persuading them to participate in the procurement program. Several procurement experts confessed that they had made mistakes in their earlier procurement projects by trying to entice car owners by offering special inducements or by trying to attract them through newspaper advertisements.

Very few car owners are willing to lend their cars for an unfamiliar procurement program merely to obtain a small monetary reward or saving. Successful contact material must be designed to apply specifically to the owner's car. Its purpose must be clear and its style professional, in order to inspire the reader with confidence in the program. Since the key to success in the solicitation is to sell the credibility of the program to car owners, the quality of contact material from the funding agency is as important as or more important than that of the contractor. During the interviews, several contractors reported that refinements in their solicitation letters or inclusion of company brochures had increased the positive response rate. Therefore, there appears to be some possibility of further improvement on the capture rate by refining ARB's solicitation letter as well as that of the contractor.

As to identification of candidate vehicles, all the contractors interviewed have pointed out that the vehicle population is so dynamic that during the 6 to 8 months lag time, which is typical for the vehicle registration records provided by Polk, Inc., as many as 25% of car owners may have either moved to other locations or sold their cars. This turnover rate seems to be higher in areas where many of the residents are transient or new comers, and lower in areas where the resident population is stable. Communities with predominantly Mexican-American populations are often characterized by a high vehicle turnover rate. Knowing of such differences in vehicle turn-over rate among communities, procurement contractors tend to concentrate their efforts so as to procure more vehicles from communities with more stable resident populations and fewer from those with more volatile populations.

The rate of non-response due to this high turnover might be reduced substantially if ARB could provide the contractor with more recent vehicle registration records or provide the contractor with a direct access to the latest registration data at the Department of Motor Vehicles (DMV). This direct access to DMV's vehicle registration data will counter the concern raised by some contractors that reliance on Polk-provided registration records, which are highly organized by manufacturer, engine type, city, etc., may actually introduce bias due to the use of an ordered list instead of a randomized list of vehicle owners. (A randomized sample of vehicle owners can still be selected from an ordered list but later adjustments of the sample size to unduly high or low capture rates are more difficult to make than when an unordered list is used.) Recognizing this possibility, EPA has recently discontinued the use of Polk-provided registration records in favor of the use of original state DMV records through its own computer program which randomly selects owners of desired types of cars for each procurement program in a statistically sound manner (for more detailed discussion, see Section 4.0).

Higher-than-normal procurement costs occur when the contractor tries to procure cars from communities which are dominated by minorities or from communities which are distant from the contractor's facility. The communities dominated by minorities are often associated with a high vehicle turnover rate; also, they present communication problems because many households there do not include any English-speaking adult. The communities distant from the contractor's facility also tend to lower the capture rate because of psychological barriers against lending their cars to an unknown, remote organization. Extra costs due to this lower-than-normal capture rate are reinforced by the added costs of transporting their cars to the facility and back.

One contractor also blamed security problems for his fewer-than-normal vehicle procurements from minority dominated communities, since some of these communities are considered high crime areas. In such communities, security is questionable, for both the drivers and the loaner cars. Thus, if possible, the contractor prefers to avoid procuring vehicles from high crime areas.

Unlike "cost-plus-fixed-fee" (CPFF) contracts of EPA, ARB's procurement contracts are "fixed-price" (FP) and related to the number of vehicles procured. Under FP contracts, cost considerations tend to deter scientific but cost-enhancing sampling practices, such as tracing non-respondents by telephoning them or by second-mailing to them, implementing PSU's or similar random sampling methods, equally covering both high and low response areas, and so forth. Requests by ARB that only vehicles of a certain manufacturer be delivered at certain times also tend to cause deviation from true random sampling because, just to meet the delivery requirement, the contractor is forced to go after a specific prospective donor, not necessarily chosen at random from all prospective donors with the desired vehicle model. This practice of using specific prospective donors reduces validity of currently employed vehicle sampling, especially when the vehicle is a rare model, or there is some special mileage stipulation.

As to vehicle screening, the primary concerns of ARB contractors are to ensure that the car is a desired-type vehicle and that the car is safe for driving and testing. Although the vehicle registration record and the returned post card information may indicate that the car is a desired-type vehicle, under-the-hood inspection is required before the car is delivered to ARB. Minor repairs which do not affect emission rates are sometimes made at contractor's expense when the expense for the repairs is judged to be less than the cost of locating and procuring another vehicle of the same type. Indications of tampering with the emission control system do not result in rejection of the car unless the car is procured for a recall program.

Surprisingly or not surprisingly, all the contractors interviewed indicated that they had very sparingly used incentives (beyond the basic compensation of providing a tankful of gas and a loaner car) for attracting candidate vehicle owners to the procurement program. They report that such an incentive as a $50 U.S. savings bond is little appreciated by most vehicle owners, while other incentives like free guaranteed tune-up attract mainly those vehicles which are totally out of proper maintenance. However, the same contractors said that nearly all prospective participants demand a loaner car which is approximately equivalent to or slightly better than the participant's car. This desire of prospective donors seems to lead to lower capture rates and added costs for procuring expensive or uncommon vehicles.

3.3 STATISTICAL TESTS FOR BIAS IN SAMPLING

3.3.1 INTRODUCTION

In order to minimize potential bias in the sample of vehicles procured, ARB specified a geographical distribution of the fleet (see Section 2.1.1) and specified that the procurement procedure should be designed to achieve an appropriate representation of socioeconomic groups within the SCAB.

Since the estimation of emissions by SSCD procedures does not depend in any way on the geographical or socioeconomic distribution of vehicles in the SCAB, it is evident that these requirements have been based on the speculative possibility that emissions might be related to location or socioeconomic status, independently of the other parameters used to specify vehicle groups. PES therefore undertook certain statistical tests to evaluate the possible importance of such relations, using CVS emissions data. Project resources did not permit an in-depth statistical analysis of LDVSP data; in any event, such an effort was not within the scope of the project as defined by ARB. Consequently, PES limited the statistical testing to a preliminary exploration of only a few of the theoretically possible relations.

Using data presented in ARB's Test Report on LDVSP 5, (ARB 1982e) together with information taken from post-test questionnaire responses, regarding 339 vehicles procured for the base program in that series, PES created a computer data base containing the following information for each car:

1. Identification number assigned in the Test Report;

2. Model year;

3. Odometer reading as given in the Test Report;

4. Hydrocarbon emissions indicated by the CVS test of each vehicle in "as received" condition;

5. Carbon monoxide emissions, CVS, "as received";

6. Nitrogen oxides emissions, CVS, "as received";

7. Miles per gallon as determined in the same test;

8. The location of the registered address of the owner of the vehicle, in terms of the Regional Statistical Areas defined by the Southern California Association of Governments (RSA's 12 to 14, 16 to 29, 37 to 42, 45 to 47, and 49);

9. The income level of the family of the owner of the vehicle, as indicated in the post-test questionnaire, coded as follows: 1, $10,000 or less; 2, $20,000 or less, but more than $10,000; 3, $30,000 or less but more than $20,000; 4, $50,000 or less but more than $30,000; 5, more than $50,000; and

10. Displacement of the engine of the vehicle, in cubic inches.

This data base, or portions of it, was used for various statistical tests intended to probe for possible signs of bias, due to possibly unrepresentative sampling, which might adversely affect the accuracy of any emissions estimates based on the test results.

3.3.2 GEOGRAPHICAL DISTRIBUTION OF VEHICLE OWNERS

The effort to assure that vehicles tested are drawn in a representative manner from all parts of the SCAB is essential if emissions from otherwise equivalent cars differ systematically in samples from different parts of the Basin. Since ARB routinely specifies that certain fractions of the sample must be obtained from the various counties in the Basin, PES divided the sample into two parts, vehicles from Los Angeles County vs. vehicles from other counties, and tested the null hypothesis that the average values of the variables in the database would be the same for each of the two subsamples.

The test was executed using a standard statistical computer routine, BMDP7D, (Dixon 1979) called "Description of Groups with Histograms and Analysis of Variance." This program calculates the required mean values for each of the subgroups and, by analysis of variance, determines the probability that a difference as large as that found is attributable to chance (and thus does not reflect any essential difference between the subgroups).

Results of this analysis are presented in Table 3-1. The subgroups consisted of 249 vehicles from Los Angeles County and 88 from other counties, after two vehicles were excluded because of incomplete information. The analysis indicated that the probability of chance differences as large as those found was more than 5 percent (a conventional statistical significance criterion) for all variables except #7, the fuel economy (MPG); for this variable, the probability corresponding to the observed difference was only 4 percent, or one chance in 25.

This test indicates that the county of origin is not a significant factor in determining the FTP mean emission factors for HC, CO and NOx (although there may be a measurable relation of average MPG to county of origin; for Los Angeles County vehicles average MPG was 16.5, while for the remaining 88 cars it was 17.6). PES concludes that deviations from proportional representation by counties within SCAB are not of great importance in sampling to determine emission factors.

3.3.3 EFFECT OF SOCIOECONOMIC STATUS OF VEHICLE OWNERS

ARB's contract stipulation regarding stratification of sampling on the basis of socioeconomic status is predicated on the perceived possibility that the attention paid to vehicle maintenance may depend on the owners' economic status and may affect emission factors. For this reason, contractors have been required (ARB 1982a) "to exercise some method of introducing a socioeconomic factor" into the sample design. ARB has indicated that this may be done by specifying appropriate geographic subareas.

In developing a list of procurable vehicles for LDVSP2, in which the latest available model year was 1977, ERM applied a rather specific selection process (described in Section 2.1.1, above) to achieve a sampling based on socioeconomic stratification. In later LDVSP procurements, however, ERM has relied on a more casually randomized selection of vehicles

Table 3-1. STATISTICS[a] ON TESTED VEHICLES BY GEOGRAPHIC ORIGIN

Group:	L.A. Co.			Other				
Variable	n	$\bar{x}$	s	n	$\bar{x}$	s	F	Prob
2. Model Year	249	76.9	2.66	88	77.0	2.62	0.18	0.67
3. Odometer(KMI)	249	40.6	28.9	88	44.4	32.6	1.02	0.31
4. HC Emissions	249	1.69	2.31	88	1.67	2.57	0.01	0.94
5. CO Emissions	249	21.3	26.9	88	18.5	28.4	0.69	0.41
6. NOx Emissions	249	2.06	1.21	88	2.15	1.38	0.34	0.56
7. Fuel Economy	249	16.5	4.36	88	17.6	4.64	4.13	0.04*
9. Income (coded)	192	3.00	1.03	72	2.94	1.20	0.11	0.74
10. Engine Size	249	227	111	88	208	112	1.94	0.16

a. Explanation of table:
Variables are described in Section 3.3.1 Units are as follows:
Variable 3, thousands of miles
Variables 4,5,6, grams per mile
Variable 7, miles per gallon
Variable 10, cubic inches of displacement
n is the number of vehicles included (in each group)
$\bar{x}$ is the mean value of the variable (for each group)
s is the standard deviation (within each group)
F is the F-statistic for Snedecor's test in the one-way analysis of variance
Prob. is the probability associated with the value of F
* indicates the difference between group means is significant at the 5 percent level by Snedecor's test.

within each county to produce a distribution which would be representative in regard to socioeconomic status.

For a preliminary indication of whether vehicle emissions are appreciably related to owners' incomes, PES applied BMDP7D to the data from LDVSP5, organized as described above (Section 3.3.2). The vehicles were classified into four categories of engine size (CID) as follows:

Group 1: CID less than 100 (cubic inches)

Group 2: CID less than 200, but not less than 100

Group 3: CID less than 300, but not less than 200

Group 4: CID equal to or greater than 300

The vehicles were also classified according to owners' income (INC) as follows:

Lower income group, not more than $20,000 (INC codes 1 and 2)

Middle income group, $20,000 to $30,000 (INC code 3)

Upper income group, more than $30,000 (INC codes 4 and 5)

Thus, 12 subgroups were defined with four engine size groups for each of three income levels. The matrix of sample sizes was as follows:

	CID	100	200	300	300 up	Total
Income group:	lower	13	29	18	32	92
	middle	17	26	15	24	82
	upper	14	29	12	35	90
	Total	44	84	45	91	264

The grand total of vehicles used in this analysis was 264, less than the full complement of 339 vehicles tested in LDVSP5 because in 75 cases, the income was not listed in the post-test questionnaire. Table 3-2 shows the means and standard deviations of emission test results on these subgroups.

The analysis of variance showed that, in this data base, none of the emission factors were significantly related to income, as shown in Table 3-3. For that matter, none of the other variables compiled were significantly related to income at the 5 percent probability level. This result strongly suggests that randomization of vehicle selection within counties is an adequate means of controlling any possible bias from the

Table 3-2. STATISTICS[a] ON EMISSIONS (LDVSP 5) BY ENGINE SIZE AND OWNER INCOME

Engine Size	<100		<200		<300		300 up	
	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
(a) HC emissions								
Lower incomes	1.17	0.91	1.75	1.90	1.15	1.65	2.40	2.59
Middle incomes	1.78	1.94	0.98	1.00	1.06	0.87	2.25	4.00
Higher incomes	1.34	1.08	1.09	1.34	1.33	1.78	3.06	4.44
(b) CO emissions								
Lower incomes	15.5	13.3	21.3	25.7	10.7	8.8	27.1	27.9
Middle incomes	18.2	19.2	12.6	12.7	16.8	21.3	18.0	13.0
Higher incomes	18.3	15.8	16.9	22.6	9.8	9.1	37.0	52.4
(c) NOx emissions								
Lower incomes	1.20	0.94	1.74	1.01	2.81	1.69	2.69	1.38
Middle incomes	2.15	0.86	1.95	1.25	1.66	0.98	2.21	1.32
Higher incomes	2.24	1.02	2.15	1.39	1.87	0.60	2.32	1.29

a. Explanation of table:
For matrix of sample sizes, see text (Section 3.3.3)
Units are grams per mile for emissions, cubic inches displacement for engine size
$\bar{x}$ is the group mean rate of emissions
s is the standard deviation (within each group)
Incomes: lower, $20,000 or less annually
middle, over $20,000 but not over $30,000
higher, over $30,000 annually.

Table 3-3. RESULTS OF TWO-WAY ANALYSIS OF VARIANCE[a] ON LDVSP 5 EMISSIONS DATA BY ENGINE SIZE AND OWNERS' INCOME

Source	DF	HC		CO		NOx	
		F	Prob	F	Prob	F	Prob
Engine Size	3	5.09	.002	3.99	.008	2.83	.039
Income	2	0.10	.90	0.46	.63	0.31	.73
Interaction	6	0.56	.76	1.15	.33	2.90	.009

a. Explanation of table:
DF is degrees of freedom for the comparison indicated
F is the F-statistic for Snedecor's test
Prob is the probability associated with the value of F
Units of emissions are grams per mile.

suggested socioeconomic source.

However, further inspection of Table 3-3 reveals that, for NOx emissions, the interaction of the two effects is significant; that is, deviations of the means of NOx emissions for particular subgroups from the values that would be predicted on the basis of strictly linear effects of two independent factors (engine size and income) are greater than are readily ascribed to chance variations.

Further scrutiny of Table 3-2c (the NOx group average emissions) shows that various anomalous effects are present. For instance, with engines in the two smaller categories, vehicles belonging to owners with relatively low incomes emit appreciably less NOx than the other vehicles; but with engines in the two larger categories, that situation is reversed. Together with the finding of significant interaction, this indicates that, for NOx emissions, owners' income (if it has any effect at all) cannot be interpreted as having a consistent effect independent of engine size.

In investigating the geographic distribution of socioeconomic status within the SCAB, PES obtained (from SCAG) computer printouts of household income distribution by ZIPcodes in the Basin (see Section 2.2.3). As a check on the performance of ERM in selecting PSU's by income level in connection with LDVSP2, PES reviewed the ZIPcode areas selected by ERM as representing low income and high income areas. Table 3-4 shows these ZIPcode areas, with two parameters derived from the CDC data: the percentage of households with incomes less than $2,500, and the percentage with incomes greater than $20,000.

The figures exhibited in Table 3-4 do, in general, confirm ERM's assignment of ZIPcode areas to relatively low or high economic status, even though ERM's assessment of this factor was done independently of the SCAG analysis. Thus the percentage of households with income below $2,500, for ZIPcodes selected as lower income areas, ranges from 4.2 to 11.6; the corresponding range for higher income ZIPcodes is from 1.4 to 3.9 percent, not overlapping the range for the former group. Again, using the $20,000 criterion (which is near the median for SCAG households), ERM's lower income ZIPcodes show less than 50 percent above the mark in every case,

Table 3-4. HOUSEHOLD INCOMES DATA [a] FOR ZIP CODE AREAS SELECTED[b] AS LOW AND HIGH INCOME AREAS FOR LDVSP2

		ZIP	n^c <2.5	n^d >20	n^e Total	% <2.5	% >20
Lower income PSU's:	1	91331	790	8400	17,500	4.5	48
	2	90022	1130	4100	16,600	6.8	25
		90023	750	2600	10,800	6.9	24
	3	90001	1180	1950	12,200	9.7	16
		90002	1160	1710	10,000	11.6	17
	4	90805	1040	9400	24,700	4.2	38
	5	90280	1150	8200	23,100	5.0	36
	6	90723	560	3600	11,000	5.1	33
Higher income PSU's:	1	90501	450	5600	12,000	3.8	47
		90502	170	2500	4,400	3.9	57
		90503	310	8700	14,000	2.2	62
		90504	270	7100	11,200	2.4	63
	2	91364	100	5200	7,000	1.4	74
		91356	330	6800	11,100	3.0	61
	3	91316	310	5900	10,600	2.9	56
	4	90049	560	10400	15,900	3.5	66
	5	91790	280	5700	11,300	2.5	59
	6	91311	230	6900	10,100	2.3	68
		91324	240	5100	8,400	2.9	61

a. Data from 1980 Census Data Printout provided by SCAG Census Data Center, Southern California Association of Governments, Los Angeles, CA.

b. As listed in Table A-1, LDVSP 2 Final Report (ARB 1980b)

c. Number of households having incomes of less than $2,500

d. Number of households having incomes more than $20,000

e. Total number of households in area.

whereas the higher income ZIPcodes are above 50 percent in all cases but one. Thus, PES concludes that the technique applied by ERM did generate a reasonably appropriate selection of geographically dispersed PSU's within Los Angeles County, in accord with the desired stratification on socioeconomic status.

3.3.4 EFFECT OF MODEL YEAR

ARB procurement contracts have required the contractors to deliver specified numbers of vehicles of each of several (usually ten) model years, in order to assure a reasonable representation of the various age groups in the population. This type of stratification of the sample is thought to be important because emission control technology as represented in models of succeeding years has been quite variable, largely in response to increasingly stringent emission control requirements. Thus it may be presumed that the distributions of emissions, as found in random samples of vehicles of different model years, should differ significantly, in general.

PES utilized the LDVSP5 data base (described in Section 3.3) to provide an estimate of the magnitude of this effect. Again applying BMDP7D, PES generated statistics for each model year from 1971 to 1980; means and standard deviations of emissions by model year are shown in Table 3-5. One-way analysis of variance showed (Table 3-6) as expected, that emissions were significantly related to model year for all three pollutants (HC, CO and NOx). As also expected, mean odometer readings, fuel economy and engine sizes differed significantly for different model years, while the mean owners' incomes did not.

As expected, the mean accumulated mileage (odometer reading) increases directly with age of the model year group. It ranges from 14,200 miles for 1980 vehicles to 99,400 miles for 1971 vehicles, suggesting an average yearly increase of nearly 9,500 miles per year of age (after the first year).

In contrast, the average fuel economy (miles per gallon) was at a minimum, 14.5 mpg, for the 1974 model year, compared with 16.7 mpg for 1971 and 18.6 for 1980.

Table 3-5. STATISTICS[a] ON EMISSIONS, ODOMETER READINGS, AND FUEL ECONOMY (LDVSP 5) BY MODEL YEAR

		HC		CO		NOx		KMI		MPG	
Model Year	n	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
1971	14	3.90	2.64	48.1	27.4	3.53	0.87	99.4	34.3	16.7	5.0
1972	17	3.84	1.75	50.1	25.5	3.00	1.06	92.1	28.1	15.2	5.1
1973	18	4.90	3.38	67.9	62.0	2.66	1.07	72.7	25.9	14.7	4.6
1974	14	4.68	4.83	38.2	15.5	3.45	1.17	65.6	23.4	14.4	4.2
1975	29	2.02	1.45	23.2	20.3	2.81	1.38	56.0	20.6	15.1	4.1
1976	38	1.41	1.19	20.6	24.5	2.52	1.05	49.3	20.5	16.1	4.6
1977	43	1.64	3.21	14.9	16.4	2.22	1.44	41.1	20.7	16.2	4.5
1978	44	0.82	0.56	10.6	10.4	1.73	0.77	30.2	11.0	17.3	4.4
1979	54	0.69	0.54	9.8	12.3	1.59	0.87	24.3	8.7	17.5	4.3
1980	68	0.59	0.89	8.5	13.4	1.09	0.70	14.2	8.1	18.6	3.6
All	339	1.68	2.37	20.5	27.2	2.08	1.25	41.5	29.9	16.8	4.4

a. Explanation of table:
Variables and their units are as follows (see Section 3.3.1):
Variable 4, HC, hydrocarbon emissions (grams per mile)
Variable 5, CO, carbon monoxide emissions (grams per mile)
Variable 6, NOx, oxides of nitrogen emissions (grams per mile)
Variable 3, Odometer readings (thousands of miles)
Variable 7, Fuel economy (miles per gallon)
n is the number of vehicles included (in each group)
x is the mean value of the variable (for each group)
s is the standard deviation (within each group)

Table 3-6. RESULTS OF ONE-WAY ANALYSIS OF VARIANCE[a] ON LDVSP 5 EMISSIONS AND OTHER CHARACTERISTICS BY MODEL YEAR GROUPS (DF=9)

Variable	MS	RES	F	Prob
4. HC emissions	70.4	3.84	18.4	0.0000*
5. CO emissions	10,200	482	21.3	0.0000*
6. NOx emissions	20.5	1.04	19.6	0.0000*
3. Odometer (KMI)	21,900	320	68.3	0.0000*
7. Fuel Economy (MPG)	63.9	18.6	3.44	0.0000*
9. Income (Coded)	0.63	1.17	0.54	0.85
10. Engine Size (CID)	29,000	12,100	2.41	0.012*

a. Explanation of table:
Variables are described in Section 3.3.1. Units are as follows:
Variable 3, thousands of miles
Variables 4,5,6, grams per mile
Variable 7, miles per gallon
Variable 10, cubic inches of displacement
MS is the mean square for the between-group effect
RES is the mean square for error
F is the F-statistic for Snedecor's test
Prob is the probability associated with the value of F
* indicates the difference in group means is significant at the 5 percent level.

The same analysis (see Table 3-7) showed that emissions of hydrocarbons (HC) and of carbon monoxide (CO) were rather strongly correlated, both for the sample fleet as a whole (0.68) and within individual model years (range, 0.34 to 0.90). However, emissions of oxides of nitrogen were substantially independent of HC and CO emissions, with overall correlation coefficients of 0.24 and 0.20 respectively.

From this analysis, PES concludes that stratification of the sample by model year, as already required in ARB procurement contracts, is appropriate and should be continued, since it contributes to the development of a data base in which the significant differences between model years can be readily tracked and utilized.

3.3.5 EFFECTS OF ENGINE SIZE

Results from the statistical analysis of vehicle emissions by model year, described in the previous section (Section 3.3.4), indicated that emissions of the three types of pollutants (HC, CO and NOx) were only weakly related to engine size. Correlation coefficients between engine size and emissions were close to 0.2 over the whole sample. (Within model years, these correlation coefficients ranged from -.22 to +.56, and not more than three of the thirty were significant at the 5 percent level.) Although these correlation coefficients may be statistically significant, the relations they reflect are not quantitatively important; furthermore, they may be caused simply by the highly significant changes in popularity of different engine size classes in different model years, over a ten-year period in which emission factors were declining because of changes in control technology.

As earlier indicated (Section 3.3.3; see Table 3-3), a two-way analysis of variance of emissions of vehicles classified by engine size as well as owners' incomes showed significant differences among the engine size categories. Inspection of the group means (Table 3-2), however, shows that the variation of emissions with engine size is not mainly due to a simple linear relation. This information appears to be compatible with the findings expressed in the previous paragraph.

Table 3-7. COEFFICIENTS OF LINEAR CORRELATION[a] BETWEEN EMISSIONS AND OTHER CHARACTERISTICS OF TESTED VEHICLES (LDVSP 5)

	Emissions of		
Variable	HC	CO	NOx
4. HC emissions	-	+.68*	+.24*
5. CO emissions	+.68*	-	+.20*
6. NOx emissions	+.24*	+.20*	-
2. Model Year	-.52*	-.54*	-.56*
3. Odometer (KMI)	+.53*	+.48*	+.45*
7. Fuel Economy (MPG)	-.25*	-.29*	-.14*
9. Income (coded)	-.03	+.06	-.02
10. Engine Size (CID)	+.20*	+.22*	+.20*

a. Explanation of table:
Variables are described in Section 3.3.1. Units are as follows:
Variable 3, thousands of miles
Variables 4,5,6, grams per mile
Variable 7, miles per gallon
Variable 10, cubic inches of displacement.
The number of data pairs is 339 for all correlations except that it is 264 for correlations involving Variable 9.
*Indicate the correlation coefficient is significantly different from zero at the 5 percent level.

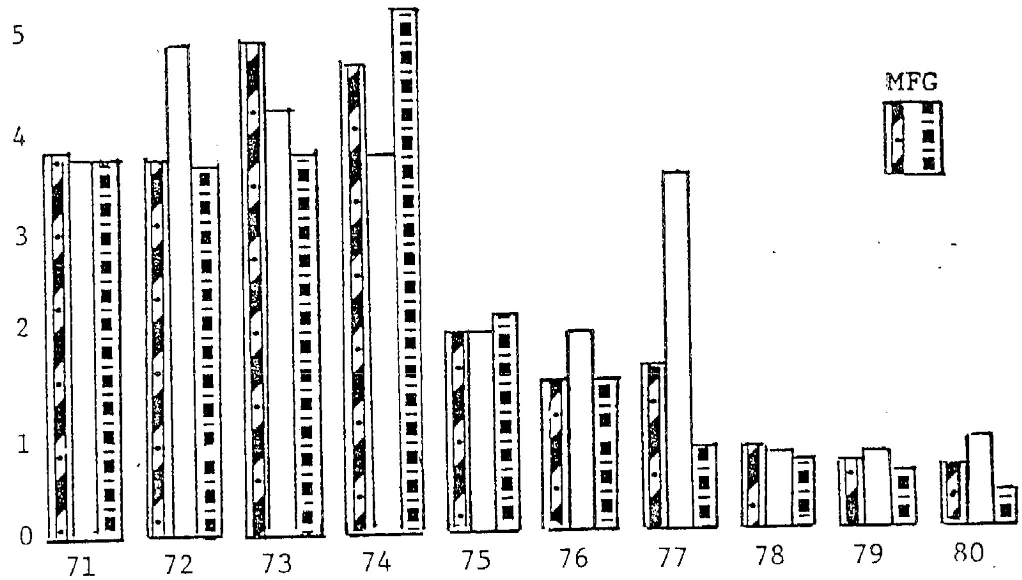
PES therefore concludes that stratification of the sample by engine sizes (CID) is not of the first importance in improving the accuracy of emission factors and regional emissions estimates.

This conclusion is supported by the results of a similar analysis regarding sampling methodology used in EPA's Emission Factors program. In summarizing these results, EPA concluded "This analysis showed a trend. . . of decreasing significance of the vehicle specific variables, to the point that manufacturer and engine size/transmission variables were not significant predictors of emission levels for HC and CO by FY '79, while manufacturer was a significant variable only for NOx in FY '75 and FY '79. Thus it appears that engine size/transmission may not be a particularly useful stratification variable."

3.3.6 EFFECT OF STRATIFICATION BY MANUFACTURER

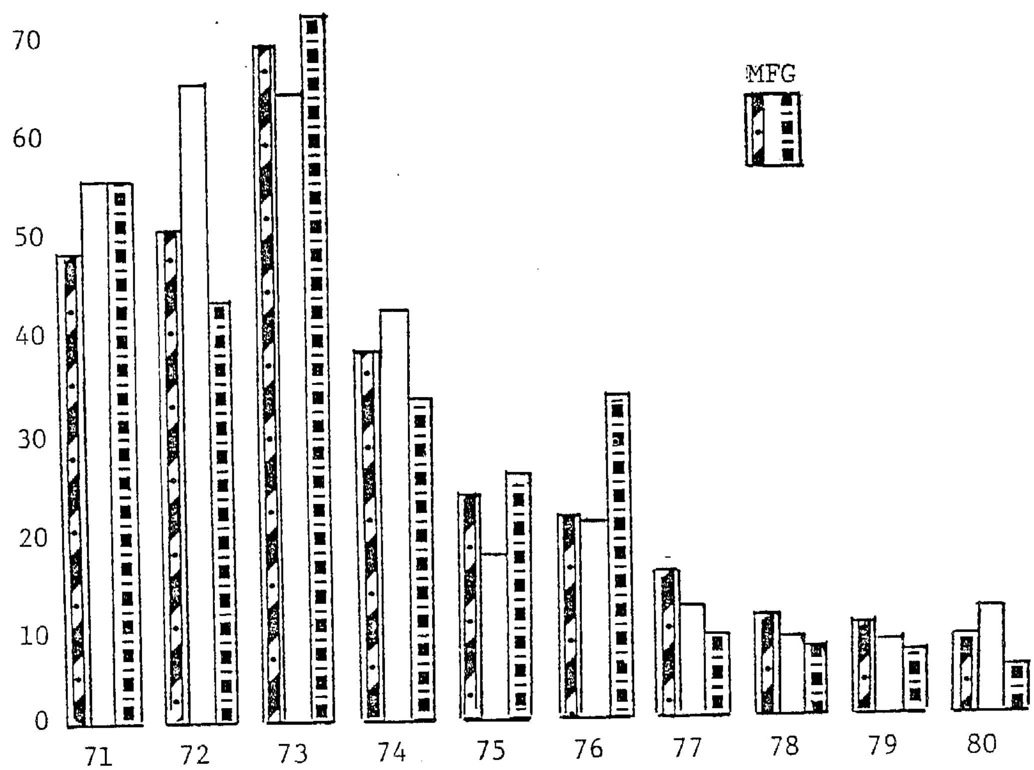
As indicated in the previous section (Sec. 3.3.5), an analysis cited by EPA (EPA 1982) concluded that the manufacturer was not a "significant predictor" of emission levels for HC and CO by FY '79, and only occasionally so for NOx. To compare the EPA findings with ARB's data, PES calculated mean emissions from LDVSP5 for each model year for vehicles produced by the two largest U.S. auto manufacturers. The results are shown in Figures 3-1, 3-2 and 3-3 for HC, CO and NOx respectively in histograms which also indicate the mean emissions for the entire model year samples.

Although, as the histograms show, there were marked deviations from the general mean for one or the other of the major manufacturers for some model years, none of these deviations were statistically significant. For example, Figure 3-1 shows that the mean HC emissions for Ford vehicles of model year 1977 were 3.6 grams per mile, more than twice the general mean of 1.6 and over four times the mean for General Motors vehicles of that model year. Nevertheless, calculation shows that this difference is not significant at the 5 percent level. Specific review of the data shows that the high average for Ford vehicles is caused mainly by a single vehicle which produced emissions of 20.2 grams per mile, thus doubling the average for the subgroup of seven vehicles.



	71	72	73	74	75	76	77	78	79	80
F	3.8	4.9	4.3	4.0	2.0	1.9	3.6	0.78	0.76	0.91
G	3.8	3.6	3.9	5.3	2.2	1.4	0.8	0.69	0.57	0.35
M	3.9	3.8	4.9	4.7	1.0	1.4	1.6	0.82	0.69	0.59

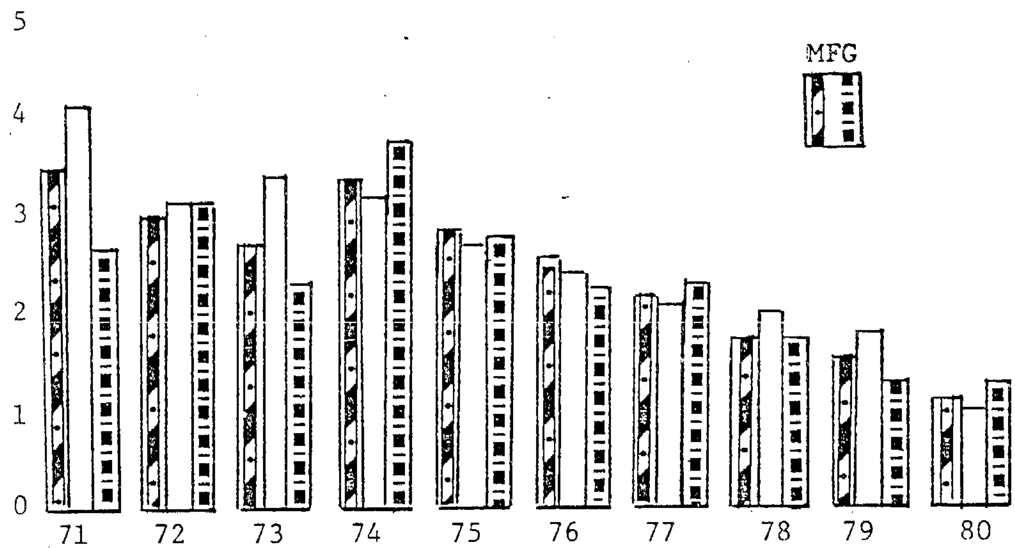
Figure 3-1. Mean HC Emissions From LDVSP 5 Model Year Subgroups



	71	72	73	74	75	76	77	78	79	80
F	55	65	64	42	17	20	12	8.5	8.1	11.4
G	55	43	72	33	25	33	9	7.7	6.6	5.3
M	48	50	69	38	23	21	15	10.6	9.8	8.5

Figure 3-2. Mean CO Emissions From LDVSP 5 Model Year Subgroups

M, mean for all vehicles of the indicated model year
F, mean for Ford vehicles of the indicated model year
G, mean for GM vehicles of the indicated model year



	71	72	73	74	75	76	77	78	79	80
F	4.1	3.1	3.4	3.7	2.5	2.4	2.1	2.0	1.8	1.0
G	2.7	3.1	2.3	3.2	2.6	2.3	2.3	1.7	1.3	1.3
M	3.5	3.0	2.7	3.4	2.8	2.5	2.2	1.7	1.6	1.1

Figure 3-3. Mean NOx Emissions From LDVSP 5 Model Year Subgroups

M, mean for all vehicles of the indicated model year
F, mean for Ford vehicles of the indicated model year
G, mean for GM vehicles of the indicated model year

PES therefore concludes that these calculations confirm EPA's contention (EPA 1982) that the identity of the manufacturer of a vehicle is not a useful variable for predicting emission levels. By the same token, stratification of vehicle sampling among manufacturers cannot be expected to improve the accuracy of determining overall emissions when using ARB's LDVSP data base. Nevertheless, other circumstances may justify the continuance of ARB's stratification procedure, as discussed elsewhere in this report.

3.4 THE INFLUENCE OF OUTLIERS ON THE ACCURACY OF EMISSION FACTORS AND CALCULATED EMISSIONS

One of the most obvious features of the data bases for HC and CO emissions from LDVSP 5 is a marked positive skewness of the frequency distribution, i.e., the existence of a very substantial fraction of the vehicles tested which give CVS emissions more than twice as high as the average vehicle. For example, although the average CO emission of the 339 vehicles in the study was only about 20 grams per mile, the largest individual value was nearly 230 g/mi and there were 40 vehicles with emissions of more than 50 g/mi.

A further important consequence of this highly skewed distribution is the fact that emissions from the vehicles at the upper end of the distribution contribute disproportionately to the emissions and emission factor estimates. Thus, the 40 vehicles which emit more than 50 g/mi of CO account for only 12 percent of the vehicles tested, but they account for nearly half (47%) of the total CO emissions from vehicles tested. Furthermore, since HC emissions are strongly correlated with CO emissions in this data base, it is likely that these same high CO emitters contribute a similarly large fraction of the HC emissions from vehicles tested.

These vehicles which give high emissions in the testing program presumably represent a fraction of the in-use vehicle population which contributes a much-higher-than-average share of CO emissions to the atmosphere in the SCAB. It follows that, in order to estimate overall CO emissions from the in-use population with acceptable accuracy, special emphasis should be placed on assuring the adequacy of representation of these high-emitting vehicles in the ARB's emission testing data base.

It is, of course, not possible to select vehicles for testing on the basis of their emissions which have not as yet been measured. However, if some sections of the vehicle population are known to include more high-emission vehicles (which we shall hereinafter call, "outliers") than other sections, it is logical to sample a relatively larger proportion of the former group than of the latter. (Since outliers contribute a large fraction of the total emissions, it is more important to know their contribution accurately than to know accurately the relatively minor contributions of low-emitting vehicles).

To try to identify vehicle subgroups having high incidence of outliers, PES again examined the data base from LDVSP 5 (ARB 1982E). Specifically, PES investigated the characteristics of the vehicles which were found to be outliers, in regard to model year, manufacturer, and engine size, as well as owners' income level, to determine whether particular groups would show significantly higher or lower incidence of outliers than the entire sample.

The results are shown in the contingency tables in Table 3-8, which also indicates the values of χ^2 and the corresponding probability level for each of the groupings tested. Each contingency table shows the two-way distribution of all vehicles tested, with one classifying factor being CO or HC emissions in each case and the other, respectively, model year, engine size, manufacturer, and reported household income.

Results of the chi-square (χ^2) test are shown below each contingency table in Table 3-8. χ^2 is a statistic based on the deviation of table entries from strict proportionality; if the value found is zero, or near zero, the series of entries in separate rows or columns are proportional or nearly proportional to each other, and there is no reason to suspect that outliers will be found predominantly in one group or another. On the other hand, if the calculated value of χ^2 is large, the probability is small that it would arise from a table based on sampling from a uniform sample. In this case, the "null hypothesis", that outlier frequency is independent of the other chosen classifying variable, must be rejected.

Thus, when the entire sample is classified by model year (see Table

Table 3-8. CONTINGENCY TABLES AND CHI-SQUARE TESTS RELATING TO DISTRIBUTION OF OUTLIERS FOR CO AND HC EMISSIONS, LDVSP 5 (Outliers defined as CO emissions above 50 g/mi or HC emissions above 5.0 g/mi)

	CO			HC		
Model Years	Outliers	Others	Total	Outliers	Others	Total
1971-1976	33	97	130	17	113	130
1977-1980	7	202	209	3	206	209
All	40	299	339	20	319	339
	$\chi^2 = 35$	$p < .001$		$\chi^2 = 20$	$p < .001$	
Engine Size						
<318 c.i.	17	231	248	8	240	248
≥318 c.i.	23	68	91	12	79	91
All	40	299	339	20	319	339
	$\chi^2 = 20$	$p < .001$		$\chi^2 = 12$	$p < .001$	
Manufacturer						
GM, Ford, Chr.	32	178	210	17	193	210
Others	8	121	129	3	126	129
All	40	299	339	20	319	339
	$\chi^2 = 5.4$	$p < .02$		$\chi^2 = 4.8$	$p < .05$	
Owner Income						
To $20,000	12	79	91	6	85	91
$20-30,000	4	79	83	3	80	83
Over $30,000	13	78	91	9	82	91
All	29	236	265	18	247	265
	$\chi^2 = 5.8$	$p \sim .06$		$\chi^2 = 2.4$	$p \sim .30$	

3-8), there are 130 vehicles of years 1971 through 1976, of which 33 (or 25 percent) are outliers with respect to CO emissions. There are 209 vehicles of years 1977 through 1980, of which 7 (or 3.3 percent) are CO outliers. χ^2 for this table turns out to be 35, which is so large that there is less than one chance in a thousand that it would be reached if the sample came from a population in which CO outliers were equally prevalent in older and newer vehicles. We therefore conclude that CO outliers mark a significantly higher proportion of vehicles of model year 1976 and earlier than of later vehicles.

In a similar way, other results shown in Table 3-8 indicate that there are significantly more CO outliers in vehicles of engine sizes 318 cubic inches and over, and in vehicles manufactured by three major domestic auto producers as compared to other vehicles, mainly imported. However, the contingency table for socioeconomic status reveals no statistically significant relation of CO outliers to owners' incomes; there is a deficiency of outliers in the middle income class ($20 to $30 thousand), as opposed to those with incomes either higher or lower, but it is not significant at the conventional 5 percent level.

Results of the χ^2 tests with respect to outliers (above 5 grams per mile) of HC emissions are generally parallel to those for CO emissions, as can be seen from Table 3-8. Again, a higher proportion of outliers is characteristic of older vehicles, larger engines, and major manufacturers, whereas owners' income status has no significant relation to the proportion of outliers.

A computer analysis using the program BMDP4F, Log-Linear Model, on 4-way frequency tables showing the distribution of outliers (HC and CO separately) against model year, engine size and owner income, confirmed the results listed in Table 3-8, showing significant relations of outliers to model year and engine size but not to income. Higher order interactions were not significant at the 5 percent level.

The indicated conclusion of these tests regarding the proportion of outliers in various subgroups of the vehicle population is that both HC and CO outliers occur predominantly among pre-1977 vehicles of major domestic makes having large engines. As a consequence, it should be possible to improve the overall accuracy of estimating areawide motor vehicle emissions by increasing the emphasis on sampling these vehicles. (The standard error of estimation is inversely proportional to the square root of the number of vehicles sampled; thus the mean for a sample of 16 vehicles is twice as accurate as for a sample of only 4.). However, this possibility will be important only if the most serious limitations on the accuracy of estimating areawide emissions do, in fact, reside in the limitations of the LDVSP data base. This possibility will be further considered in Section 5.

It may, at first glance, appear that there is some contradiction between the results presented in this section and those reported in Sections 3.3.5 and 3.3.6, where it was shown that engine size and manufacturer were not significantly related to emissions (within model years) and were not significant predictors of emission levels, according to an analysis done for EPA (EPA 1982). The contradiction is not real, however. The comments in EPA's memorandum were directed especially toward prediction of emission factors for newer vehicles, as shown by the wording: "were not significant predictors . . . by FY '79." These newer vehicles have been subject to increasingly stringent emission requirements, which have stimulated the development of increasingly effective control technologies, resulting in much smaller emission levels and less variability between vehicles of various makes and engine sizes. In this situation it is understandable that the statistically most significant effects on emissions are those found on dividing the vehicle population by model years.

4.0 EVALUATION OF EPA VEHICLE COMPLIANCE PROGRAMS

U.S. Environmental Protection Agency (EPA) designs, conducts (through contractors) and implements the vehicle recall program to ensure that current in-use motor vehicles are meeting the emission standards throughout the 5-year or the 50,000 mile emission warranty period. This chapter first reviews both EPA's and ARB's recall enforcement programs and then evaluates the EPA program as to which aspects may be applicable for improvement of ARB procurement programs in the future.

EPA uses a cluster sampling method in obtaining in-use vehicles for emission testing. Specifically, EPA selects three or four metropolitan areas in different parts of the country, and procures and tests vehicles obtained from those study areas. Since motor vehicles in the cluster sampling areas constitute only a small fraction of the national vehicle population, the study areas must be carefully chosen so that the combined vehicle population over the study areas will reasonably approximate the national vehicle population mix. In order to implement proportional sampling throughout the clusters, the sample plan must be tailored to each study area by taking into consideration the vehicle population mix in the area.

ARB's in-use vehicle surveillance program can also be viewed as a sort of cluster sampling, since the vehicles tested are procured only from the SCAB, ordinarily. However, the cluster represented by the SCAB constitutes about a half of the statewide vehicle population. Therefore, a sample representative of the SCAB vehicle population is, in practical terms, representative of the California vehicle population as well, unless the vehicle mix of the SCAB is markedly different from that of the overall California vehicle population.

PES has reviewed an EPA guideline document (EPA 1981), which is given to EPA contractors to ensure a consistent procurement operation from one procurement program to another. Attachments II through IV of the guideline document spell out all major steps involved in the procurement operation. Such steps are diagrammatically illustrated in Figure 4-1. For easy comparison, a similar diagram for ARB procurement is shown in Figure 4-2.

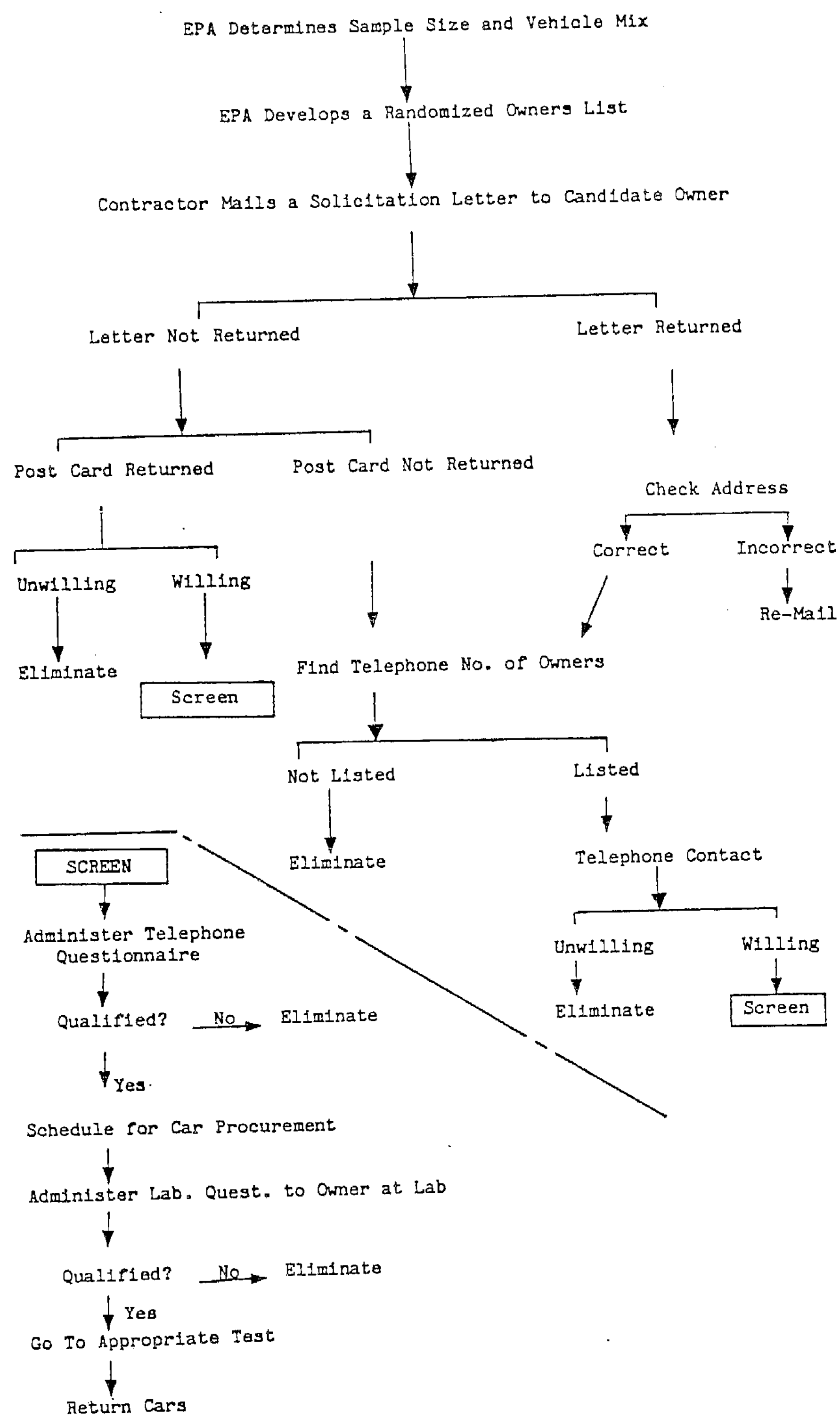


Figure 4-1. EPA Procurement Diagram

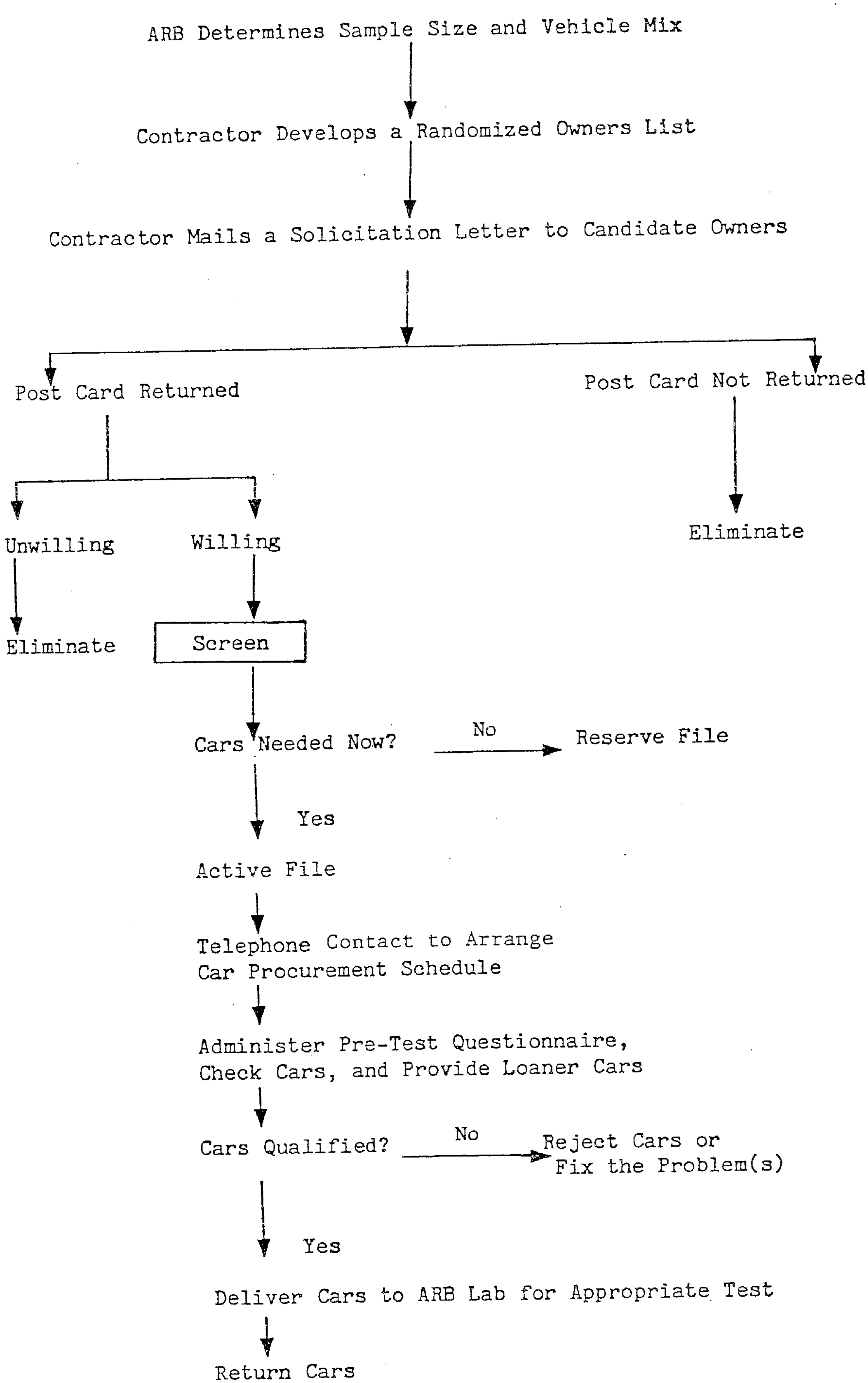


Figure 4-2. ARB Procurement Diagram

Major differences between ARB and EPA procurement procedures are:

1. While ARB lets its contractor select candidate owners randomly from vehicle registration records, EPA provides the contractor with an owners list already randomized in a statistically sound manner.

2. While ARB's contractor contacts only those owners who have returned the reply post card indicating willingness to participate, EPA's contractor contacts, in addition, those owners who have not returned the reply post card.

3. EPA's contractor is not allowed to procure a candidate vehicle until all candidates above it in the list have been either determined to be acceptable for testing or eliminated.

4. EPA's owner-screening criteria appear to be stricter than those of ARB, and are implemented rigorously by its contractor.

As to the first difference, the EPA method eliminates uncertainty as to how well the contractor designs and implements its sampling plan. "Random selection" is an objective, unbiased, statistically sound method of selecting study items without any preference. Random selection, however, is often misinterpreted as a synonym for "arbitrary selection", which is usually subjective and biased by the selector's habits and preferences. Present ARB procedures leave much more discretional choice to the contractor in selecting candidate vehicle owners.

As to the second difference, the EPA method increases considerably the cost of identifying "willing" candidates but it helps little to improve the representativeness of samples so obtained or to raise the final capture rate. According to an EPA contractor in Virginia, 50% to 60% of the candidates to whom the solicitation letter is mailed are eventually telephoned. In spite of such an intensive effort, the final capture rate is about 5%. (This figure would be higher if the contractor did not reject owners whose vehicles fail to meet EPA criteria for the testing program. These owners account for 20% to 50% of the "willing" candidates.)

As to the third difference, the EPA method imposes on the contractor a burdensome inflexibility in the vehicle procurement operation. EPA's intent is to eliminate an unknown bias which may be caused by behavioral differences between so-called "eager beavers" and those who are more wary of any kind of solicitation.

The fourth difference is only conjectural and has not been substantiated with documentary evidence. However, the rejection rates of "willing" candidates seem to be considerably higher among EPA contractors than among ARB contractors. To be qualified for the EPA recall program, the vehicle owner must have adequate records of manufacturer-specified maintenance for his car and must also be willing to appear in court as a witness, if required to do so.

5.0 DISCUSSION AND BASIS OF RECOMMENDATIONS

The overall purpose of this study, as defined by ARB, was to evaluate the adequacy of vehicle sample selection methodologies in relation to three major ARB activities: (1) establishing emission factors for mobile source emission inventories; (2) setting emission standards for new motor vehicles; and (3) investigating the durability of emission control systems. In particular, methods used by ARB's procurement contractors were to be critically reviewed and the methods used by EPA in vehicle compliance testing were to be considered, in order to generate recommendations for possible improvements in selection of vehicles for ARB's "in-use" vehicle surveillance program.

ARB has defined the objective of the procurement contractor as, "to select a random sample of vehicles that is representative of the population of vehicles of each engine size, manufacturer and model year being driven in California." Thus, a particular emphasis of this study has been to evaluate the procurement contractor's success in achieving this objective.

On first examination of the existing arrangements for procurement of vehicles, two limitations became apparent. First, it would not be feasible for any procurement contractor to achieve a scientifically random sampling of motor vehicles in the South Coast Air Basin. Second, of the ARB activities which depend on the "in-use" vehicle testing program, the establishment of vehicle emission factors for the mobile source inventories is the one which most critically depends upon obtaining a data base that would be adequately representative of the California motor vehicle population.

5.1 RANDOM SAMPLING AND PROBLEMS OF BIAS

Scientifically, a random sample can be defined as a sampling of items from a population, in which each item in the population has an equal chance of being selected and each item is selected independently of each other item. In the ARB "in-use" motor vehicle surveillance program, it is impossible to achieve a random sample of this sort because some components of the population are excluded from the sample. In the first place, certain types of vehicles are ineligible by ARB policy, e.g., rental cars, fleet vehicles, and law enforcement vehicles. In the second place, dependence on a voluntary participation system automatically excludes from the sample all vehicles whose owners decline to take part in the program.

These exclusions from the sample universe may or may not have an appreciable effect on the quantities estimated for use in emission inventories (i.e., on emission factors). If the excluded vehicles and those withheld from the program by owners have essentially the same emission distribution as those actually obtained for the sample, then no error will be made by relying on the sample to represent the population. Unfortunately, there is no apparent method for determining the emissions of these unsampled vehicles, and there is no convincing rationale for assuming that they are the same as those sampled. Thus, even if the procurement contractor were to deliver a truly random sample of the available population universe, that sample could not be guaranteed to yield reliable statistics for the entire in-use vehicle population.

Implicitly recognizing these difficulties, ARB has directed its procurement contractors to use a stratified sampling system, with quotas for sample sizes specified by model year, manufacturer and engine size, and to apply random sampling techniques within the specified vehicle categories ("strata"). This eliminates any requirement for the contractor to randomize the sampling of vehicles as between model years etc., but it leaves unchanged the problems of bias due to excluded classes of vehicles and due to the large fraction of non-participating vehicle owners.

Thus, given the constraints imposed upon the procurement contractor by ARB and by real world conditions, the contractor cannot guarantee to achieve a scientifically random unbiased sample of vehicles which will accurately represent the actual vehicle population.

5.2 REQUIREMENTS FOR ACCURACY OF DATA BASE

Reviewing the major activities for which ARB intends to use data from the motor vehicle testing, we must consider what are the necessary characteristics of the data base to serve each purpose to best effect.

First, to establish emission factors for developing mobile source emission inventories, an adequate sampling of all classes of vehicles which constitute the in-use population must somehow be achieved.

Second, to provide adequate information for setting emission standards for new motor vehicles, it would appear that the accuracy required would be easily assured if reasonable confidence limits for the first objective were achieved.

Third, for investigating the durability of emission control systems, each such system to be assessed could be the subject of a specially designed survey protocol, defining numbers of vehicles required as well as any special conditions for the procurement. If durability information is to be derived from the same data base used for the previous two objectives, its accuracy would be constrained by the proportion of the data base devoted to vehicles equipped with the system in question.

Adequacy in this context may be arbitrarily defined as the achievement of 95 percent confidence limits no more than 10 percent above and below the emissions index which is ultimately to be estimated. This implies that the standard error of the emissions estimate should be not more than 5 percent of the mean.

This is in fact a very stringent goal. It may not even be currently feasible, in view of some technical factors as well as resource constraints. Nevertheless, an emission inventory total that is intended for use in air quality management planning, but cannot meet the criterion of $\pm$ 10 percent accuracy, will not be generally perceived as fully satisfactory. Therefore, we have assumed that such accuracy, even if not currently attainable, should ultimately be a goal of ARB.

5.2.1 EMISSION FACTORS

Since the emission factors constitute only one part of the input required to calculate total emissions, 10 percent accuracy in determining emission factors does not automatically assure 10 percent accuracy in estimating emission inventory totals. There is little doubt that, in most emission inventory calculations, the activity factors (VMT, numbers of trips, etc.) are no more accurately known than the emission factors. However, in ARB practice, determination of activity factors is the responsibility of a separate organizational division (the Stationary Source Control Division, SSCD) and is outside the scope of this study. We therefore assume that the desired accuracy in determination of emission factors is such that a surrogate emissions total or "composite emission factor," based on the vehicle testing data base and appropriately weighted by activity factors assumed known <u>without uncertainty</u>, will be accurate to within ± 10 percent.

It is true that the composite emission factors actually applied by SSCD are not the emission factors supplied by MSCD but are derived by a detailed process (see Section 2.3.2) in which the values furnished by MSCD are adjusted for the effects of road conditions and weighted by VMT and trip estimates corresponding to local conditions in whatever area is under study. It is also true that MSCD furnishes its emission factors not in the form of individual estimates for individual matrix elements (calendar year, model year, odometer mileage, etc.) but in the form of linear equations (see Section 2.3.2) having coefficients whose accuracy is unspecified. However, the chosen approach, involving consideration of accuracy in terms of a surrogate emissions total, obviates much complicated analysis and the necessity for several additional arbitrary assumptions in favor of one general assumption: that the vehicle testing data base will be adequate if it provides for standard accuracy when applied to any vehicle distribution of the 1980's. Thus, in this report, we discuss procurement problems in the light of their effect on the accuracy of total emissions (or, what is mathematically the equivalent, the overall composite emission factors) for such a surrogate distribution.

This study has not focused on the problem of determining the durability of emission control systems. Information relating to this endeavor has been published (ARB 1981d) (not by MSCD, but by SSCD) in the form of tables of new-vehicle emission rates and deterioration rates for 11 classes of vehicles, to be used in the estimation of area-wide motor vehicle emissions for emission inventories. An example of such a table is shown as Table 5-1. The tabulated coefficients are derived from the

Table 5-1. EXHAUST EMISSION RATES-ARB[1]

LIGHT-DUTY PASSENGER CARS

POLLUTANT	MODEL YEAR	A (GM/MILE) NEW VEHICLE EMISSION RATE	B (GM/MILE) DETERIORATION RATE (PER 10,000 MILES)
HC	PRE-1966	9.580	0.180
HC	1966-1967	4.840	0.250
HC	1968-1969	4.310	0.250
HC	1970-1971	3.700	0.370
HC	1972-1974	3.020	0.170
HC	1975-1976	0.480	0.350
HC	1977-1979	0.280	0.210
HC	1980-1982	0.330	0.200
HC	1983-1984	0.290	0.210
HC	1985+	0.250	0.210
CO	PRE-1966	79.310	2.250
CO	1966-1967	57.410	2.550
CO	1968-1969	53.750	2.550
CO	1970-1971	46.430	3.130
CO	1972-1974	33.290	2.440
CO	1975-1976	7.730	2.780
CO	1977-1979	4.320	2.220
CO	1980	5.720	2.440
CO	1981-1982	5.440	2.440
CO	1983-1984	4.750	2.100
CO	1985+	3.900	1.760
NOx	PRE-1966	2.930	0.000
NOx	1966-1970	4.060	0.000
NOx	1971-1973	3.400	0.040
NOx	1974	2.200	0.040
NOx	1975-1976	2.030	0.060
NOx	1977-1979	1.510	0.070
NOx	1980	0.760	0.120
NOx	1981-1982	0.560	0.100
NOx	1983+	0.350	0.080

[1]Source: Supplement 2 to Procedure and Basis for Estimating On-Road Motor Vehicle Emissions, June 1981. State of California Air Resources Board, Stationary Source Control Division, Emission Inventory Branch, Motor Vehicle Emissions and Projections Section. (ARB 1981d)

vehicle testing data base by linear regression analysis.

It is not clear that the practices of the vehicle procurement contractors affect the accuracy of these deterioration factors except in the most general sense, that a biased sample of any population is likely to yield inaccurate statistics.

5.3 THEORETICAL OPTIONS FOR SAMPLING DESIGN

To provide the data needed for developing emission inventories, the LDVSP data base will be most effective if it contains information covering a statistically representative set of vehicles. The principal difficulties in obtaining such a data set may be listed as:

(1) exclusion of certain segments of the vehicle population for administrative or logistic reasons

(2) costs of sampling and testing adequate numbers of vehicles

(3) the problem of representing a continually shifting population.

The principal theoretical options for implementing a sampling program may be listed as:

(1) continuous random sampling

(2) stratified sampling (e.g., present ARB practice)

(3) incremental quota sampling

(4) intensive sampling of population segments (e.g., to study mileage dependence of emissions).

Following is a brief discussion of the advantages, disadvantages and status of each option.

5.3.1 CONTINUOUS RANDOM SAMPLING

We have noted (see Section 2.2.2) that completely random sampling of the vehicle population is administratively difficult, probably prohibitively expensive, and perhaps practically unfeasible under current conditions. It is also inconsistent with ARB's current policy of excluding certain classes of vehicles (or vehicle owners) from participating in the testing program.

Nevertheless, in principle, continuous random sampling offers the most direct way of obtaining accurate representation of the real on-road vehicle population. If implemented and continued over a long enough period of time, it could conceivably yield an emissions data base for a sample population which would reflect the actual daily activity of the various types of vehicles being driven in a study area, as well as trends in the vehicle population over time.

We suggest that the implementation of the currently mandated California Motor Vehicle Inspection Program may provide an opportunity for ARB to implement such a program. If all California vehicles, or all in a region such as the South Coast Air Basin, are required to undergo emissions inspection on a regular basis, it should be possible for ARB to institute a systematic survey of the entire vehicle population and to select a given proportion, on a random basis, for the more elaborate testing done by MSCD. If participation in the extended testing sequence could be made automatic, the problem of possible bias because of non-participating owners would vanish. Otherwise, perhaps incentives to participate could be set high enough to ensure a large capture rate and thus minimize the likelihood of appreciable bias due to non-participation.

Another approach would be to mount a continuing survey program to compile a random list of license plate numbers observed in traffic, then obtain by solicitation each vehicle thus selected, or an acceptable surrogate vehicle. For example, a list of principal traffic links could be compiled and one randomly selected for visiting on each day of the year; on each visit, one or two passing vehicles could be selected, using a standard protocol to assure a random choice; then a procurement contractor would identify the owner and persuade him to participate in the ARB testing program. On failing to do so, the procurement contractor would seek to obtain, as an alternate, some similar vehicle from a more willing owner. This procedure would not eliminate the problem of bias due to voluntary participation, and it would be substantially more costly than the current procurement program. It would offer the advantage that the sample obtained would be directly representative of vehicles in use, eliminating the problem of setting quotas for various vehicle population segments (which is one source of uncertainty in the current program).

5.3.2 STRATIFIED SAMPLING

Stratified sampling provides an approximation to the results of random stratified sampling and is based on essentially the same rationale. If the entire vehicle population can be divided into a series of subpopulations ("strata") whose numbers are known with reasonable accuracy, then the population mean can sometimes be determined more conveniently by independently determining the means for the various strata and applying standard statistical algorithms to combine them. In particular, if the variances of the various subpopulations differ substantially from each other, it may be possible to enhance the efficiency and improve the accuracy of the desired overall mean by appropriately adjusting the proportions to be sampled for each subpopulation.

ARB, in its vehicle procurement contracts, attempts to specify the sample numbers appropriate to the subpopulations rather than to leave those numbers to the contractor's discretion (ARB 1983a). It is intended that, within each stratum, vehicle selection will be random. It is also intended that each stratum will be represented by the same proportion of the total population of such vehicles in the study area (at least for the subpopulations which compose the population of vehicles of any one model year; quotas for different model years have apparently been formulated on a more intuitive basis).

In its LDVSP procurement contracts, ARB has typically specified quotas for as many as 200 strata, based on manufacturer, model year, vehicle type (car or truck) and engine size. Some quotas in LDVSP 5 were for as many as seven vehicles, but most were for only one or two vehicles. The quotas are determined in proportion to estimates of the real subpopulations, obtained from State registration data and manufacturers' reports of production of various engine sizes.

A similar sampling methodology described by EPA (EPA 1982) has been criticized by Survey Research Center of the Institute for Social Research, which commented (EPA 1982) "no one really knows how well quota samples reflect the population of interest . . . actual bias is highly likely if there is any relationship between quota criteria and other factors influencing vehicle emissions."

While stratified sampling imposes a requirement for accurately assessing the populations of the strata, it also opens options for optimizing the accuracy and efficiency of sampling in relation to the various purposes which may be served by the data base. Appendix A offers a brief discussion of some options for optimizing sampling quotas in the VSP programs.

The power of random stratified sampling (see discussion, Section 2.3.1) diminishes when the number of items sampled from each stratum becomes too small. In ARB's sampling scheme, many individual quotas are only one or two vehicles. This degree of stratification seems likely to be counterproductive, as it interferes with the potential randomness of the sampling procedures and, in effect, substitutes an arbitrarily defined sample for a random sample of the population.

Stratification of a population for sampling purposes confers no improvement in accuracy of estimating the mean unless the individual strata have distributions that differ significantly from that of the overall population. In particular, if the true means of the strata do not differ significantly, stratification will not be advantageous as compared to random sampling.

When sample quotas for individual strata are too small, the conditions for demonstrating significant differences between strata means become difficult to meet, because the confidence intervals surrounding those means are often inordinately wide. However, the problem of overstratification can be avoided or alleviated by combining non-essential strata and thus reducing the number of strata. For example, the nearly 340 vehicles allotted for testing in LDVSP 5, if stratified by model year only, would allow for average sample size of 34. Any further substratification would begin to be excessive if it produced more than 5 strata within each model year. Thus, one reasonable substratification would divide the vehicles of each model year into 5 groups: perhaps one for each of three major US manufacturers, one for European and other American manufacturers, and one for Japanese manufacturers. However, a stratification into five sets of manufacturers and two ranges of engine size would be excessive, because this would produce 100 strata (10 model years x 5 manufacturers x 2 engine size ranges) and an average sample size less than 4 vehicles.

Another way to reduce the number of strata, so as to facilitate the statistical interpretation of vehicle emissions test data, is to sample only vehicles of every second model year, instead of all model years, in each annual surveillance program. This strategy can also lead to a new system of estimating both areawide composite emission factors and annual rates of deterioration of those emission factors, with certain advantages over the current system.

To sample in this manner, for example, in 1983, vehicles are sampled from the half of the vehicle population consisting of model year vehicles of 1983, 1981, 1979, 1977 and 1975. In the next year, 1984, vehicles will be sampled from the other half of the vehicle population consisting of 1984, 1982, 1980, 1978, and 1976. Thus, the sample size for each model year will be doubled in each surveillance program, compared to the currently used sampling scheme.

The above sampling scheme provides emissions test data for 1-year-old cars, 3-year-old cars, 5-year-old cars and so on. The lack of even-age cars will not cause any significant data loss; for the purpose of determining emission deterioration factors, the spread of accumulated miles among those odd-age cars would be as large as that among every-age cars. Increased sample sizes in model year strata might enable ARB to derive an emission deterioration rate for given-age cars directly from each model year stratum instead of going through the conversion of a regression relationship between emissions and accumulated mileage on individual cars into the less direct relationship for ages of vehicles.

5.3.3 INCREMENTAL QUOTA SAMPLING

It is a point of some importance that, in ARB's LDVSP data base, the yearly quota samples do not constitute the sole or even principal source of the data which must be utilized in providing emission factors for use by SSCD. Rather, the new information each year constitutes an increment to the data base; it yields the first tentative evaluation of the emissions from new vehicles reaching the highways, and it bolsters the determination of emission factors and deterioration factors for previous years by further increasing the body of testing data available for statistical interpretation. Viewed in this light, it appears perhaps less critical to assure accurate representation of the extant motor vehicle population of the current year than to assure that the entire data base, accumulated over several years, contains the data required to permit accurate estimates of emissions and trends in emissions during those years for emission inventory and planning purposes. If there are important sources of bias, as there may be, which affect the annual data increments year after year, then it is to be expected that the integrated data base is subject to a similar bias. However, for the sake of maintaining continuity in the ability to assess trends, it may be important to evaluate any important source of bias thoroughly before deciding whether and how to eliminate it.

What may be called "incremental" quota sampling is envisioned as a system which would operate in essentially the same way as ARB's current quota sampling system. However, the selection of strata to be represented and of quotas to represent them would be specifically adjusted to optimize the usefulness of the resulting cumulative data base for emission inventory purposes.

We do not wish to imply that ARB has neglected the cumulative aspect of the data base in previous vehicle procurement quotas. Although total quotas for various model years are said to be determined arbitrarily, it is evident from the tables given in the contract solicitations (and in the final reports of each testing series) that there is in fact always a greater emphasis on testing more recent model years than older ones. (See, for example, A-2 in Appendix A.)

It is, of course, fully appropriate that the quotas for new vehicles be larger than for older ones, since the cumulative data base already contains much information regarding the older vehicles, but none regarding the newest. The purpose of incremental quota sampling would be to capitalize on the same principle by explicitly considering the information already in the data base in order to calculate quotas which would supplement that information most efficiently.

To apply these principles on an incremental quota basis, the quotas would be calculated for the entire data base (rather than for one year's procurement as shown in the example) and quotas for the year would be determined by difference between the calculated total quotas and the sample numbers already in the data base. However, the examples discussed in Appendix A illustrate that the required quota adjustments might very well be major adjustments.

5.3.4 INTENSIVE STUDIES OF SPECIFIC SUBPOPULATIONS

For some purposes it may be desirable not to rely solely on random or quasi-random sampling of the general vehicle population, but to develop a special sample designed to provide particular technical information. For example, we have seen that ARB would like to study the durability of control systems; but we have also observed that the regression analysis approach, applied to routine testing data accumulated in the LDVSP data base, may yield uncertain results and, therefore, questionable interpretations. Again, in order to decide on the desirability or necessity of requiring a manufacturer to recall and repair vehicles of a certain cohort, it may be necessary to sample that cohort much more intensively than the general vehicle population.

In assessing the durability of a specific control system, an analysis based on repeated testing of a relatively small set of vehicles is likely to be much more productive than an accumulation of single tests on randomly selected cars. While random sampling and statistical evaluation can furnish a very good data base for characterizing the emissions of a wide-ranging population with a multiplicity of vehicle types and ages, it is really of very little use in investigating the behavior of individual vehicles and the changes in their emissions with accumulated mileage or time. Such investigations are more likely to profit from an emphasis upon detailed observation and experimentation to illuminate the physical causes of system deterioration.

The statistical approach minimizes precisely these detailed aspects of vehicle behavior which might be used to create a useful mathematical model of system deterioration. Further, such a model is very likely not to conform to the assumptions which are implicit in the use of linear regression analysis. In short, an investigation of control system durability is a type of study to which simple survey procedures may be unsuited.

On the other hand, sampling for recall testing is a case in which statistical procedures will usually be important. However, the most appropriate statistical procedures for recall testing may be a different set from those most useful in estimating areawide emissions. In planning for such testing, therefore, numbers of vehicles sampled should be predicated on these statistical aspects. From the fact that ARB procurement quotas for recall testing purposes ("special sampling") (ARB 1982b,c,d,e) have been substantially higher than quotas of the same type of vehicles for surveillance testing, it appears that ARB is aware of these considerations and has been applying them. We therefore offer no further discussion of the statistical aspects of recall sampling at this point.

5.4 IMPLICATIONS FOR VEHICLE PROCUREMENT SYSTEMS

5.4.1 EFFECT OF CAPTURE RATE ON BIAS IN DATA BASE

In Section 5.1 we have argued that vehicle procurement contractors cannot guarantee to provide unbiased samples of the existing vehicle population because (a) some vehicles are excluded by policy and (b) some vehicles will not be made available by their owners. However, bias introduced by these conditions may not be serious, depending on two factors: (1) how sharply each excluded subpopulation differs from the remaining population which is adequately sampled, and (2) how large the unsampled population is in relation to the sampled population.

For example, if the population of vehicles excluded for policy reasons is small, and if there is no reason to suppose that such vehicles predominantly have excessive emissions, then it can be concluded that the effect on the composite emission factors due to excluding this subpopulation will be small. If this is the case, very likely there is no reason to be concerned about this effect.

On the other hand, the population of vehicles not made available by their owners can only be considered large, in view of the small capture rates reported by procurement contractors. In both ARB and EPA vehicle surveillance programs, available information indicates that capture rates exceeding 10 percent are uncommon. From this we may infer that the population of unavailable vehicles is of the order of magnitude of 90 percent. In this case, a relatively minor difference in the distribution of emissions between the sampled and unsampled populations could have very serious effects on the accuracy of estimating areawide emissions.

Low capture rate is certainly a source of many unknown biases in the samples obtained in the past programs or to be obtained in future programs. However, there is no practical prospect of raising the capture rate to a level that would ensure the removal of such biases. Thus, rather than seeking a definitive solution to the problem of low capture rates, it seems more useful to examine how current procurement operations affect and are affected by the low capture rates.

The procurement contractor is required to procure a given number of vehicles of specified types under financial and time constraints, regardless of what the final capture rate may be. Under such a condition, it is natural for the procurement contractor to use his best knowledge and available resources to minimize the cost of procuring the desired vehicles. Suppose that there are two areas: one where the capture is somehow consistently high and the other where the capture rate is somehow consistently low. Assuming that other conditions are practically same for the two areas, the procurement contractor would naturally focus his effort on obtaining vehicles from the area with high capture rate rather than the other area with low capture rate. Thus, his practice of minimizing the cost naturally tends toward raising the capture rate.

However, such a practice may bring a more imminent problem of, say, oversampling middle income surburban communities where people perhaps tend to maintain their cars more carefully and undersampling low-income midurban communities where people tend not to maintain their cars. Such a systematic difference in car maintenance practice between communities could cause more serious bias in the sample than the difference in maintenance practice between people willing to participate and those not willing because the former would produce a real bias whereas effects of the latter would more likely be random.

Although it is more or less conjectural, the above discussion clearly indicates a need for mandatory sample allocations over different socioeconomic areas.

Of the model year strata, the earliest in the data base are most likely to constitute significant sources of bias. As the data from LDVSP 5 show (see Section 3.3.3), average emissions of both HC and CO are much larger for these older vehicles than for the overall data base; for model years 1971 through 1974, the averages (in grams per mile) are HC 4.3, CO 52 compared with overall averages HC 1.7, CO 21. Since the variance associated with these larger averages is also much larger than the variance for later model years, the averages for the older vehicles are less accurately known. Also, as has been discussed (Section 3.4), the proportion of outliers is significantly higher for the earlier model years.

Thus, even with the best random sampling procedures, the small numbers of vehicles sampled for these model years lead to a large uncertainty in estimating their emissions and a correspondingly large uncertainty in overall emissions. When it is also considered that the procurement contractors may be reluctant to deal with such vehicles (many of which may be disqualified for reasons of safety and driveability), it appears that a substantial bias from this source is a distinct possibility.

Such bias, if it exists, is most likely to cause underestimation of areawide emissions (because existing excessive emitters will not be proportionately represented). A mitigating factor is that the relative populations of vehicles of earlier model years is clearly declining, so that bias for those years may become less important. However, the problem will no doubt increase in later model year cohorts as more of those vehicles outlast their warranty periods and their control systems deteriorate.

5.4.2 APPROACHES FOR CONTROLLING BIAS

It is basically not possible to eliminate bias from any survey or sampling operation, since the possible sources of bias are legion and many are highly speculative in nature. To reduce possible biases to an acceptable level, the designers of a survey must endeavor to identify the possible important factors capable of causing bias; knowing these, the principal options are (1) to eliminate them by redesigning the sampling operation and (2) to account for them by evaluating them and appropriately adjusting survey results.

Fortunately, many conceivable sources of bias are of little consequence in the execution of a program, but it is important to evaluate the possibilities at an early stage in program planning. In the case of ARB's in-use vehicle surveillance programs, data are already available, from several years of testing, which can facilitate judgments as to the probable importance of various sources of bias. Some of the statistical procedures which are useful in this context have been illustrated by example in Section 3.3, using the LDVSP 5 data. They indicate, in a preliminary manner, that biases due to inaccurate representation of model year subpopulations are potentially serious; that biases in sampling by manufacturer and engine size may exist but are probably less important; that biases due to inadequate geographical and socioeconomic sampling appear to be minor.

In view of these considerations, it is logical to recommend that the accumulated data base from the in-use vehicle surveillance program be statistically reviewed annually, not only to determine the appropriate quotas for sampling by model year, but also to monitor the importance of recognizable sources of bias and to devise an appropriate partition of model year quotas into quotas for the various subgroups. An analysis of the error bounds associated with the statistical parameters (emission factors and deterioration factors) reported to SSCD for use in estimating areawide emissions should also be performed. Comparison of such results from year to year would indicate whether the program could achieve consistent improvement in the efficiency of the sampling design and overall accuracy of the results.

Besides the simple exploration of the data base and annual monitoring of its quality, other means of evaluating and adjusting for possible bias in sampling exist. The imminent implementation of a California vehicle inspection and maintenance program may provide an opportunity to compile relevant statistics on the emission status of practically all vehicles which belong to specified categories. The inspection systems used in the MVIP program will, of course, not be equivalent to the CVS testing done at the MSCD laboratory, but the distribution of failures in the MVIP program should show a rather high correlation to statistics which can be derived from the VSP data bases. It would be advantageous, if possible, to provide for a continuous and fairly intensive statistical monitoring of the consistency between data from the two sources. Exploration of the relations and discrepancies should help to evaluate suspected sources of bias and, perhaps, reveal some unexpected ones.

Finally, to deal forthrightly with the problem of bias due to low capture rates, various experimental approaches can be devised. Incentives to participate can be varied, and they should cause some variations in capture rates. If parallel testing programs could be run based on different incentives and therefore different capture rates, there would be an opportunity to determine the sensitivity of the estimated emission factors to the capture rates.

Other sources of bias can also be investigated by special surveys and experimental approaches. It is likely, however, that for most practical purposes such approaches will be too expensive and time-consuming to be of much use in connection with the in-use vehicle surveillance program.

5.5 USE OF INCENTIVES TO ENCOURAGE PARTICIPATION

A low response rate is one of the principal concerns expressed by many vehicle procurement personnel whom PES has interviewed under the present study. The reason for the concern is the effect of low response rates on sample bias (the lower the response rate, the greater the chance of introducing a significant bias into the sample population of procured vehicles).

The Institute of Social Reseach (ISR) at University of Michigan conducted, for EPA, the telephone interviews of both participants and non-participants in the EPA's FY '80 Emission Factor program in San Antonio. Many of the ISR interviewees have indicated that an independent report on the mechanical condition of the vehicle, or a guaranteed tune-up, could be more valuable to them than a $50 savings bond or a loaner vehicle. However, PES interviews with the procurement contractors revealed that such alternative incentives had already been tried in the SCAB. The solicitation letter of one procurement contractor implies that upon request, such an independent report on the mechanical condition and the test result will be furnished. This offer seems to increase the rate of positive response to mail solicitation.

Several procurement contractors have tried offering prospective participants a free tune-up service as an incentive. According to their experience, such an incentive tends to attract vehicles which are totally out of proper maintenance and thus introduce a bias toward poorly maintained vehicles. This incentive is also quite risky for the contractor because the cost of fulfilling the promised service varies widely from car to car and is difficult to control.

The ISR survey also identified the three major reasons for non-participation: not recalling the solicitation letter; doubt about the legitimacy of the program; and inconvenience of participating in the program. The first and second reasons reinforce the importance of contact material and follow-up contact, while the third reason indicates a need of reasonable compensation for the perceived inconvenience. One solution to this perceived inconvenience is to provide the owner with a loaner vehicle, which is practically equivalent to his own car. However, this type of incentive is quite costly because the contractor would have to acquire a wide variety of loaner vehicles in order to meet all requests.

In California, the mandatory inspection/maintenance (I/M) program is scheduled to start in March, 1984. Under this program, owners will be required to have their vehicles inspected every other year at established emission test stations. Such requirement for regular, mandatory inspection will be considered inconvenient and annoying by many motorists. Therefore, it may be suggested that a waiver for the I/M program be offered to vehicle owners as an incentive to participate in the VSP program. The feeling of inconvenience avoided might prove to be a substantial stimulus to owners in various categories to make their vehicles available for testing.

5.6 SUMMARY

In this chapter we have reviewed the objectives of this project in relation to ARB's stated information needs. The sampling system described by ARB does not correspond to a "random" or "proportional" sampling design, but rather to a "stratified" sampling system. In such a system, the procurement contractor is given a list of types of vehicles with numbers of each type to be procured, and is instructed to sample randomly within each type until its quota is reached.

Although there are substantial elements of randomicity involved in this procedure, there are two main impediments to the achievement of true random sampling, even within the strata defined by the quotas. These arise from (1) the exclusion of certain vehicles by ARB policy; (2) the unwillingness of most vehicle owners to make their vehicles available for testing. An unknown amount of bias in the assessment of emission factors arises from each of these causes, but the second cause appears the most important, in view of low vehicle capture rates (5 to 10 percent).

In relation to conventional accuracy criteria (95 percent confidence limits at + 10 percent about an estimated mean), it is seen that such accuracy will be very difficult to achieve, gauged by distributions of emission factors for HC and CO as observed in LDVSP 5. Further difficulties are imposed by the use of a linear regression approach to represent the relation of emissions to miles driven and by the application of uncertain activity factors (VMT) in the ultimate calculation of areawide emissions.

Theoretical options for the modification of sampling design include:

1. continuous random sampling
2. quota sampling similar to present practice
3. incremental quota sampling based on needs for supplementing the accumulated data base
4. intensive sampling of population segments where special information is required.

These options are not at the disposal of the procurement contractor, but should be considered by ARB in designing solicitations for the procurement contract.

Study of the data from LDVSP 5 indicates that the procurement contractor is not in a position to reduce known sources of bias to any major extent, since it appears that differences in emissions as related to geographic and socioeconomic factors are, at worst, minor. On the other hand, potential bias due to the unavailability factor could be somewhat reduced by improving capture rates, perhaps by increasing incentives to owners; but this would be an expensive option. Furthermore, since this bias cannot at present be estimated, any benefit from reducing it must be considered highly speculative.

Improvement of the adequacy of vehicle sampling appears to depend much more directly on the improvement of the plan than on possible improvements in the technical procedures of the contractors. A quota sampling plan, like that in use at present, may be the only practical option for the in-use vehicle surveillance program, in view of existing constraints on resources. But there are good prospects of improving the overall accuracy of emission factors determined in the program, by directing due attention to elements of statistical design when formulating the sample quotas.

Most importantly, the annual sampling plan should be designed with explicit reference to the status of the existing data base. The accuracy (i.e., the standard deviation) associated with each of the recognized subpopulations in the data base should be determined, and the annual incremental quotas should be calculated to place the greatest emphasis on those subpopulations which contribute most to emission totals and those which are least adequately characterized by the existing data. Since both the emissions and the standard deviations are especially large for older vehicles, there may be a need for further emphasis on sampling of such vehicles.

Even though the exploratory statistical review performed as part of this study provides no firm evidence that geographical location is a potential source of bias in the VSP data base, reasonable geographical dispersion should be maintained at least until that potential can be more fully evaluated. Socioeconomic status is another factor which has been viewed as a potential source of bias; again, no evidence of bias has been seen, but reasonable dispersion of samples (consistent with costs of implementation) is desirable. For both factors, statistical review should be performed annually to evaluate upper bounds on the magnitude of these effects.

6.0 REFERENCES

ARB 1978
Memorandum, K.D. Drachand to A. Donnelly. Subject, Proposed Test Plan for Light Duty Vehicle Surveillance Test Program, 3rd Series, CVS (with subsequent associated correspondence).
State of California, Air Resources Board, May 22, 1978, et seq.

ARB 1979
MS-80-002, Final Report of the High Mileage Catalyst Vehicle Surveillance Test Program, 1st Series, CVS.
State of California, Air Resources Board, December 14, 1979.

ARB 1980a
Procedure and Basis for Estimating On-Road Motor Vehicle Emissions.
State of California, Air Resources Board, Technical Services Division, Data Processing and Emission Branch, Motor Vehicle Emissions and Projections Section, January 1980.

ARB 1980b
MS-80-08, Final Report of the Light-Duty Vehicle Surveillance Test Program, 2nd Series, CVS (LDVSP-II).
State of California, Air Resources Board, February 1980.

ARB 1980c
California Assembly-Line Test Procedures for 1980 Model Year Passenger Cars, Light-Duty Trucks and Medium-Duty Vehicles.
State of California, Air Resources Board, Adopted November 16, 1978, Amended January 30, 1979, Amended May 9, 1979, Amended March 5, 1980.

ARB 1980d
California Assembly-Line Test Procedures for 1981 Model Year Passenger Cars, Light-Duty Trucks and Medium-Duty Vehicles.
State of California, Air Resources Board, Adopted December 19, 1979, Amended March 5, 1980, Amended March 26, 1980.

ARB 1980e
MS-80-024, Test Report of the Light Duty Vehicle Surveillance Program, Series 3 (LDVSP-III).
State of California, Air Resources Board, November 28, 1980.

ARB 1981a
MS 81-008, Test Report of the Medium Duty Vehicle Surveillance Program, Series 3 (MDVSP-III).
State of California, Air Resources Board, March 1981.

ARB 1981b
MS 81-006, Test Report of the Motorcycle Surveillance Test Program, Series 2 (MCSTP-II).
State of California, Air Resources Board, March 13, 1981.

ARB 1981c
California Assembly-Line Test Procedures for 1982 and Subsequent Model Year Passenger Cars, Light-Duty Trucks and Medium-Duty Vehicles.
State of California, Air Resources Board, Adopted April 14, 1981.

ARB 1981d
Supplement 2 to Procedure and Basis for Estimating On-Road Motor Vehicle Emissions.
State of California, Air Resources Board, Stationary Source Control Division, Emission Inventory Branch, Motor Vehicle Emissions and Projections Section, June 1981.

ARB 1981e
MS 81-010, Test Report of the San Francisco Surveillance Program, Series 1, CVS.
State of California, Air Resources Board, June 3, 1981.

ARB 1981f
A9-138-62 Am. 1, Contract Amendment, Excerpts (Contractor, Environmental Resource Management) in accordance with Contractor's Proposal entitled Acquisition of Light Duty Vehicles for ARB Fifth Surveillance Tests, CVS, dated March 20, 1980, etc. 15 July 1981.

ARB 1981g
MS 82-03, Test Report of the Light Duty Vehicle Surveillance Program, Series 4 (LDVSP-IV).
State of California, Air Resources Board, November 1981.

ARB 1981h
California Assembly-Line Test Procedures for 1983 and Subsequent Model Year Passenger Cars, Light-Duty Trucks and Medium-Duty Vehicles.
State of California, Air Resources Board, Adopted November 14, 1981.

ARB 1981i
Statement of Work for Procurement of Approximately 400 Vehicles for Light Duty Vehicle Surveillance Testing, 6th Series, CVS.
State of California, Air Resources Board.

ARB 1982a
Statement of Work for Procurement of Approximately 400 Vehicles for Light-Duty Vehicle Surveillance Testing, 7th Series, CVS.
State of California, Air Resources Board, February 16, 1982.

ARB 1982b
Memorandum, R.H. Cross to A. Donnelly. Subject: Special Testing of 15 - 1980 Mitsubishi Cars - LDVSP 6 (Project 2S81C1).
State of California, Air Resources Board, April 22, 1982.

ARB 1982c
Memorandum, R.H. Cross to A. Donnelly. Subject: Special Testing of 14 -- 1980 Ford Cars -- LDVSP 6 (Project 2S81C1).
State of California, Air Resources Board, May 26, 1982.

ARB 1982d
Memorandum, R.H. Cross to A. Donnelly. Subject: Special Testing of 26 1980 and 1981 Nissan Cars - LDVSP 6 (Project 2S81C1).
State of California, Air Resources Board, June 22, 1982.

ARB 1982e
MS 82-08, Test Report of the Light-Duty Vehicle Surveillance Program, Series 5 (LDVSP V).
State of California, Air Resources Board, July, 1982.

ARB 1982f
Mail Out No. 82-17. To: All Light-, Medium-, and Heavy-Duty Engine and Vehicle Manufacturers and Motorcycle Manufacturers.
State of California, Air Resources Board, July 23, 1982.

ARB 1982g
Memorandum, M. Hostak to R.H. Cross. Subject, Special Testing of 1980 Ford Vehicles (Engine Family 4.2/5.0/BJC).
State of California, Air Resources Board, September 16, 1982.

ARB 1983a
Privately furnished by M. Hostak of ARB. Examples illustrating sources of data, compilations and computations supporting quota determination for Light-Duty Vehicle Surveillance Programs (83 pages).

ARB 1983b
Privately furnished by M. Hostak of ARB. Regression analysis printouts for "as-received" 1975-76 cars (SAE 800395) and "Repaired" 1975-76 cars (SAE 800395).

CARLSON, 1977
EPA-460/3-77-004, Study of Emissions From 1966 Through 1976 Model Year Light-Duty Vehicles in Los Angeles, St. Louis, and Washington, D.C.
Richard Carlson, Olson Laboratories, Inc., August 1977.
(Prepared for EPA, Contract No. 68-03-2379).

CROSS, R.H., 1980
SAE Technical Paper Series 800395, Emissions From Catalyst Equipped California Cars in Customer Service.
Society of Automotive Engineers, Inc., Congress and Exposition, Detroit, February 25-29, 1980.

CROSS 1981
Emissions From In-Use Catalyst Equipped California Passenger Cars - An Update.
Proceedings - Institute of Environmental Sciences 41-45, May 1981.

DIXON, 1979
W.J. Dixon, M.B. Brown, "Biomedical Computer Programs, P-Series."
University of California Press, Los Angeles, 1979.

EPA 1981
"Scope of Work for EPA Recall Testing Programs." EPA Guideline Document transmitted to the vehicle procurement contractors on May 26, 1981. (U.S. Environmental Protection Agency).

EPA 1982
Memorandum, J.F. Powers, Jr., to Division Directors. Subject: In-Use Vehicle Sampling Methodology.
U.S. Environmental Protection Agency, Office of Air, Noise, and Radiation (Ann Arbor, Michigan), November 10, 1982.

ERM 1983a
Private communication. Protocol of Instructions in Soliciting Candidates for Vehicle Procurement (ARB Surveillance Testing). Attachments.
Environmental Resources Management, January 21, 1983.

ERM 1983b
Private communication. Computerized Vehicle Selection Procedure, J. Stormes.
Environmental Resources Management, February 2, 1983.

HORIE, 1981
Y. Horie and J. Cassmassi, "A Study of Driving and Storage Patterns of Non-Primary Private Vehicles in the South Coast Air Basin."
Pacific Environmental Services, Inc. Prepared for California Air Resources Board, Contract A0-073-31, August 1981.

OLI 1975
Plan of Performance, FY 1975 LDV and LDT Emission Measurement Program at Washington, D.C.; St. Louis, MO; and Los Angeles, CA. (Excerpts).
Olson Laboratories, Inc. (Submitted to EPA; Contract 68-03-2379).

OLI 1977a
Final Report, The California Consumer-Owned Catalyst-Equipped Emission Study (EPA Contract No. 68-02-2232). Excerpts.
Olson Laboratories, Inc. (Submitted to EPA).

OLI 1977b
Final Report, Basic Testing Support for the Emission Control Technology Division, EPA Contract 68-03-2411, Task Order 6, Procurement of 19 Vehicles.
Olson Laboratories, Inc. (Submitted to EPA).

SCI 1978
EED-79-178, A Proposal for Procurement of 300 Vehicles for ARB Fourth Surveillance Series of Tests.
Systems Control, Inc., Environmental Engineering Division, December 5, 1978. (Submitted to EPA).

SCI 1979
7657-100, Procurement of Motorcycles for Air Resources Board Surveillance Testing (Final Report).
Systems Control, Inc., June 1, 1979. (Submitted to ARB).

APPENDIX A. SUGGESTED APPROACHES FOR OPTIMIZING VSP SAMPLING QUOTAS

It is beyond the scope of this report to pursue in detail the algorithms and mathematical approaches which could be applied to optimizing stratum quotas so as to minimize the standard error of estimation of total annual emissions or of population average emission rates. A few important principles will be mentioned here, however.

First, we observe that the importance of obtaining an accurate estimate of emissions for any stratum (subpopulation) of vehicles from the general vehicle population is in proportion to the amount of emissions contributed by that stratum. That is, a given relative error in determining the emission factor for a group would have a greater effect on the error for the total population if that group contributed, say, half the total emissions than if it contributed only one tenth. Consequently, the quota for each group should be adjusted for this effect.

Second, the number of vehicles needed in a sample for accurate determination of the emission factor for a particular stratum depends on the standard deviation within that stratum. Quotas could be adjusted for this effect by requiring a greater total sample for strata which have larger standard deviation.

According to ARB's motor vehicle emissions manual (ARB 1980a), the fractional contribution of given-age vehicles to the total vehicle population or to the total VMT in a typical calendar year is distributed non-monotonically, as shown in Table A-1. The greatest fractional contribution to both the vehicle population and the total VMT is made by two-year-old cars instead of one-year-old cars. (These results are derived from vehicle populations in years from 1968 through 1975).

For a more recent year, a vehicle use survey conducted in the SCAB (Horie and Cassmassi 1981) indicates that the greatest fractional contribution to the vehicle population is made by four-year-old vehicles in all LDVs and by three-year-old vehicles in primary LDVs, which are used frequently (4 days per week or more) and are driven over long distance (7.5 miles each way or longer). This shift of the maximal contribution from the 2nd year to the 3rd or 4th year may have been caused by diminished auto sales in recent years.

Table A-1. AGE DISTRIBUTIONS OF MOTOR VEHICLES, VMT, AND SAMPLE POPULATIONS BY MODEL YEAR

Vehicle Age	Light Duty Passenger Cars (Statewide Population)[a]	Light Duty Passenger Cars (Statewide VMT)[b]	All Light Duty Passenger Cars (Sample Population)[c]	Primary Light Duty Vehicles (Sample Population)[d]	Vehicles Procured in LDVSP 5 (Sample Population)	Vehicles Procured in LDVSP 6 (Sample Population)
1	.0647	.0944	.042	.052	.201	.225
2	.0986	.1357	.079	.117	.159	.188
3	.0970	.1246	.098	.128	.130	.154
4	.0945	.1136	.108	.109	.127	.138
5	.0920	.1030	.078	.074	.112	.113
6	.0875	.0907	.064	.068	.086	.092
7	.0820	.0775	.041	.044	.041	.092
8	.0745	.0642	.076	.069	.053	0
9	.0665	.0518	.065	.055	.050	0
10	.0565	.0394	.073	.060	.041	0
>10	.1922	.1051	.275	.224	0	0

[a]Based on 1968-1975 National Vehicle Registration Service data for California
[b]Based on the same NVRS data and annual mileage accumulation rate taken from EPA document, Mobile Source Emission factor
[c]Based on 1981 survey of 1335 randomly selected vehicles in the SCAB
[d]Based on 1981 survey of 635 randomly selected primary vehicles in the SCAB

Table A-1 also presents the age distributions of motor vehicles procured in LDVSP 5 (ARB 1982e) and 6 (ARB 1981i). Unlike the nonmonotonic distributions described above, the age distributions of vehicles procured in the two surveillance programs are a monotonically decreasing function of vehicle age. Furthermore, the fraction of latest model year vehicles in the total sample population is 0.201 in LDVSP 5 and 0.225 in LDVSP 6. These values are more than twice as large as the fraction of 1-year-old cars in all the four age distributions of vehicle population and VMT listed in Table A-1.

The above discussion indicates that the currently used allocation scheme for the in-use surveillance program does not agree either with the age distribution or the VMT distribution of the vehicle population. However, it is interesting that the age distribution of vehicles in LDVSP 5 approximates the distribution of VMT by age of light duty passenger cars* (Table A-1, column 3), except for the two latest model years and for vehicles over 10 years old.

However, even if quotas were made roughly proportional to VMT for the various model years, this would not necessarily be the best sampling scheme for minimizing error either in overall average emission rates per vehicle or in estimated on-the-road emissions based on the VSP data base.

To illustrate how these principles can be applied in setting quotas, we consider again the data from LDVSP 5. Suppose that a total of 340 vehicles were to be procured and that each of the 10 model years from 1971 to 1980 were to be taken as an independent stratum. Taking the prior estimates of the activity (VMT) by model year from Table A-1, activity-based quotas are obtained by normalizing; i.e., $Q_i = 340\ A_i/\Sigma A$, where Q stands for quota, A for activity and i for model year. Results of this

---

*It appears that MSCD may have intended to specify quotas roughly proportional to VMT, at least for the earlier model years of each VSP. Thus, in ARB 1980b, "Vehicle Selection Methodology" is described as follows: "To obtain a sample of vehicles representative of the California state vehicle population, statistical blocks of vehicle types were chosen for each model year. The sample within each block was chosen to represent the actual population by vehicle manufacturer, model, and engine size. More statistical blocks were chosen for later model years (1974-1977), since these vehicles travel more miles per year than older vehicles."

calculation are shown in Table A-2, (Column 4).

Where prior estimates of emission factors for these model years are available, these can also be factored in; i.e., $Q_i = 340\ A_iE_i/\Sigma A_iE_i$ where E stands for emission factor. Taking as an example the mean CO emission factors as found for various model years in LDVSP 5, emission-based sample quotas for successive model years are also shown in Table A-2. Thus, taking into account the probable emissions as well as the activities for each model year, it appears that quotas for the earlier model years would be two to three times as large as for the later. This is a complete reversal of the relation initially postulated.

The situation is similarly altered when the effect of standard deviation of vehicle emissions is incorporated. An adjustment for this effect, proposed by Snedecor & Cochran (1980), may be based on direct proportionality to the standard deviation. This gives quotas $Q_i = 340\ A_iS_i/\Sigma A_iS_i$ (normalized to the total of 340). As a rough approximation, sample standard deviations as shown in LDVSP 5 were about 27 for model years 1971 to 1973, 20 for 1974 to 1977, and 12 for 1978 to 1980. Use of these more aggregated estimates for standard deviation would improve the sequence of quotas shown in Table A-2 by smoothing it.

Even the adjustment for standard deviation (discussed in the previous paragraph) is insufficient to optimize the set of quotas in relation to the contributions of the strata to on-the-road emissions. The standard-deviation adjustment is designed to optimize quotas for estimating the true population mean, which, in the data base under discussion, would be an average of emissions per mile over all vehicles in the defined population. However, the contribution of any stratum to on-the-road emissions is not in simple proportion to the number of vehicles in that stratum, nor even in proportion to the activity of those vehicles; in fact, it also involves the typical emission factor for the stratum.

PES has not developed an algorithm or formula for optimizing stratum quotas in view of the rather large differences in emission factors found for vehicles of different model years (shown in Section 3.3.4). However, it appears likely that such optimization would require an even greater

Table A-2. QUOTA ADJUSTMENTS ILLUSTRATED

Model Year	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980
ARB Quota[a]	19	19	19	20	30	35	40	44	51	60
Activity Fraction[b]	.039	.052	.064	.078	.091	.103	.114	.125	.136	.094
Activity Quota[c]	15	20	25	29	34	39	43	47	52	36
Emission Factor (CO)[d]	48.1	50.1	67.9	38.2	23.1	20.6	14.9	10.6	9.8	8.5
Emission Fraction[e]	.089	.123	.205	.141	.099	.100	.080	.062	.062	.037
Emission Quota[f]	30	42	70	48	34	34	27	21	21	13
Standard Deviation[g]	27.4	25.5	62.0	15.5	20.3	24.5	16.4	10.4	12.3	13.3
Activity x S.D.[h]	1.07	1.33	3.97	1.21	1.85	2.52	1.87	1.30	1.67	1.25
(Act.xS.D.)Fraction[i]	.059	.074	.220	.067	.103	.140	.104	.072	.093	.069
(Act.xS.D.) Quota[j]	20	25	74	23	35	48	35	24	32	24

a: ARB quotas for LDVSP 5 are shown, with a total of 337 vehicles. Numbers of vehicles actually procured differed from the quotas, which represented procurement goals.
Alternative quotas suggested herein are adjusted to a total of 340 vehicles.

b: From Table A-1, for light-duty passenger cars, statewide.

c: Obtained by partitioning 340 by activity fraction given above.

d: From results of LDVSP 5, calculated by PES.

e: Obtained by multiplying emission factor by activity fraction, then normalizing the product to a total of 1.00.

f: Obtained by partitioning 340 by the emission fraction.

g: From results of LDVSP 5, calculated by PES.

h: Product of activity fraction by standard deviation.

i: Obtained by normalizing previous product to a total of 1.00.

j: Obtained by partitioning 340 by the fraction in previous line.

emphasis on sampling the older model years than is suggested by the standard-deviation adjustment alone.

In connection with these suggestions, it should be noted that the total of emissions measured over the entire sample cannot be an unbiased estimator of on-the-road emissions of the sampled population unless activity-based quotas are used. If any other quota system is used, the emissions of each model year sample must be weighted by the ratio of the activity-based quota to the quota used, before totaling for an estimate of population emissions. Despite this slight increase in the complexity of the calculations, however, it appears that some of these alternate quota systems could yield real improvements in accuracy for estimating on-the-road emissions based on the MSCD data base.

GLOSSARY

ACE
: Automotive Control Engineering

Activity Factor
: (See "use factor")

AESI
: Automotive Environmental Systems, Inc.

Analysis of variance
: A statistical technique for determining the relation of variations within a body of data to suggested underlying factors

ARB
: State of California, Air Resources Board

Bias
: Systematic error in estimates, due to factors known or unknown to an investigator

BMDP7D
: A statistical computer routine, titled "Description of Groups with Histograms and Analysis of Variance," part of the BMD statistical package (ref.)
(W.J. Dixon, M.B. Brown, "Biomedical Computer Programs P-Series", University of California Press, Los Angeles, 1979.)

Candidate vehicle
: A vehicle listed, prior to agency acceptance or prior to owner contact, for possible procurement by a contractor

CBD
: Central business district

CEF
: Composite emission factor

CID
: Cubic inches displacement, a measure of engine size

Cluster sampling
: Gathering items for a sample within selected small regions from a large population

CO
: Carbon monoxide emissions

Compliance testing
Measurement of vehicle emissions by EPA under its compliance testing program

Composite emission factor
An estimate of emissions per VMT for a composite vehicle population, based on appropriate weighting of emission factors of individual types of vehicles

Computerized vehicle profile
A randomization system devised by ERM

Confidence interval
The end-points of an interval having a specified probability of covering a selected estimate. E.g., a 95% confidence interval about a calculated mean, $\bar{x}$ may be expressed as $\bar{x} \pm I$, which means that the "true" value of the mean in 95 cases out of 100 will be between $\bar{x} + I$ and $\bar{x}-I$

Correlation coefficient
A measure of the degree of association between two measured characteristics of individuals (e.g. vehicles) in a population. A correlation coefficient near 1.0 indicates very close association

CVS
Constant-volume sampling procedure (ARB Laboratory)

Deterioration factor
In the context of this report, a factor determined by regression analysis indicating the average rate of increase of an emission factor per mile of accumulated vehicle travel (VMT)

DOL
U.S. Department of Labor

Emission factor
A value in grams per mile, measured in the surveillance program, characterizing the average rate of emission of a particular pollutant from a particular type of vehicle

Emission inventory
A compilation of estimated total rates of emissions of pollutants within a defined area, such as SCAB

Emissions
In the context of this report, primarily hydrocarbon (HC), carbon monoxide (CO), and oxides of nitrogen (NOx)

Emission standards
: Limits for emissions from motor vehicles as specified pursuant to law by ARB and EPA

EPA
: U.S. Environmental Protection Agency

ERM
: Environmental Resource Management

Evaporative emissions
: Emissions of hydrocarbon vapors due to evaporation from fuel tank, crankcase, carburetor, and other unsealed escape points

Fleet composition
: The detailed composition of the sample of vehicles procured for a surveillance program, in terms of numbers by vehicle class, model year, manufacturer, engine displacement and, occasionally, other vehicle parameters

FTP
: Federal test procedure

HC
: Hydrocarbon emissions

I/M
: Vehicle inspection and maintenance

Incentive
: Monetary or other inducements offered to vehicle owners to encourage participation in a testing program

Index of determination
: The square of the correlation coefficient, q.v. When expressed in terms of percent, it is equivalent to "percent of variance explained" (pve)

In-use vehicle surveillance program
: An ARB program in which emissions from consumer-owned vehicles are measured

ISR
: Institute of Social Research, University of Maryland

LDV
: Light-duty vehicle

LDVSP
: Light-duty Vehicle Surveillance Program

Linear equation
: Any equation of the form
  $y = a + bx$,
  where x and y are variables
  and a and b are coefficients

Matrix
: A tabular arrangement of sample quotas specified by various characteristics of desired vehicles

MDVSP
: Medium-duty Vehicle Surveillance Program

MSCD
: Mobile Source Control Division, State of California, Air Resources Board

MVIP
: California Motor Vehicle Inspection Program

Normalize
: Given a list of numbers, to compute a second list proportional to the first but having a different, preselected, sum

NOx
: Emissions of oxides of nitrogen

OLI
: Olson Laboratories, Inc.

PES
: Pacific Environmental Services, Inc.

Polk
: R.L. Polk Company

Probability level
: In statistical testing, the chance that a particular statistical result might have arisen solely because of random sampling fluctuations. A comparison is usually said to be significant if its probability level is 5 percent or less

Quota
: A number of items specified to represent a subpopulation; i.e., a specified subsample size

Randomization

Any system for sampling items from a list in such a manner that systematic error is avoided

Random sample

A sample selected from a specified population by a procedure wherein choices are dictated by chance

Random selection

A sampling method in which items are chosen by chance; also called "probability sampling"

Recall programs

ARB and EPA programs for testing emissions of motor vehicles, oriented toward the possibility of requiring manufacturers to recall and repair or improve vehicles with unsatisfactory emission control performance

Regression analysis (linear)

A statistical procedure for fitting the best straight-line relation to a set of bivariate data

RSA

Regional Statistical Area (as defined by SCAG)

Sample

In the context of this report, usually a set of motor vehicles procured for emissions testing

Sample selection

The process of choosing vehicles to be tested in a vehicle surveillance or compliance program

SCAB

South Coast Air Basin

SCAG

Southern California Association of Governments

SCI

System Control, Inc.

SES

Socioeconomic status

Solicitation

A request by ARB or EPA for proposals by prospective contractors

SSCD

Stationary Source Control Division (California Air Resources Board)

SSU
    Secondary sampling unit

Statistically valid
    Conforming to conventional criteria for significance in statistical tests; for example, less than a 5 percent chance of incorrectly identifying an effect

Stratification
    Definition of strata in stratified sampling

Stratified sampling
    Specification of a sample as a set of mutually exclusive subsamples

Stratum (pl.,strata)
    In the context of this report, any vehicle subpopulation utilized in stratified sampling, or for which a sampling quota is specified

Surveillance program
    Any program in which emissions from motor vehicles are measured

Use factor
    Any measure of the extent of use of vehicles in a given time; for example, annual or daily VMT

VMT
    Vehicle miles travelled; a measure of vehicular activity used in estimating mobile source emissions

VSP
    In-use Vehicle Surveillance Program (ARB)