CRC Project E-72

Consumer Response to MIL Illumination

FINAL REPORT

Prepared for:

Coordinating Research Council

Prepared by:

Eastern Research Group, Inc. & NuStats

April 15, 2005

ERG No.: 3457.00.001.001

CRC Project E-72
Consumer Response to MIL Illumination

FINAL REPORT

Prepared for:

The Coordinating Research Council
3650 Mansell Road, Suite 140
Alpharetta, GA 30022

Prepared by:

Timothy H. DeFries
Eastern Research Group, Inc.
5608 Parkcrest Drive Suite 100
Austin, TX 78731

Aylin Sertkaya
Eastern Research Group, Inc.
110 Hartwell Avenue
Lexington, MA 02421

Della Santos
NuStats
3006 Bee Caves Road, Suite A300
Austin, TX 78746

April 15, 2005

ACKNOWLEDGMENTS

This project was conducted by the Coordinating Research Council (CRC) Real World Vehicle Emissions and Emissions Modeling Group operating under the direction of the CRC Emissions Committee. The project leader was Hannah Murray of Toyota. We gratefully acknowledge all of their contributions to this report. We also would like to thank Brent Bailey of CRC for his cooperation, communication, and assistance at every stage of this project.

K. Helmer, EPA
J. J. Jetter, Honda
J. Koupal, EPA
D. R. Lawson, NREL
D. H. Lax, API
J. Lindner, EPA
J. R. Long, CARB
H. Maldonado, CARB
M. M. Maricq, Ford Motor
S. R. Mazure, DaimlerChrysler
E. McCauley, CARB
F. Minassian, SCAQMD
H. Murray, Toyota
E. K. Nam, EPA
M. T. Noorman, ExxonMobil
D. W. Raney, American Honda
C. H. Schleyer, ExxonMobil
A. Schuettenberg, ConocoPhillips
S. A. Shimpi, Cummins, Inc.
J. H. Somers, EPA
E. Tierney, EPA
B. Trestrail, Int'l. Truck & Engine
J. P. Uihlein, BP
W. Vance, CARB
M. I. Watkins, ExxonMobil
J. S. Welstand, ChevronTexacoProds.
K. J. Wright, ConocoPhillips

Table of Contents

List of Tables

List of Figures

PLEASE REFER TO THE GLOSSARY AT THE END FOR THE EXPLANATIONS OF VARIOUS TERMS, ACRONYMS, AND ABBREVIATIONS USED THROUGHOUT THE REPORT.

1.0 Executive Summary

Under contract to the Coordinating Research Council (CRC), Eastern Research Group, Inc. (ERG) and its subcontractor, NuStats, conducted a national study to characterize the response of vehicle owners/drivers in non-inspection/maintenance (non-I/M) areas to illuminated Malfunction Indicator Lights (MILs) on their 1996 and newer vehicles – those equipped with federally-mandated onboard diagnostic (OBD) systems. Data on these vehicles and owners/drivers were obtained by a telephone survey using a questionnaire developed specifically for this project. The vehicles of all respondents were registered in counties without an I/M program that included an emissions test. The impetus for the project was to gather and analyze data that could be used to update the non-I/M response to illuminated MILs currently used by Environmental Protection Agency's (EPA) highway vehicle emission factor model, MOBILE6.2.[1] This report documents the findings of this study, including its design and implementation.

1.1 Study Background and Objectives

Currently, EPA's highway vehicle emission factor model, MOBILE6.2, incorporates several assumptions related to OBD systems that address the effectiveness of these systems and how vehicle owners will respond to them. In non-I/M areas, the model approximates consumer response to illuminated MILs as given in Table 1-1.

Table 1-1. Estimated MIL Repair Rates Used in MOBILE6.2 for Non-I/M Areas

Odometer Reading	Percent of Vehicles with Illuminated MILs That are Repaired
Less than 36,000 miles	90%
36,000 to 80,000 miles	10%
More than 80,000 miles	0%

These assumptions were formed in the absence of related data and hence may result in over- or under-estimation of vehicle emission inventories. For example, in year 2020 MOBILE6.2 estimates that, on average, vehicles will produce 0.3 g/mile of NO_x emissions. If

[1] Note that this study was specifically focused on the response to MIL illuminations on vehicles in non-I/M areas. Presumably, data from I/M program vehicle information databases (VIDs) could be used to determine responses to MIL illuminations on vehicles participating in I/M programs – at least for response rates at the time of inspections. If estimates of the response rates for MILs that illuminate between regular I/M inspections are desired, some modifications to this questionnaire would need to be made – at least to properly screen respondents to ensure that they are in I/M areas, to find out if they participate in I/M or not, etc.

however, there is a 45 percent response to MIL illuminations resulting in repair, vehicles will produce 0.25 g/mile of NO_x emissions, the equivalent of a 17 percent reduction.

In light of the above, the primary objective of this study was to characterize how vehicle owners in non-I/M areas respond to the illumination of the MIL present on 1996 and newer vehicles (i.e., those equipped with federally mandated onboard diagnostic systems) via a survey of vehicle owners. The survey generated statistically valid information on whether vehicle owners get their vehicles repaired in response to MIL illuminations and the degree to which this response changes as mileage increases, ownership changes, and warranties or extended service contracts expire. We then used the collected data to build models of response rates as a function of relevant factors, such as vehicle mileage, warranty status, vehicle model year, etc.

1.2 Survey of Consumer Response to MIL Illumination

The survey used to collect data utilized standard computer-aided telephone interviewing (CATI) technology and was based on a random digit dial (RDD) sample of households with land-based telephone numbers. While the survey universe excluded those households with no telephones and those with cellular telephones only, such noncoverage is expected to have a negligible impact on the survey results. First, only 2 percent of households do not have land-based telephone service according to data from the 2000 US Census. Second, it is estimated that only about 3 percent of households in the US are serviced exclusively by cell phones.[2]

The survey work commenced with the design of the questionnaire in March 2004, followed by cognitive testing of the instrument in late June 2004. Pretesting of the CATI questionnaire occurred in mid-August 2004. Following minor revisions to the questionnaire, full data collection ran from September 7, 2004 to October 10, 2004. In total, 14,500 telephone numbers were called in the study. Of the 3,415 known households in the sample, 1,406 households indicated they operated a 1996 or newer vehicle and saw lights on the dashboard displayed but could not confirm if it was a “check engine” related light. These were considered “short interviews” because they simply collected information on the respondent’s perception of the “check engine light.” Another 606 households said they reported “episodic” MIL experiences, and were administered a longer survey instrument (“long interview”). The overall response rate to the survey was 43 percent. Table 1-2 provides a summary of completed interviews by household type and source.

[2] The list-assisted RDD design ensures a very high coverage of households in the U.S. regardless of whether or not they have unlisted numbers. Thus, households with unlisted telephones are not problematic for this survey design.

Table 1-2. Summary of Completed Interviews by Household Type and Source

Source	Survey	Pretest	Total
Completed Interview Type:			
With MIL experience	569	37	606
Without MIL experience	1,330	76	1,406
Total	**1,899**	**113**	**2,012**

1.3 Statistical Analysis of Survey Data and Response Function Development

Upon completion of the telephone survey, we performed an analysis of the data gathered to determine the factors, such as vehicle mileage, ownership status, warranty status, that are likely to affect the response of vehicle owner/drivers to an illuminated MIL. We also modeled the functional relationship of these factors to vehicle owner/driver response using regression analysis. Because some of the questions in the survey were slightly modified as a result of the experience gained during the survey pretest, we limited the data used in the analysis only to those collected during the full-scale survey (569 out of 606).

The key findings of this analysis are summarized below:

- Overall, of the 569 respondents who experienced a MIL illumination episode on one of their 1996 to 2004 vehicles had a positive Response, that is they claimed to have taken the vehicle for repairs or tried to repair it themselves.[3]
- In addition, after reviewing the open-ended answers to the questions, we found that 91 percent of the 569 respondents had a positive Response to the MIL illumination, indicating that the owner either took the vehicle for repairs, repaired the vehicle himself, or the MIL went out on its own. A positive Response does not necessarily mean the action caused the MIL to be extinguished.
- From a MOBILE modeling perspective, the most statistically significant variable that affected the overall Response rate was the presence of a warranty or service contract.
- The odometer reading at the time of the MIL illumination had a significant effect on the Response rate at the 87 percent confidence level. However, when the existence of a warranty was considered first, the effect of the odometer reading at MIL illumination became insignificant.
- Whether a vehicle was company owned or privately owned had a statistically significant effect on the Response. However, we did not perform further investigation of this subset as only 15 vehicles in the dataset were company owned.

[3] In this study, "repair" means work was done and does not necessarily mean the MIL was extinguished.

- Claimed knowledge of the meaning of a MIL illumination also had a statistically significant effect on the Response at the 86 percent confidence level. Those who claimed to know what a MIL meant were less likely to have a positive Response.
- The gender of the respondent to the survey had a significant effect on the Response rate at the 97 percent confidence level. Males were less likely to have a positive Response than females.
- No statistical significance for Response was found for vehicle model year, whether the vehicle was obtained new or used, or whether the current owner had previously experienced an illuminated MIL on the vehicle .
- Table 1-3 shows a summary of the effect of warranty status on the Response to the most recent illuminated MIL. The table also shows the modeled fraction of MILs that were still on after the vehicle was driven 5,000 miles after the illuminated MIL was first seen. (The models for MILs-still-on after any mileage driven after a MIL illumination are provided in the report.)
- For example, 94 percent of the owner/drivers of vehicles that were under warranty that saw an illuminated MIL made some sort of positive Response. Many took their vehicles for repairs, a few attempted to repair the vehicle themselves, and some MILs went out on their own. However, the action taken was not always effective at extinguishing the MIL. For this group, after driving 5,000 miles after the MIL illumination episode, 15 percent of the vehicles still had illuminated MILs.
- The difference between the rate of positive Response is not strongly dependent on the presence of a warranty (even though it is statistically significant).
- There is a large and important difference between the MIL-still-on rates for vehicle owners with a positive and negative Response – whether or not the vehicle is under warranty.

Table 1-3. An Illustration of the Major Factors Affecting Response to a MIL Illumination and MIL-Still-On Rates

	Type of Response to Illuminated MIL	Rate of Response (Percent)	Percent of MILs-Still-On After Vehicle Driven 5,000 Miles After MIL Illumination
Warranty	Positive	94%	15%
	Negative	6%	70%
No Warranty	Positive	89%	20%
	Negative	11%	70%

It is important to recognize that the information analyzed was obtained by administering a questionnaire to vehicle owners and/or drivers. The data obtained in this study cannot be used to independently validate their responses.

Because respondents who had never seen any dashboard light come on were screened out of the survey (see question SQ4 in the survey), the data obtained cannot be used to make statistically valid determinations of MIL illumination rates for comparison to reported ones. Cursory examination of the data collected, however, show that MIL illuminations are more common in older vehicle model years, as would be expected. The distribution of the types of vehicles by make in the survey data also correlate well to independently reported vehicle market shares. Combined with the fact that our survey design ensured a reasonable representation of households from non-I/M areas, we think that the vehicle data collected is representative of the vehicle universe.

1.4 Implications of Study Findings

The survey results indicate that motorists in non-I/M areas respond to an illuminated MIL at substantially higher rates than assumed in EPA's MOBILE6.2 model. Table 1-1 above presents how EPA has estimated illuminated-MIL repair rates in non-I/M areas for three odometer categories in the current version of MOBILE6.2.

EPA had to develop these estimates in the late 1990s for inclusion in the MOBILE6.2 model, at a time when there were no data available on real-world response to OBD2 MIL illumination. Therefore, EPA assumed that response rates would vary as a function of warranty status, with full-warranty vehicles (i.e., vehicles with less than 36,000 miles) having the highest response rates, and vehicles fully out of warranty (i.e., vehicles with greater than 80,000 miles) having no response. The results of this study, however, indicate that the MILs of 82 percent [=(94%)(100%-15%) + (6%)(100%-70%)] of vehicles under warranty and 75 percent [=(89%)(100%-20%) + (11%)(100%-70%)] of vehicles not under warranty were extinguished within 5,000 miles after a MIL illumination. Further, the study shows that these response rates continue to increase beyond 5,000 miles.

In non-I/M areas, there is no requirement that motorists repair illuminated MILs – they can drive their vehicles indefinitely with the MIL on, as long as the vehicle keeps operating. In I/M areas, however, vehicles with illuminated MILs are required to obtain repairs until the MIL is extinguished. Therefore, the effective "response rate" to MILs in I/M areas is very high (equivalent to I/M program compliance rates) and independent of warranty status.

The relative difference in MIL response rates that EPA has assumed between I/M and non-I/M areas has a significant impact on the emissions benefit of I/M programs, as modeled by MOBILE6.2. Given current MOBILE6.2 assumptions, adopting an I/M program will significantly increase the fraction of vehicles with illuminated MILs receiving repairs. However, based on this study results, response rates to illuminated MILs are already high in the base case (i.e., in non-I/M areas). *Therefore adopting an I/M program will have relatively little incremental improvement in MIL response rates and hence, in repairs and emission reductions among OBD2 vehicles.*

We acknowledge that the rates determined in this study may be biased upward given that they are based on self-reported behaviors of vehicle drivers/owners. However, our findings strongly support the notion that when drivers/owners see an important light illuminate on their dashboard, they are more likely to get the problem fixed than to drive the vehicle for years with the light on. After all, anticipated voluntary response to MIL illumination is one of the main reasons that MILs on the dashboard were mandated in the first place.

If the results in this study accurately characterize behavior in non-I/M areas, then the MOBILE6.2 estimate of the benefit of the MIL illumination component of an I/M program may be greatly over-estimated for vehicles with more than 36,000 miles. For example, MOBILE6.2 might estimate the benefit at 90 percent (100 percent in I/M minus 10 percent in non-I/M) whereas the true benefit could be 25 percent (100 percent in I/M minus 75 percent in non-I/M). In addition, emissions inventories for non-I/M areas developed using MOBILE6.2 may overestimate emissions somewhat given the low default MIL response rates assumed in the model.

2.0 Survey Methodology

In Project No. E-72, CRC wanted to determine the response rates of vehicle owners to illuminated MILs, to find the variables that affect the response rates, and to build models that estimate the response rates as a function of the important variables. Additionally, CRC requested that the data for the study be collected via a statistical survey of owners/drivers of 1996 or newer vehicles. Thus, under contract to CRC, Eastern Research Group, Inc. (ERG) and its subcontractor, NuStats undertook a national study of consumer response to MIL illumination.

This section describes the survey methodology, including analytical objectives, questionnaire design and sampling. The section also provides the overall sample disposition (i.e., complete account of the number of phone interviews as well as interview attempts) for the survey.

2.1 Survey Objectives

The design of the survey commenced with a review of concepts/items that would satisfy analytical requirements of the study. The following were key elements to the analytical objectives:

Respondent Eligibility

- Person 18 and older
- Responsible for the maintenance of vehicles

Household Location

- Non-I/M county or Non-I/M city[4]
- Not adjacent to an I/M county

Vehicle Information

- All 1996 or newer vehicles registered in non-I/M areas[5]
- Year
- Make

[4] The state of Virginia has independent cities that do not belong to any county.

[5] While areas with I/M programs were out of scope and hence not included in our sample, the survey still had to screen for those respondents residing in non-I/M areas but had their vehicles registered in an I/M area.

- Model
- Check engine light illuminated on the dashboard while being driven
- Owned or leased
- Year and month acquired
- Purchased/leased new or used
- Mileage at the time of purchase/leased
- Present mileage

MIL Illumination Episodes

- Most recent episode
- Mileage
- Existence of other episodes

Respondent's Knowledge and Behavior of MIL Illumination

- Awareness and understanding
- Sources of information

Vehicle Repair actions

- Actions taken
- Location of vehicle repair

2.2 Questionnaire Design

The survey design entailed the crafting of questions to address the objectives of the study mainly 1) determination of the response rates of vehicle owners to illuminated MILs, 2) collection of information on those variables, such as mileage and warranty status, believed to impact response rates, and hence would be needed in building models that estimate the response rates as a function of most relevant variables.

One of the most important issues was the design of survey questions to ensure that the respondent knew which dashboard light was the MIL. We anticipated that since the MIL and the maintenance light can have similar text, respondents could be confused about which type of light they saw illuminate. Accordingly, we spent considerable effort working on the design of questions to address this issue. The questionnaire used for the cognitive interviews[6] contained several exploratory questions to determine if a respondent was able to properly recall which light was the MIL, which was referred to as the "check engine light." (See pages 10 and 11 of the cognitive interview report in Appendix A.) However, all six of the respondents who reported MIL illuminations correctly identified the MIL, which was verified by examination of each vehicle's dash after each cognitive interview. Ultimately, in the final questionnaire used for the survey, we asked respondents a separate illumination question about each of the two lights. (See questions VQ11 and VQ12 and their introductory text on pages B-4 and B-5 of Appendix B.) We believe that those two questions were able to increase the likelihood that we were getting detailed information only from respondents that had seen MIL illuminations. Nevertheless, some responses to Question MQ23 (see Table 4-24: Sample Number 2703 and Table 4-27: Sample Numbers 1660, 10215, 12443, 13365) indicate that some respondents that witnessed maintenance light illuminations rather than MIL illuminations may have slipped into the dataset.

It was essential that the survey design be sensitive to the accurate retrieval of information from memory to respond to the request for information. For this reason, we conducted a series of cognitive interviews with owners/drivers of vehicles that met the study requirements. The objective of cognitive testing was to assess the respondents' ability to generate a response by examining their 1) comprehension of the questions; 2) their ability to retrieve relevant information from memory; and 3) assess the adequacy of the questionnaire structure and flow design.

The result of the cognitive tests provided insights on MIL incidence and response behaviors in a way that clearly established how to screen for qualified respondents and collect detailed information about specific and unique MIL episode events. There were several design recommendations that resulted from the cognitive interview tests that improved the quality of the survey instrument content and flow, reduced the complexity of some questions and, overall, lessened respondent burden. Appendix A provides the full Cognitive Interview Report submitted to CRC in August 2004.

Prior to formal data collection, a pretest of the revised survey questionnaire took place in mid-August. The purpose of the pretest was to evaluate the performance of the survey on a small

[6] A cognitive interview is a preliminary test of the draft survey questionnaire with persons that possess the similar characteristics of the survey's intended audience and involves in-person interviewing. The testing objectives are related to the question-answering process of potentially complex questions in that they assess the respondents' ability to generate a response by examining their comprehension of questions and their ability to retrieve relevant information from memory. Cognitive interviews are also used to assess the adequacy of the questionnaire flow (structure and design) (see Appendix A for the complete Cognitive Interview Report for the study).

portion of the actual study sample. Because the pretest is generally a component of interviewer training, we also conducted interviewer debriefings after the pretest. The results of the pretest contributed to a number of improvements to the survey instrument (see Appendix B) and the interviewer-training manual (see Appendix C). For example, the pretest revealed that respondents were unable to clearly recall MIL illuminations in more than one vehicle (the vehicle with which they were most familiar), and due to frustration with repetitive questions/test, some respondents were unwilling to participate in a complete interview. As a result, we modified the survey instrument to only collect data on one vehicle per household and made minor edits to the test to eliminate respondent frustration with the interview.

2.3 Sampling Design

The *population* of inference for this study was composed of households in a specific geographic coverage area that had one or more vehicles manufactured between 1996 and 2004 that have experienced a MIL illumination episode by fall of 2004. The geographic coverage area consisted of all non-I/M counties and non-I/M independent cities in the U.S. that do not border I/M counties.

The survey objectives for this study called for a sample of 600 MIL event households. MIL households were to be determined through screening questions that were asked as a precursor to the telephone interview. Only one interview per household was involved.

We drew a stratified, list-assisted[7] random digit dial (RDD) sample of households for this study. Under this design, households with land-based telephone numbers[8] have an equal chance of selection (apart from the small proportion that have more than one land-based telephone

---

[7] "List-assisted" is a survey research industry term that pertains to how the sampling frame was developed. In the telephone survey world, it is very inefficient to simply draw a random sample of all possible telephone numbers in the U.S. Only about 1 in 5 such numbers from such a sample turns out to be a household. To improve the "hit rate," sampling statisticians developed a "list assisted" Random Digit Dial technique. It starts by taking the pool of all possible telephone numbers in the universe, and allocating them to corresponding "banks" of 100 consecutive telephone numbers – e.g., 512 306-9065 would belong to bank 512 306-90XX where XX ranges 00 to 99. These banks represent 100% of the land based households in the U.S. Next, a file containing all published telephone numbers in the U.S. is used to identify the subset of all banks that contain at least one published (i.e., listed in the phone book) household telephone number. The collection of these banks is called 1+ banks. Research has shown that 1+ banks contain about 97 percent of all telephone numbers of households (with land based telephones) in the U.S. A list-assisted RDD design draws a random sample from all telephone numbers (listed, unlisted, business/government, nonworking, etc) that are members of the 1+ banks. It turns out that the household hit rate for such a design increases to about 50 percent (relative to the crude RDD design). Plus, coverage of households in the U.S. is very high whether or not they have unlisted numbers.

[8] Census 2000 data show that only 2 percent of households do not have telephone service. And about 3 percent of all households in the US are serviced exclusively by cell phones. Such low levels of noncoverage would have negligible impact on survey results.

number). A proportionate stratified sample was drawn from the collection of counties that comprised our coverage area. The stratification we used incorporated state and county. In the survey, household eligibility was determined by asking a number of screening questions to the person most knowledgeable about auto maintenance within the household. We drew and called a total of 14,500 telephone numbers for this survey.

2.4 Survey Disposition

Table 2-1 presents a summary of the final numbers of completes for this study. A total of 606 completed MIL illumination event household interviews were obtained, with 569 of these coming from the survey data collection and the balance (n=37) originating from the survey pretest. Among the non-MIL illumination experience households with vehicles built between 1996 and 2004, a total of 1,406 completed interviews were obtained, of which 1,330 originated from the sample survey and the remainder from the pretest.

Table 2-1. Summary of Completed Interviews by Household Type and Source

Source	Survey	Pretest	Total
Completed Interview Type:			
With MIL experience	569	37	606
Without MIL experience	1,330	76	1,406
Total	**1,899**	**113**	**2,012**

The overall disposition of the sample (i.e., complete account of the number of phone interviews as well as interview attempts) is provided in Table 2-2. Note that for the sake of reporting survey performance, the 113 pretest cases are intentionally *excluded* from our response rate analysis. Their inclusion would mask the disposition profile of the survey sample (making response rates appear higher than they should be), thereby giving an optimistically skewed view of survey performance. Thus, the response rates are based on a total of 1,899 completes (for the sake of reporting survey performance). While Table 2-2 documents the final sample disposition, it does not provide easy access to response rates due to the inclusion of the sample that was not in the universe (e.g., disconnected numbers). It is provided here for the sake of completeness.

Table 2-3 provides the screening response rate for the study. We attained a screening response rate of 48 percent. In other words, we were able to successfully determine the eligibility status of just under half the purchased RDD sample. This level is on par with (or better than) that of the survey research industry for random digit dialing surveys requiring screening to determine eligibility.

Table 2-4 presents the interview response rates among the set of households found to be eligible for interviewing. For these households, a 90 percent interview response rate was achieved. It appears that once the subjects answered the screening questions, they were almost always willing to continue with the longer interview (be it the "short" set of questions for the households without MIL experience, or the longer set of questions for those with MIL experience).

Table 2-2. Final Disposition of the Sample

Households with Known Eligibility Status	Number	Percent
Completed Interviews among Eligible Households (HHs)		
With MIL experience	569	3.9%
Without MIL experience	1,330	9.2%
Unable to Interview Eligible HH	178	1.2%
Refusal of Eligible HH to be Interviewed	38	0.3%
Ineligible HH		
Invalid State/County	79	0.5%
No Vehicles/Only Older Vehicles in HH	1,036	7.1%
Language Barrier	185	1.3%
Undetermined Household Status		
Answering Machines	1,490	10.3%
Busy	317	2.2%
No Answer	1,935	13.3%
Hang-ups	1,733	12.0%
Caller ID devices [a]	23	0.2%
Refusal to Screening Questions	2,094	14.4%
Not a Household		
Business	593	4.1%
Disconnections	2,116	14.6%
Faxes	784	5.4%
Total	**14,500**	**100%**

[a] The existence of a Caller ID device is detectable by the household line. What the calling center receives is a message indicating that the call is not an acceptable number to be received by that household telephone.

Table 2-3. Screening Response Rates for the MIL Household Survey

Screening Disposition	Number	Percent
Telephone # Screened	**6,908**	**48%**
Screened Eligible HH	2,115	
Screened Ineligible HH	1,300	
Telephone Number is a Non-HH	3,493	
Telephone # Not Screened	**7,592**	**52%**
Answer machines	1,490	
Busy	317	
No Answer	1,935	
Hang-ups	1,733	
Caller Id devices [a]	23	
Refusals	2,094	
Total	**14,500**	**100%**

[a] The existence of a Caller ID device is detectable by the household line. What the calling center receives is a message indicating that the call is not an acceptable number to be received by that household telephone.

Table 2-4. Interview Response Rates for the MIL Household Survey

Interview Response Rates	Number	Percent
Completed Interview	1,899	90%
With MIL experience	569	
Without MIL experience	1,330	
Not Interviewed	216	10%
Total	**2,115**	**100%**

The RDD sample of households comprises an equal probability sample of households with telephones (ignoring the small fraction of households with more than one land line). As such, weighting is not required for the analysis of this data set. We judged that post-stratification weighting was not appropriate given the unique definition of the population. There are no available population totals that could be used to fine-tune the sample distributions to population distributions. (Census 2000 data are unavailable.) Thus, the survey data were analyzed without the use of weights.

Appendix D presents the detailed frequency tabulations of survey responses.

2.5 Concluding Remarks

The survey provides a reasonable representation of households from non-I/M states, counties and independent cities. The study's methodological approach included careful questionnaire design, cognitive testing of the instrument, a scientific, probability based sample design, a pretest, the CATI data collection and finally, the processing of survey data for analysis.

The study employed an RDD sample design for household selection and incorporated a three-stage screening process to determine household eligibility. The first stage determined if the household was located within a non-I/M county, the second stage affirmed that there was a household member eighteen and older who was familiar with the maintenance of household vehicles manufactured between 1996 and 2004. The third stage of screening required that all vehicles had to be registered in the county where the household resided and at least one of those vehicles had to have experienced a “check engine light going on” while it was being driven. This tri-stage screening process ensured the capture of MIL illumination occurrences desired for the study.

Overall, the survey data captured useful aspects of public awareness on the diagnostic functionality of the check engine lights. The data provide insights about public perceptions of a MIL illumination experience, the actions individuals take when it happens, and their source of information about the MIL.

The survey results can be used for statistical analyses to determine the factors that affect the response of vehicle owners/drivers to an illuminated MIL and to build models around the most important of those factors for the purposes of discovering the important trends in vehicle owner/driver response to an illuminated MIL.

3.0 Preparation of Survey Data for Statistical Analysis

The survey database consists of three related data tables. Appendix E contains the Data Matrix Listing, which describes all of the variables in each of the data files and follows the flow of the questionnaire. MAIN is the data table that contains household information. ROST1 is the data table that describes the model year, make, and model of each 1996 through 2004 vehicle that is in each household. ROST2 is the data table containing additional information on vehicles that have had dash lights go on. All of the information pertaining to the most recent MIL illumination episodes is included in ROST2; however, ROST2 also includes information on some vehicles where a MIL was not illuminated. Each observation in MAIN is given a sample number (Sampn)[9] for each household. Each vehicle in ROST1 is given the same sample number and a vehicle number (Vehn1). Each observation in ROST2 also has a sample number and a vehicle number. Consequently, ROST2 is linked to ROST1 by sample number and vehicle number, and both ROST1 and ROST2 are linked to MAIN by sample number alone.

The first step in the preparation of the data for analysis was to restrict the observations to those households that were in the main part of the survey. The survey was conducted in five batches: Batch 0, 1, 2, 3, and 4. Batch 0 was a pretest after which some of the questions in the survey were modified as a result of the experience gained during the pretest telephone survey. Consequently, for the data analysis, we eliminated Batch 0 from further consideration.

For MAIN, the variables for county and city were combined to create a new variable called Location. MAIN contained 1,899 observations.

For ROST1, we created a new variable, Make, that spelled out the make of the vehicle based on the make code variable used during CATI administration. In addition, for those instances where the respondent gave an open-ended response to the question on vehicle model, we combined the menu-driven model responses with the open-ended responses to ensure that all values of Model had the same spellings. ROST1 contained 1,186 observations.

For ROST2, a number of new variables were added to the data file to assist in subsequent data analysis. The first area had to do with vehicle odometer readings. Respondents were asked in questions MQ16, MQ18, and MQ17 to give odometer readings for their vehicles when they first acquired their car, for the most recent MIL illumination episode, and the current odometer reading. For the majority of cases, respondents provided an answer that was an actual mileage.

[9] In the text of this report, the dataset variable names begin with an upper case letter. See the Data Matrix Listing in Appendix E for the text of the corresponding questionnaire question.

However, some respondents, in some cases, could only estimate the range of the odometer reading, as provided by the questions in the questionnaire.

To facilitate analysis of the MIL illumination data, we needed to put all of the odometer readings on an actual mileage basis. Thus, we examined the individual observations for all vehicles where the owner reported an illuminated MIL and considered the odometer information for each vehicle separately. In many cases, owners could remember their actual odometer readings, whereas some were only able to recall the approximate range of their odometer readings. For each vehicle that had at least one range estimate, we used the actual odometer readings reported to estimate specific odometer readings for each response with a range. The odometer reading estimates assigned to each of the 89 vehicles requiring an estimate are shown in Appendix F.

Here is an example of how these odometer readings were estimated. Consider the first observation in Appendix F; i.e., Sampn=275 and Vehno = 2. For this particular vehicle, the respondent indicated acquiring the vehicle at 25,000 miles and that the vehicle currently had 42,000 miles on it, but he was only able to estimate the mileage of the MIL illumination episode as falling in the range of 25,000 to 50,000 miles. In this case, to estimate an actual mileage for the MIL illumination episode, we started with the range of 25,000-50,000 miles and then considered the other exact mileages that were reported. In this case, if he got the vehicle at 25,000 miles, the lower limit on the range of 25,000 miles stands. On the other hand, if the current mileage is 42,000, the upper limit on the mileage range for the episode cannot be 50,000 miles. Therefore, the odometer reading range for the MIL illumination episode for this vehicle was restricted to 25,000 to 42,000 miles. To arrive at a single mileage for the episode, we averaged the low and high values to arrive at an estimated mileage of 33,500 miles for the MIL illumination episode of this vehicle.

In nine cases this procedure was not able to produce specific odometer readings because of conflicts or ambiguities in the information provided by the respondent. For example, consider the observation in Appendix F for Sampn=12683 and Vehno = 1. For this particular vehicle, the respondent indicated acquiring the vehicle at 38,000 miles and having a MIL illumination at 38,050 miles. However, the respondent also indicated that the vehicle currently had less than 25,000 miles on the odometer, which is impossible since the earlier odometer readings have higher values. For this situation, and the other eight like it, the final odometer assignment for the current odometer reading was designated as “Not Possible” in Appendix F and was set to missing in the dataset.

After this procedure was completed for the 89 vehicles that had odometer reading ranges in the survey, we examined the results to ensure that all mileages were monotonically increasing from the mileage at initial acquisition, the mileage at MIL illumination episode, and the current mileage. Where mileages were not monotonically increasing, we changed all three odometer readings to missing values. This was done for the ten vehicles listed by sample number and vehicle number in Table 3-1.

Table 3-1. Vehicles with Odometer Readings Set to Missing

Sample Number	Vehicle Number
1296	2
1334	1
2911	1
3065	1
3354	1
5296	1
7178	1
9243	1
11299	1
11773	1

Next, we used these odometer readings to calculate new variables for analysis. The number of miles that the current owner had put on the vehicle between original acquisition and the most recent MIL episode (Mile_owner) was calculated as the odometer at the episode minus the odometer at acquisition. The number of miles put on the vehicle by the owner since the MIL illumination episode (Mile_since) was calculated as the current odometer reading minus the odometer reading at the time of the episode. For the purposes of the analysis and graphical presentation of the results, the different odometer readings and differences between odometer readings were binned in various ways. The current age of the vehicle (Vehage) was also calculated by rounding the number of years between January 1 of the model year of the vehicle until October 1, 2004, which was the approximate date of the telephone survey.

While a large number of respondents reported that they made repairs on vehicles, there was also a significant number of other vehicles that were not repaired but that had the MIL go out on its own. Accordingly, we created a new variable called Response that had values of Positive and Negative. Response was assigned to be Positive in all cases where respondents claimed to have made a repair (MQ22E: Whtd5 = 1).[10] For all other cases, where the respondents did not claim to have made a repair (MQ22E: Whtd5 = 2), we considered the

[10] In this report, we occasionally give in parentheses the question, variable name, and the coded response, as defined in the Data Matrix Listing in Appendix E, to describe specifically the meaning of the text.

response to Question MQ23 ([Whyh*] – "Why did you choose not to take the vehicle to be repaired/serviced?") for each vehicle individually.

Based on the answer to MQ23, we decided to assign Response as Positive or Negative. Response was set to Positive only if the respondent indicated that the MIL went out on its own, that he took the vehicle in for repair, or that he was given guidance (such as tightening the gas cap) by a mechanic that resulted in the MIL going out. The specific answers to MQ23 for these positive assignments are shown later in Table 4-24.

For any other answers to Question MQ23, Response was set to Negative. This included having an appointment, having had the problem diagnosed but not repaired, knowing what the problem was but not having it repaired, claiming that the light just came on, stating that the car runs fine, being too busy, and not being willing to pay for repairs – all of which indicate that the lit MIL had not been addressed. The specific answers to MQ23 for these negative assignments are shown later in Table 4-27.

Appendix G shows the assignments of Response made to the 569 observations where vehicles had reported MIL illuminations.

4.0 Preliminary Analysis of the MIL Illumination Dataset

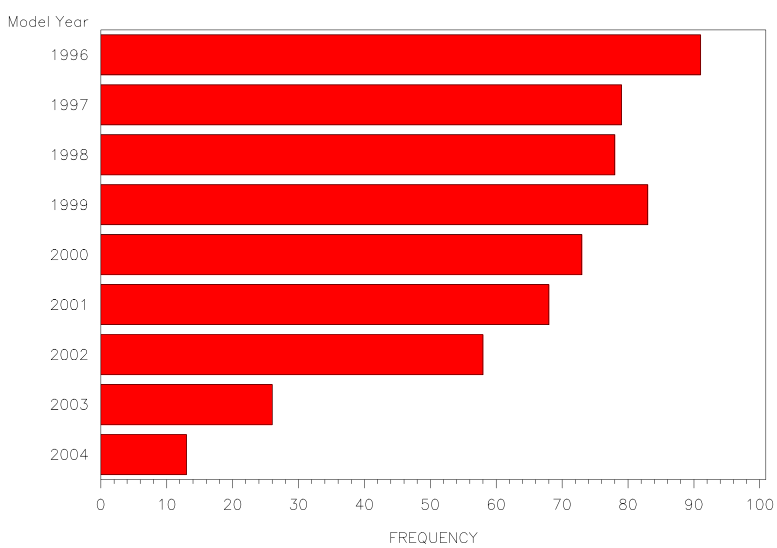
As noted previously, the occurrences of recent MIL illuminations were reported on 569 vehicles. Appendix G gives the sample number, vehicle number, model year, make, and model of these 569 vehicles. Since these data will be used to evaluate trends and build models of MIL illumination response rates, we initially examined the characteristics of this dataset.

The distribution of the model years of the vehicles is shown in Figure 4-1. The figure shows a generally decreasing trend in the number of reported MIL illuminations as model years become newer. Figure 4-2 shows the distribution of the odometer readings on these vehicles at the time of the telephone survey. Odometer readings are presented in bins with midpoints every 25,000 miles. Approximately 90 percent of the vehicles had less than 150,000 miles.

Figure 4-3 shows the distribution of odometer readings at the time of the MIL illumination episodes for the vehicles. Approximately 95 percent of the vehicles had MIL illumination episodes at mileages less than 150,000 miles. The median mileage for the most recent MIL illumination episode was approximately 65,000 miles.

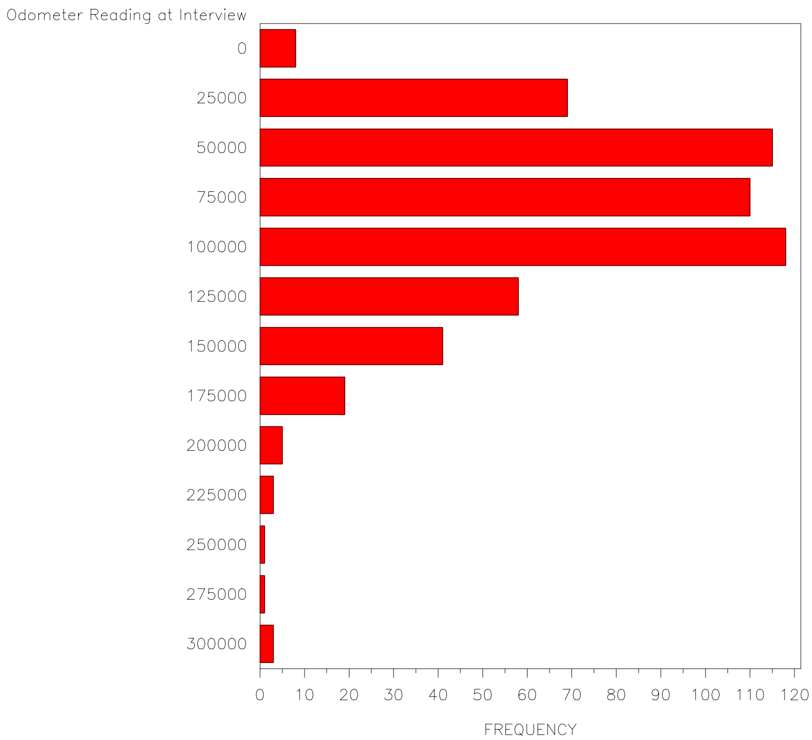
Figure 4-4 shows the distributions of miles that the vehicle owners had put on their vehicles between the time of the MIL illumination episode and the time of the telephone survey. About 42 percent of the respondents reported their most recent MIL episode as less than 7,500 miles ago. Approximately 95 percent of the vehicles had their most recent MIL illumination episode within the last 50,000 miles.

Figure 4-1. Distribution of Model Years



Model Year	FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
1996	91	91	15.99	15.99
1997	79	170	13.88	29.88
1998	78	248	13.71	43.59
1999	83	331	14.59	58.17
2000	73	404	12.83	71.00
2001	68	472	11.95	82.95
2002	58	530	10.19	93.15
2003	26	556	4.57	97.72
2004	13	569	2.28	100.00

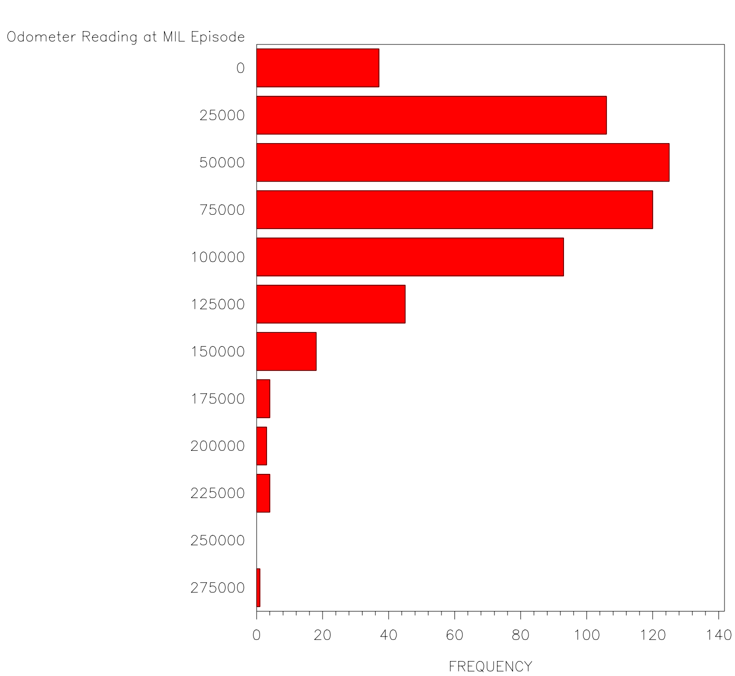
Figure 4-2. Distribution of Odometer Readings at the Time of the Interview



FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
8	8	1.45	1.45
69	77	12.52	13.97
115	192	20.87	34.85
110	302	19.96	54.81
118	420	21.42	76.23
58	478	10.53	86.75
41	519	7.44	94.19
19	538	3.45	97.64
5	543	0.91	98.55
3	546	0.54	99.09
1	547	0.18	99.27
1	548	0.18	99.46
3	551	0.54	100.00

Note: The bin labels are the midpoint values for the ranges. For example, the bin labeled 25,000 contains observations with values between 12,500 and 37,500.

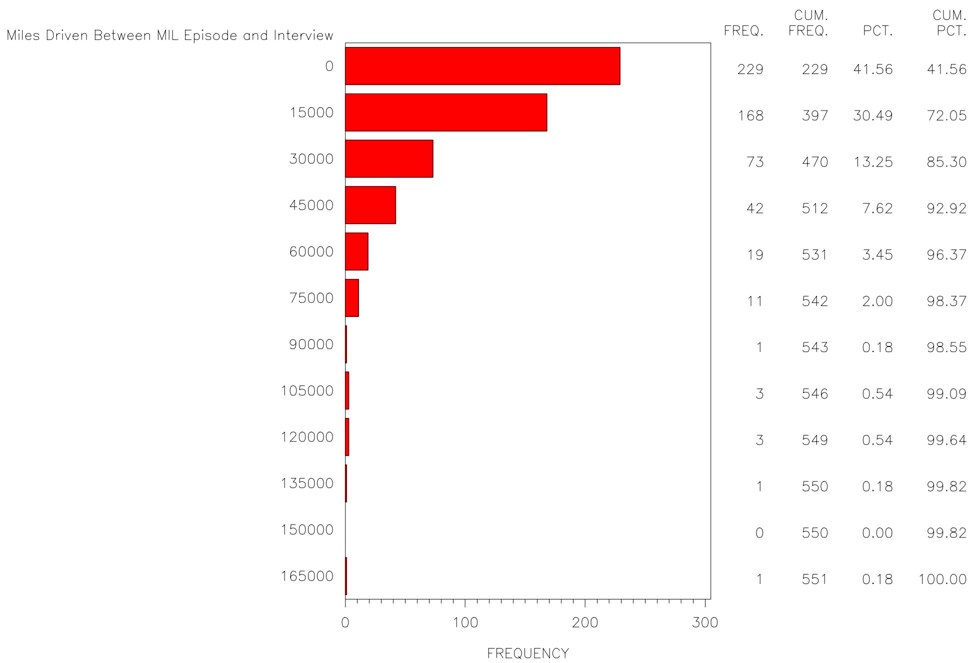
Figure 4-3. Distribution of Odometer Readings at the Time of the MIL Illumination



FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
37	37	6.65	6.65
106	143	19.06	25.72
125	268	22.48	48.20
120	388	21.58	69.78
93	481	16.73	86.51
45	526	8.09	94.60
18	544	3.24	97.84
4	548	0.72	98.56
3	551	0.54	99.10
4	555	0.72	99.82
0	555	0.00	99.82
1	556	0.18	100.00

Note: The bin labels are the midpoint values for the ranges. For example, the bin labeled 25,000 contains observations with values between 12,500 and 37,500.

Figure 4-4. Distribution of Miles Driven Between the MIL Illumination Episode and the Interview



Note: The bin labels are the midpoint values for the ranges. For example, the bin labeled 25,000 contains observations with values between 12,500 and 37,500.

Table 4-1 shows that for the 569 vehicles that reported an illuminated MIL, 85 percent of the respondents reported that they either had the vehicle repaired or repaired the vehicle themselves. We know from our examination of the dataset, that most respondents took Question MQ22 to ask if the vehicle had work performed on it. Respondents did not necessarily mean that the work had caused the MIL to be extinguished. If the sample of 569 vehicles is a random subset of the fleet, then the 95 percent confidence interval is 82 to 88 percent. Table 4-2 shows the same analysis when positive versus negative Response is considered. Out of the 569 vehicles, 517 or 91 percent of the vehicles had owners that claimed to have either repaired their vehicles or had taken action which we judged to be a positive Response to the illuminated MIL. If the 569 vehicles are representative of the fleet, then the 95 percent confidence interval for a positive Response is 88 to 94 percent.

Table 4-1. Claimed Repair Rates

Repair	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No	85	14.94%	85	14.94%
Yes	484	85.06%	569	100.00%

Table 4-2. Response Rates

Response	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Negative	52	9.14%	52	9.14%
Positive	517	90.86%	569	100.00%

One of the questions (MQ31) asked of all respondents who reported an illuminated MIL was whether the MIL still came on at the time of the interview. This question was asked regardless of whether the respondent reported a repair. Table 4-3 shows that of the 569 vehicles, 26 percent of the MILs still came on at the time of the telephone survey. In some cases, the respondents reported that their MILs had just come on. Accordingly, we present Table 4-4, which shows the counts of MILs that were illuminated at the time of the survey as a function of the miles driven since the MIL illumination episode. The mileage bins were created so that 50 observations were in the lowest mileage bin, and the higher mileage bins were divided at mileages to create approximately 100 observations in each bin. The table shows that for the 50 vehicles with less than 460 miles since the MIL illumination episode, 48 percent of the MILs were still illuminated. However, this percentage drops for the next four bins at higher mileages since the episode. The bin for 17,500 to 30,500 miles has the lowest percentage of MILs still illuminated with a value of 15 percent. The highest mileage bin with greater than 30,500 miles has a larger percentage of MILs still illuminated with a value of 42 percent. This increased value at high mileages may be an indication of MILs being re-illuminated during a new episode even though the interview questions were supposed to be directed toward the most recent MIL illumination episode.

Table 4-3. Overall MIL-Still-On Rate

STILL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
MIL Still Comes On (1)	148	26.01%	148	26.01%
MIL Off (2)	416	73.11%	564	99.12%
DK/RF (9)	5	0.88%	569	100.00%

[a] STILL is the variable name that corresponds to question MQ31 (Does the “check engine” light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-4. MIL-Still-On Rate vs. Miles Driven Since MIL Illumination

Frequency **Percent**	Mile_since_bin (miles)						
STILL	**0-460**	**460-4,100**	**4,100-8,700**	**8,700-17,500**	**17,500-30,500**	**>30,500**	**Total**
MIL Still Comes On (1)	24 **48.00**	29 **27.88**	17 **17.71**	18 **18.00**	16 **15.38**	41 **42.27**	145 26.32
MIL Off (2)	26 **52.00**	74 **71.15**	78 **81.25**	81 **81.00**	88 **84.62**	56 **57.73**	403 73.14
DK/RF (9)	0 **0.00**	1 **0.96**	1 **1.04**	1 **1.00**	0 **0.00**	0 **0.00**	3 0.54
Total	50 9.07	104 18.87	96 17.42	100 18.15	104 18.87	97 17.60	551 100.00

Frequency Missing = 18

[a] STILL is the variable name that corresponds to question MQ31 (Does the “check engine” light still come on when the vehicle is driven?).

[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.

[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

[d] The “frequency missing” corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

In the three subsections below, we characterize the answers to questions in the questionnaire for three groups of vehicles of the 569 that had reported MIL illuminations:

- **The repaired-vehicle subset** has 484 vehicles that claimed to have gotten a repair. These vehicles were assigned a value of Positive for Response.
- **The not-repaired positive-response subset** has 33 vehicles that did not claim to get a repair, but there was evidence that the MIL soon went out on its own so that there was no reason to seek a repair. All of these observations were assigned a value of Positive for Response.
- **The negative-response subset** has 52 vehicles that did not claim to get a repair and yet there was no evidence that the MIL went out. All of these observations were assigned a value of Negative for Response.

4.1 Repaired-Vehicle Subset (N =484)

As mentioned above, this dataset contains all of those observations where the respondent reported seeing a MIL illuminate and for which they claim to have gotten the vehicle repaired or repaired it themselves as indicated by their answer to Question MQ22E (Whtd5 = 1). We believe the respondents took the question to mean that work was done on the vehicle and not necessarily to mean that the work caused the MIL to be extinguished. This subsection reports a series of frequency counts of the responses in the questionnaire to observations in this data subset.

Table 4-5 shows the distribution of the responses to Question MQ24: "Which of the following best describes where you had the vehicle repaired?" Almost half of the vehicles that had MIL illuminations were repaired at dealerships.

Table 4-5. Vehicle Repair Locations for the Repaired-Vehicle Subset (N=484)

WHREP	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Self (1)	76	15.70%	76	15.70%
Friend/Household Member (2)	23	4.75%	99	20.45%
Independent Shop (3)	138	28.51%	237	48.97%
Dealership (4)	240	49.59%	477	98.55%
Other (7)	6	1.24%	483	99.79%
DK/RF (9)	1	0.21%	484	100.00%

[a] WHREP is the variable name that corresponds to question MQ24 (Which of the following best describes where you had the vehicle repaired).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-6 shows the responses for the data subset to Question MQ25: "Which of the following best describes the time period between when the 'check engine' light was first noticed and when the vehicle was eventually taken in for service and repaired?" Almost 80 percent of the respondents claimed they took the vehicles in for repair within a day or within a week.

Table 4-6. Time Between Illumination and Repair Work for the Repaired-Vehicle Subset (N=484)

TIMEP	Frequency	Percent	Cumulative Frequency	Cumulative Percent
< 1 day (1)	172	35.54%	172	35.54%
< 1 week (2)	210	43.39%	382	78.93%
< 2 weeks (3)	36	7.44%	418	86.36%
< 1 month (4)	37	7.64%	455	94.01%
< 3 months (5)	18	3.72%	473	97.73%
< 6 months (6)	3	0.62%	476	98.35%
< 1 year (7)	3	0.62%	479	98.97%
Other (97)	3	0.62%	482	99.59%
DK/RF (99)	2	0.41%	484	100.00%

[a] TIMEP is the variable name that corresponds to question MQ25 (Which of the following best describes the time period between when the "check engine" light was first noticed and when the vehicle was eventually taken for service and repair?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-7 shows a distribution of responses to the relative cost Question MQ29: "Which of the following best describes how much it cost to get it repaired?" The table indicates that 25 percent of the vehicles were repaired under warranty and that 62 percent of the vehicles were repaired inexpensively or moderately expensively.

Table 4-7. Relative Repair Cost for the Repaired-Vehicle Subset (N=484)

COST	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Not Expensive (1)	152	31.40%	152	31.40%
Moderately Expensive (2)	146	30.17%	298	61.57%
Very Expensive (3)	49	10.12%	347	71.69%
Under Warranty (4)	119	24.59%	466	96.28%
DK/RF (9)	18	3.72%	484	100.00%

[a] COST is the variable name that corresponds to question MQ29 (Which of the following best describes how much it cost to get it repaired?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-8 shows the distribution of responses to the specific dollar Question MQ30: "What was the cost to get this repair fixed?" Almost 50 percent of the responses was that it was less than $100.

Table 4-8. Dollar Repair Cost for the Repaired-Vehicle Subset (N=484)

TOTCT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
< $100 (1)	229	47.31%	229	47.31%
$100-200 (2)	83	17.15%	312	64.46%
$200-500 (3)	79	16.32%	391	80.79%
> $500 (4)	49	10.12%	440	90.91%
DK/RF (9)	44	9.09%	484	100.00%

[a] TOTCT is the variable name that corresponds to question MQ30 (Was the cost to get this repair fixed less than $100, between $100 to $200, between $200 to $500, or greater than $500?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-9 shows the distribution of responses to Question MQ31: "Does the 'check engine' light still come on when the vehicle is driven?" This table shows that even though repairs were claimed to have been made on these vehicles, about 22 percent of the vehicles in this repaired vehicle dataset reportedly have MILs that still come on. When we examine the dataset for the distribution of miles that were driven since the MIL illumination episode, we see in Table 4-10 that for MIL illumination episodes of less than 460 miles ago, about 41 percent of the MILs still come on. However, the MIL-still-on percentage drops to about 13 percent for MIL illumination episodes that occurred between 17,500 and 30,500 miles prior to the interview. This decrease in the fraction of MILs-still-on may be the result of the repairs that were made or may be the result of MILs that extinguished themselves. For the 81 vehicles where the MIL illumination episode was more than 30,500 miles before the interview, about 36 percent of the vehicles still had MILs that came on. This might be caused by a new episode that has occurred after the episode that was the subject of the interview.

Table 4-9. MIL-Still-On Rates for the Repaired-Vehicle Subset (N=484)

STILL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
MIL Still Comes On (1)	106	21.90%	106	21.90%
MIL Off (2)	375	77.48%	481	99.38%
DK/RF (9)	3	0.62%	484	100.00%

[a] STILL is the variable name that corresponds to question MQ31 (Does the “check engine” light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-10. MIL-Still-On Rates vs. Miles Driven Since MIL Illumination for the Repaired-Vehicle Subset (N=484)

Frequency **Percent**	**Mile_since_bin (miles)**						
STILL	**0-460**	**460-4,100**	**4,100-8,700**	**8,700-17,500**	**17,500-30,500**	**>30,500**	**Total**
MIL Still Comes On (1)	12 **41.38**	20 **24.69**	15 **17.24**	15 **15.79**	12 **12.63**	29 **35.80**	103 22.01
MIL Off (2)	17 **58.62**	60 **74.07**	71 **81.61**	79 **83.16**	83 **87.37**	52 **64.20**	362 77.35
DK/RF (9)	0 **0.00**	1 **1.23**	1 **1.15**	1 **1.05**	0 **0.00**	0 **0.00**	3 0.64
Total	29 6.20	81 17.31	87 18.59	95 20.30	95 20.30	81 17.31	468 100.00

Frequency Missing = 16
[a] STILL is the variable name that corresponds to question MQ31 (Does the “check engine” light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.
[d] The “frequency missing” corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

Question MQ26: “Did the ‘check engine’ light remain off after it was first repaired?” gets to the effectiveness of the repair process. Table 4-11 shows that about 80 percent of the 484 vehicles that were repaired had their MILs stay off. However, for 19 percent of the vehicles, the MILs did not remain off following the repair.

Table 4-11. MIL-Stay-Off Rates for the Vehicle-Repaired Subset (N=484)

REMAN	Frequency	Percent	Cumulative Frequency	Cumulative Percent
MIL Remained Off (1)	389	80.37%	389	80.37%
MIL Came Back On (2)	92	19.01%	481	99.38%
DK/RF (9)	3	0.62%	484	100.00%

[a] REMAN is the variable name that corresponds to question MQ26 (Did the "check engine" light remain off after it was first repaired?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

At this point, we divided the remaining analyses for the 484 vehicles into two further subsets: those 389 vehicles where the MIL remained off and those 95 vehicles where the MIL did not remain off or the respondent didn't know if the MIL remained off.

First let's consider the 95 vehicles where the MIL did not remain off or the respondent did not remember whether the MIL remained off or not. Question MQ27: "How many times did you have to return the vehicle for repairs for this episode?" was asked of these 95 respondents. The result is shown in Table 4-12. Some respondents apparently did not understand Question MQ26 because 17 of them said they did not have to return. If we eliminate those 17 observations from the vehicles that had to return for repairs, the median number of times for return was between two and three times. One person claimed to have returned 35 times for repair. This respondent was called back to confirm the 35 times answer and he confirmed it. However, we believe that this single response was just a reflection of the frustration probably encountered during the repair process. Table 4-13 shows the distribution of responses for the 95 vehicles to Question MQ28: "Was the vehicle finally repaired?" The table indicates that even after returns for additional repair work, only about half of the illuminated MILs were successfully extinguished. For this subset of 95 vehicles, Tables 4-14 and 4-15 show the responses for the relative and absolute cost for the repairs. Table 4-16 shows the response to Question MQ31: "Does the 'check engine' light still come on when the vehicle is driven?" The table indicates that 60 percent of the respondents said that the MIL light still came on at the time of the interview. Table 4-17 shows the distribution of responses to Question MQ31 as a function of the miles that the vehicle has been driven between the MIL illumination episode and the time of the interview. As we have seen for other data subsets, the fraction of MILs-still-on in the lowest and highest miles-since-MIL-on bins are higher than for the middle mileage bins.

Table 4-12. Number of Returns for Vehicles that Returned for Repairs (N=95)

RETUN	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Did not have to return (0)	17	17.89%	17	17.89%
Returned 1 time (1)	17	17.89%	34	35.79%
Returned 2 times (2)	19	20.00%	53	55.79%
Returned 3 times (3)	21	22.11%	74	77.89%
Returned 4 times (4)	8	8.42%	82	86.32%
Returned 5 times (5)	5	5.26%	87	91.58%
Returned 6 times (6)	2	2.11%	89	93.68%
Returned 10 times (10)	1	1.05%	90	94.74%
Returned 35 times (35)	1	1.05%	91	95.79%
DK/RF (99)	4	4.21%	95	100.00%

[a] RETUN is the variable name that corresponds to question MQ27 (How many times did you have to return the vehicle for repairs for this episode?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-13. Success Rates for Vehicles that Returned for Repairs (N=95)

FINAL	Frequency	Cumulative Percent	Cumulative Frequency	Percent
Vehicle was finally repaired (1)	45	47.37%	45	47.37%
Vehicle was not finally repaired (2)	49	51.58%	94	98.95%
DK (9)	1	1.05%	95	100.00%

[a] FINAL is the variable name that corresponds to question MQ28 (Was the vehicle finally repaired?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-14. Relative Repair Cost for Vehicles that Returned for Repairs (N=95)

COST	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Not Expensive (1)	19	20.00%	19	20.00%
Moderately Expensive (2)	26	27.37%	45	47.37%
Very Expensive (3)	16	16.84%	61	64.21%
Under Warranty (4)	27	28.42%	88	92.63%
DK/RF (9)	7	7.37%	95	100.00%

[a] COST is the variable name that corresponds to question MQ29 (Which of the following best describes how much it cost to get it repaired?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-15. Dollar Repair Cost for Vehicles that Returned for Repairs (N=95)

TOTCT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
< $100 (1)	42	44.21%	42	44.21%
$100-200 (2)	10	10.53%	52	54.74%
$200-500 (3)	16	16.84%	68	71.58%
> $500 (4)	18	18.95%	86	90.53%
DK/RF (9)	9	9.47%	95	100.00%

[a] TOTCT is the variable name that corresponds to question MQ30 (Was the cost to get this repair fixed less than $100, between $100 to $200, between $200 to $500, or greater than $500?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-16. MIL-Still-On Rates for Vehicles that Returned for Repairs (N=95)

STILL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
MIL Still On (1)	57	60.00%	57	60.00%
MIL Off (2)	37	38.95%	94	98.95%
DK/RF (9)	1	1.05%	95	100.00%

[a] STILL is the variable name that corresponds to question MQ31 (Does the “check engine” light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-17. MIL-Still-On Rates vs. Miles Driven Since MIL Illumination for Vehicles that Returned for Repairs (N=95)

Frequency **Percent**	**Mile_since_bin (miles)**						
STILL	**0-460**	**460-4,100**	**4,100-8,700**	**8,700-17,500**	**17,500-30,500**	**>30,500**	**Total**
MIL Still Comes On (1)	10 **83.33**	7 **41.18**	8 **57.14**	9 **64.29**	5 **35.71**	17 **77.27**	56 60.22
MIL Off (2)	2 **16.67**	10 **58.82**	5 **35.71**	5 **35.71**	9 **64.29**	5 **22.73**	36 38.71
DK/RF (9)	0 **0.00**	0 **0.00**	1 **7.14**	0 **0.00**	0 **0.00**	0 **0.00**	1 1.08
Total	12 12.90	17 18.28	14 15.05	14 15.05	14 15.05	22 23.66	93 100.00

Frequency Missing = 2
[a] STILL is the variable name that corresponds to question MQ31 (Does the “check engine” light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.
[d] The “frequency missing” corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

At this point, we went back and considered the same distributions for the 389 vehicles where a single repair was sufficient to extinguish the MIL. Tables 4-18 and 4-19 show the distribution of responses to the relative cost Question MQ29 and the absolute cost Question MQ30. Comparison of these two tables with Tables 4-14 and 4-15 for the set of vehicles where multiple returns for repair were required show that lower costs were experienced when only one repair was needed. Table 4-20 shows the distribution of results to Question MQ31: "Does the 'check engine' light still come on when the vehicle is driven?" For this set where vehicles required only a single repair to extinguish the MIL, only 13 percent of the 389 vehicles still had MILs on at the time of the interview. This compares with 60 percent shown in Table 4-16 for the dataset where multiple repairs were necessary. Table 4-21 shows the distribution of MILs on for single-repaired vehicles as a function of the miles driven between the MIL episode and the interview. The distribution shows the now familiar trend of higher percentages of vehicles with MILs-still-on at low miles and at high miles driven since the episode. The fraction of MILs-still-on in the 8,700 to 17,500 mile bin was the lowest at 7 percent. A comparison of the percent of MILs-still-on shown in Table 4-21 with those shown in Table 4-17 indicates that while the overall trend versus miles driven since the MIL episode is the same, the individual MIL-still-on rates are much lower for vehicles that required only a single repair.

Table 4-18. Relative Repair Cost for the Single-Repair Vehicles (N=389)

COST	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Not Expensive (1)	133	34.19	133	34.19
Moderately Expensive (2)	120	30.85	253	65.04
Very Expensive (3)	33	8.48	286	73.52
Under Warranty (4)	92	23.65	378	97.17
DK/RF (9)	11	2.83	389	100.00

[a] COST is the variable name that corresponds to question MQ29 (Which of the following best describes how much it cost to get it repaired?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-19. Dollar Repair Cost for the Single-Repair Vehicles (N=389)

TOTCT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
< $100 (1)	187	48.07	187	48.07
$100-200 (2)	73	18.77	260	66.84
$200-500 (3)	63	16.20	323	83.03
> $500 (4)	31	7.97	354	91.00
DK/RF (9)	35	9.00	389	100.00

[a] TOTCT is the variable name that corresponds to question MQ30 (Was the cost to get this repair fixed less than $100, between $100 to $200, between $200 to $500, or greater than $500?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-20. MIL-Still-On Rates for the Single-Repair Vehicles (N=389)

STILL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
MIL Still Comes On (1)	49	12.60	49	12.60
MIL Off (2)	338	86.89	387	99.49
DK/RF (9)	2	0.51	389	100.00

[a] STILL is the variable name that corresponds to question MQ31 (Does the "check engine" light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-21. MIL-Still-On Rates vs. Miles Driven Since MIL Illumination for Single-Repair Vehicles (N=389)

Frequency **Percent**	**Mile_since_bin (miles)**						
STILL	**0-460**	**460-4,100**	**4,100-8,700**	**8,700-17,500**	**17,500-30,500**	**>30,500**	**Total**
MIL Still Comes On (1)	2 **11.76**	13 **20.31**	7 **9.59**	6 **7.41**	7 **8.64**	12 **20.34**	47 12.53
MIL Off (2)	15 **88.24**	50 **78.13**	66 **90.41**	74 **91.36**	74 **91.36**	47 **79.66**	326 86.93
DK/RF (9)	0 **0.00**	1 **1.56**	0 **0.00**	1 **1.23**	0 **0.00**	0 **0.00**	2 0.53
Total	17 4.53	64 17.07	73 19.47	81 21.60	81 21.60	59 15.73	375 100.00

Frequency Missing = 14
[a] STILL is the variable name that corresponds to question MQ31 (Does the "check engine" light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.
[d] The "frequency missing" corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

4.2 Not-Repaired Positive-Response Subset (N = 33)

This subset is for those vehicles that did not claim repairs in Question MQ22 (Whtd5 = 2) but where the respondent indicated that the MIL went out by itself. Because the respondent did not claim that he had the vehicle repaired, questions about where the vehicle was repaired (MQ24), how soon the vehicle was repaired (MQ25), if the MIL stayed off after the repair (MQ26), how many times the vehicle had to be returned for repair (MQ27), if the vehicle was finally repaired (MQ28), and costs for the repair (MQ29 and MQ30) were not asked of the respondent. Table 4-22 shows that only 15 percent of the 33 vehicles still had illuminated MILs at the time of the interview. This compares with 22 percent in Table 4-9 for the vehicle subset (N=484) that received repairs. Table 4-23 shows the responses to Question MQ31: "Does the 'check engine' light still come on when the vehicle is driven?" versus the miles driven between the MIL illumination episode and the interview. For the bins in this table, the precision of the estimates is low because of the low number of counts. Nevertheless, the table still shows that in the lowest mileage bin, 37 percent of the MILs were still on while the middle and upper mileage bins show much lower MIL-still-on rates.

Table 4-22. MIL-Still-On Rates for the Not-Repaired Positive-Response Subset (N=33)

STILL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
MIL Still Comes On (1)	5	15.15	5	15.15
MIL Off (2)	27	81.82	32	96.97
DK/RF (9)	1	3.03	33	100.00

[a] STILL is the variable name that corresponds to question MQ31 (Does the "check engine" light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-23. MIL-Still-On Rates vs. Miles Driven Since MIL Illumination for the Not-Repaired Positive-Response Subset (N=33)

Frequency **Percent**	**Mile_since_bin (miles)**						
STILL	**0-460**	**460-4,100**	**4,100-8,700**	**8,700-17,500**	**17,500-30,500**	**>30,500**	**Total**
MIL Still Comes On (1)	3 **37.50**	0 **0.00**	1 **25.00**	0 **0.00**	1 **25.00**	0 **0.00**	5 15.63
MIL Off (2)	5 **62.50**	13 **100.00**	3 **75.00**	1 **100.00**	3 **75.00**	2 **100.00**	27 84.38
DK/RF (9)	0 **0.00**	0 **0.00**	0 **0.00**	0 **0.00**	0 **0.00**	0 **0.00**	0 0.00
Total	8 25.00	13 40.63	4 12.50	1 3.13	4 12.50	2 6.25	32 00.00

Frequency Missing = 1
[a] STILL is the variable name that corresponds to question MQ31 (Does the “check engine” light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.
[d] The “frequency missing” corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

Table 4-24 shows the responses to Question MQ23: “Why did you choose not to take the vehicle to be repaired/serviced?” including the open-ended responses for the 33 vehicles in the dataset. These responses demonstrate that the vehicle owners were acting properly when they saw the MIL illuminate. In most cases, the MIL either extinguished itself or the owner contacted a mechanic to determine what the problem was.

4.3 Negative-Response Dataset (N=52)

The Responses for 52 of the vehicles of the 569 that had illuminated MILs were assigned as Negative to indicate that the owners had not had the vehicle repaired and the MILs had not soon gone out by themselves. Just as for the dataset described in Section 4.2, because no repairs were made, the questions about repairs were not asked of these respondents. However, there is information about the MILs that were still illuminated and the reasons that vehicles were not repaired.

Table 4-25 shows that, at the time of the interview, 71 percent of the MILs were still coming on for these 52 vehicles. This compares with only 15 percent MILs-still-on for the non-repaired positive-response subset in Table 4-22 and the 22 percent MILs-still-on for the 484 repaired vehicles shown in Table 4-9. Table 4-26 shows the distribution of miles driven between the MIL illumination episode and the interview. We can combine mileage bins to increase the number of observations in bins and thus, increase the precision of the MIL-still-on rates. If we combine the second and third bins, the MIL-still-on rate for 460 to 8,700 miles would be 67 percent (=10/15). If we combine the fourth and fifth bins, the MIL-still-on rate for 8,700 to 30,500 miles would be 67 percent (=6/9). That modified table shows that for bins up to 30,500 miles, the MIL-still-on rates are in the 60 and 70 percent range.

Table 4-24. Reasons Given in MQ23 for Not Repairing Vehicles for the Not-Repaired Positive-Response Dataset (N=33)

Sample Number	Response	Explanation for the "Other" Response
1393	Just happened MIL went out	
1474	MIL went out	
1642	MIL went out	
1647	MIL went out	
1714	MIL went out	
1972	Too costly MIL went out	
1974	Other	The light was triggered by something that was self-correcting.
2491	Other	Called then took it in.
2703	Other	I thought it was time for service on the car to be worked on. And when I took it in that was what it was.
3524	MIL went out	
4265	MIL went out	
5494	MIL went out	
5756	MIL went out Other	What is said in the owner's manual.
6895	MIL went out	
7079	MIL went out	
8426	MIL went out Other	Mechanic's advice: absent of other indications; it is unnecessary to bring it in.
8957	Other	Didn't think it was a big deal and let the light go out by itself.
8969	MIL went out	
9195	Other	Like the respondent explained earlier; the problem was just due to the gas cap being loose; tech said to try screwing it.
9627	Too costly Car runs fine MIL went out	
9759	MIL went out	
9810	Other	Took it to the dealer.
10189	MIL went out Other	It was the gas cap and the light went out.
10483	MIL went out Other	I called the people who work on the car and they said not to worry unless it happened again.
11073	MIL went out	
12678	MIL went out Other	Called someone about the car and they told me it might be a bad sensor because everything seemed fine.
12860	MIL went out	
13667	MIL went out	

Sample Number	Response	Explanation for the "Other" Response
14262	MIL went out	
14323	MIL went out	
14590	Other	Because I did not think it was anything serious and it went off and did not come back on.
14749	Other	Because the dealership told me not to.
15190	MIL went out	

Table 4-25. MIL-Still-On Rates for the Negative-Response Subset (N=52)

STILL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
MIL Still Comes On (1)	37	71.15	37	71.15
MIL Off (2)	14	26.92	51	98.08
DK/RF (9)	1	1.92	52	100.00

[a] STILL is the variable name that corresponds to question MQ31 (Does the "check engine" light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.

Table 4-26. MIL-Still-On Rates vs. Miles Driven Since MIL Illumination for the Negative-Response Subset (N=52)

Frequency **Percent**	**Mile_since_bin (miles)**						
STILL	**0-460**	**460-4,100**	**4,100-8,700**	**8,700-17,500**	**17,500-30,500**	**>30,500**	**Total**
MIL Still Comes On (1)	9 **69.23**	9 **90.00**	1 **20.00**	3 **75.00**	3 **60.00**	12 **85.71**	37 72.55
MIL Off (2)	4 **30.77**	1 **10.00**	4 **80.00**	1 **25.00**	2 **40.00**	2 **14.29**	14 27.45
DK/RF (9	0 **0.00**	0 **0.00**	0 **0.00**	0 **0.00**	0 **0.00**	0 **0.00**	0 0.00
Total	13 25.49	10 19.61	5 9.80	4 7.84	5 9.80	14 27.45	51 100.00

Frequency Missing = 1
[a] STILL is the variable name that corresponds to question MQ31 (Does the "check engine" light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.
[d] The "frequency missing" corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

Table 4-27 shows the menu and open-ended responses to Question MQ23: "Why did you choose not to take the vehicle to be repaired/serviced?" An examination of the table indicates that typical reasons ranged from the car runs fine without the repair, to being too busy, to too expensive, to claiming that the MIL is just a moneymaker for dealerships. There are some instances where the respondent claimed that an appointment had been made. However, we placed these observations in this data subset because the car had not yet been repaired.

Table 4-27. Reasons Given in MQ23 for Not Repairing Vehicles for the Negative-Response Subset (N=52)

Sample Number	Response	Explanation for the "Other" Response
1103	Car runs fine Other	It is just an annoyance put on by the manufacturers. The oil is changed and nothing is wrong with it. Usually the light comes on due to a dirty connection. The service engine light is a joke.
1134	Other	It wasn't a big deal having the light on.
1270	Car runs fine Other	People don't know what's wrong with this vehicle.
1577	Other	Took it to get diagnosed; but never got it repaired; was getting other things on the vehicle repaired.
1609	Too busy Too costly	
1657	Other	It was unnecessary.
1660	Other	It was due for service.
2193	Other	She took the car in to be repaired.
2201	Too costly	
2799	Have appointment	
3160	Have appointment	
3182	Too busy	
3719	Car runs fine	
3821	Just happened	
3999	Other	Does not affect the mechanics of the car.
4080	Too costly	
4084	Other	It was just a sensor; that can be fairly expensive; I don't know.
4231	Other	Consider the light to be a fraud just a moneymaker for the dealerships.
4244	Other	The light was not a factory problem. It is a performance issue to enhance the power of the vehicle. Vehicle was not taken in because respondent put more horsepower in the car, which triggered oxygen gauges to send the signal to light up check.
4269	Have appointment	
4337	Car runs fine	
4482	Other	Took to Auto Zone but was not repaired.
4553	Car runs fine	
4699	Car runs fine	
5662	Too costly	
6004	Car runs fine	
6093	No warranty	
6252	Other	It was 30 below zero when the light can on; and I was convinced that it was a faulty reading because of the cold.
6677	Too costly	
6720	Too costly	
7453	Other	Not a lot of free time and the drive is far; and with the way he works (on call) he can't get an appointment because he might get paged.
7542	Too busy	Figured I could fix it myself; but after I got the parts I realized that I

Sample Number	Response	Explanation for the “Other” Response
	Other	couldn't fix the car myself so just kept putting it off.
8600	Have appointment	
9317	Have appointment	
9476	Other	I know what the problem is and it is nothing to worry about. It is the catalytic converter sensor.
9870	Car runs fine	
9923	Just happened	
10215	Other	Had hit one of those mileage things where the dealer wants you to bring it in for service.
10780	Other	Knew that it would go out on its own.
10845	Car runs fine	
11367	Too busy Other	Too dependent of the vehicle.
11713	Car runs fine	
12387	Too busy	
12435	Don’t know	
12443	Car runs fine Other	Think it comes on at prescribed intervals and does not mean anything.
12845	Other	Trusted the owner’s manual.
13365	Car runs fine Other	A friend explained to her that this was the time to get the vehicle maintenance and that the light won't go off until it’s done.
13977	Just happened	
14024	Car runs fine	
14425	Have appointment Too busy	
14662	Other	At 6000 miles I knew everything was OK and Chevy's are known for having their check engine lights come on. The mechanic said that it was running fine plus it cost $300 to get it fixed and I did not want to pay that.
14936	Car runs fine	

5.0 Statistical Analysis for Response Functions

The primary objective of this study is to determine the rate at which vehicle owners respond to MIL illuminations. In addition, we want to be able to determine the most important factors that influence the response rate and the functional relationship of those factors that the response rate follows.

In Section 5.1 we examine the statistical effects of many independent variables on two dependent variables of interest:

- **Repair –** this variable was provided directly by Question MQ22 Part E: "What did you finally do to address the light being on? Get the vehicle repaired/repaired the vehicle yourself." In the results presented in the previous section, we saw that respondents for 484 out of the 569 vehicles with MIL illuminations claimed to have gotten their vehicles repaired.
- **Response –** this is the same variable that we described in the previous section. It includes all 484 vehicles that got repairs plus the 33 vehicles that we judged as having a positive response even though the vehicles did not get repairs. For practical purposes, a positive Response was assigned for repaired vehicles and for vehicles where the MIL went out soon after it was seen to be on.

Both the Repair and Response variables have binary values. Repair is either yes or no, and Response is either positive or negative. Because of this, we used logistic regression to determine the statistical significance of independent variables on Repair and Response and to develop models. Appendix H provides an example to demonstrate the logistic regression method.

In Section 5.2, we proceed to the development of specific response functions that could be used in EPA MOBILE models to model the Response of non-I/M area drivers to illuminated MILs. In addition, we modeled the rate at which MILs are extinguished as a function of the miles driven after they are first illuminated. We believe that when the Response model and the MIL-Still-On models are used together, they will estimate the fraction of MILs-still-on after MILs are first illuminated.

5.1 Statistical Analysis of Repair and Response

In this section, we analyze the variables Repair and Response for influence by other variables in the dataset. The dataset of vehicles that reported MIL illuminations had 569 observations. Of these observations, 484 reported getting repairs and 517 were assigned a Response of positive. Accordingly, the Repair rate was 85 percent and the positive Response rate was 91 percent. Either of those values could be used in MOBILE6 to model the behavior of non-I/M drivers to MIL illuminations. While these percentages by themselves are new information, we further investigated to see if other factors influence the MIL response rate.[11]

The first step in the analysis was to create tables of values of Repair and Response against binned values of the independent variables to get a visual indication of the trends. These tables are shown in Appendix I. While the tables show the counts of Repair and Response in each bin, they do not provide a measure of statistical significance. For this, we performed logistic regressions using each of the variables shown in Table 5-1 as a single independent variable in each regression.

Table 5-1. Summary of Single Independent Logistic Regressions Against Repair and Response

Variable Description	Question	Variable Name	Variable Type	Pr > |t|	
				Repair	**Response**
Odometer at MIL-On	MQ18	Mile_epi_war	Class	0.3512	0.4520
Odometer at MIL-On	MQ18	Mile_epi_bin	Class	**0.0982**	**0.1346**
Odometer at MIL-On	MQ18	Mile_epi	Continuous	0.2033	**0.1304**
Vehicle Age (Model Year)	VQ8	Vehage	Continuous	0.7744	0.7473
Miles Owner Put On	MQ18-MQ16	Mile_owner_bin	Class	0.6784	0.3962
Miles Owner Put On	MQ18-MQ16	Mile_owner	Continuous	0.5448	0.5079
Private/Company Owned/Leased	MQ13	Aleas	Class	**0.0365**	**0.0122**
Acquired New or Used	MQ15	Purc	Class	0.7451	0.4992
Previous MIL Episodes	MQ19	Morec	Class	0.9036	0.4214
Knowledge of MIL Meaning	MQ20	Knwll	Class	**0.0050**	**0.1377**
Respondent's Age	MQ39	Age	Continuous	0.4627	0.2038
Gender	MQ40	Gend	Class	**0.0471**	**0.0259**
Education Level	MQ38	Educa	Class	**0.1113**	0.5496
Maintenance Familiarity	SQ5	Respp	Class	0.8778	0.9808
Warranty Status	MQ32	Warrn	Class	0.2270	**0.0816**

Table 5-1 shows a description of the variable, the question that is the source of the independent variable value, the name of the variable as it appears in Appendix I, the type of variable as it was used in the logistic regression, and the probabilities for regression against Repair and Response that the size of the effect could have occurred by chance alone.

To demonstrate the difference between the variable type of class and the variable type of continuous, consider the results in Table 5-1 for Mile_epi_bin, which was a class regression and

[11] Specifically, the request for proposals for the E-72 project asks to determine how "this response changes as mileage increases, ownership changes, and warranty or extended service contracts expire" and specifies that "the key product of the analysis of the survey data shall be response rates as a function of vehicle mileage and vehicle model year."

Mile_epi, which is a continuous regression. For both regressions, the raw data were the same; that is, the independent variable was the odometer reading at the time of the MIL illumination. For the class regression, we binned the odometer readings as shown in Appendix I Tables I-2a and I-2b by rounding them to the nearest 25,000 miles. In addition, any odometer reading for a MIL episode that was greater than 175,000 miles was placed in the 175,000 mile bin. The class regression was run on the response using the eight bin labels as classes. Accordingly, the logistic regression determines whether there is a significant difference among the response rates for the eight bins. The order of the response rates from the low mileage bin to the high mileage bin is not considered in the class regression. Consequently, a class regression reveals whether or not the average values of the separate bins are significantly different.

In the continuous regression, the independent variable was the actual odometer mileage at the time of the MIL illumination. A linear function of the odometer reading is used for the argument in the logistic regression (see Appendix H). So, in this case, the odometer readings are not binned and the order from low odometer readings to high odometer readings is considered. Consequently, a continuous regression reveals whether or not the coefficient of the independent variable (in this case, Mile_epi) is significantly different from zero.

Because we used only linear continuous functions in the logistic regressions, the trend of the dependent variable on the independent variable was monotonic. An examination of the tables in Appendix I for the appropriate binned independent variable helps give us a feel for whether the trends are monotonic or not.

The logistic regression results shown in Table 5-1 can be used to determine which independent variables, taken one at a time, have a significant influence on Repair and Response. Suppose we use the rather low confidence level of 80 percent; the $Pr > |t|$ results in the table that are significant at the 80 percent confidence level have the probability values in bold. They are the odometer reading at the time of the MIL episode (Mile_epi_bin and Mile_epi), whether the vehicle was privately owned or leased or company owned or leased (Aleas), if the respondent claimed to know what an illuminated MIL meant (Knwll), the gender of the respondent (Gend), the education level of the respondent (Educa), and the warranty status at the time of the MIL episode (Warrn). The variables that are notable in their lack of significant influence on Repair and Response are the mileage at the time of the MIL episode divided into bins related to warranty (Mile_epi_war), the vehicle age (vehage) which is directly related to model year, and whether the vehicle was originally obtained new or used (Purc). The trends for each of the variables in Table 5-1 are discussed separately below.

Odometer at MIL illumination episode – The current version of MOBILE6 uses repair rates for vehicles in areas without I/M programs of:

- 90 percent of vehicles with less than 36,000 miles;
- 10 percent of vehicles between 36,000 and 80,000 miles; and
- 0 percent of vehicles with more than 80,000 miles.

In the dataset, the variable Mile_epi_war bins the odometer readings at MIL illumination in the same manner. Table 5-2a and 5-2b show that the responses in the survey do not drop much at all as the mileage in these three bins increases. Table 5-2a shows that the claimed repair response rate drops from 88 percent to 82 percent and Table 5-2b shows that the positive response rate drops from 92 percent to 89 percent. The Response rates for the 36,000 to 80,000 mile bin and the greater than 80,000 mile bin are much larger than the rates (shown above) currently used in MOBILE6.2. Table 5-1 indicates that logistic regression found no significant difference among these three bins in influence on Repair and Response.

Table 5-2a. Claimed Repair Rates by Mile_epi_war (Three Warranty-Related Bins)

Frequency **Percent**	**Mile_epi_war**			
Repair	**0-36,000**	**36,000-80,000**	**> 80,000**	**Total**
No	17 **11.97**	33 **14.86**	34 **17.71**	84 **15.11**
Yes	125 **88.03**	189 **85.14**	158 **82.29**	472 **84.89**
Total	142 25.54	222 39.93	192 34.53	556 100.00

Frequency Missing = 13
[a] The "frequency missing" corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

Table 5-2b. Response Rates by Mile_epi_war (Three Warranty-Related Bins)

Frequency **Percent**	**Mile_epi_war**			
Response	**0-36,000**	**36,000-80,000**	**> 80,000**	**Total**
Negative	11 **7.75**	19 **8.56**	22 **11.46**	52 **9.35**
Positive	131 **92.25**	203 **91.44**	170 **88.54**	504 **90.65**
Total	142 25.54	222 39.93	192 34.53	556 100.00

Frequency Missing = 13
[a] The "frequency missing" corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

Tables 5-3a and 5-3b show that when the mileage at MIL illumination is divided into more bins, there is no obvious trend of the response rate with the odometer reading for both Repair and Response. The logistic regression result shown in Table 5-1 for Mile_epi_bin indicates significant differences for Repair and Response at the 90 percent and 87 percent confidence level, respectively. This indicates that there may be some significant difference at those confidence levels between the binned odometer readings when order of odometer readings is not considered.

Table 5-3a. Claimed Repair Rates by Mile_epi_bin (Eight Odometer Bins)

Frequency **Percent**	**Mile_epi_bin**								
Repair	**0-12,500**	**12,500-37,500**	**37,500-62,500**	**62,500-87,500**	**87,500-112,500**	**112,500-137,500**	**137,500-162,500**	**>162,500**	**Total**
No	5 **13.51**	12 **11.32**	14 **11.20**	21 **17.50**	22 **23.66**	4 **8.89**	5 **27.78**	1 **8.33**	84 **15.11**
Yes	32 **86.49**	94 **88.68**	111 **88.80**	99 **82.50**	71 **76.34**	41 **91.11**	13 **72.22**	11 **91.67**	472 **84.89**
Total	37 6.65	106 19.06	125 22.48	120 21.58	93 16.73	45 8.09	18 3.24	12 2.16	556 100.00

Frequency Missing = 13
[a] The "frequency missing" corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

Table 5-3b. Response Rates by Mile_epi_bin (Eight Odometer Bins)

Frequency **Percent**	**Mile_epi_bin**								
Repair	**0-12,500**	**12,500-37,500**	**37,500-62,500**	**62,500-87,500**	**87,500-112,500**	**112,500-137,500**	**137,500-162,500**	**>162,500**	**Total**
Negative	3 **8.11**	8 **7.55**	6 **4.80**	15 **12.50**	13 **13.98**	2 **4.44**	4 **22.22**	1 **8.33**	84 **15.11**
Positive	34 **91.89**	98 **92.45**	119 **95.20**	105 **87.50**	80 **86.02**	43 **95.56**	14 **77.78**	11 **91.67**	472 **84.89**
Total	37 6.65	106 19.06	125 22.48	120 21.58	93 16.73	45 8.09	18 3.24	12 2.16	556 100.00

Frequency Missing = 13
[a] The "frequency missing" corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

When the odometer readings at MIL illumination are considered as a continuous variable (Mile_epi), the logistic regressions found significant trends at only the 80 percent and 87 percent confidence level for repair and response, respectively. For both of these regressions, the logistic regression procedure found no lack of fit indicating that the linear relationship of Mile_epi describes the trend in the data with Mile_epi adequately. Both of these regression results indicate the tendency of the owner to respond to an illuminated MIL at a slightly lower rate at high odometer readings then at low odometer readings. Thus, the survey data indicate that the direction of the tendency in MOBILE6.2 is correct. However, the magnitude of the change with odometer readings is much smaller according to the survey.

Vehicle age (model year) – Tables I-1a and I-1b show the effect of vehicle age on Repair and Response. A vehicle age of 1 corresponds to a 2004 model year and a vehicle age of 9 corresponds to the 1996 model year. This can be seen by comparing the counts in the Appendix I tables with Figure 4-1. The tables do not show any definite trend in the rates of Repair or Response. In addition, the logistic regression results for vehicle age, which are shown in Table 5-1, indicate that when vehicle age is treated as a continuous variable, there is no significant trend of Repair or Response with vehicle age.

Miles driven by the current owner –An owner may respond to MIL illuminations at different rates according to the miles he has put on the vehicle regardless of whether he originally obtained the vehicle new or used. Accordingly, we created a variable, Mile_owner, which was calculated as a difference in odometer readings between when the current owner acquired the car and when the most recent MIL illumination occurred. The binned value of this variable was used to make Tables I-2a and I-2b for Repair and Response. An examination of these tables indicate no particular trend for these variables with the number of miles a current owner put on his vehicle. In addition, Table 5-1 shows no significant trend when Mile_owner was treated as a continuous variable or as a class variable in the logistic regressions.

Private/Company Owned/Leased Vehicle Status – The acquisition status of the vehicle potentially has an influence on the rates of Repair and Response. Tables I-3a and I-3b show these rates for the four different classifications of acquisition status. 95 percent of the vehicles were privately owned. Therefore, the rates for Repair and Response are much more precise for this level than for the other three levels. The rates of Repair and Response for the privately owned and privately leased vehicles are similar. However, the 15 vehicles that were company owned have apparent Repair rates and Response rates that are substantially lower than for the private vehicles. The logistic regression results shown in Table 5-1 for the acquisition variable indicate there is a significant difference among the four levels of the variable with 96 percent confidence for Repairs and 99 percent confidence for Response. We re-ran the logistic regression keeping only the two levels for privately owned and privately leased vehicles in the regression and found no significant difference between these two levels. Therefore, we conclude that the reason for the significant differences shown in Table 5-1 are a result of the substantially different rates for the company owned vehicles.

Vehicle acquisition new or used – Tables I-4a and I-4b show the influence of whether the vehicle was acquired new or used by the current owner on the rates of Repair and Response. The tables indicate only a small difference in the rates and the logistic regression results shown in Table 5-1 show no significant difference.

Previous MIL Illumination – It seemed reasonable that vehicle owners might respond to an illuminated MIL at higher rates if it was the first time they had seen an illuminated MIL on this vehicle. Therefore, Question MQ19 asked if the most recent episode was the only episode or if there were earlier MIL illumination episodes on this vehicle. Tables I-5a and I-5b show there were only very slight differences in Repair and Response for this variable. In addition, Table 5-1 shows that the logistic regression found no significant difference.

Prior knowledge of the meaning of an illuminated MIL – Another factor that might influence the response rate was if the vehicle owner had a prior knowledge of the meaning of an illuminated MIL. Question MQ20 asked: "When the light first came on, did you already know what it meant or represented?" Tables I-6a and I-6b indicate that if the respondent knew what the MIL represented when the MIL came on, he was less likely to repair the vehicle or respond to the illuminated MIL positively. Table 5-1 shows that the confidence levels for the difference between the effects of knowing and not knowing the meaning of the MIL were significant at the 99 percent and 86 percent confidence levels for Repair and Response, respectively. The tables also indicate that approximately half of the respondents claimed to know in advance what the MIL meant and the other half of the respondents did not know in advance what it meant.

Respondent's age – Tables I-7a and I-7b show the trend of age on the Repair and Response rates. No obvious trend is apparent. Table 5-1 shows that a continuous logistic regression of these dependent variables on the age of the owner were not significant.

Respondent's gender – Since we had the gender of the respondent in the database, we decided to see if there was a difference in the rates by gender for Repair and Response. Tables I-8a and I-8b indicate that men were less likely to claim a repair or produce a positive response to an illuminated MIL than women were. The respondents were approximately half men and half women. The logistic regression results shown in Table 5-1 indicate that gender made a significant difference with 95 percent confidence for Repair and 97 percent confidence for Response.

Education level – Since education level was one of the demographic variables obtained in the survey, we examined the data to see if there was a trend in rates of Repair and Response by education level. Tables I-9a and I-9b show the results. There seems to be a slight tendency for the lowest and highest education levels to have slightly lower rates of Repair and positive Responses to MIL illuminations. Table 5-1 indicates that differences among education levels for Repairs was significant at the 89 percent confidence level but that for Responses there was no significant difference among the levels.

Familiarity with vehicle maintenance – As part of the screening process, Question SQ5 asked potential respondents how familiar they were with the maintenance of their 1996 and newer vehicles. They were to rank their familiarity with maintenance on a scale of 1 to 5 with 1 being no familiarity and 5 being extremely familiar. Only respondents that rated themselves a 3, 4, or a 5 were accepted for the survey unless they were the only person living in the household. Tables I-10a and I-10b show the influence of maintenance familiarity on the rates of Repair and Response. The tables indicate a variable trend and the logistic regression results in Table 5-1 show no significant difference among the five different levels.

Warranty status – Question MQ32 asked: "Was the vehicle covered under warranty for repairs because of the 'check engine' light going on?" This question was asked because it is suspected that vehicles that are covered by a warranty or service contract may be more likely to be repaired to fix a MIL illumination. The results of the analysis tend to support this trend. Tables 5-4a and 5-4b show the rates for Repair and Response. In both cases, vehicles that are under warranty have about a 4 percent higher action rate than those vehicles that are not under warranty. In the dataset, approximately one-third of the vehicles were under warranty or service contract and two-thirds were not. The results of the logistic regression shown in Table 5-1 show an insignificant difference for Repair but a significant difference for Response at the 92 percent confidence level.

Table 5-4a. Claimed Repair Rates by Warranty Status

Frequency **Percent**	**Warranty Status (WARRN)**			
Repair	**Covered**	**Not Covered**	**Don't Know**	**Total**
No	26 **12.56**	58 **16.34**	1 **14.29**	85 **14.94**
Yes	181 **87.44**	297 **83.66**	6 **85.71**	484 **85.06**
Total	207 36.38	355 62.39	7 1.23	569 100.00

Table 5-4b. Response Rates by Warranty Status

Frequency **Percent**	Warranty Status (WARRN)			
Repair	**Covered**	**Not Covered**	**Don't Know**	**Total**
Negative	13 **6.28**	38 **10.70**	1 **14.29**	85 **14.94**
Positive	194 **93.72**	317 **89.30**	6 **85.71**	484 **85.06**
Total	207 36.38	355 62.39	7 1.23	569 100.00

Following the regression investigations on variables one at a time, which were discussed above, we proceeded to investigate models using multiple independent variables and two factor interactions in different logistic regressions. We were not able to find any models that were substantially better than the models developed using a single independent variable.

5.2 Development of Response Functions

The examination of the survey data for Repair and Response performed in the previous subsection separated single independent variables that had significant influences on the rates of Repair and Response from other independent variables that did not demonstrate any significant influence. In this section, we revisit the variables that demonstrated significant levels of influence on the Repair and Response variables with the purposes of developing candidate response functions for use in MOBILE models.

At this point, odometer reading at the time of the MIL illumination episode (Mile_epi) still needs to be considered even though the significance levels shown in Table 5-1 for this variable are quite low. Vehicle model year will no longer be considered. The acquisition status (Aleas) for private/company owned/leased status showed a significant level of influence on Repair and Response. However, the significance was provided almost entirely by the 15 (out of 569) vehicles in the dataset that were company owned. Accordingly, for the purposes of this report, we eliminated this variable from further consideration. Whether a vehicle is acquired new or used (Purc) was eliminated from further consideration as was the variable (Morec) that revealed whether the MIL illumination episode that was investigated was the first for this vehicle or not. While advanced knowledge of the meaning of a MIL (Knwll) and the gender (Gend) of the respondent had significant influences on the rate of Repairs and Response, we do not anticipate that MOBILE models would use these independent variables. The last variable for the existence of warranty (Warrn) demonstrated a significant effect for response in the trends for both Repair and Response and was in the direction that was expected – the vehicles with warranties were more likely to have illuminated MILs addressed. Thus, we believe that the response functions could be reasonably modeled by considering only mileage at the time of MIL illumination episode and the warranty status of the vehicle.

Another important observation of the data in the survey is that even though respondents reported that repairs were made to their vehicles, repairs were not always successful at extinguishing the MILs. Therefore, the Repair rates or Response rates alone should not be used in determining the effect of the repairs on emissions reductions. The effectiveness of the repairs must also be considered. Revisiting the MIL-still-on rates for the three subsets of the data discussed in Section 4, we produced Table 5-5. The three datasets were made up of the 484 vehicles that got repaired and were judged to have a positive Response, the set of 33 vehicles that did not receive repairs but were judged to have a positive Response, and the 52 vehicles that did not receive a repair and were judged to have a negative Response. Table 5-5 shows the average MIL-still-on rates, which were taken from Tables 4-9, 4-22, and 4-25. If we calculate the 90 percent confidence intervals for each of these MIL-still-on rates, which are shown in Table 5-5, we see that the MIL-still-on rates for the two datasets with positive Response are not significantly different. On the other hand, the MIL-still-on rate for the negative Response vehicles is substantially and statistically significantly higher than for the vehicles with a positive Response. We, therefore, suggest that this large difference in MIL-still-on rate be used as a criterion for selecting Response as the key variable to be modeled. This causes the first two subsets in Table 5-5 to be combined since both are Positive for Response. This grouping of the first two subsets differentiates them from the third subset, which has a significantly higher MIL-still-on rate.

Table 5-5. Comparison of Rates of MILs Still Illuminated at Interview

Dataset	N	MIL-Still-On	90% Confidence Interval
Repair = Yes Response = Positive	484	22%	19% to 25%
Repair = No Response = Positive	33	15%	5% to 25%
Repair = No Response = Negative	52	71%	61% to 81%

Now that we can focus only on Response as the variable of interest, we can determine how Response varies with warranty status and/or odometer reading at the time of MIL illumination. Tables 5-6a and 5-6b show Response as a function of warranty status and odometer reading at MIL illumination. Table 5-6a shows that the positive Response rate when warranties are in force is about 94 percent and is quite steady through the odometer reading bins.

Table 5-6b shows that when no warranty is in force, the response rate is relatively steady at about 89 percent through the odometer reading bins. Comparison of the total counts for corresponding bins in Tables 5-6a and 5-6b shows that, as expected, more counts are present at low odometer bins when warranties are in force and more counts at higher odometer bins when warranties are not in force. This is simply a consequence of warranties expiring at certain odometer readings.

We also performed logistic regressions for Response against the continuous variable for odometer reading at MIL illumination (Mile_epi) for the data subset with warranty and separately for the data subset with no warranty. These regressions indicated that for these two warranty subsets, there was no significant trend of odometer reading at MIL illumination on Response. In other words, if the response rate is modeled as 94 percent for vehicles with warranty and 89 percent for vehicles without warranty, the addition of odometer reading at MIL illumination adds nothing to the predicting ability of the model.

Thus, we conclude that the best model assigns a Response of 94 percent for vehicles under warranty and a Response of 89 percent for vehicles that are not under warranty. Additionally, our analysis of the survey data has found that model year, odometer reading, and new or used vehicle acquisition do not provide any additional information to predict owner response to an illuminated MIL.

Table 5-6a. Response vs. MIL Illumination Odometer Reading for Vehicles With Warranty

Frequency **Percent**	**Mile_epi_bin (miles)**								
Response	**0-12,500**	**12,500-37,500**	**37,500-62,500**	**62,500-87,500**	**87,500-112,500**	**112,500-137,500**	**137,500-162,500**	**>162500**	**Total**
Negative	2 **6.45**	4 **5.48**	2 **4.08**	1 **4.00**	3 **15.79**	0 **0.00**	1 **25.00**	0 .	13 6.40
Positive	29 **93.55**	69 **94.52**	47 **95.92**	24 **96.00**	16 **84.21**	2 **100.00**	3 **75.00**	0 .	190 93.60
Total	31 15.27	73 35.96	49 24.14	25 12.32	19 9.36	2 0.99	4 1.97	0 0.00	203 100.00

Frequency Missing = 4

[a] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.

[b] DK/RF means that the respondent did not know the answer or refused to answer the question.

[c] The "frequency missing" corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

Table 5-6b. Response vs. MIL Illumination Odometer Reading for Vehicles Without Warranty

Frequency **Percent**	Mile_epi_bin (miles)								
Response	**0-12,500**	**12,500-37,500**	**37,500-62,500**	**62,500-87,500**	**87,500-112,500**	**112,500-137,500**	**137,500-162,500**	**>162500**	**Total**
Negative	1 **16.67**	4 **13.33**	4 **5.33**	13 **14.13**	10 **13.51**	2 **4.65**	3 **21.43**	1 **8.33**	38 10.98
Positive	5 **83.33**	26 **86.67**	71 **94.67**	79 **85.87**	64 **86.49**	41 **95.35**	11 **78.57**	11 **91.67**	308 89.02
Total	6 1.73	30 8.67	75 21.68	92 26.59	74 21.39	43 12.43	14 4.05	12 3.47	346 100.00

Frequency Missing = 9

[a] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.

[b] DK/RF means that the respondent did not know the answer or refused to answer the question.

[c] The "frequency missing" corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

The last important feature to model is the rate that MILs are still on after they have been discovered by the driver. We have found that these MIL-still-on rates can be modeled as a function of the miles driven since the MIL was illuminated. To develop models for MIL-still-on, we break the dataset into those vehicles where Response was positive and those where Response was negative.

In the case of positive Responses, we have seen by the statistical comparison of the average MIL-still-on rate in Table 5-5, that the 484 vehicle dataset of repairs and positive responses and the 33 vehicle dataset of the no repairs and positive Responses had no significant difference in the overall MIL-still-on rates. In addition, if we look at the dependences of the MIL-still-on rates for these two datasets in Table 4-10 and Table 4-23 as a function of miles since the MIL was illuminated, we see a similar downward trend as miles are driven after the MIL was first illuminated. The exception to this trend is for observations with miles driven greater than 30,500 miles following the MIL illumination. Table 4-10 shows an increase in the MIL-still-on rate. We suggest that this increased rate in vehicles driven more than 30,500 miles after the MIL was first illuminated could actually be caused by a second, different, and more recent MIL illumination episode. We use this suggestion to support dropping the observations that are more than 30,500 miles after the MIL illumination. This then provides a dataset (N=414) made up of all observations that had a positive Response and that had driven less than 30,500 miles since the MIL was illuminated. This will include observations from the set that was repaired and had a positive Response (N=484) and the set that was not repaired had a positive Response (N=33).

Table 5-7 shows the distribution of MILs-still-on and MILs off as a function of the miles driven since the MIL was first illuminated for the resulting dataset. The table shows that the rate of MILs-still-on is monotonically decreasing as the miles driven since MIL illumination increase. In addition, the drop in MIL-still-on rates is larger at the lower mileage bins. We believe this decrease in MIL-still-on may be a consequence of the repair efforts of the vehicle owners since they have demonstrated a positive attitude toward the illuminated MIL.

Table 5-7. MIL-Still-On Rates vs. Miles Driven Since MIL Illumination for Vehicles with Positive Response (N=414)

Frequency **Percent**	**Mile_since_bin (miles)**					
STILL	**0-460**	**460-4,100**	**4,100-8,700**	**8,700-17,500**	**17,500-30,500**	**Total**
MIL Still Comes On (1)	15 **40.54**	20 **21.51**	16 **17.78**	15 **15.79**	13 **13.13**	79 19.08
MIL Off (2)	22 **59.46**	73 **78.49**	74 **82.22**	80 **84.21**	86 **86.87**	335 80.92
Total	37 7.44	93 18.71	90 18.11	95 19.11	99 19.92	414 100.00

Frequency Missing = 16
[a] STILL is the variable name that corresponds to question MQ31 (Does the “check engine” light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.
[d] The “frequency missing” corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

We performed a logistic regression on the data that was used to create Table 5-7. In the regression, we modeled the 414 observations of MIL-still-on as a function of the continuous variable representing miles since the MIL was illuminated (Mile_since). When we used Mile_since linearly we found that while it was statistically significant, the model had a significant lack of fit. In the final model, we used the natural log of the miles since MIL illumination. In this model, there was no lack of fit and the coefficient of natural log of Mile_since was significantly different from zero with 95 percent confidence. For vehicles with positive response, the probability that the MIL is still on is given by:

$$P\ (\text{MIL-still-on}) = [\exp(\text{arg})] / [1 + \exp(\text{arg})]$$

arg (combined Warranty and No Warranty) $= 0.4334 - 0.2268 * \ln(M)$

where M = number of miles driven since MIL illumination

In addition to the fact that this function fits the survey data by describing the probability that the MIL will still be on after discovery by a person with a positive Response, the function also has well-behaved properties at the extreme ends of its range. At 0 miles since MIL illumination, the predicted value of the probability is 1. In other words, the function predicts that when the driver sees the MIL come on, the probability of the MIL being on is 100 percent. At the other extreme, the function predicts that after an infinite number of miles since the MIL illumination occurs, the probability of the MIL still being illuminated is 0. In other words, even the most procrastinating, but positively inclined owner will eventually get the MIL extinguished.

It also makes sense that the rate at which illuminated MILs are still on after they have been seen is related not only to the number of miles (or time) since the MIL was first seen but whether the vehicle is under warranty or service contract. Tables 5-8 and 5-9 show the distribution of MIL-still-on rates as a function of the miles driven since the MILs were first seen for the vehicles under warranty and not under warranty. The tables show that the rates have a similar behavior as a function of miles. We built separate logistic regression models using these data for the cases of warranty and no warranty. Both models showed a downward trend in the probability of MIL-still-on as miles since lit MIL increased. Both models had no significant lack of fit. However, the Mile_since coefficient for the warranty model was not significant and the coefficient of Mile_since for the no warranty model was significant at 90 percent confidence. The models for positive Responses are given by:

$$P\,(\text{MIL-still-on}) = \exp(\text{arg})/(1 + \exp(\text{arg}))$$

$$\text{arg (Warranty)} \quad = \quad 0.1472 - 0.2181 * \ln(M)$$

$$\text{arg (No Warranty)} \quad = \quad 0.5437 - 0.2258 * \ln(M)$$

where M = Number of miles driven since MIL illumination

Table 5-8. MIL-Still-On Rates vs. Miles Driven Since MIL Illumination for Vehicles with Positive Response and Under Warranty

Frequency **Percent**	**Controlling for WARRN=1**					
	Mile_since_bin (miles)					
STILL	**0-460**	**460-4,100**	**4,100-8,700**	**8,700-17,500**	**17,500-30,500**	**Total**
MIL Still Comes On (1)	6 **54.55**	7 **21.88**	5 **14.71**	5 **11.11**	5 **11.90**	28 17.07
MIL Off (2)	5 **45.45**	25 **78.13**	29 **85.29**	40 **88.89**	37 **88.10**	136 82.93
Total	11 5.82	32 16.93	34 17.99	45 23.81	42 22.22	164 100.00

[a] STILL is the variable name that corresponds to question MQ31 (Does the "check engine" light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.

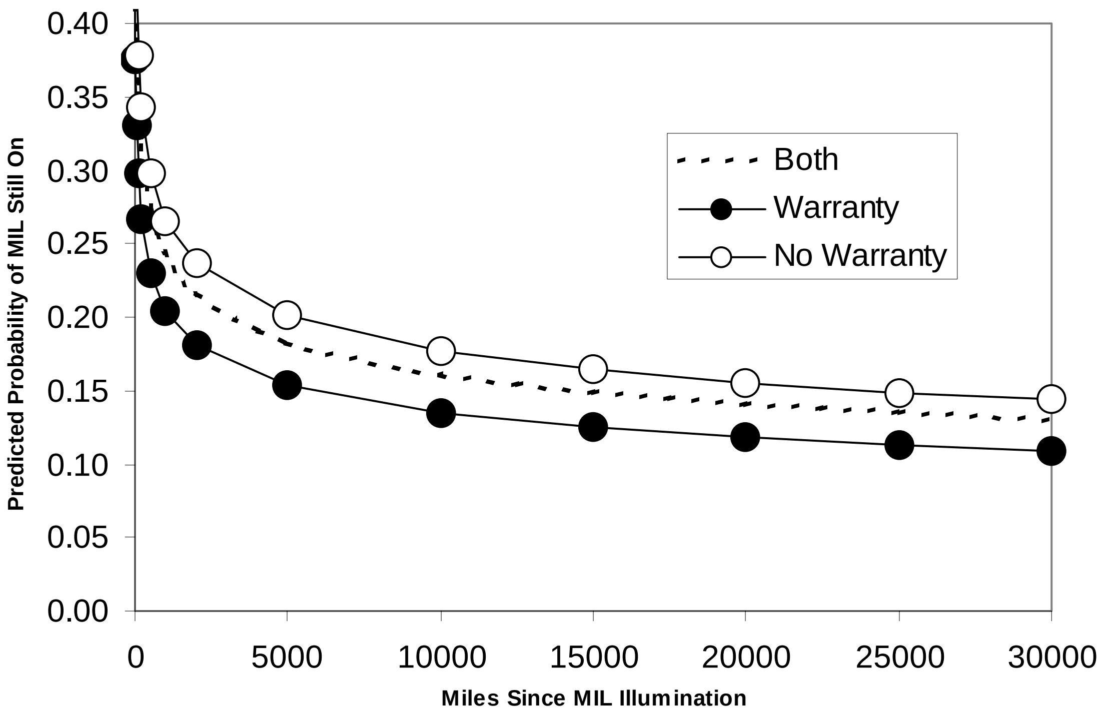
Table 5-9. MIL-Still-On Rates vs. Miles Driven Since MIL Illumination for Vehicles with Positive Response and Not Under Warranty

Frequency **Percent**	**Controlling for WARRN=2**					
	Mile_since_bin (miles)					
STILL	**0-460**	**460-4,100**	**4,100-8,700**	**8,700-17,500**	**17,500-30,500**	**Total**
MIL Still Comes On (1)	9 **34.62**	13 **21.31**	11 **19.64**	10 **20.83**	7 **12.96**	50 20.41
MIL Off (2)	17 **65.38**	48 **78.69**	45 **80.36**	38 **79.17**	47 **87.04**	195 79.59
Total	26 8.61	61 20.20	56 18.54	48 15.89	54 17.88	245 100.00

[a] STILL is the variable name that corresponds to question MQ31 (Does the "check engine" light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.

Figure 5-1 shows a plot of the predicted probabilities from all three positive-Response models as a function of miles since MIL illumination. The dashed line shows the model that includes both warranty and no warranty data. Comparison of the dashed curve with MIL-still-on values in Table 5-7 show good agreement. The plot shows that at a given mileage since MIL illumination, vehicles under warranty (filled dots) have a lower probability of having MILs still on than vehicles that are not under warranty (open circles). In addition, the curves also show that for all miles in the plot, the difference between the probabilities of MILs-still-on for warranty and no warranty vehicles is relatively small and is on the order of a 5 percent difference. The similar shape of these three curves and the fact that the no-warranty curve is above the warranty curve are attractive to us. We, therefore, recommend using the separate equations for warranty and no-warranty MIL-still-on rates for positive Responses even though the Miles_since coefficient for the warranty model was not statistically significant.

Figure 5-1. Modeled Trends of MILs-Still-On vs. Miles Driven Since MIL Illumination for Vehicles with Positive Responses



We also looked at the data to determine if the presence of a warranty had an effect on the rate of MIL-still-on for the vehicles where Response was negative. In the case where warranties were in force and the number of miles driven since the MIL illumination was less than 30,500 miles, 6 out of 10 or 60 percent of the vehicles still had MILs on. For the case where warranties were not in force and the vehicles had been driven less than 30,500 miles since the MIL was illuminated, 18 out of 26 vehicles or 69 percent still had MILs on. With this small number of observations, the difference between 60 percent and 69 percent is not statistically significant and it is not practically important. Therefore, we combined the data for negative Responses for observations with warranties and without warranties and examined the trend of the MIL-still-on rate versus miles driven since illumination. This is actually the distribution we have already seen in Table 4-26. Table 5-10 combines some of the mileage bins of Table 4-26 to increase the number of observations in the combined bins and hence reduce the uncertainty in the MIL-still-on rates. While the most probable estimate of the 0 to 460 mile bin is 69 percent, given the low number of counts (13) in the bin, we are 90 percent confident that the population value is between 43 and 89 percent. While the most probable estimate of the 460 to 8,700 mile bin is 67 percent, given the low number of counts (15) in the bin, we are 90 percent confident that the population value is between 43 and 84 percent. Finally, while the most probable estimate of the 8,700 to 30,500 mile bin is 67 percent, given the low number of counts (9) in the bin, we are 90 percent confident that the population value is between 33 and 92 percent.

Table 5-10. MIL-Still-On Rates vs. Miles Driven Since MIL Illumination for the Negative-Response Subset (N=52)

Frequency **Percent**	**Mile_since_bin (miles)**				
STILL	**0-460**	**460-8,700**	**8,700-30,500**	**>30,500**	**Total**
MIL Still Comes On (1)	9 **69.23**	10 **66.67**	6 **66.67**	12 **85.71**	37 72.55
MIL Off (2)	4 **30.77**	5 **33.33**	3 **33.33**	2 **14.29**	14 27.45
DK/RF (9	0 **0.00**	0 **0.00**	0 **0.00**	0 **0.00**	0 0.00
Total	13 25.49	15 29.41	9 17.64	14 27.45	51 100.00

Frequency Missing = 1
[a] STILL is the variable name that corresponds to question MQ31 (Does the "check engine" light still come on when the vehicle is driven?).
[b] The numbers in parentheses correspond to the programming codes for the responses to the questions. These codes are also defined in Appendix E.
[c] DK/RF means that the respondent did not know the answer or refused to answer the question.
[d] The "frequency missing" corresponds to the number of observations for which the value of the mileage is missing. Accordingly, these observations could not be binned.

The logistic regression for this negative Response data indicated that the coefficient for Mile_since was significant at the 88 percent confidence level and there was no lack of fit. The model for negative Responses for MIL-still-on regardless of warranty status can be described by:

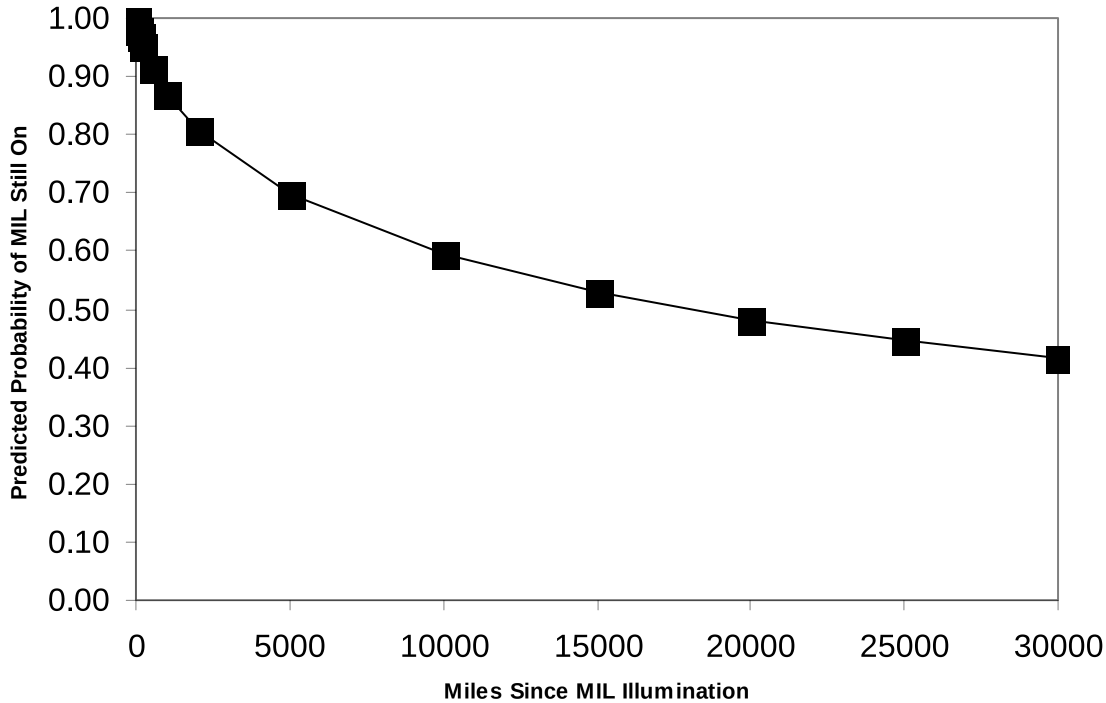
$$P\ (\text{MIL-still-on}) = \exp(\text{arg})/(1 + \exp(\text{arg}))$$

$$\text{arg} = 6.4219 - 0.6554 * \ln(M)$$

where M = Number of miles driven since MIL illumination

The plot in Figure 5-2 shows the shape of this function. However, because of the small number of observations, the location of the curve in Figure 5-2 and the coefficients in the above equation are very uncertain. This can be seen by comparing the curve in Figure 5-2 with the 90% confidence intervals given above.

Figure 5-2. Modeled Trend of MILs-Still-On vs. Miles Driven Since MIL Illumination for Vehicles with Negative Responses



Overall, we believe that the predicted probability curves shown in Figures 5-1 and 5-2 and their models are a reasonable representation of the rates of MILs-still-on after drivers see the MILs illuminate. For the case where drivers have a positive Response to a MIL illumination, we believe that the rapid drop in the predicted probabilities of MILs-still-on as shown in Figure 5-1 indicate the rapidity that drivers with positive attitudes respond to illuminated MILs. For the case where drivers have a negative Response to a MIL illumination, we believe that the slow drop in MILs-still-on as shown in Figure 5-2 indicate the slow effects of other vehicle repairs that may ultimately cause MILs to go out.

The overall results of this modeling exercise are shown in Table 5-11. If the vehicle is under warranty, 94 percent of drivers will make an effort to contact a service professional to get the vehicle fixed. For these vehicles, MILs will stay on according to the equation in the last column, which is described graphically by the curve with the solid dots in Figure 5-1. This shows that after driving about 5,000 miles only about 15 percent of the MILs will still be on. For vehicles under warranty, about 6 percent of the drivers will have a negative response and do little to address the illuminated MIL. The table shows the equation that describes the fraction of MILs still on after they were seen by the driver. The equation is described by the curve in Figure 5-2.

At 5,000 miles, 70 percent of the MILs of these vehicles will still be on. In the case of vehicles that are not under warranty, 89 percent of the vehicles will have a positive response. Note that this level of response is statistically significantly lower than the positive response of 94 percent for vehicles under warranty, while it is not greatly lower. As owners have these vehicles repaired, the probability of MILs staying on follows the curve with the open circles in Figure 5-1 and is given by the equation in the table. Finally, for vehicles that are not under warranty, 11 percent of the drivers will have a negative response and will do little to get the vehicle repaired. We have modeled the rate of MILs-still-on following MIL illumination for these vehicles to be the same as for negative responders of vehicles under warranty. The curve is shown in Figure 5-2 and is described by the equation in the table.

Table 5-11. Owner/Driver Response Rates to MIL Illumination and MIL-Still-On Rates After MIL Illumination

	Type of Response to Illuminated MIL	Rate of Response (Percent)	Probability of MIL-Still-On After Vehicle Driven M Miles Since Lit MIL Seen by Driver
Warranty	Positive	94%	P = exp(arg)/(1+exp(arg))
			Arg = 0.1472 – 0.2181 * ln (M)
	Negative	6%	P = exp(arg)/(1+exp(arg))
			Arg = 6.4219 – 0.6554 * ln (M)
No Warranty	Positive	89%	P = exp(arg)/(1+exp(arg))
			Arg = 0.5437 – 0.2258 * ln (M)
	Negative	11%	P = exp(arg)/(1+exp(arg))
			Arg = 6.4219 – 0.6554 * ln (M)

It is important to note that the data obtained for this analysis were based on a telephone survey. Therefore, it relies on the recollection and honesty of the respondents for the accuracy of the information.

Glossary

AGE. The variable name that corresponds to question MQ39 (How old were you on your last birthday?).

ALEAS. The variable name that corresponds to question MQ13 (Is this vehicle owned or leased by you or anyone in your household, or is it company owned or company leased?).

Cognitive interview. A cognitive interview is a preliminary test of the draft survey questionnaire with persons that possess the similar characteristics of the survey's intended audience and involves in-person interviewing. The testing objectives are related to the question-answering process of potentially complex questions in that they assess the respondents' ability to generate a response by examining their comprehension of questions and their ability to retrieve relevant information from memory. Cognitive interviews are also used to assess the adequacy of the questionnaire flow (structure and design) (see Appendix A for the complete Cognitive Interview Report for the study).

Computer-aided telephone interviewing (CATI). Interviewer-administered telephone surveying using a computer-based questionnaire where the questions appear on a computer screen, and the questioning is directed to some degree by computer. With CATI, respondent answers are entered directly into a computer database during the telephone interview.

Confidence interval. A range of values constructed around a point estimate that makes it possible to state that an interval contains the population parameter between its upper and lower confidence limits. The most frequently used confidence interval is the 95% confidence interval. This can be interpreted as there is only a 5% chance that the sample is so extreme that the 95% confidence interval calculated will not cover the population mean.

COST. The variable name that corresponds to question MQ29 (Which of the following best describes how much it cost to get it repaired?).

DK/RF. Don't know or refused.

EDUCA. The variable name that corresponds to question MQ38 (What is the highest education level you have completed?).

FINAL. The variable name that corresponds to question MQ28 (Was the vehicle finally repaired?).

GEND. Respondent's gender recorded by the interviewer.

Independent city. An independent city is a city in the United States of America that does not belong to any county, but rather interacts directly with the state government. Because counties have historically been a strong institution in local government in most of the United States, independent cities are relatively rare outside of Virginia, whose state constitutions make them special cases.

KNWLL. The variable name that corresponds to question MQ20 (When the light first came on, did you already know what it meant or represented?).

Logistic regression. A generalization of linear regression that is used for predicting a binary variable (with values such as yes/no or 0/1). An example of its use is modeling the odds that a borrower will default on a loan based on the borrower's income, debt and age.

MAIN. The data table that contains household information in the final survey database.

Mile_epi. The odometer reading of the vehicle at the time of the MIL episode.

Mile_epi_bin. The range (i.e., bin) for the odometer reading of the vehicle at the time of the MIL episode.

Mile_epi_war. The odometer reading of vehicle at the time of the MIL episode classified using warranty-related bins of MOBILE6.2.

MIL episode. In the context of the survey, a MIL episode starts when the driver first notices that his/her MIL is on and ends when the light has been off for so long that the driver is convinced that the light will stay off. During a given episode, the MIL may be on continuously or it may go on and off.

Mile_owner. The number of miles that the current owner had put on the vehicle between original acquisition and the most recent MIL episode.

Mile_owner_bin. The range (i.e., bin) for the number of miles that the current owner had put on the vehicle between original acquisition and the most recent MIL episode.

Mile_since. The number of miles put on the vehicle by the owner since the MIL illumination episode.

MOREC. The variable name that corresponds to question MQ19 (Would you say that this most recent episode was the only episode or were earlier episodes on this vehicle?).

PURC. The variable name that corresponds to question MQ15 (Did you purchase or lease the vehicle new or used?).

Random digit dial (RDD) technique. A method to give all phone numbers in a region an equal chance of being dialed.

REMAN. The variable name that corresponds to question MQ26 (Did the “check engine” light remain off after it was first repaired?).

Response. Vehicle owner/driver response to MIL illumination episode, Response was assigned to be Positive in all cases where respondents claimed to have made a repair and negative otherwise.

RESPP. The variable name that corresponds to question SQ4 (How familiar are you with the maintenance of this/these vehicles? Rank your level of familiarity on a scale of 1 to 5, with 1 being "no familiarity" and 5 being "extremely familiar"?).

RETUN. The variable name that corresponds to question MQ27 (How many times did you have to return the vehicle for repairs for this episode?).

ROST1. The data table that describes the model year, make, and model of each 1996 through 2004 vehicle that is in the each household in the final survey database.

ROST2. The data table containing additional information on vehicles that have had dashboard lights go on in the final survey database.

Sample disposition. A complete account of the number of phone interviews as well as interview attempts.

SAMPN. A unique sample number assigned to each household.

STILL. The variable name that corresponds to question MQ31 (Does the "check engine" light still come on when the vehicle is driven?).

Survey instrument (i.e., questionnaire). A record of the questions to be asked of a respondent during an interview, with appropriate instructions indicating which questions are to be asked, and in which order.

TIMEP. The variable name that corresponds to question MQ25 (Which of the following best describes the time period between when the "check engine" light was first noticed and when the vehicle was eventually taken for service and repair?).

TOTCT. The variable name that corresponds to question MQ30 (Was the cost to get this repair fixed less than $100, between $100 to $200, between $200 to $500, or greater than $500?).

Vehage. The current age of the vehicle computed by rounding the number of years between January 1 of the model year of the vehicle until October 1, 2004, which was the approximate date of the telephone survey.

VEHNO. A unique vehicle number assigned to each vehicle.

WARRN. The variable name that corresponds to question MQ32 (Was the vehicle covered under warranty for repairs because of the "check engine" light going on?).

WHREP. The variable name that corresponds to question MQ24 (Which of the following best describes where you had the vehicle repaired).

WHTD5. The variable name that corresponds to question MQ22e (What did you finally do to address the light being on? Did you get the vehicle Repaired or repaired the vehicle yourself?).

WHYH. The variable name that corresponds to question MQ23a through k (Why did you choose not to take the vehicle to be repaired/serviced?).