PART B OF THE SUPPORTING STATEMENT

Determine Percentage of High Evaporative Emissions
Vehicles in the On-Road Fleet

OMB Control Number 2060-NEW

USEPA Agency Form Number 2292.01

October 16, 2008

Table of Contents

List of Tables

List of Figures

1 SURVEY OBJECTIVES, KEY VARIABLES AND OTHER PRELIMINARIES

Evaporative emissions from gasoline vehicles have been evaluated and regulated since the early 1970s. Gasoline vehicles have evaporative emissions control systems that control excessive evaporative emissions, which are essentially gasoline vapors. When these systems or the gasoline delivery system of a vehicle malfunction, excessive evaporative emissions can be emitted. The mass of evaporative emissions from individual vehicles has been quantified in previous studies [1, 2, 3], but the frequency of vehicles in such a state in the general population has been estimated based on limited data [1, 3, 4].

1(a) Survey Objectives

This data collection effort is a survey designed to estimate the prevalence of high evaporative emissions vehicles in the on-road fleet of gasoline-powered passenger cars and light-duty trucks. Specifically, the primary question of this study is:

What fraction of the fleet is made up of high evaporative emissions vehicles?

Evaporative emissions are made up of various types (diurnal, hot-soak, running losses, resting losses, fugitives, and gross liquid leaks). Official measurements of the evaporative emissions of a vehicle are made in a "Sealed Housing Evaporative Determination" (SHED) enclosure in a laboratory setting. However, these measurements are expensive and require several days for each vehicle tested. This study requires a field method that is quick and inexpensive and which correlates with official laboratory SHED results. The August 2008 Denver pilot study demonstrated that high-emitting evaporative emissions vehicles can be identified by the levels of ambient hydrocarbon compounds (HC) that accumulate in a portable SHED (PSHED) after a vehicle is enclosed and its engine is turned off. We expect that this technique can serve as the field method to answer the question: "Is this vehicle a High Evap?" By using the portable SHED to measure the evaporative emissions of a stratified sample of vehicles in the in-use fleet, the fraction of "High Evaps" in the fleet can be estimated.

1(b) Key Variables

Variables to be surveyed or measured include:

- Vehicle identifiers: License plate and Vehicle Identification Number
- Vehicle description : Model year, make, model, and odometer reading
- Vehicle usage and maintenance history through a vehicle owner survey

- Time trace of HC concentration of the air inside the portable SHED after a vehicle's engine is turned off and the portable SHED doors are sealed.
- Measured values of high evaporative emissions vehicle screening methods:
  - Remote Sensing Device measurements
  - Modified California Method (Under-hood visual inspection and electronic HC vapor leak detector inspection)

1(c)   Statistical Approach

We have selected a statistical approach for this effort for two reasons:

- While a census or partial census would be ideal, the effort and expense required is prohibitive.
- To meet the objectives for use of these data, it is necessary to draw valid and defensible inferences from sets of equipment surveyed or measured to equipment populations at wider scales, such as the county or state or nation. This requirement in itself rules out non-probabilistic approaches.

1(d)   Feasibility

Obstacles to Participation. We do not anticipate substantial obstacles to participation. We plan to locate measurement sites at several locations to facilitate travel by participants. Participation does impose some burden, as participants must modify their schedules for several hours. However, the burden is less than for previous efforts that required respondents to leave their vehicles overnight and make two trips to the study site. Thus, we anticipate that in conjunction with incentives, solicitation will prove no more difficult than and probably somewhat easier than in previous studies.

Availability of Funds. At present we expect to have adequate funds available to conduct the survey as designed. However, if funding shortfalls occur, we can take measures to reduce sampling costs. One possibility would be to reduce the number of vehicles in the study.

2   SURVEY DESIGN

2(a)   Target Population and Coverage

The target population includes the fleet of gasoline-powered passenger cars and light-duty trucks. Passenger cars are light-duty vehicles with gross vehicle weight ratings of less than 6,000 lbs. Light-duty trucks are trucks with gross vehicle weight ratings of less than 8,500 lbs. Passenger cars and light-duty trucks form the majority of the on-road motor vehicle fleet.

Nationally, they account for 96.6% of the on-road vehicle fleet and 89.0% of the total on-road vehicle miles traveled. Heavy-duty vehicles account for the remainder of the on-road vehicle fleet and the on-road vehicle miles traveled.

Coverage for this study will comprise the fleet of vehicles operating in the area of San Antonio, located in Bexar County, Texas.

The San Antonio population will broadly represent a fleet whose evaporative emissions are uninfluenced by the presence of an I/M program. Given the assumptions that (1) a substantial fraction of vehicles in the current fleet have been manufactured since 1996 and contain on-board diagnostic systems (OBD), (2) the OBD systems detects fuel leaks and trips the malfunction indicator light, (3) the OBD system is scanned during I/M inspections, and (4) drivers are required to perform satisfactory repairs, usually under pain of registration denial, we can plausibly conclude that the prevalence of fuel-system leaks should be somewhat lower than would be expected in the absence of the program. Thus, a study performed in a non-I/M area such as San Antonio should represent a base-line or upper-bound situation in which drivers' maintenance and repair behavior is unaffected by outside influence.

Several additional questions remain regarding the representativeness of this study to the national fleet as a whole. Results obtained in San Antonio will not reflect the extent to which differing climate, temperature and humidity would affect the emission of fuel hydrocarbons by vehicles. Additionally, these results will not reflect differences in driver or maintenance behavior in different regions, nor the extent to which climate or altitude would affect the wear patterns of vehicles or the tendency to develop leaks.

A final practical consideration is that the southerly location allows ready measurement of HC emissions during late fall and early spring. The availability of an extended field season substantially increases the feasibility of the study.

2(b)   Sample Design

2(b)(i) Sample Frame

As described below, we have evidence that the hydrocarbon channel of certain remote-sensing instruments (RSD) may be sensitive to vehicles with high evaporative emissions levels. By oversampling vehicles with higher RSD measurements, we believe that the effort can capture "High Evaps" more efficiently than by completely randomized sampling from the fleet as a whole.

The sample frame will be constructed from the set of vehicles that have received screening RSD measurements. All screened vehicles will have a chance of selection for participation in the study.

When vehicles are screened, their license plates will be captured, which allows the screening measurements to be merged with descriptive and ownership information obtained from a vehicle registration file for the study area. Given the fact that the registration data will be somewhat out of date, and allowing for errors in one or more data sources, we can expect that some fraction of screened vehicles will be excluded from the sample frame. An as yet undetermined incidence of frame blanks can also be expected in cases where ownership information proves to be out of date or incorrect.

To address the loss of potential participants due to frame blanks, the registration data as a whole can be used to assess the representativeness of the frame on typical descriptive variables such as vehicle type (car vs. truck), model year, or even vehicle manufacturer.

Screening Measurements. Since most vehicles are not excessive emitters of evaporative hydrocarbons , screening based on RSD measurements will be used to obtain an index of evaporative hydrocarbons (RSD Evap Index) for purposes of sampling. Stratification based on the RSD Evap Index, which is under development based on the data collected in the Denver pilot study[9], will be used to enrich the sample with vehicles that are potentially high evaporative emitting. In the survey we will use an RSD Evap Index to screen vehicles into four RSD Evap Index categories. As described below, the RSD Evap Index can identify vehicles that are more likely to be "High Evaps". By preferentially sampling more vehicles from the higher RSD Evap Index bins, we believe that the effort can capture a larger fraction of "High Evaps" in the 1000-vehicle survey than could be captured by completely randomized sampling from the fleet as a whole.

The RSD instrument uses a light beam shining across the roadway to measure pollutants in a vehicle's tailpipe plume. The instrument has HC, CO, NO, and CO_2 channels. When a vehicle drives past the instrument, the light beam shines through the emissions plume behind the vehicle and takes 50 10-millisecond-spaced measurements for each of these channels. The data collected during the Denver pilot study revealed that for vehicles with zero evaporative emissions, the 50 data points in the HC-versus-CO_2 plot all fall on a straight line. The exhaust HC emissions concentration is proportional to the slope of the line. This method is a standard method for calculating exhaust HC emissions and has been known for many years. However, the Denver pilot study also found that for vehicles with high evaporative emissions, the 50 data

points in the HC-versus-CO_2 plot do not fall on a straight line. We believe this is a consequence of HC evaporative emissions plume, which is produced by non-tailpipe sources on the vehicle, wafting into the light beam at the same time as the tailpipe plume is in the light beam. Thus, the degree to which the 50 data points deviate from a straight line in the HC-vs-CO_2 plot is a measure of the amount of evaporative emissions produced by the vehicle at the instant that it passes through the RSD light beam.

We are developing several measures of the characteristics of the deviations from the straight line as they relate to evaporative emissions. The statistical measures that we are using include the average deviation from the straight line, correlation coefficient, principal component analysis, and spectral analysis. No single approach outperforms all of the others. Each approach has advantages and disadvantages. Before actual vehicle selection begins in this 1000-vehicle study, we will likely settle on an RSD Evap Index that combines more than one approach so that the performance of the index is superior to any single approach. Since that development work is not complete, for the purposes of writing this ICR, we will use the average deviation from the straight line as the RSD Evap Index. Later, when development of the comprehensive index is complete, we will use the same statistical techniques described here to re-calculate the desired allocations among the strata.

The RSD Evap Index was developed using the 175 observations of the Denver pilot study. Each observation is a paired measurement of RSD and PSHED. Each RSD measurement produces the 50-sample-point data from the vehicle's plume which in turn is used to calculate the RSD Evap Index value for the observation. The PSHED is a portable 10' x 20' x 8' sealed enclosure. To determine the evaporative emissions produced by a vehicle, the vehicle is warmed-up by driving it for about 7 miles, is then immediately placed in the PSHED with the engine off, and the enclosure is sealed. During the next 15 minutes the HC concentration in the air inside the PSHED is measured as the evaporative emissions leave the vehicle. At the end of 15 minutes the HC concentration is used to calculate the mass of HC that has been emitted by the vehicle. This PSHED mass is the quantity that is used to determine whether or not the vehicle is a High Evap.

Figure B.1 shows a plot of the 175 observations from the Denver pilot. The vertical axis is the PSHED result, and the horizontal axis is the RSD Evap Index. The plot shows widely scattered data with positively skewed values on both axes. The skewed nature of the PSHED values is derived from the fact that most vehicles have low evaporative emissions; only a few vehicles are "High Evaps". In spite of the scatter of the data points, the plot shows that at low values of the RSD Evap Index, most vehicles have low PSHED emissions. In addition, the group of data points with the highest PSHED values increase with increasing RSD Evap Index. Thus, it

is reasonable to conclude that vehicles with higher RSD Evap Index tend to have higher PSHED values. The relationship between RSD Evap Index and PSHED value needs to be quantified so that it can be used to develop the allocations for the 1000-vehicle study.

Figure B.1. RSD Evap Index and PSHED Values for the Denver Pilot Data

/proj1/EPA_High_Evap/DenverData/Pilot/Analysis/pshed_rsd_v2.sas 09OCT08 12:45
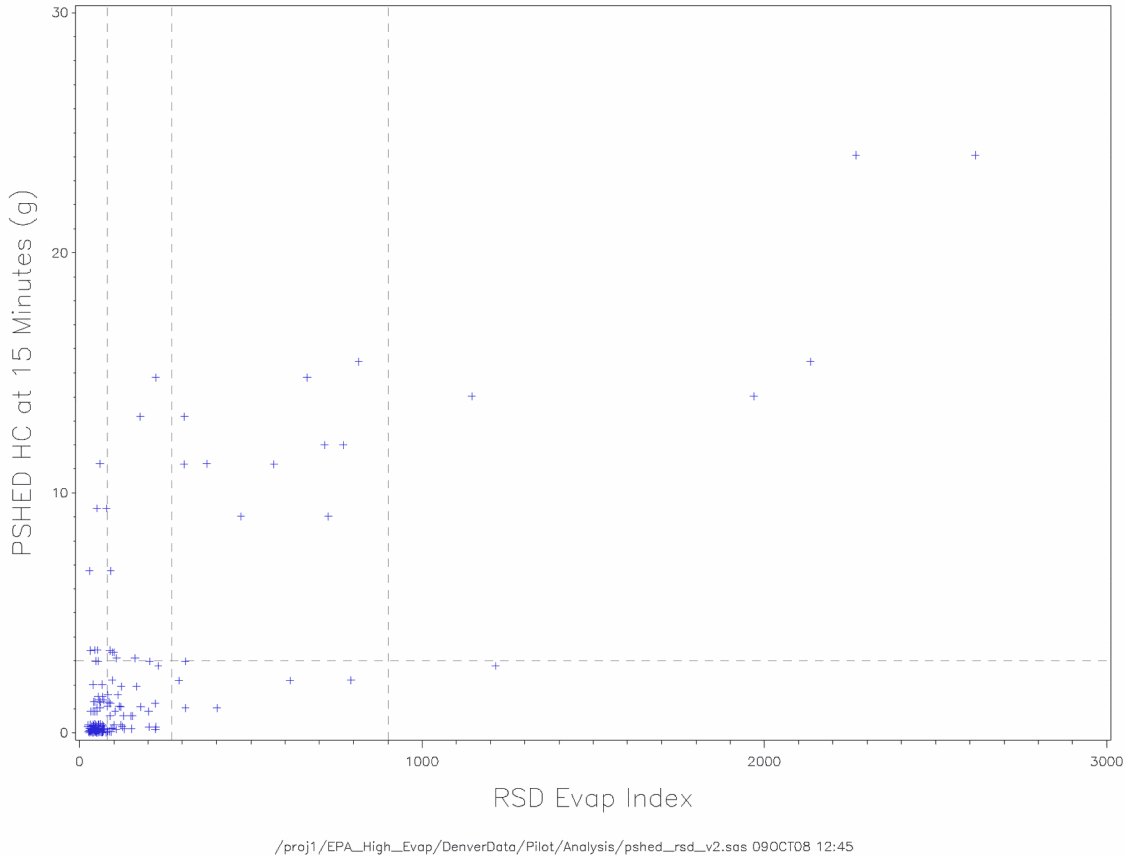
Based on an independent evaluation of evaporative emissions mechanisms, and trends, a PSHED value of about 3 grams can serve as the standard for "High Evaps" when measured in the PSHED. Accordingly, Figure B.1 has a dashed horizontal reference line at 3 grams. Points above this line represent "High Evaps".

Rather than attempt to predict the PSHED value based on the RSD Evap Index, we choose to calculate the probability that a vehicle with a given RSD Evap Index is likely to be a High Evap. To help visualize the trend in probability, vertical dashed lines at 80, 270, and 900 in Figure B.1 are used to divide the plot into four RSD Evap Index bins. The count of the number of observations within each of these bins and the number of "High Evaps" within each of the bins is given in Table B.1. The table shows a monotonically increasing High Evap fraction as the RSD Index increases.

6

Table B.1 Counts of High Evap Designations for the Denver Pilot Data

RSD Evap Index		Count		Measured High Evap (fraction)	Modeled High Evap (fraction)
Low	High	PSHED > 3g	Total		
0	80	7	112	0.06	0.054
80	270	8	41	0.20	0.225
270	900	10	16	0.63	0.568
900	Inf	5	6	0.83	0.856

To further quantify the relationship, we performed a logistic regression to predict the probability that the PSHED value would be larger than 3 grams using RSD Evap Index as a predictor. First, we found that the natural log of the RSD Evap Index was an advantageous transformation. It caused the values of the RSD Evap Index to be more nearly homogeneous. The log transformation also was able to predict the PSHED failure probability from a simple expression with no lack of fit:

$$P_{fail} = \exp(arg) / (1 + \exp(arg))$$

Where

P_{fail} =   Probability that the PSHED test has a value above 3 grams,

arg   =   -7.5262 + 1.2582 * ln(RSD Evap Index)

During model development, we found that the model had only a very small dependence on exhaust emissions level. This means that the RSD Evap Index is largely independent of exhaust emissions influence and therefore a given RSD Evap Index value has the same meaning whether the vehicle has high exhaust emissions or low exhaust emissions. Since older model year vehicles tend to have higher exhaust emissions, this means that RSD Evap Index is independent of model year.

The model development also revealed that alternative High Evap definitions (other than 3 grams) could also be used to develop a logistic regression models with the same favorable properties. The concordance[1] for the logistic regression was 79.6%.

---

[1] Concordance is a statistic that evaluates the agreement between the predicted probabilities of a logistic regression model and the pass and fail values of the individual observations in the training set. Concordance can have a value from 0% to 100%. If the predicted probabilities completely agree with the pass and fail values of the response variable, then the concordance is 100%.

The PSHED failure probabilities predicted by the model at the center of the ln(RSD Evap Index) bins are given in the last column of Table B.1. The values calculated by the model are reasonably close to the values derived from the counts in Table B.1.

2(b)(ii) Stratification Variables

Two stratification variables will be used in this study:

1)   Vehicle Model Year Group, and
2)   RSD Evaporative Emissions Index.

When taken together, these two stratification variables create twelve strata.

Vehicle Model Year Group – Based on considerations of vehicle evaporative emissions control technology, we have chosen three vehicle model year groups:

- 1980-1995,
- 1996-1999, and
- 2000-2009.

The 1980-1995 group represents older evaporative emissions control technologies that were designed for the older evaporative emissions certification standards. About 12% of the light-duty vehicle fleet falls into this group. The 1996-1999 group represents a transition from the older to the newer evaporative emissions certification standards. For this group, some vehicles were designed to meet the older standards and some vehicles were designed to meet the newer standards. About 20% of the light-duty vehicle fleet falls into this group. The 2000-2009 group represents the newest evaporative emissions control technologies that were designed for the newest evaporative emissions certification standards. The evaporative emissions certification standards have been constant throughout this range of model years. About 68% of the light-duty vehicle fleet falls into this group. Vehicles with model years older than 1980 will not be part of this study.

RSD Evaporative Emissions Index – Based on considerations of the values of RSD Evap Index seen in the Denver PreTesting data, the Denver Pilot data, and the Colorado RSD fleet data, we have chosen four RSD Evap Index groups:

- 0 to 80,
- 80 to 270,

- 270 to 900, and
- 900 and larger.

The 0-to-80 group represents the group of vehicles where the RSD Evap Index is probably dominated by noise in the calculation of the index. We believe that the noise is produced by variability in the raw RSD measurements taken in the plume of a vehicle. Thus, even when the vehicle has zero evaporative emissions, the RSD Evap Index could be as high as 80. The 900-and-larger group represents vehicles that have profoundly elevated evaporative emissions. This group might include vehicles that have gross liquid leaks or that were manufactured before 1972, which is approximately the year that vehicles were first equipped with evaporative emission controls. The intervening 80-to-270 and 270-to-900 groups take up the space between the low and high RSD Evap Index groups. As shown by Table B.1, these four RSD Evap Index groups cover a wide range of high evaporative emission probability.

2(b)(iii) Sample Sizes

The primary goal of this study is to determine the fraction of vehicles in a non-IM fleet that have high evaporative emissions. A secondary goal is to collect data to develop and validate an RSD Evap Index that can be used in future studies and locations to estimate the fraction of "High Evaps" in other fleets. As discussed above, since we have developed an RSD Evap Index based on the data collected in the Denver pilot study, we plan to select approximately 1,000 survey vehicles from the fleet using the RSD Evap Index as a screening tool. We know from past experience that remote sensing can screen many more vehicles than the required 1,000 vehicles in a relatively short time. Therefore, using remote sensing to select survey vehicles is feasible.

Designing a stratified sampling strategy that can achieve the desired precision (see Section 2(c)(i)) requires an estimate of the abundance of "High Evaps" in the fleet as a function of the stratifying variables: model year group and RSD Evap Index. We have estimated the abundance of "High Evaps" in Colorado by applying the RSD Evap Index developed from the Denver pilot data to on-road RSD data collected by Colorado as part of their regular RSD monitoring program. We obtained a sample of Colorado's RSD dataset. The sample contains model year and RSD data on 1,485 1980 to 2009 light-duty vehicles collected on one day (August 13, 2008) and from one RSD instrument (#8093) in Colorado. Table B.2 shows the distribution of the vehicles in the twelve bins defined by model year group and RSD Evap Index.

Table B.2. Observed Distribution of Colorado Vehicles

Model Year Group	RSD Evap Index			
	000-080	080-270	270-900	900-Inf
1980-1995	180	5	0	0
1996-1999	283	3	2	0
2000-2009	1008	2	2	0
Total	1471	10	4	0

The table shows no observations in four of the bins. To do an optimal allocation for a stratified sample, we must have non-zero fractions of the fleet in all bins. Therefore, we supplemented the counts in Table B.2 with our engineering judgment to estimate fractions of the fleet in all twelve bins. The result is Table B.3.

Table B.3. Estimated Distribution of Colorado Vehicles

Model Year Group	RSD Evap Index			
	000-080	080-270	270-900	900-Inf
1980-1995	0.1209	0.0034	0.0013	0.0004
1996-1999	0.1902	0.0020	0.0013	0.0002
2000-2009	0.6773	0.0013	0.0013	0.0003
Total	0.9884	0.0067	0.0039	0.0009

Table B.3 shows that 99% of the vehicles falls into the lowest RSD Evap Index bin. Because of the predominance of the vehicles in the lowest RSD Evap Index bin, the optimum allocation caused almost all allocated vehicles to go into the lowest RSD Evap Index bin for each model year group. The resulting allocation was therefore close to the allocation that would arise from random sampling of the fleet without using RSD or its index at all. While this allocation would meet the primary goal of the study (estimating the fraction of "High Evaps" in the fleet), it would not meet the needs of the secondary goal at all. Therefore, we used a different approach.

First, we determined the optimal allocation using only model year group, as shown in Table B.4.

Table B.4 Optimal Allocation for Model Year Alone.

Model Year Group		Colorado fleet sample	Population Fraction (based on Colorado RSD Data) (W_h)	Fraction of Vehicles that are "High Evaps" (based on Colorado RSD Data) (p_h)	Std Dev of the High Evap Probability (s_h)	Weighted Standard Deviation (W_h*s_h)	Expected Number of Optimally Stratified Vehicles (N_h)
Lower	Upper						
1980	1995	185	0.1246	0.051	0.220	0.0274	127
1996	1999	288	0.1939	0.052	0.222	0.0431	199
2000	2009	1012	0.6815	0.048	0.214	0.1457	674
All		1485	1.0000			0.2161	1000

Then, we distributed the vehicles in each model year bin so that when the allocations were calculated across just RSD Evap Index (that is, ignoring the model year groups), the allocations produced a constant standard error of the mean for the top three RSD Evap Index bins. (See Section 5(b)(iii) for estimates of the standard errors of the mean.) This is justified since we know that the RSD Evap Index is independent of model year. The resulting allocation considering both model year group and RSD Evap Index is shown in Table B.5.

Table B.5 Allocation for Model Year and RSD Evap Index.

Bin	A	Model Year Group		RSD Evap Index		Population Fraction (based on Colorado RSD Data) (W_h)	Fraction of Vehicles that have PSHED>3g (p_h)	Std Dev of the High Evap Probability (σ_h)	Expected Number of Vehicles (N_h)
		Lower	Upper	Lower	Upper				
1				0	80	0.1209	0.054	0.225	96
2		1980	1995	80	270	0.0034	0.225	0.418	10
3				270	900	0.0013	0.568	0.495	14
4				900	∞	0.0004	0.856	0.351	7
5				0	80	0.1902	0.054	0.225	152
6		1996	1999	80	270	0.0020	0.225	0.418	15
7				270	900	0.0013	0.568	0.495	21
8				900	∞	0.0002	0.856	0.351	11
9				0	80	0.6773	0.054	0.225	513
10		2000	2009	80	270	0.0013	0.225	0.418	52
11				270	900	0.0013	0.568	0.495	73
12				900	∞	0.0003	0.856	0.351	36
All						1.0000			1000

The resulting allocation is a compromise between the goals of the project. It provides for 761 vehicles sampled in the lowest (0-80) RSD Evap Index bin so that near optimum allocation for determining the fleet High Evap fraction can be measured. It also provides for 239 vehicles sampled in the upper three (80-270, 270-900, 900-∞) RSD Evap Index bins so that a large amount of data will be available to improve and validate new RSD Evap Indices and evaluate the significance of an interaction between RSD Evap Index and model year group. Finally, the allocations provide for limiting the risk of large instabilities in the estimate of the fleet High Evap fraction caused by statistical fluctuations in the number of "High Evaps" found in Bins 1, 5, and especially 9.

Since the allocation of vehicles among the RSD Evap Index groups is a compromise, there is no allocation to optimally achieve the two goals of the study. However, using our best judgment, we believe the allocation in Table B.5 will meet the needs of the study. Specifically, we believe that the allocations in Bins 3, 4, 7, 8, 11, and 12 provide sufficient over-sampling to estimate the connection between RSD Evap Index and High Evap probability. To make a

substantial reduction in the uncertainty in that connection would require a massive increase in the number of vehicles allocated to the upper three RSD Index bins. Since the total number of vehicles sampled in the study must be kept at or below 1,000, this would result in a large decrease in the number of vehicles allocated to the lowest RSD Evap Index bins. This would greatly increase the risk of large instabilities in the estimate of the fleet High Evap fraction.

2(b)(iv) Sampling Methods

Vehicles will be sampled using a stratified random design. The sample will be stratified with optimal allocation among vehicle model year group bins and, within each of those bins, allocated among RSD Evap Index bins so that the standard error of the predicted PSHED failure probability is constant. Simple random sampling will be used for the strata defined by model year group and RSD Evap Index.

2(c)   Precision Requirements

2(c)(i) Precision Targets

The main quantity to be determined in this study is the fraction of the fleet vehicles that are high evaporative emitters ("High Evaps"). We would like to know with high confidence that this quantity has an uncertainty of no more than ±30% of the value. Specifically, the precision target is that the half width of the 95% confidence interval of the fraction is no more than 30% of the fraction.

2(c)(ii) Non-Sampling Error

2(c)(ii)(1)   *Frame-coverage error*

This error is defined as potential bias in key variables resulting from imperfections in the sample frame. The central issue is incomplete coverage, in which members of the target population are simply absent from the frame. The bias that may result from incomplete coverage may reduce the representativeness of the sample in a way analogous to that from whole-survey non-response. We have incorporated measures in the survey plan to detect and reduce the effects of these errors on the survey results.

2(c)(ii)(2)   *Non-response error*

As in any survey, non-response is one of the most important potential sources of error in final results. Survey non-response occurs when no response at all is obtained from a potential participant in the study, whereas item-nonresponse occurs when a respondent provides responses to some but not all items. Survey non-response occurs if a respondent refuses to participate.

Item-nonresponse may occur in a number of ways. A respondent may answer some items but refuse others, or may break off an interview for unrelated reasons. A form of item-nonresponse detrimental to emissions measurement but unrelated to the respondent could occur in cases where equipment malfunction or measurement errors make emissions datasets unsuitable for subsequent analysis.

2(c)(ii)(3)   Measurement error

The measurement of hydrocarbon emissions in the PSHED involves the use of complex instrumentation in a non-laboratory environment. The potential for measurement error in PSHED measurements have been evaluated during the Denver pilot and are described below in Section 4(c).

Calibration. All instruments including the remote sensor, handheld HC detection device, HC instrument in the portable SHED will be calibrated regularly.

2(c)(ii)(4)   Equipment malfunction

Following the measurements based on the various instruments, quality assurance measures will be undertaken to verify that the instruments operated correctly and that the results are reliable for further analysis. The QA process will involve the use of computer programs that automatically scan the time-series for patterns that may suggest instrument error, combined with graphic presentation of the data to allow case-by-case visual inspection.

2(c)(ii)(5)   Respondent error

The emphasis on collection of key information for the survey through direct inspection and instrumentation involves a conscious decision to reduce reliance on human memory to the maximum extent possible. A primary example is the use of electronic dataloggers to measure vehicle emissions. As much as possible, we have restricted interview items to general questions that can be easily answered without involved or detailed estimation and without heavy reliance on human memory.

2(c)(ii)(6)   Data entry error

Emissions results and other data collected electronically will not be input manually. Data files will be downloaded directly from the measurement instrument and transferred to the database, following quality-assurance procedures.

2(d)   Questionnaire

Vehicle owners that participate in the study will be presented a questionnaire when they arrive at the testing facility. We will collect recent vehicle usage and maintenance history information at the testing site by conducting a personal interview using questions in the questionnaire in Appendix B.   The questions request information on the number of miles driven by the vehicle owner and how the vehicle is parked at the owner's home. Specific maintenance questions are also included which request information on the quality and frequency of vehicle maintenance.

3   PRETESTS AND PILOT TESTS

3(a)   Pretests

The recruitment methods of using RSD measurements to screen and solicit participants for testing in San Antonio will be pre-tested in a pilot study this fall.   The basic plan is to first collect remote sensing measurements in various sites across the city of San Antonio. We will access the vehicle and owner information for the vehicles we measure in the city. Then, we will use the Denver study results to sample vehicles in the Model Year Groups and RSD Evap Index bins identified in Section 2. These vehicles/owners will then be solicited to come to a central location to participate in evaporative emissions testing. We will use the portable SHED developed in the Denver study as the primary method to measure evaporative emissions. In addition, we will conduct a visual and handheld sniffer to identify any potential vapor or liquid leaks on participating vehicles.

3(b)   Pilot Tests

3(b)(i) Pilot Test Objectives

For the purposes of refining procedures for the main study, we will conduct a pilot study. The following questions will be answered by the pilot study:

a)   How can we recruit a sample of vehicles to represent a non-IM fleet geographically and socio-economically?

b)   What is the approximate fraction of the fleet that is made up of high evaporative emissions vehicles?

c)   Based on our previous work, can we stratify the sample to get enough vehicles to estimate the fraction of high evaporative emitters?

d)   Our previous work has helped in developing a relationship between RSD Evap Index and high evaporative emissions. Can we confirm our methodology based on a new San Antonio sample?

e)   What refinements does the pilot suggest for the design of the larger study?

3(b)(ii) Pilot Test Design

A pilot test will be run in the San Antonio area to test vehicle screening and solicitation procedures and field evaporative emissions measurement techniques and to obtain initial measurements of the occurrence of "High Evaps" in the San Antonio fleet. To reconfirm the pilot Denver results, we plan on screening approximately 10,000 – 20,000 valid RSD measurements over a two – four week period using the sampling methodology described in this ICR to solicit and recruit vehicles. The non-IM pilot calls for us to recruit about 50 vehicles over a three week period . The vehicles will be sampled across the model year and RSD Evap Index bins as defined in Section 2.

3(b)(iii) Sampling for Measuring Portable SHED Precision and Accuracy

Because the main study will use the portable SHED to determine the "true" High Evap status of the 1,000 vehicles selected for testing, we must determine its precision and accuracy in the pilot study so that we can be confident that the portable SHED can be relied upon for the main study.

Precision. A subset of the vehicles participating in the portable SHED measurements will receive duplicate portable SHED measurements. Approximately 10 – 20 % of the vehicles will be randomly selected to receive duplicate testing and will be selected so that they span a range of portable SHED emissions results. This is necessary so that the precision can be estimated across the range of portable SHED emissions that are produced. Vehicles will be driven before the second PSHED to make all assurances that the vehicle has been reconditioned to the same "testing state."

Accuracy. To estimate the accuracy of the portable SHED, a known amount of propane vapor will be injected into the portable SHED. A measurement will be conducted every second for 15 minutes or more after the known quantity is injected into the portable SHED. A stabilized reading will reveal how much of the vapor is recovered in the SHED and the time trace over 15 minutes will reveal the quantity of vapor that is recovered in the SHED. We expect a retention and recovery rate of 95% in the SHED. The Denver Pilot saw an average initial recovery of

97.6% with a standard deviation of 3.3%, and a retention after 15 minutes of 95.7% with a standard deviation of 2.3%.

3(b)(iii) Incentives/Compensation experiment

During the pilot phase, we propose to offer a compensation to participants who bring their vehicles to the measurement site. During solicitation, vehicle owners will be informed that they will receive the initial $10 incentive regardless of their participation. In addition, participants bringing vehicles to the study site will receive a reimbursement for the inconvenience caused to them. We propose to experiment with allowing for this amount to be $50 and $100 during the pilot phase. For a total number of 100 respondents, the minimum detectable difference in the response rate with 90% confidence and 80% power should be about 20%.

4   COLLECTION METHODS AND FOLLOW-UP

As described above, the testing program will include the following elements:

- RSD screening measurements across San Antonio
- Connection of vehicle RSD screening measurements to vehicle owner and contact information
- Selection of target vehicles – random sample in stratified Model Year Group and RSD Evap Index bin
- Solicitation of vehicles through recruiting partner
- Setting up of appointments for vehicles agreeing to participate
- Vehicles show up at testing center
- Conduct RSD measurements, portable SHED measurements, and visual/functional leak check on participating vehicle

4(a)   Screening Methods

RSD vans will be deployed across the city of San Antonio to collect screening measurements on the vehicle population. Sites will be selected to be representative of the population and to provide good quality RSD readings. For the main study sites will be identified across the geographical area and socio-economic levels. Criteria for a good site required primarily to capture evaporative emissions would include low vehicle speed (10-20 mph), moderate acceleration, mild positive slope, single lane of traffic, safe RSD van area, and a moderate level of vehicle traffic.

All vehicles will have the emissions plume scanned by an RSD instrument to measure emissions concentrations. RSD instruments perform these measurements by shining a light beam across the roadway. Associated equipment will also simultaneously determine other quantities. All of these measurements will be performed without notifying the vehicle driver that they are being taken. For each vehicle the following quantities will automatically be taken as the vehicle drive past the RSD instrument at a low speed (10-20 mph):

- Item 1: DateTime: The date, hour, minute, and second of the RSD measurement.

- Item 2: Speed and Acceleration: The speed and acceleration of the vehicle.

- Item 3: RSD Emissions Absorbances: The absorbances (concentration * pathlength) of HC, CO, NO, and CO_2 in the vehicle's plume.

- Item 4: License Plate: A digital image of the rear of the vehicle so that the license plate may be determined.

- Item 5: Owner Information: The License plate will be converted to an alpha-numeric field and matched with the vehicle registration records to obtain the vehicle (including model year) and owner information.

Selection of target vehicles: Based on the information presented in Section 2, vehicle candidates will be selected for solicitation. Within each Model Year Group and RSD Evap Index bin, vehicles will be selected randomly.

4(b) Solicitation Methods

Based on the screening RSD Evap Index, the model year, and the number of vehicles desired for each bin in the stratified random plan, a sample of passenger cars and light-duty trucks will be selected for solicitation.

- Item 1: Selected vehicles will be transmitted to call center for solicitation.

- Item 2: Vehicle owners will be sent information about the study in the mail which will be followed by phone. Sample will be released in a series of replicates, each of which will be thoroughly processed before subsequent replicates are released. Interviewer will make multiple attempts at contact with each respondent.

- Item 3: Offer incentive: During the call the solicitor will offer a $10 incentive whether or not the vehicle owner participates.

- Item 4: Introduction: Explain that an emissions study is being conducted. Explain that a mechanic would like to perform measurements that would take about one hour. The solicitor will explain that the measurements would involve driving the

car past the RSD unit, performing under-hood inspections, and testing the air in the PSHED after the vehicle has sat in it for a few minutes.

- Item 5: Ownership: Ask if the driver owns the vehicle. Only vehicles with their owners driving will be eligible for participation.

- Item 6: The vehicle owner will be offered $50-$100 (experiment in fall pilot, (see section 3) as compensation for their inconvenience which they will receive when they show up and leave their vehicle for testing. Following assessment of fall pilot results a final amount will be determined and applied for the duration of the study. In addition, we will also have rental vehicles to reduce the time burden on the participants.

- Item 7: Model Year: Verify the model year of the vehicle.

- Item 8: If owner agrees to participate, solicitor makes appointment with vehicle owner to bring the selected vehicle to testing facility.

4(c)   Measurement Methods

The vehicles whose owners agree to participate and arrive for their appointment at the test site, the following information will be collected:

Vehicle owner survey – A solicitor will collect recent vehicle usage and maintenance history information at the testing site by conducting a personal interview.

Vehicle information – A technician will visually examine the vehicle to collect the vehicle identity information listed in Appendix C. Some items will be documented with a camera to reduce the chance of transcription errors.

The vehicle would then undergo the following tests:

- Item 1: RSD Emissions: A technician will drive the vehicle twice past the RSD unit at a low speed (10-20 mph). The same type of data will be recorded as for the screen drive as described in 4(a)

- Item 2: PSHED Emissions: The vehicle will be driven up to a small tent (PSHED) with non-permeable walls, turned off, and pushed into the PSHED. After closing the tent, the HC emission concentration of the air inside the tent will be measured after 15 minutes. This will serve as an estimate of the true evaporative emissions of the vehicle. Following the measurement, the vehicle will be driven out of the tent, and the air in the tent will be vented.

- Item 3: Look and smell inspection: A visual and olfactory inspection of evaporative emission control systems to look for missing, malfunctioning, damaged, or disconnected components.

- Item 4: Handheld electronic HC sniffer inspection: Detection of high evaporative emissions using a handheld HC vapor detector. The small probe of the detector will be moved around components, fittings, and hoses to try to find escaping HC vapors.   A picture will be taken on known (visible leaks) where possible.

- Item 5: To confirm precision, a small subset (~10%) of the vehicles participating in the portable SHED measurements will receive duplicate portable SHED measurements.

PSHED Characteristics from Denver Pilot Study – To develop methods to quantify evaporative emissions quickly and inexpensively, a number of methods were explored during the Denver Pilot project. One of these was the use of a portable SHED (Sealed Housing for Evaporative Determination) or PSHED to perform a short hot-soak emission test. The development of this technique included performing regular propane retention and recovery tests to establish the integrity of the PSHED. Additionally, paired hot-soak data on 23 vehicles was obtained to allow the comparison of PSHED measurements to the traditional laboratory SHED (LSHED) measurements. It should be noted that other evaporative emissions test are commonly performed during the course of EPA's certification testing, such as running loss and diurnal hot-soak testing; however, these are considerably more complex than a hot-soak test and it was not possible to perform such tests using the PSHED equipment. Pictures of the PSHED used are provided below in Figures B.2 and B.3.

Figure B.2 PSHED – side view

Figure B.3   PSHED – front view

PSHED Recovery and Retention Data – A recovery test consists of injecting a known amount of propane into the enclosure and then measuring the amount of propane detected in the PSHED after a short period of time. This test verifies that the analytical equipment is working and establishes a baseline level of propane above the ordinary ambient background level. The retention test is the measurement of the propane level in the PSHED after 15 minutes and it establishes the integrity of the PSHED ensuring that there are no major leaks. Both recovery and retention values are presented as percent difference values and will likely be negative because it is far more likely the propane will be lost from the PSHED (or any enclosure) than introduced into the enclosure by the ambient background or some other mechanism.

The average and standard deviation PSHED retention and recovery measurements are presented below. This data was collected at the start of each day and ensured that the analytical equipment was functioning properly and the PSHED had not developed any significant leaks. It can be seen from these values that the unit's performance was good and in fact exceeded expectations as no one anticipated a PSHED unit costing less than $200 plus instrumentation and gas costs could provide such accurate and precise recovery and retention values.

Average Recovery   97.6%
Recovery Std. Dev.   3.3%

Average Retention   95.7%
Retention Std. Dev.   2.3%

PSHED vs. LSHED Hot-Soak Data – The PSHED performed in the IM lane consisted of a 15 minute test that was performed after a test vehicle had been taken on a pre-conditioning drive over a prescribed 7-mile route. The pre-conditioning drive was important because LSHED hot soak evaporative emissions are measured following a standardized drive cycle on a dynamometer. Therefore, to mimic this procedure with the understanding that future PSHED measurements would likely be performed without any access to a dynamometer, the test vehicles were pre-conditioned on the road. In addition to the standard dynamometer pre-conditioning drive cycle, an LSHED is also one hour test, with strict temperature and fuel level controls. Given the objective of developing a quick evaporative test, the PSHED was only 15 minutes in duration and there was no attempt made to control the ambient temperature or the fuel level of the vehicle.

The results of the paired tests are illustrated in the Figure B.4. The results are encouraging; however, closer agreement between the PSHED and LSHED measurements was hoped for given the PSHED retention and recovery results. Vehicles were tested at the LSHED, which was at a different site from the PSHED, after many hours or the next day. The state of the vehicle, like the fuel tank temperature or other conditions associated with evaporative emissions could be different during the two tests. Evaporative emissions, and especially hot-soak emissions, are variable and after influenced by ambient temperature, prior vehicle driving, and the temperature of the various vehicle components such as fuel tank, fuel delivery systems, and the temperature of the vehicle fuel.

LSHED vs PSHED
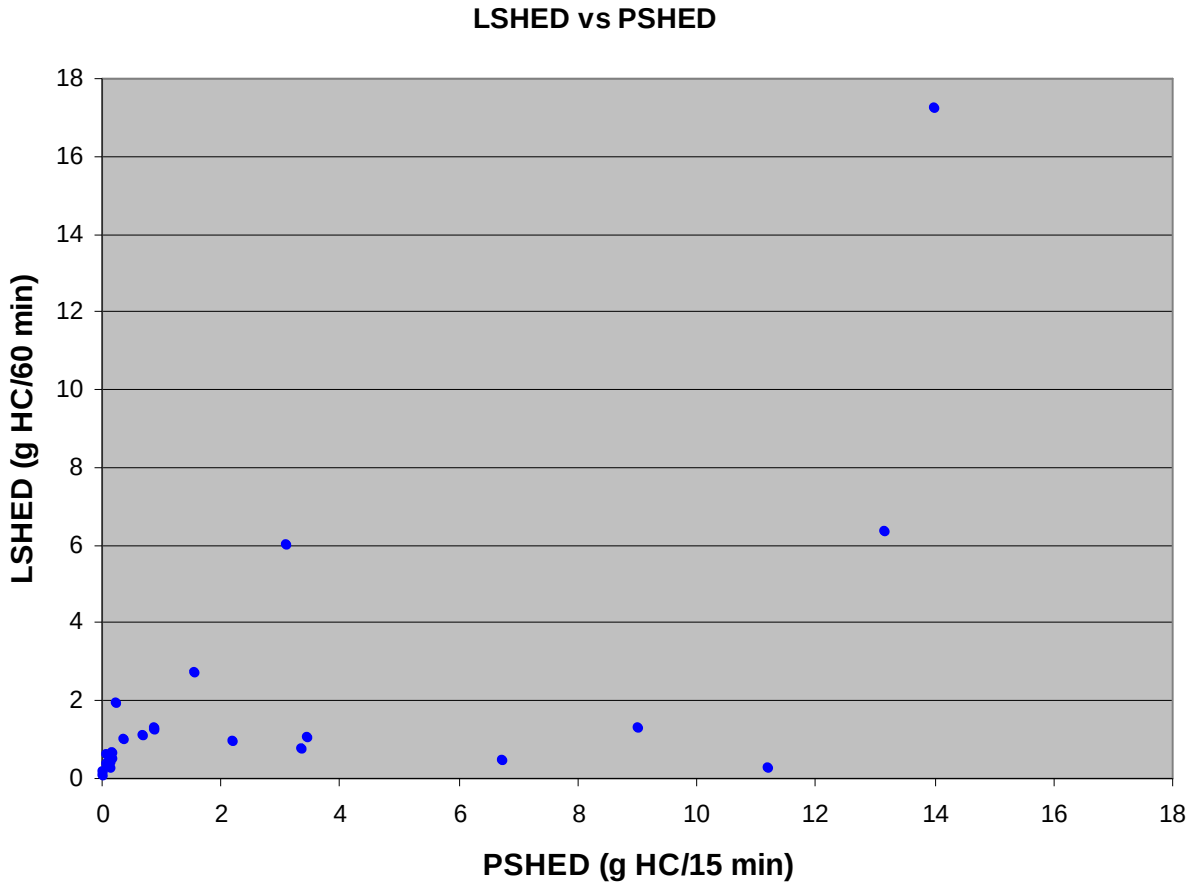
Figure B.4 Laboratory and Portable SHED correlation

5   ANALYZING AND REPORTING SURVEY RESULTS

5(a)   Data Preparation

5(a)(i) De-Stratifying the Results

Because the field tests will be performed following a stratified random design, the results need to be de-stratified to determine the measured characteristics on the basis of an application fleet. The application fleet is the fleet of interest for any particular calculation. For example, the application fleet may be the national fleet or the fleet of a particular state. De-stratification needs to be applied to the portable SHED evaporative emissions results to determine the High Evap fraction of the fleet.

5(b)   Data Analysis

5(b)(i) Non-response Analysis

Following data collection, a number of data items will be available for both respondents and non-respondents to enable analysis to response patterns.

A first step in analysis of response will be to assess response patterns in relation to characteristics of respondents or their vehicles. Detailed demographic data on vehicle owners, such as household size, age, educational level or income will not be available. However, vehicle characteristics such as number of vehicles owned, model year, vehicle type, and vehicle manufacturer will be readily available. Additional items may be constructed to incorporate information such as maintenance level and occurrence of accidents. As an index of burden, it will also be possible to construct additional measures such as driving distances from the respondents' home addresses to the study site. Based on these characteristics, response patterns can be assessed through response cells or logistic regression. This level of analysis may identify patterns of interest, may suggest potential for non-response bias, and serve as the basis for non-response weighting, but is not sufficient to confirm or estimate bias. An advantage of having screened all vehicles in the sample frame is that it should be possible to assess response patterns using the screening measure as an index of HC emissions, the response variable for the study. The success of this analysis will depend on the degree of correlation between the screening measure and the "truth" measurement. If this correlation is fairly high to high it should be possible to relate participation directly to emissions, allowing what may amount to a definitive assessment of the existence and magnitude of non-response bias.

5(b)(ii) Fraction of High Evaporative Emissions Vehicles in the Fleet

Table B.6 demonstrates how the results of the study can be de-stratified for a hypothetical application population. Of the 1000 vehicles to be sampled, 165 vehicles are expected to be "High Evaps" as shown at the bottom of Column I. The primary High-Evap-occurrence results of the study are the twelve High Evap fractions shown in Column J. With these twelve values and population fractions for any application population (such as the nationwide fleet), the overall fraction of "High Evaps" in the application population can be calculated as described in Appendix A and as demonstrated by Table B.6. For demonstration purposes, this sample calculation uses the Colorado fleet sample population fractions shown in Column F. To estimate the High Evap fraction in an application population, the Column F values must be changed to the actual population fractions for the application fleet in question.

The de-stratification results in Table B.6 show that with 1000 vehicles sampled and 165 "High Evaps" detected, the design can determine the High Evap fraction of an application fleet that is near the actual fleet High Evap fraction. The table shows that the half width of the 95% confidence interval is expected to be 28% of the estimated fraction of "High Evaps" in the fleet. This precision value of 28% meets the precision target of 30% that the 1000-vehicle design was intended to achieve.

Table B.6 De-Stratification for the Expected Study Results

A	B	C	D	E	F	G	H	I	J	K
	Bin Definitions				*Application* Vehicle Population	Expected Stratified Design Results			De-Stratifying Calculations	
	Model Year Group		RSD Evap Index			Size of Screening Sample Needed to Fill Bin	Number of Vehicles Sampled (N_h)	Number of "High Evaps" (PSHED >3g) Expected	Fraction of Vehicles that are "High Evaps" (p_h)	Std Dev of the High Evap Probability (s_h)
Bin	Lower	Upper	Lower	Upper	Population Fraction (W_h)					
1	1980	1995	0	80	0.1209	794	96	5	0.052	0.222
2			80	270	0.0034	2,977	10	2	0.200	0.400
3			270	900	0.0013	11,112	14	8	0.571	0.495
4			900	∞	0.0004	16,668	7	6	0.857	0.350
5	1996	1999	0	80	0.1902	799	152	8	0.053	0.223
6			80	270	0.0020	7,441	15	3	0.200	0.400
7			270	900	0.0013	15,627	21	12	0.571	0.495
8			900	∞	0.0002	65,483	11	9	0.818	0.386
9	2000	2009	0	80	0.6773	757	513	28	0.055	0.227
10			80	270	0.0013	38,695	52	12	0.231	0.421
11			270	900	0.0013	54,321	73	41	0.562	0.496
12			900	∞	0.0003	107,154	36	31	0.861	0.346
All					1.0000		1,000	165		

Estimated % "High Evaps" in Fleet:	Half-Width of the 95% Confidence Interval:
5.8%	1.6%

Half Width of the 95% Confidence Interval as a Percent of the Mean:	
	28%

5(b)(iii) Improved RSD Evap Index

When the data is collapsed around the four RSD Evap Index bins, they will provide information for the development of improved RSD evaporative emissions indices. As an example, Table B.7 shows the number of observations for the current RSD Evap Index. The table shows that the standard errors of the mean for the highest three RSD Evap Index bins are approximately equal. The numbers of observations in each of these bins were chosen to achieve this equality and to provide a relatively even spread of observations across the full range of evaporative emissions. Such a dataset should be valuable for developing improved indices.

Table B.7 Distribution of Planned Observations Across RSD Evap Index Bins

A	B	C	D	E	F	G
	RSD Evap Index		Number of Vehicles Sampled (N_h)	Number of "High Evaps" (PSHED>3g) Expected	Fraction of Vehicles that are "High Evaps" (p_h)	Std Error of the Mean High Evap Fraction $(sqrt(p_h*q_h/N_h))$
Bins	Lower	Upper				
1, 5, 9	0	80	761	41	0.054	0.008
2, 6, 10	80	270	77	17	0.221	0.047
3, 7, 11	270	900	108	61	0.565	0.048
4, 8, 12	900	∞	54	46	0.852	0.048
All			1,000	165		

REFERENCES

1.	D. McClement, "Raw Fuel Leak Survey in I/M Lanes," prepared for American Petroleum Institute and Coordinating Research Council, prepared by Automotive Testing Laboratories, Mesa, Arizona, June 10, 1998.

2.	H.M. Haskew, T.F. Liberty, D. McClement, "Fuel Permeation from Automotive Systems," CRC Project No. E-65, prepared for California Air Resources Board and Coordinating Research Council, prepared by Harold Haskew & Associates and Automotive Testing Laboratories, September 2004.

3.	D. Amlin, R. Carlisle, S. Kishan, R.F. Klausmeier, H. Haskew, "Evaporative Emissions Impact of Smog Check," prepared for California Bureau of Automotive Repair, prepared by California Bureau of Automotive Repair, Eastern Research Group, de la Torre Klausmeier Consulting, Harold Haskew & Associates, September 15, 2000.

4.	L.C. Landman, "Evaporative Emissions of Gross Liquid Leakers in MOBILE6," M6.EVP.009, U.S. Environmental Protection Agency, EPA420-R-01-024, April 2001.

5.	A.D. Burnette, S. Kishan, T.H. DeFries, "Evaluation of Remote Sensing for Improving California's Smog Check Program (Version 15 final)," Final Report, ARB-080303, prepared for California Air Resources Board and California Bureau of Automotive Repair, prepared by Eastern Research Group, Austin, Texas, March 3, 2008.

6.	R.F. Klausmeier, P. McClintok, "Virginia Remote Sensing Study – Final Report," prepared for Virginia Department of Environmental Quality, prepared by ESP, Tucson, Arizona, February 2003.

7.	G.W. Snedecor and W.G. Cochran, Statistical Methods, Sixth Edition, 1967, pp. 216-217.

8.	R.O. Gilbert, Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold, New York, 1987.

9.	Denver Pilot data, August 2008.

Appendix A
Stratified Sampling

The equations pertaining to stratified sampling discussed in this section are presented by

$$n_h = \frac{n \, W_h \, \sigma_h}{\sum_{h=1}^{L} W_h \, \sigma_h}$$

Gilbert [8]. The equation for the optimal sample size for a given stratum is:

where

n_h =   the sample size in stratum number h,
n =   the total sample size for all strata,
W_h =   the fraction of the actual population that falls in stratum h,
L =   the number of strata, and
σ_h =   the standard deviation of the distribution from which the individual data values in stratum h are sampled.

This equation follows conceptual guidelines. The number of points taken from a stratum is directly proportional to the fraction of the population comprised of that stratum (the fraction is W_h). Also, the number of points from a stratum is directly proportional to σ_h, which is a measure of the variability in the stratum.

The estimate of the population mean, $\overline{x}_{pop}$, is the weighted mean of the stratum means, $\overline{x}_h$:

$$\overline{x}_{pop} = \sum_{h=1}^{L} W_h \, \overline{x}_h$$

The point here is that the strata are not sampled proportionately to their actual representation in the population. If a simple arithmetic average of the complete stratified sample were calculated, the different strata would be weighted disproportionately to their representation in the population, and a biased average would result. The weighting scheme in the calculation of $\overline{x}_{pop}$ accounts for the nature of the sample and produces an unbiased estimate of the population mean. The formulation here produces the unbiased estimate of the population mean with the

minimum error variance, given the total sample size, n. The standard error of the mean is the square root of its error variance. The standard error of this weighted mean estimate is as follows:

$$\sigma_{\overline{x}_{pop}} = \sqrt{\sum_{h=1}^{L} W_h^2 \frac{\sigma_h^2}{n_h}(1-f_h)}$$

where f_h is the number of data points in stratum h divided by the population size of this stratum.

The factor $(1-f_h)$ accounts for the finitude of the population in stratum h. If the sample sizes are small compared to the sizes of the strata in the population, this factor can be ignored. The result is somewhat conservative (larger) estimates of the standard errors for the stratified results. The factor $(1-f_h)$ has been ignored (set to 1) in the calculations presented below, so the standard errors for the stratified analysis are somewhat conservative.

In practice the true standard deviations, σ_h, are not known and are estimated on the basis of historical data that exist before the planned stratified sampling effort. The sample standard deviation, s_h, based on a sample, $x_{h,i}$, i=1 to m, is:

$$s_h = \sqrt{\frac{\sum_{i=1}^{m}(x_{h,i} - \overline{x}_h)^2}{m-1}}$$

where $\overline{x}_h$ is the arithmetic mean.

When the individual data values are dichotomous, for example, 1 for a vehicle with high evap and 0 for a vehicle with low evap, then the standard deviation can also be expressed using the probability p_h that the vehicle has high evap:

$$s_h = \sqrt{p_h q_h}$$

where:

p_h is the probability that a vehicle in stratum h has high evap, and
q_h is the probability that a vehicle in stratum h has low evap ($q_h = 1 - p_h$).

Appendix B
Questionnaire

Questionnaire for Vehicle participants in the San Antonio Study

Name of Participant:

Vehicle ID :

Name of ERG Personnel:

Date :

1. Can you tell us approximately how many miles you drive in a given year?

2. Do you park this vehicle inside a garage or outside at night?

3. When was the last time you fueled your vehicle?

4. When was the last time you changed the oil in this vehicle?

5. Have you had any other maintenance performed on the vehicle in the last year?

6. Have you ever had a gasoline smell around your vehicle? If yes, could you describe the circumstance.

7. If yes, have you done anything to fix it?

8. How long have you owned your car?

9. Has the car ever been in an accident to your knowledge?

10. Have you ever replaced the gas cap?

Appendix C
Vehicle Information

1) Vehicle Identification Number

2) Make/Model/ModelYear

3) Odometer reading

4) License plate

5) Evaporative emissions control family

6) Engine displacement

7) Transmission type (manual vs. automatic)

8) Date and Time

9) IM code number (delete #9 it is not needed in an non-I/M area)

10) A picture of the car (front quarter view)

11) A close-up picture of the license plate

12) A close-up picture of the VIN (could be at the windshield location or some other like the door frame)

13) A close-up picture of the under-hood VECI label

14) Picture of found evaporative leak or cause of leak (add this)

Tracking page to document test completed

RSD
Modified CA method
Portable SHED