MOBILE SOURCE EMISSIONS ANALYSIS FOR CALIFORNIA

VOLUME I

Contract No. A2-065-32

Prepared for:

CALIFORNIA AIR RESOURCES BOARD
Sacramento, California

Prepared by:

R. Crawford, K. Duleep and R. Dulla

ENERGY AND ENVIRONMENTAL ANALYSIS, INC.
1655 North Fort Myer Drive
Arlington, Virginia 22209

T. Austin and G. Rubenstein

SIERRA RESEARCH, INC.
1009 14th Street
Sacramento, California 95814

June 1985

PREFACE

This report presents the results of a major research study on mobile source emissions from California vehicles. That effort was divided into four task areas:

Task 1: Analysis of Pre-1980 Model Year Light-and-Medium Duty Vehicle Emissions

Task 2: Analysis of Post-1979 Model Year Light-Duty and Medium-Duty Vehicle Emissions

Task 3: Analysis of Heavy Duty Vehicle Emissions

Task 4: Analysis of Regulatory Issues

A total of 14 separate reports were produced under the contract. This volume contains all of the reports* produced under the first three tasks of the contract. They are:

Review of Alternative Emission Factor Specifications for 1972 to 1979 Automobiles

Assessment of Misfueling Trends for California Vehicles

Temperature Corrections Factors for California's Motor Vehicle Emissions Model

Speed Correction Factors for California's Motor Vehicle Emissions Model

Critique of the EPA I/M Benefits Model for 1980 and Older Model Cars

---

*The report addressing the "Development of California's I/M Credits Model" was jointly funded by tasks 2 and 4 and is contained in Volume II.

Forecast of Emission Control Technology and Strategy for Light-Duty Vehicles

Emission Factors for 1980 and Later Model Year California Passenger Cars and Light-Duty Trucks

Review and Critique of Current Heavy-Duty Truck Emission Factors

All of the reports produced under Task 4, Analysis of Regulatory Issues, are contained in Volume II. For an overview of all of the reports produced under the contract, the reader is referred to:

Executive Sumary of Work Produced Under ARB Contract
"Mobile Source Emissions Analysis for California"

REVIEW OF ALTERNATIVE EMISSION FACTOR SPECIFICATIONS FOR 1972 to 1979 MODEL YEAR LIGHT-DUTY VEHICLES

Prepared for:

CALIFORNIA AIR RESOURCES BOARD
Sacramento, California 95812

Contract No. A2-065-32

Prepared by:

ENERGY AND ENVIRONMENTAL ANALYSIS, INC.
1655 North Fort Myer Drive
Arlington, Virginia 22209

June 1985

TABLE OF CONTENTS

1. INTRODUCTION

The Air Resources Board (ARB) and the Environmental Protection Agency (EPA) conduct periodic testing programs to measure the emissions of regulated pollutants from representative samples of in-use California vehicles (both cars and light duty trucks). The information collected in these "surveillance programs" provides insight into the durability of emission control systems, frequency of tampering and owner maintenance practices. These insights support a variety of regulatory activities that range from the formulation of certification standards, to the specification of I/M programs, to the enforcement of emissions warranty requirements.

Emission factors are used to estimate the influence of all of these regulatory activities on the pollutants emitted from in-use vehicles. The surveillance program provides a proportional sample of emissions measurements from the in-use fleet affected by these programs. Regression analysis reduces the data set into equations or emission factors, that can be used to estimate the emissions of a vehicle at any point in its life. The accepted method of emission factor development is to use linear regression to estimate emissions as a function of the odometer reading. The linear specification guarantees that the estimated emissions increase with rising mileage, consonant with the knowledge that emission control effectiveness deteriorates with age.

For a well-maintained vehicle, (one maintained to manufacturers specifications) this deterioration is termed "normal deterioration." Several studies have confirmed that normal deterioration of "engine out" emissions is insignificant, thus tailpipe emissions deteriorate primarily due to losses in catalyst efficiency in a well maintained vehicle. Deterioration also occurs as a result of a component or system malfunction. These malfunctions can be related to maladjustment, tampering or failure, and unlike normal deterioration, can occur in both engine emission control components and the catalyst. An accurate estimate of in-use emissions must consider both the effects of normal deterioration and component malfunction; the current linear emission factor equations do not explicitly account for these effects.

The purpose of this report is to examine the validity of the current approach used to estimate emission factors for California vehicles. This requires a review of the surveillance data and the range of odometer readings that it contains. It also requires a review of the linear specification and its accuracy in predicting the emissions performance of high mileage vehicles. Alternative specifications are tested, including both more complex linear models that incorporate component malperformances and non-linear models. The analysis focuses on the performance of 1975 to 1979 model year cars and provides a cursory review of 1972-1974 models.

The remainder of this report is organized into four sections. In Section 2 all ARB surveillance data and EPA emission factor data for 1972-1979 model year vehicles is used to develop linear emission factors for the relevant model year and emission standard groups. These emission factors are calculated for the EPA and ARB separately; a method is presented for combining the data to estimate a composite emission factor. Section 3 addresses the problem of finding the correct specification to represent the emissions performance of high mileage vehicles in light of the poor sample of high mileage vehicle data. Alternative emission factor specifications incorporating malperformance variables are testing in Section 4. Section 5 presents the conclusions.

2. DEVELOPMENT OF LINEAR EMISSION FACTORS

2.1 INTRODUCTION

Both EPA and ARB have used the data collected from their individual surveillance programs to independently estimate emission factors for California vehicles. While there has been general agreement between ARB and EPA estimates, there have been some differences, particularly for CO. The purpose of this section is to explore the differences in the composition of these data bases, (both have been expanded with addition of data from recent surveillance programs) and to prepare new emission factors for 1972 to 1979 California vehicles. In order to highlight the differences in the data bases, emission factors are estimated from EPA and ARB data separately. A method of integrating the information from the two data sets is proposed and emission factors are estimated from a combined ARB/EPA data set.

2.2 COMPARISON OF ARB AND EPA SURVEILLANCE PROGRAM DATA

2.2.1 ARB Surveillance Program Data

The data supplied by ARB for this analysis included data from the first through sixth surveillance programs plus data collected from special testing programs. The special testing data focus on vehicles of specific interest to ARB (three-way catalyst vehicles and vehicles with possible warranty and recall defects). Inclusion of this data with the surveillance program data would alter the representativeness of the sample and bias the estimation of emission factors. For this reason, the special testing is excluded from the ARB data set used in this report.

Data from the sixth surveillance program[1/] has only recently become available. It includes emissions measurements from 365 gasoline powered passen-

---

[1/] "Test Report of the Light-Duty Vehicle Surveillance Program, Series 6 (LDVSPVI)," State of California Air Resources Board, MS #83-08, June 1983.

ger cars and light-duty trucks of model years 1975 through 1981. Emission factor estimates will benefit from the incorporation of this most recent data describing the in-use performance of these vehicles, particularly from the inclusion of vehicles with higher odometer readings from the earlier model years.

The manufacturer representation within the ARB data set by model year is presented in Table 2-1. All of the vehicles selected for use in ARB's surveillance program come from the South Coast Air Basin. They are selected to be representative of California vehicle sales according to manufacturer, model and engine size. Twenty-nine manufacturers are represented in ARB's data base from 1972 through 1979. A review of the table indicates that the larger manufacturers, GM, Ford, and Chrysler, appear in rough proportion to their sales within the State, and that these proportions do not vary substantially across the model years.

Table 2-2 provides the distribution of the emission control technologies represented within each model year of the ARB data set. While numerous definitions of emission control system are possible, all definitions within this section are based on the type of catalyst and related control system. The four following categories were employed:

- Three-way catalyst
- No catalyst
- Catalyst only
- Catalyst with air (either air pump or pulse air)

These categories will be used to explore the comparability of the EPA and ARB data sets. Because the ARB data set is proportional to vehicle sales in California, the technology distribution displayed in Table 2-2 will be used to weight technology-specific emission factors (for any combination of EPA and ARB data) together to form aggregate emission factors for appropriate model years and pollutant standards (i.e., 1975-1976 and 1977-1979).

TABLE 2-1

Calif ARB Data

```
                         TABLE OF MFR BY MY

MFR       MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |72      |73      |74      |75      |76      |77      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+--------+--------+--------+--------+
AMC      |      4 |      4 |      3 |      7 |      6 |      5 |      4 |      2 |     35
         |   0.36 |   0.36 |   0.27 |   0.63 |   0.54 |   0.45 |   0.36 |   0.18 |   3.14
         |  11.43 |  11.43 |   8.57 |  20.00 |  17.14 |  14.29 |  11.43 |   5.71 |
         |   4.49 |   4.35 |   3.45 |   3.61 |   2.48 |   2.70 |   2.99 |   2.15 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
AUDI     |      1 |      0 |      0 |      1 |      2 |      0 |      0 |      2 |      6
         |   0.09 |   0.00 |   0.00 |   0.09 |   0.18 |   0.00 |   0.00 |   0.18 |   0.54
         |  16.67 |   0.00 |   0.00 |  16.67 |  33.33 |   0.00 |   0.00 |  33.33 |
         |   1.12 |   0.00 |   0.00 |   0.52 |   0.83 |   0.00 |   0.00 |   2.15 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
BL       |      0 |      0 |      0 |      0 |      1 |      0 |      0 |      0 |      1
         |   0.00 |   0.00 |   0.00 |   0.00 |   0.09 |   0.00 |   0.00 |   0.00 |   0.09
         |   0.00 |   0.00 |   0.00 |   0.00 | 100.00 |   0.00 |   0.00 |   0.00 |
         |   0.00 |   0.00 |   0.00 |   0.00 |   0.41 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
BMW      |      0 |      1 |      0 |      1 |      2 |      0 |      0 |      0 |      4
         |   0.00 |   0.09 |   0.00 |   0.09 |   0.18 |   0.00 |   0.00 |   0.00 |   0.36
         |   0.00 |  25.00 |   0.00 |  25.00 |  50.00 |   0.00 |   0.00 |   0.00 |
         |   0.00 |   1.09 |   0.00 |   0.52 |   0.83 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
CHRY     |     12 |     10 |      9 |     18 |     23 |     19 |     14 |     10 |    115
         |   1.08 |   0.90 |   0.81 |   1.61 |   2.06 |   1.70 |   1.25 |   0.90 |  10.30
         |  10.43 |   8.70 |   7.83 |  15.65 |  20.00 |  16.52 |  12.17 |   8.70 |
         |  13.48 |  10.87 |  10.34 |   9.28 |   9.50 |  10.27 |  10.45 |  10.75 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
F-CAPRI  |      0 |      1 |      1 |      0 |      0 |      0 |      0 |      0 |      2
         |   0.00 |   0.09 |   0.09 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.18
         |   0.00 |  50.00 |  50.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
         |   0.00 |   1.09 |   1.15 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
F-COUR   |      1 |      0 |      0 |      0 |      0 |      0 |      0 |      0 |      1
         |   0.09 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.09
         | 100.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
         |   1.12 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
FIAT     |      0 |      0 |      0 |      2 |      3 |      1 |      1 |      2 |      9
         |   0.00 |   0.00 |   0.00 |   0.18 |   0.27 |   0.09 |   0.09 |   0.18 |   0.81
         |   0.00 |   0.00 |   0.00 |  22.22 |  33.33 |  11.11 |  11.11 |  22.22 |
         |   0.00 |   0.00 |   0.00 |   1.03 |   1.24 |   0.54 |   0.75 |   2.15 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
```

TABLE 2-1 (con't)

Calif ARB Data

```
                              TABLE OF MFR BY MY

MFR          MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |72      |73      |74      |75      |76      |77      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+--------+--------+--------+--------+
FORD     |     18 |     18 |     23 |     47 |     57 |     42 |     36 |     21 |    262
         |   1.61 |   1.61 |   2.06 |   4.21 |   5.11 |   3.76 |   3.23 |   1.88 |  23.48
         |   6.87 |   6.87 |   8.78 |  17.94 |  21.76 |  16.03 |  13.74 |   8.02 |
         |  20.22 |  19.57 |  26.44 |  24.23 |  23.55 |  22.70 |  26.87 |  22.58 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
FUJI     |      0 |      0 |      0 |      0 |      3 |      3 |      0 |      2 |      8
         |   0.00 |   0.00 |   0.00 |   0.00 |   0.27 |   0.27 |   0.00 |   0.18 |   0.72
         |   0.00 |   0.00 |   0.00 |   0.00 |  37.50 |  37.50 |   0.00 |  25.00 |
         |   0.00 |   0.00 |   0.00 |   0.00 |   1.24 |   1.62 |   0.00 |   2.15 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
GM       |     35 |     38 |     28 |     63 |     75 |     66 |     40 |     30 |    375
         |   3.14 |   3.41 |   2.51 |   5.65 |   6.72 |   5.91 |   3.58 |   2.69 |  33.60
         |   9.33 |  10.13 |   7.47 |  16.80 |  20.00 |  17.60 |  10.67 |   8.00 |
         |  39.33 |  41.30 |  32.18 |  32.47 |  30.99 |  35.68 |  29.85 |  32.26 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
GM-OPEL  |      0 |      0 |      1 |      0 |      0 |      0 |      0 |      0 |      1
         |   0.00 |   0.00 |   0.09 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.09
         |   0.00 |   0.00 | 100.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
         |   0.00 |   0.00 |   1.15 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
HONDA    |      0 |      0 |      1 |      8 |     12 |      5 |      7 |      4 |     37
         |   0.00 |   0.00 |   0.09 |   0.72 |   1.08 |   0.45 |   0.63 |   0.36 |   3.32
         |   0.00 |   0.00 |   2.70 |  21.62 |  32.43 |  13.51 |  18.92 |  10.81 |
         |   0.00 |   0.00 |   1.15 |   4.12 |   4.96 |   2.70 |   5.22 |   4.30 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
ISUZU    |      0 |      0 |      0 |      0 |      1 |      3 |      3 |      0 |      7
         |   0.00 |   0.00 |   0.00 |   0.00 |   0.09 |   0.27 |   0.27 |   0.00 |   0.63
         |   0.00 |   0.00 |   0.00 |   0.00 |  14.29 |  42.86 |  42.86 |   0.00 |
         |   0.00 |   0.00 |   0.00 |   0.00 |   0.41 |   1.62 |   2.24 |   0.00 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
JRT      |      0 |      0 |      0 |      0 |      2 |      2 |      0 |      0 |      4
         |   0.00 |   0.00 |   0.00 |   0.00 |   0.18 |   0.18 |   0.00 |   0.00 |   0.36
         |   0.00 |   0.00 |   0.00 |   0.00 |  50.00 |  50.00 |   0.00 |   0.00 |
         |   0.00 |   0.00 |   0.00 |   0.00 |   0.83 |   1.08 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
M-BENZ   |      0 |      1 |      0 |      0 |      0 |      0 |      0 |      0 |      1
         |   0.00 |   0.09 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.09
         |   0.00 | 100.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
         |   0.00 |   1.09 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
```

TABLE 2-1 (con't)

Calif ARB Data

```
                         TABLE OF MFR BY MY

MFR          MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |72       |73       |74       |75       |76       |77       |78       |79       |  TOTAL
---------+---------+---------+---------+---------+---------+---------+---------+---------+
MAZDA    |       1 |       0 |       0 |       0 |       0 |       0 |       0 |       0 |      1
         |    0.09 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |   0.09
         |  100.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |
         |    1.12 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
MERCEDES |       0 |       0 |       0 |       1 |       0 |       0 |       0 |       0 |      1
         |    0.00 |    0.00 |    0.00 |    0.09 |    0.00 |    0.00 |    0.00 |    0.00 |   0.09
         |    0.00 |    0.00 |    0.00 |  100.00 |    0.00 |    0.00 |    0.00 |    0.00 |
         |    0.00 |    0.00 |    0.00 |    0.52 |    0.00 |    0.00 |    0.00 |    0.00 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
MITSU    |       0 |       1 |       1 |       5 |       9 |       7 |       4 |       4 |     31
         |    0.00 |    0.09 |    0.09 |    0.45 |    0.81 |    0.63 |    0.36 |    0.36 |   2.78
         |    0.00 |    3.23 |    3.23 |   16.13 |   29.03 |   22.58 |   12.90 |   12.90 |
         |    0.00 |    1.09 |    1.15 |    2.58 |    3.72 |    3.78 |    2.99 |    4.30 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
NISSAN   |       3 |       2 |       3 |       9 |      14 |       9 |       8 |       4 |     52
         |    0.27 |    0.18 |    0.27 |    0.81 |    1.25 |    0.81 |    0.72 |    0.36 |   4.66
         |    5.77 |    3.85 |    5.77 |   17.31 |   26.92 |   17.31 |   15.38 |    7.69 |
         |    3.37 |    2.17 |    3.45 |    4.64 |    5.79 |    4.86 |    5.97 |    4.30 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
PEUGEOT  |       0 |       0 |       1 |       1 |       1 |       0 |       0 |       0 |      3
         |    0.00 |    0.00 |    0.09 |    0.09 |    0.09 |    0.00 |    0.00 |    0.00 |   0.27
         |    0.00 |    0.00 |   33.33 |   33.33 |   33.33 |    0.00 |    0.00 |    0.00 |
         |    0.00 |    0.00 |    1.15 |    0.52 |    0.41 |    0.00 |    0.00 |    0.00 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
PORSCHE  |       0 |       0 |       0 |       2 |       1 |       1 |       1 |       0 |      5
         |    0.00 |    0.00 |    0.00 |    0.18 |    0.09 |    0.09 |    0.09 |    0.00 |   0.45
         |    0.00 |    0.00 |    0.00 |   40.00 |   20.00 |   20.00 |   20.00 |    0.00 |
         |    0.00 |    0.00 |    0.00 |    1.03 |    0.41 |    0.54 |    0.75 |    0.00 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
RENAULT  |       0 |       0 |       0 |       0 |       0 |       0 |       2 |       0 |      2
         |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.18 |    0.00 |   0.18
         |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |  100.00 |    0.00 |
         |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |    1.49 |    0.00 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
SUBARU   |       0 |       0 |       1 |       0 |       0 |       0 |       0 |       0 |      1
         |    0.00 |    0.00 |    0.09 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |   0.09
         |    0.00 |    0.00 |  100.00 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |
         |    0.00 |    0.00 |    1.15 |    0.00 |    0.00 |    0.00 |    0.00 |    0.00 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
```

TABLE 2-1 (con't)

Calif ARB Data

```
                         TABLE OF MFR BY MY

MFR       MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |72      |73      |74      |75      |76      |77      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+--------+--------+--------+--------+
TOYOK    |      2 |      4 |      1 |      0 |      0 |      0 |      0 |      0 |      7
         |   0.18 |   0.36 |   0.09 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.63
         |  28.57 |  57.14 |  14.29 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
         |   2.25 |   4.35 |   1.15 |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
TOYOKOGY |      0 |      1 |      1 |      3 |      1 |      2 |      0 |      1 |      9
         |   0.00 |   0.09 |   0.09 |   0.27 |   0.09 |   0.18 |   0.00 |   0.09 |   0.81
         |   0.00 |  11.11 |  11.11 |  33.33 |  11.11 |  22.22 |   0.00 |  11.11 |
         |   0.00 |   1.09 |   1.15 |   1.55 |   0.41 |   1.08 |   0.00 |   1.08 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
TOYOTA   |      4 |      4 |      5 |     10 |     15 |      9 |      7 |      4 |     58
         |   0.36 |   0.36 |   0.45 |   0.90 |   1.34 |   0.81 |   0.63 |   0.36 |   5.20
         |   6.90 |   6.90 |   8.62 |  17.24 |  25.86 |  15.52 |  12.07 |   6.90 |
         |   4.49 |   4.35 |   5.75 |   5.15 |   6.20 |   4.86 |   5.22 |   4.30 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
VOLVO    |      0 |      0 |      1 |      3 |      2 |      1 |      0 |      2 |      9
         |   0.00 |   0.00 |   0.09 |   0.27 |   0.18 |   0.09 |   0.00 |   0.18 |   0.81
         |   0.00 |   0.00 |  11.11 |  33.33 |  22.22 |  11.11 |   0.00 |  22.22 |
         |   0.00 |   0.00 |   1.15 |   1.55 |   0.83 |   0.54 |   0.00 |   2.15 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
VW       |      8 |      7 |      7 |     13 |     12 |     10 |      7 |      5 |     69
         |   0.72 |   0.63 |   0.63 |   1.16 |   1.08 |   0.90 |   0.63 |   0.45 |   6.18
         |  11.59 |  10.14 |  10.14 |  18.84 |  17.39 |  14.49 |  10.14 |   7.25 |
         |   8.99 |   7.61 |   8.05 |   6.70 |   4.96 |   5.41 |   5.22 |   5.38 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
TOTAL          89       92       87      194      242      185      134       93     1116
             7.97     8.24     7.80    17.38    21.68    16.58    12.01     8.33   100.00
```

TABLE 2-2

Calif ARB Data

```
                         TABLE OF TECH BY MY

TECH          MY

FREQUENCY    |
 PERCENT     |
 ROW PCT     |
 COL PCT     |72      |73      |74      |75      |76      |77      |78      |79      |  TOTAL
-------------+--------+--------+--------+--------+--------+--------+--------+--------+
 3-WAY  CAT  |      0 |      0 |      0 |      0 |      0 |      1 |      2 |      7 |     10
             |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.09 |   0.18 |   0.63 |   0.90
             |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |  10.00 |  20.00 |  70.00 |
             |   0.00 |   0.00 |   0.00 |   0.00 |   0.00 |   0.54 |   1.49 |   7.53 |
-------------+--------+--------+--------+--------+--------+--------+--------+--------+
NO CAT       |     89 |     92 |     86 |     20 |     30 |     17 |      9 |      8 |    351
             |   7.97 |   8.24 |   7.71 |   1.79 |   2.69 |   1.52 |   0.81 |   0.72 |  31.45
             |  25.36 |  26.21 |  24.50 |   5.70 |   8.55 |   4.84 |   2.56 |   2.28 |
             | 100.00 | 100.00 |  98.85 |  10.31 |  12.40 |   9.19 |   6.72 |   8.60 |
-------------+--------+--------+--------+--------+--------+--------+--------+--------+
CAT ONLY     |      0 |      0 |      1 |     34 |     41 |     11 |     12 |     13 |    112
             |   0.00 |   0.00 |   0.09 |   3.05 |   3.67 |   0.99 |   1.08 |   1.16 |  10.04
             |   0.00 |   0.00 |   0.89 |  30.36 |  36.61 |   9.82 |  10.71 |  11.61 |
             |   0.00 |   0.00 |   1.15 |  17.53 |  16.94 |   5.95 |   8.96 |  13.98 |
-------------+--------+--------+--------+--------+--------+--------+--------+--------+
CAT W/  AIR  |      0 |      0 |      0 |    140 |    171 |    156 |    111 |     65 |    643
             |   0.00 |   0.00 |   0.00 |  12.54 |  15.32 |  13.98 |   9.95 |   5.82 |  57.62
             |   0.00 |   0.00 |   0.00 |  21.77 |  26.59 |  24.26 |  17.26 |  10.11 |
             |   0.00 |   0.00 |   0.00 |  72.16 |  70.66 |  84.32 |  82.84 |  69.89 |
-------------+--------+--------+--------+--------+--------+--------+--------+--------+
TOTAL              89       92       87      194      242      185      134       93     1116
                 7.97     8.24     7.80    17.38    21.68    16.58    12.01     8.33   100.00
```

As shown in Table 2-2 three-way catalyst vehicles did not achieve a significant penetration of the market until 1979. The dominant form of emission control system for the 1975 to 1979 model year vehicles was the catalyst with air, with penetrations ranging from approximately 70 to 85 percent during that time.

Table 2-3 presents a distribution of cylinder count by model year. The increasing demand for fuel economy exhibited through the decade of the 1970s is reflected in the trends in this table. The market penetration of eight cylinder engines steadily declines from 56 percent in 1972 to 38 percent in 1979. Conversely, the penetration of four cylinder engines increases slowly from 29 to 37 percent during the same time. Six cylinder engines remain relatively stable except for a sharp increase from 16 to 23 percent from 1978 to 1979.

Table 2-4 displays the market penetration of transmissions in the model years of interest. The penetrations of automatic and manual transmissions are relatively stable throughout the time period with exception of the sharp increase in manual transmissions, almost doubling from 1978 to 1979; possibly reflecting the sharp increase in gasoline prices and decrease in fuel availability experienced in that calendar year.

Figure 2-1 through 2-3 display frequency distributions or histograms for odometer readings for each of the model year/emission standard categories. For the sake of display, each vehicle's odometer readings has been divided by 10,000; the resulting values which range from 0 to 12 are then organized into mile categories or bins labeled midpoint VMTX. The following example, illustrates this organization: those vehicles with a midpoint of 1 have odometer readings that range from 5001 to 14,999. As would be expected, Figure 2-1 (model years 1972-1974) shows a wide distribution of odometer readings. The median value is in 50,000 mile bin; approximately 24 percent of these vehicles have an odometer reading greater than 65,000 miles.

TABLE 2-3

Calif ARB Data

```
                              TABLE OF CYL BY MY

CYL          MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |72       |73       |74       |75       |76       |77       |78       |79       |  TOTAL
---------+---------+---------+---------+---------+---------+---------+---------+---------+
   2     |       1 |       2 |       1 |       0 |       0 |       0 |       0 |       1 |      5
         |    0.09 |    0.18 |    0.09 |    0.00 |    0.00 |    0.00 |    0.00 |    0.09 |   0.45
         |   20.00 |   40.00 |   20.00 |    0.00 |    0.00 |    0.00 |    0.00 |   20.00 |
         |    1.12 |    2.17 |    1.15 |    0.00 |    0.00 |    0.00 |    0.00 |    1.08 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
   4     |      26 |      26 |      31 |      69 |     100 |      62 |      57 |      34 |    405
         |    2.33 |    2.33 |    2.78 |    6.18 |    8.96 |    5.56 |    5.11 |    3.05 |  36.29
         |    6.42 |    6.42 |    7.65 |   17.04 |   24.69 |   15.31 |   14.07 |    8.40 |
         |   29.21 |   28.26 |   35.63 |   35.57 |   41.32 |   33.51 |   42.54 |   36.56 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
   5     |       0 |       3 |       0 |       2 |       0 |       0 |       0 |       2 |      7
         |    0.00 |    0.27 |    0.00 |    0.18 |    0.00 |    0.00 |    0.00 |    0.18 |   0.63
         |    0.00 |   42.86 |    0.00 |   28.57 |    0.00 |    0.00 |    0.00 |   28.57 |
         |    0.00 |    3.26 |    0.00 |    1.03 |    0.00 |    0.00 |    0.00 |    2.15 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
   6     |      12 |       9 |      17 |      34 |      45 |      30 |      21 |      21 |    189
         |    1.08 |    0.81 |    1.52 |    3.05 |    4.03 |    2.69 |    1.88 |    1.88 |  16.94
         |    6.35 |    4.76 |    8.99 |   17.99 |   23.81 |   15.87 |   11.11 |   11.11 |
         |   13.48 |    9.78 |   19.54 |   17.53 |   18.60 |   16.22 |   15.67 |   22.58 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
   8     |      50 |      52 |      38 |      89 |      97 |      93 |      56 |      35 |    510
         |    4.48 |    4.66 |    3.41 |    7.97 |    8.69 |    8.33 |    5.02 |    3.14 |  45.70
         |    9.80 |   10.20 |    7.45 |   17.45 |   19.02 |   18.24 |   10.98 |    6.86 |
         |   56.18 |   56.52 |   43.68 |   45.88 |   40.08 |   50.27 |   41.79 |   37.63 |
---------+---------+---------+---------+---------+---------+---------+---------+---------+
TOTAL           89        92        87       194       242       185       134        93     1116
              7.97      8.24      7.80     17.38     21.68     16.58     12.01      8.33   100.00
```

TABLE 2-4

Calif ARB Data

```
                          TABLE OF TRANS BY MY

TRANS        MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |72      |73      |74      |75      |76      |77      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+--------+--------+--------+--------+
       . |     61 |     62 |     60 |     79 |    101 |    119 |    130 |     78 |      .
         |      . |      . |      . |      . |      . |      . |      . |      . |      .
         |      . |      . |      . |      . |      . |      . |      . |      . |
         |      . |      . |      . |      . |      . |      . |      . |      . |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
MANUAL   |      6 |      8 |      6 |     29 |     46 |     15 |      1 |      7 |    118
         |   1.41 |   1.88 |   1.41 |   6.81 |  10.80 |   3.52 |   0.23 |   1.64 |  27.70
         |   5.08 |   6.78 |   5.08 |  24.58 |  38.98 |  12.71 |   0.85 |   5.93 |
         |  21.43 |  26.67 |  22.22 |  25.22 |  32.62 |  22.73 |  25.00 |  46.67 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
 AUTO    |     22 |     22 |     21 |     86 |     95 |     51 |      3 |      8 |    308
         |   5.16 |   5.16 |   4.93 |  20.19 |  22.30 |  11.97 |   0.70 |   1.88 |  72.30
         |   7.14 |   7.14 |   6.82 |  27.92 |  30.84 |  16.56 |   0.97 |   2.60 |
         |  78.57 |  73.33 |  77.78 |  74.78 |  67.38 |  77.27 |  75.00 |  53.33 |
---------+--------+--------+--------+--------+--------+--------+--------+--------+
TOTAL          28       30       27      115      141       66        4       15      426
             6.57     7.04     6.34    27.00    33.10    15.49     0.94     3.52   100.00
```

FIGURE 2-1
Calif ARB Data

```
                                            CARGRP=MY 72-74

                                          FREQUENCY BAR CHART
MIDPOINT
 VMTX                                                                                  FREQ  CUM.   PERCENT    CUM.
                                                                                             FREQ            PERCENT
        |
   0    |                                                                                 0     0     0.00     0.00
        |
   1    |                                                                                 0     0     0.00     0.00
        |
   2    |********                                                                         4     4     1.49     1.49
        |
   3    |****************************                                                    14    18     5.22     6.72
        |
   4    |**************************************************************************      37    55    13.81    20.52
        |
   5    |********************************************************                        28    83    10.45    30.97
        |
   6    |******************************************************************              33   116    12.31    43.28
        |
   7    |************************************************************************************  42   158    15.67    58.96
        |
   8    |********************************************************************************************  46   204    17.16    76.12
        |
   9    |************************************                                            18   222     6.72    82.84
        |
  10    |********************************                                                16   238     5.97    88.81
        |
  11    |******************                                                               9   247     3.36    92.16
        |
  12    |******************************************                                      21   268     7.84   100.00
        |
        ----+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
            2   4   6   8  10  12  14  16  18  20  22  24  26  28  30  32  34  36  38  40  42  44  46

                                             FREQUENCY
```

FIGURE 2-2
Calif ARB Data

```
                                        CARGRP=MY 75-76

                                      FREQUENCY BAR CHART
MIDPOINT
  VMTX                                                                                                FREQ   CUM.   PERCENT     CUM.
                                                                                                             FREQ             PERCENT
         |
    0    |********************************                                                            32     32      7.37      7.37
         |
    1    |**********************************************************************************          82    114     18.89     26.27
         |
    2    |***********************************************************************************         83    197     19.12     45.39
         |
    3    |*******************************************************************************************  91    288     20.97     66.36
         |
    4    |**********************************************************                                  58    346     13.36     79.72
         |
    5    |************************************                                                        36    382      8.29     88.02
         |
    6    |******************************                                                              30    412      6.91     94.93
         |
    7    |************                                                                                12    424      2.76     97.70
         |
    8    |******                                                                                       6    430      1.38     99.08
         |
    9    |***                                                                                          3    433      0.69     99.77
         |
   10    |*                                                                                            1    434      0.23    100.00
         |
   11    |                                                                                             0    434      0.00    100.00
         |
   12    |                                                                                             0    434      0.00    100.00
         |
         -----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+-
              5   10   15   20   25   30   35   40   45   50   55   60   65   70   75   80   85   90

                                                  FREQUENCY
```

FIGURE 2-3

Calif ARB Data
CARGRP=MY 77-79

```
                                   FREQUENCY BAR CHART
MIDPOINT
  VMTX                                                                                  FREQ  CUM.   PERCENT    CUM.
                                                                                              FREQ              PERCENT
        |
    0   |*********                                                                        17    17     4.14      4.14
        |
    1   |******************************************************                          108   125    26.28     30.41
        |
    2   |****************************************************************** 132   257    32.12     62.53
        |
    3   |**************************************                                           75   332    18.25     80.78
        |
    4   |**********************                                                           43   375    10.46     91.24
        |
    5   |*********                                                                        17   392     4.14     95.38
        |
    6   |*****                                                                            10   402     2.43     97.81
        |
    7   |****                                                                              8   410     1.95     99.76
        |
    8   |                                                                                  0   410     0.00     99.76
        |
    9   |                                                                                  0   410     0.00     99.76
        |
   10   |*                                                                                 1   411     0.24    100.00
        |
   11   |                                                                                  0   411     0.00    100.00
        |
   12   |                                                                                  0   411     0.00    100.00
        |
        -----+----+----+----+----+----+----+----+----+----+----+----+----+-
            10   20   30   40   50   60   70   80   90  100  110  120  130

                                       FREQUENCY
```

Figure 2-2 (model years 1975-1976) shows a substantially different distribution of odometer readings. In this figure, 45 percent of the vehicles have odometer readings under 25,000 miles. The median value is just over 35,000 miles and only 12 percent of the vehicles have odometer readings over 55,000 miles.

Figure 2-3 displays an even younger distribution of vehicles (model years 1977-1979). The median value for this group of vehicles is under 25,000 miles. Less than five percent of the data has an odometer reading over 55,000 miles. The data displayed in these figures corresponds well with analyses[2/] that document the relation between odometer readings and age of vehicle.

2.2.2 EPA Surveillance Program Data

The Environmental Protection Agency conducts an annual or biennial vehicle emissions testing program, the Emission Factor Program (EFP), for the purpose of estimating the average emissions from a nationally representative sample of in-use vehicles. The above description of EPA's emission factor program is accurate for 49-State cars; EPA's data collection efforts for California have not been nearly as rigorous as ARB's. Nevertheless, EPA has accumulated a large body of data detailing the performance of California vehicles. That information was accessed for this study, although not all of it was used. For example, EPA conducted a special program to evaluate the emissions of 1979 model year California vehicles that had accumulated between 40,000 and 50,000 miles. Unfortunately, all of the vehicles were screened for proper use and maintenance; vehicles that had not received proper maintenance and those with major control system disablements were rejected from the program. Because of the unrepresentativeness of the

---

[2/] "The Highway Fuel Consumption Model, Ninth Quarterly Report," Prepared for U.S. Department of Energy, Contract Number DE-AC01-79PE-70045, Task 21, Prepared by Energy and Environmental Analysis, Inc., February 25, 1983.

vehicles included in that program, these were not included in this analysis. Conversations with EPA personnel indicated that EPA calculations of emission factors for California vehicles have in the past used a mixture of EPA and ARB data. The exact proportions of the data used in these calculations is not known or documented, therefore it is impossible to replicate either the data base or the emission factors. The remainder of this section presents a description of EPA EFP data for California vehicles; the tables used to describe the EPA data are similar to those used to describe the ARB data.

Table 2-5 presents the distribution of manufacturers by model year in the EPA data. The manufacturer representation is substantially smaller than that of ARB with only 12 manufacturers included. The stability of manufacturer representation by model year is also unlike the ARB data; GM's market penetration ranges from 23 to 75 percent of the data. Essentially, no data was collected for 1978 model year vehicles.

The distribution of technologies by model year is displayed in Table 2-6. No information was collected for catalyst only vehicles and there is a disproportionate representation of three-way catalyst vehicles. The 95 percent representation of three-way catalyst vehicles in the 1979 model year compares with the 7.5 percent in the ARB data. The reason for this misrepresentation was EPA's need to understand the in-use performance of these systems. Because of their introduction in California several years prior to their 49-State introduction, EPA's California data collection efforts focused on these vehicles. Due to the bias shown in Table 2-6 toward three-way catalyst vehicles, they have been excluded from all tables including Table 2-5 in this section. Similar tables and figures describing the distribution of the three-way catalyst vehicles are displayed in Appendix A.

The distribution of engine size (as measured by cylinder count) by model year is displayed in Table 2-7. As in the previous EPA tables there are no definitive trends in the data; the proportions of engine size vary with no

TABLE 2-5

```
                              EPA DATA

                      (3 WAY CAT DATA EXCLUDED)

                          TABLE OF MFR BY MY

MFR           MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |75      |76      |77      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+--------+
BMW      |      0 |      1 |      0 |      0 |      0 |      1
         |   0.00 |   0.52 |   0.00 |   0.00 |   0.00 |   0.52
         |   0.00 | 100.00 |   0.00 |   0.00 |   0.00 |
         |   0.00 |   1.09 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+
AMC      |      2 |      4 |      0 |      0 |      0 |      6
         |   1.04 |   2.08 |   0.00 |   0.00 |   0.00 |   3.13
         |  33.33 |  66.67 |   0.00 |   0.00 |   0.00 |
         |   3.08 |   4.35 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+
CHRYSLER |      4 |     14 |      1 |      0 |      2 |     21
         |   2.08 |   7.29 |   0.52 |   0.00 |   1.04 |  10.94
         |  19.05 |  66.67 |   4.76 |   0.00 |   9.52 |
         |   6.15 |  15.22 |   7.69 |   0.00 |  10.00 |
---------+--------+--------+--------+--------+--------+
DATSUN   |      3 |      3 |      0 |      0 |      0 |      6
         |   1.56 |   1.56 |   0.00 |   0.00 |   0.00 |   3.13
         |  50.00 |  50.00 |   0.00 |   0.00 |   0.00 |
         |   4.62 |   3.26 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+
FIAT     |      2 |      0 |      0 |      0 |      0 |      2
         |   1.04 |   0.00 |   0.00 |   0.00 |   0.00 |   1.04
         | 100.00 |   0.00 |   0.00 |   0.00 |   0.00 |
         |   3.08 |   0.00 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+
FORD     |     22 |     35 |      1 |      1 |      2 |     61
         |  11.46 |  18.23 |   0.52 |   0.52 |   1.04 |  31.77
         |  36.07 |  57.38 |   1.64 |   1.64 |   3.28 |
         |  33.85 |  38.04 |   7.69 |  50.00 |  10.00 |
---------+--------+--------+--------+--------+--------+
```

TABLE 2-5 (con't)

```
                          EPA DATA

                  (3 WAY CAT DATA EXCLUDED)

                     TABLE OF MFR BY MY

MFR            MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |75      |76      |77      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+--------+
GM       |     15 |     24 |     10 |      1 |     15 |     65
         |   7.81 |  12.50 |   5.21 |   0.52 |   7.81 |  33.85
         |  23.08 |  36.92 |  15.38 |   1.54 |  23.08 |
         |  23.08 |  26.09 |  76.92 |  50.00 |  75.00 |
---------+--------+--------+--------+--------+--------+
HONDA    |      6 |      4 |      1 |      0 |      0 |     11
         |   3.13 |   2.08 |   0.52 |   0.00 |   0.00 |   5.73
         |  54.55 |  36.36 |   9.09 |   0.00 |   0.00 |
         |   9.23 |   4.35 |   7.69 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+
MAZDA    |      0 |      0 |      0 |      0 |      1 |      1
         |   0.00 |   0.00 |   0.00 |   0.00 |   0.52 |   0.52
         |   0.00 |   0.00 |   0.00 |   0.00 | 100.00 |
         |   0.00 |   0.00 |   0.00 |   0.00 |   5.00 |
---------+--------+--------+--------+--------+--------+
MERCEDES |      1 |      0 |      0 |      0 |      0 |      1
         |   0.52 |   0.00 |   0.00 |   0.00 |   0.00 |   0.52
         | 100.00 |   0.00 |   0.00 |   0.00 |   0.00 |
         |   1.54 |   0.00 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+
TOYOTA   |      6 |      6 |      0 |      0 |      0 |     12
         |   3.13 |   3.13 |   0.00 |   0.00 |   0.00 |   6.25
         |  50.00 |  50.00 |   0.00 |   0.00 |   0.00 |
         |   9.23 |   6.52 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+
VW       |      4 |      1 |      0 |      0 |      0 |      5
         |   2.08 |   0.52 |   0.00 |   0.00 |   0.00 |   2.60
         |  80.00 |  20.00 |   0.00 |   0.00 |   0.00 |
         |   6.15 |   1.09 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+
TOTAL          65       92       13        2       20      192
            33.85    47.92     6.77     1.04    10.42   100.00
```

TABLE 2-6

```
                         EPA DATA

                        (ALL DATA)

                   TABLE OF TECH BY MY

TECH         MY

FREQUENCY   |
 PERCENT    |
 ROW PCT    |
 COL PCT    |75      |76      |77      |78      |79      |  TOTAL
------------+--------+--------+--------+--------+--------+
 3-WAY  CAT |      4 |      1 |      0 |    172 |    423 |    600
            |   0.51 |   0.13 |   0.00 |  21.72 |  53.41 |  75.76
            |   0.67 |   0.17 |   0.00 |  28.67 |  70.50 |
            |   5.80 |   1.08 |   0.00 |  98.85 |  95.49 |
------------+--------+--------+--------+--------+--------+
NO CAT      |      6 |      6 |      1 |      1 |      3 |     17
            |   0.76 |   0.76 |   0.13 |   0.13 |   0.38 |   2.15
            |  35.29 |  35.29 |   5.88 |   5.88 |  17.65 |
            |   8.70 |   6.45 |   7.69 |   0.57 |   0.68 |
------------+--------+--------+--------+--------+--------+
CAT W/  AIR |     59 |     86 |     12 |      1 |     17 |    175
            |   7.45 |  10.86 |   1.52 |   0.13 |   2.15 |  22.10
            |  33.71 |  49.14 |   6.86 |   0.57 |   9.71 |
            |  85.51 |  92.47 |  92.31 |   0.57 |   3.84 |
------------+--------+--------+--------+--------+--------+
TOTAL             69       93       13      174      443      792
                8.71    11.74     1.64    21.97    55.93   100.00
```

TABLE 2-7

EPA DATA

(3 WAY CAT DATA EXCLUDED)

```
                    TABLE OF CYL BY MY

CYL           MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |75       |76       |77       |78       |79       |  TOTAL
---------+---------+---------+---------+---------+---------+
  4      |      30 |      29 |       1 |       2 |       5 |     67
         |   15.63 |   15.10 |    0.52 |    1.04 |    2.60 |  34.90
         |   44.78 |   43.28 |    1.49 |    2.99 |    7.46 |
         |   46.15 |   31.52 |    7.69 |  100.00 |   25.00 |
---------+---------+---------+---------+---------+---------+
  6      |       8 |      16 |       1 |       0 |       6 |     31
         |    4.17 |    8.33 |    0.52 |    0.00 |    3.13 |  16.15
         |   25.81 |   51.61 |    3.23 |    0.00 |   19.35 |
         |   12.31 |   17.39 |    7.69 |    0.00 |   30.00 |
---------+---------+---------+---------+---------+---------+
  8      |      27 |      47 |      11 |       0 |       9 |     94
         |   14.06 |   24.48 |    5.73 |    0.00 |    4.69 |  48.96
         |   28.72 |   50.00 |   11.70 |    0.00 |    9.57 |
         |   41.54 |   51.09 |   84.62 |    0.00 |   45.00 |
---------+---------+---------+---------+---------+---------+
TOTAL           65        92        13         2        20      192
             33.85     47.92      6.77      1.04     10.42   100.00
```

consistency and bear no relations to sales trends. The same observations are valid for the transmission categories displayed in Table 2-8. The EPA data is notable in that manual transmissions are distinguished by the number of forward gears; nevertheless, the time trends observed in the ARB data are not evident in the EPA data.

Figures 2-4 and 2-5 display histograms for odometer readings for 1975-1976 and 1977-1979 model year cars respectively. The EPA data represents a higher mileage distribution of vehicles than the ARB data; the median value in Figure 2-4 occurs in the 55,000 to 65,000 mile bin as opposed to the 35,000 to 45,000 mile bin for the ARB data. Over 50 percent of the EPA data has an odometer reading of over 55,000 miles. Figure 2-5 shows an almost polar distribution of odometer readings with a concentration at low mileages and high mileages as compared to the ARB data. The median values occur within the same mileage bin of 25,000 to 35,000 miles. The EPA data however has a substantially greater percentage of vehicles with odometer readings of over 50,000 miles.

2.3 DEVELOPMENT OF EMISSION FACTORS FOR ARB AND EPA DATA SETS

All emission factors presented in this section will be based on linear regressions employing the following model:

$$FTP_i = a_i + b_i X$$

where X represents for accumulated mileage in units of tens of thousands of miles and the index i successively represents each of the pollutants: HC, CO, NO_x and evaporative emissions. This model assumes that the intercept "a" represents the new vehicle emission rate and the slope "b" represents the deterioration rate (per 10,000 miles) in the performance of the emission control system.

Table 2-9 presents a comparison of the emission factors derived from the first through the fifth surveillance data sets (published by ARB) and the

TABLE 2-8

```
                       EPA DATA

               (3 WAY CAT DATA EXCLUDED)

                 TABLE OF TRANS BY MY

TRANS       MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |75      |76      |77      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+--------+
 AUTO    |     43 |     74 |     12 |      1 |     15 |   145
         |  22.40 |  38.54 |   6.25 |   0.52 |   7.81 |  75.52
         |  29.66 |  51.03 |   8.28 |   0.69 |  10.34 |
         |  66.15 |  80.43 |  92.31 |  50.00 |  75.00 |
---------+--------+--------+--------+--------+--------+
 3 SP    |      0 |      1 |      0 |      0 |      0 |     1
         |   0.00 |   0.52 |   0.00 |   0.00 |   0.00 |  0.52
         |   0.00 | 100.00 |   0.00 |   0.00 |   0.00 |
         |   0.00 |   1.09 |   0.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+--------+
 4 SP    |     17 |     12 |      0 |      1 |      4 |    34
         |   8.85 |   6.25 |   0.00 |   0.52 |   2.08 | 17.71
         |  50.00 |  35.29 |   0.00 |   2.94 |  11.76 |
         |  26.15 |  13.04 |   0.00 |  50.00 |  20.00 |
---------+--------+--------+--------+--------+--------+
 5 SP    |      5 |      5 |      1 |      0 |      1 |    12
         |   2.60 |   2.60 |   0.52 |   0.00 |   0.52 |  6.25
         |  41.67 |  41.67 |   8.33 |   0.00 |   8.33 |
         |   7.69 |   5.43 |   7.69 |   0.00 |   5.00 |
---------+--------+--------+--------+--------+--------+
TOTAL          65       92       13        2       20     192
            33.85    47.92     6.77     1.04    10.42  100.00
```

FIGURE 2-4

```
                                    EPA DATA

                           (3 WAY CAT DATA EXCLUDED)
                                CARGRP=MY 75-76

                              FREQUENCY BAR CHART
MIDPOINT
  VMTX                                                                           FREQ   CUM.   PERCENT      CUM.
                                                                                         FREQ             PERCENT
        |
    0   |****************************************                                   20     20     12.74     12.74
        |
    1   |**********************************************************                 29     49     18.47     31.21
        |
    2   |**************************                                                 13     62      8.28     39.49
        |
    3   |******                                                                      3     65      1.91     41.40
        |
    4   |**                                                                          1     66      0.64     42.04
        |
    5   |********************************************************                   28     94     17.83     59.87
        |
    6   |************************************************************************   36    130     22.93     82.80
        |
    7   |**************************                                                 13    143      8.28     91.08
        |
    8   |****************                                                            8    151      5.10     96.18
        |
    9   |******                                                                      3    154      1.91     98.09
        |
   10   |**                                                                          1    155      0.64     98.73
        |
   11   |**                                                                          1    156      0.64     99.36
        |
   12   |**                                                                          1    157      0.64    100.00
        |
        ----+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
               2   4   6   8  10  12  14  16  18  20  22  24  26  28  30  32  34  36

                                    FREQUENCY
```

FIGURE 2-5

EPA DATA

(3 WAY CAT DATA EXCLUDED)
CARGRP=MY 77-79

```
                                FREQUENCY BAR CHART
MIDPOINT
  VMTX                                                               FREQ  CUM.  PERCENT    CUM.
                                                                           FREQ            PERCENT
        |
     0  |************************************************************  12    12    34.29     34.29
        |
     1  |*************************                                      5    17    14.29     48.57
        |
     2  |***************                                                3    20     8.57     57.14
        |
     3  |                                                               0    20     0.00     57.14
        |
     4  |*****                                                          1    21     2.86     60.00
        |
     5  |***********************************                            7    28    20.00     80.00
        |
     6  |***************                                                3    31     8.57     88.57
        |
     7  |********************                                           4    35    11.43    100.00
        |
     8  |                                                               0    35     0.00    100.00
        |
     9  |                                                               0    35     0.00    100.00
        |
    10  |                                                               0    35     0.00    100.00
        |
    11  |                                                               0    35     0.00    100.00
        |
    12  |                                                               0    35     0.00    100.00
        |
        -----+----+----+----+----+----+----+----+----+----+----+----+
             1    2    3    4    5    6    7    8    9   10   11   12

                                 FREQUENCY
```

TABLE 2-9

COMPARISON OF EMISSION FACTORS DERIVED FROM ARB DATA FOR PASSENGER CARS

		Based on Data from the First Through Fifth Surveillance Programs*		Calculated from Data from the First Through Sixth Surveillance Programs	
Pollutant	Model Year	New Vehicle Emission Rate (gm/mile)	Deterioration Rate Per 10,000 Miles (gm/mile)	New Vehicle Emission Rate (gm/mile)	Deterioration Rate Per 10,000 Miles (gm/mile)
HC	1972-1974	3.020	0.170	2.659	0.314
	1975-1976	0.480	0.350	0.624	0.231
	1977-1979	0.280	0.210	0.298	0.184
CO	1972-1974	33.290	2.440	42.383	1.477
	1975-1976	7.730	2.780	7.446	2.690
	1977-1979	4.320	2.220	4.623	2.124
NO_x	1971-1973	3.400	0.040	2.869**	0.018**
	1974	2.200	0.040	2.223	0.112
	1975-1976	2.030	0.060	2.101	0.053
	1977-1979	1.510	0.070	1.477	0.081

*Presented in a Memorandum from Bob Cross to Rich Bradley, Subject: Emission Factors 2nd Revision Dated November 30, 1982.

**Based on data from 1972-1973 only.

first through the sixth surveillance data sets, calculated by EEA. For convenience these data sets will be referred to as data set 5 (first through fifth) and data set 6 (first through sixth). The information is organized by the applicable pollutant standard and model year. While there is general agreement in the results from the different data sets, there remain some notable differences. For 1972 to 1974 vehicles there is a substantial increase in the magnitude of the deterioration rate for HC between data sets 5 and 6; this is offset somewhat by the decline in the magnitude of the intercept. Nevertheless, this implies a 23 percent increase in HC emissions over 100,000 miles of operation (termed lifetime emissions).

At first glance it appears that there is a substantial reduction in CO emissions for 1972-1974 vehicles because of the sharp decline in the deterioration rate. However, the offsetting increase in the intercept leads to essentially no change in lifetime CO emissions for these vehicles. A similar shift appears for HC emissions for 1975-1976 vehicles, however due to the reduction in the slope there is almost a 50 percent reduction in lifetime HC emissions for these vehicles.

With the exception of the 1971-1974 vehicles there is essentially no change in NO_x emission factors. Both the 1971-1973 and the 1974 NO_x relations indicate a substantial increase in the slope or deterioration rate. The addition of the sixth surveillance program data appears to cause an increase in deterioration rates for both HC and NO_x in the older vehicles.

A similar comparison of EPA derived emission factors is presented in Table 2-10. The California emission factors employed in MOBILE2 are listed for comparison with the factors that EEA derived from the EPA data. As previously discussed the MOBILE2 factors are based on a mixture of EPA and ARB data. In addition, EPA incorporated a misfueling rate of 8 percent into their emission factors. The influence of the misfueling rate on emissions over the life of the vehicle was calculated and added to the emission factor intercept derived from the data. Therefore, the intercepts compared

TABLE 2-10

COMPARISON OF EMISSION FACTORS DERIVED FROM EPA DATA FOR PASSENGER CARS

		Emission Factor Presented in MOBILE2*		Calculated from EPA Data	
Pollutant	Model Year	New Vehicle Emission Rate (gm/mile)	Deterioration Rate Per 10,000 Miles (gm/mile)	New Vehicle Emission Rate (gm/mile)	Deterioration Rate Per 10,000 Miles (gm/mile)
HC	1972-1974	3.02	0.17	-	-
	1975-1976	0.55	0.35	0.352	0.277
	1977-1979	0.32	0.21	0.671	0.037
CO	1972-1974	33.29	2.44	-	-
	1975-1976	8.39	2.78	5.021	3.549
	1977-1979	4.68	2.22	9.895	0.137
NO_x	1971-1973	3.40	0.04	-	-
	1974	2.20	0.04	-	-
	1975-1976	2.03	0.06	1.634	0.123
	1977-1979	1.51	0.07	1.357	0.105

*"User's Guide to Mobile 2 (Mobile Source Emissions Model)," U.S. Environmental Protection Agency, February 1981.

in Table 2-10 should not be considered comparable; the MOBILE2 values are generally greater than those calculated by EEA. Because of the disproportionate share of the three-way catalyst vehicles contained in the EPA data, they have been excluded from the data set used by EEA to calculate an EPA emission factor. While this only affects the 1977-1979 category of vehicles it nevertheless adds an additional complication to the comparison.

The range of caveats outlined for the comparison of the emission factors displayed in this table makes it clear that little time should be spent examining the information presented. The foregoing notwithstanding, there is surprising agreement in the factors presented. Generally, the EEA calculated regressions exhibit lower intercepts as expected but higher deterioration rates. The poor technology representation of the EPA data set may have a substantial influence on differences in the deterioration rates.

Up to this point, all of the comparisons have focused on the differences between previous emission factors and those developed from the data assembled by EPA and ARB, Table 2-11 presents a detailed comparison of regression statistics for the EPA and ARB calculations. The most striking feature of this comparison is the difference in the sample sizes available for analysis. The elimination of approximately 600 three-way catalyst vehicles does have an impact on this comparison but their disproportionate representation would only bias the results. Owing to the large sample size, the ARB parameter estimates are uniformly significant.

The ARB deterioration rates for all of the 1975-1976 pollutants are lower than comparable EPA estimates; by the same token the intercepts for the ARB regressions are uniformly higher. Nevertheless, the ARB emission factors exhibit lower lifetime emissions for each of the pollutants by 7, 17, and 20 percent for HC, CO, and NO_x respectively.

The small sample size of the 1977-1979 EPA data set produces coefficients of dubious significance for the slopes of HC and CO. A comparison of the

TABLE 2-11

COMPARISON OF REGRESSION STATISTICS FOR AGGREGATE EMISSION FACTORS DERIVED FROM ARB & EPA DATA SETS

		EPA				ARB			
Model Year/ Pollutant		Parameter Estimate	T Value	Prob > /T/	# of observations	Parameter Estimate	T Value	Prob > /T/	# of observations
1975-1976									
HC	Intercept	0.352	2.171	0.0315	157	0.624	2.644	0.0085	434
	VMTX	0.277	8.524	0.0001		0.231	3.474	0.0006	
CO	Intercept	5.020	1.700	0.0912	157	7.446	3.958	0.0001	434
	VMTX	3.549	6.003	0.0001		2.690	5.069	0.0001	
NO_x	Intercept	1.634	9.903	0.0001	157	2.101	20.176	0.0001	434
	VMTX	0.123	3.711	0.0003		0.053	1.805	0.0718	
1977-1979									
HC	Intercept	0.671	3.243	0.0027	35	0.298	3.970	0.0001	411
	VMTX	0.037	0.694	0.4927		0.183	6.888	0.0001	
CO	Intercept	9.895	3.527	0.0013	35	4.623	3.159	0.0017	411
	VMTX	0.137	0.190	0.8506		2.124	4.086	0.0001	
NO_x	Intercept	1.357	7.040	0.0001	35	1.477	16.002	0.0001	411
	VMTX	0.105	2.113	0.0423		0.081	2.467	0.0140	

EPA and ARB relations for this category is not useful. The sample size of the EPA data must be increased before such a comparison is worthwhile. One approach is to find a way to incorporate the three-way catalyst data.

The approach selected is to integrate the EPA and ARB data sets together and to evaluate the performance of each of the technology categories. This approach is based on the belief that dissimilar emission control systems are likely to have substantial differences in deterioration rates because of the differences in the catalyst employed. Therefore, regressions were run for each of the technology categories and model year groups.

Table 2-12 presents the results of the regression analysis for 1975-1976 vehicles. Only three technology categories were in-use at that time, three-way catalyst vehicles were not introduced into California until the 1977 model year. As shown in the table, this approach reduces the sample size by disaggregating the data across the technology categories, however, the combination of the EPA and ARB data mitigates the effect. The small sample size for No Catalyst and Catalyst Only vehicles leads to some non-significant parameter estimates. Overall, the data distribution across the technology categories is in line with that shown in the ARB data; most of the vehicles had catalysts with air pumps.

Numerous comparisons across the technology categories can be made. The reader is left to make many of these because most of the parameter estimates are comparable. The CO estimate for the Catalyst Only vehicle does stand out because of its unusually high slope; it is twice as great as the Catalyst with Air and five times as great as the No Catalyst vehicles. The cause for this is not apparent but is addressed later in this section.

Table 2-13 presents a similar set of regression statistics for the 1977-1979 vehicles. Here EPA's large volume of three-way catalyst data is finally used. Again, the small sample sizes for the No Catalyst and Catalyst Only vehicles reduces the confidence of several of the coefficients calculated

TABLE 2-12

REGRESSION STATISTICS FOR EMISSION FACTORS FOR SEPARATE TECHNOLOGY CATEGORIES DERIVED FROM THE COMBINED ARB & EPA DATA SETS

Model Years 1975-1976

Technology Category	Pollutant		Parameter Estimate	T Value	Prob > /T/	# of Observations
No Catalyst	HC	Intercept	0.894	4.392	0.0001	61
		VMTX	0.104	2.105	0.0396	
	CO	Intercept	9.306	3.801	0.0003	61
		VMTX	1.013	1.703	0.0938	
	NO_x	Intercept	2.039	9.839	0.0001	61
		VMTX	0.001	0.021	0.9837	
Catalyst Only	HC	Intercept	0.749	1.887	0.0632	74
		VMTX	0.345	3.109	0.0027	
	CO	Intercept	7.083	1.193	0.2367	74
		VMTX	6.453	3.886	0.0002	
	NO_x	Intercept	2.125	6.785	0.0001	74
		VMTX	0.047	0.534	0.5949	
Catalyst With Air	HC	Intercept	0.460	2.226	0.0265	456
		VMTX	0.256	4.989	0.0001	
	CO	Intercept	5.277	3.156	0.0017	456
		VMTX	3.081	7.420	0.0001	
	NO_x	Intercept	1.961	20.490	0.0001	456
		VMTX	0.087	3.642	0.0003	

TABLE 2-13

REGRESSION STATISTICS FOR EMISSION FACTORS FOR SEPARATE TECHNOLOGY CATEGORIES DERIVED FROM THE COMBINED ARB & EPA DATA SETS

Model Years 1977-1979

Technology Category	Pollutant		Parameter Estimate	T Value	Prob > /T/	# of Observations
Three-way Catalyst	HC	Intercept	0.517	6.918	0.0001	595
		VMTX	0.146	2.930	0.0035	
	CO	Intercept	9.555	6.955	0.0001	595
		VMTX	0.459	0.503	0.0154	
	NO_x	Intercept	1.041	23.377	0.0001	595
		VMTX	0.051	1.736	0.0831	
No Catalyst	HC	Intercept	0.485	1.743	0.0896	39
		VMTX	0.173	2.054	0.471	
	CO	Intercept	10.177	2.294	0.0276	39
		VMTX	1.684	1.250	0.2191	
	NO_x	Intercept	1.111	6.218	0.0001	39
		VMTX	0.135	2.494	0.0172	
Catalyst Only	HC	Intercept	0.702	2.396	0.222	36
		VMTX	0.195	1.987	0.0550	
	CO	Intercept	11.715	1.578	0.1234	36
		VMTX	2.988	1.201	0.2379	
	NO_x	Intercept	1.467	3.449	0.0015	36
		VMTX	0.145	1.020	0.3150	
Catalyst With Air	HC	Intercept	0.350	4.743	0.0001	361
		VMTX	0.136	5.269	0.0001	
	CO	Intercept	5.162	3.879	0.0001	361
		VMTX	1.276	2.743	0.0064	
	NO_x	Intercept	1.505	16.523	0.0001	361
		VMTX	0.075	2.364	0.0186	

for these technology categories. Generally, the results are similar across the technology categories. The Catalyst Only category is again notable due to the size of its slope for CO, which is roughly double the value of the next highest technology category. The primary cause for this increase must be considered the rich drift (i.e., carburetor malperformance). This is not considered a failure but occurs in all open loop carburetors to a varying degree. This issue will be addressed in detail in a later section of this report.

Generally, the Three-way Catalyst and the Catalyst With Air exhibit lower levels of pollutants than the other two technologies. The feedback control of the three- way systems appears to perform well under in-use conditions. However, a review of Appendix A indicates that these vehicles have very low mileages, over 70 percent of these vehicles had odometer readings under 15,000 miles. Therefore, the influence of malperformances is unlikely to be seen in this data. The Catalyst with Air represents a more thoroughly tested system and appears to perform well under in-use conditions.

Since the emission factors calculated by technology category are unlikely to be useful to emission inventory modelers, the technology distribution presented in the ARB surveillance data will be used to weight the individual, technology-specific emission factors into aggregate emission factors. Table 2-14 displays the percent distribution by technology and model year grouping. As previously discussed, Catalyst with Air represents the dominant category. Table 2-15 contains the results from weighting the technology categories together. When compared with the ARB emission factors presented in Table 2-9 the results are generally favorable (i.e., lower predicted in-use emission levels) with the exception of CO for 1975-1976. There is a substantial increase in the deterioration rate when compared with the ARB results.

An increase in the slope for CO for Catalyst Only vehicles was noted earlier. Table 2-16 presents a detailed comparison of ARB and EPA technology specific emission factors for CO for 1975-1976 model year cars. It is

TABLE 2-14

PERCENT DISTRIBUTION OF TECHNOLOGIES EXHIBITED IN THE ARB SUVEILLANCE DATA ORGANIZED BY EMISSION STANDARD

Technology Category	1975-1976	1977-1979
Three-Way Catalyst	-	2.4
No Catalyst	11.5	8.3
Catalyst Only	17.2	8.7
Catalyst with Air	71.3	80.6

TABLE 2-15

EMISSION FACTORS DERIVED THROUGH THE WEIGHTING OF PARAMETER ESTIMATES FROM INDIVIDUAL TECHNOLOGY CATEGORIES BASED ON THE COMBINED ARB AND EPA DATA SETS

Model Year	Pollutant	New Vehicle Emission Rate	Deterioration Rate (Per 10,000 Miles)
1975-1976	HC	.56	.25
	CO	6.05	3.42
	NO_x	2.00	.07
1977-1979	HC	.40	.14
	CO	6.25	1.44
	NO_x	1.46	.09

TABLE 2-16

COMPARISON OF EPA AND ARB TECHNOLOGY SPECIFIC CO EMISSION FACTORS FOR 1975-1976 MODEL YEAR CARS

		ARB			EPA			ARB & EPA		
	LDT	N	a	b	N	a	b	N	a	b
No Catalyst	0.115	50	9.75	0.744	12	17.0	-0.03	61	9.31	1.01
Catalyst	0.172	75	7.08	6.45	0	-	-	74	7.08	6.45
Catalyst + Air	0.713	311	7.05	2.10	145	4.96	3.77	456	5.28	3.08
	1.000	436	7.37	2.69	157	5.02	3.55	591	6.05	3.42

difficult to make relevant comparisons between the EPA and ARB categories due to the paucity of EPA data. The only relevant comparison can be made between Catalyst With Air vehicles. In that case, the EPA data shows a substantially higher slope than the ARB value. That EPA technology category appears to be responsible for the increase in the deterioration rate for CO reported for the combined data sets.

The determination of the appropriateness of the EPA data included in that technology category is considered to be outside the scope of this task. The evaluation of the representativeness of the data could not be exclusively reviewed for CO but rather must be reviewed for all of the technology categories. The relatively poor manufacturer representation in the EPA data could possibly bias the representativeness of any of the EPA technology categories. If sufficient differences exist it might be necessary to exclude the EPA data from the analysis altogether. Because of the funding limitations of this task and the general agreement of the results between the ARB emission factors and the combined ARB and EPA factors, the detailed analysis of EPA's technology representativeness has not been pursued.

3. EMISSION FACTOR NON-LINEARITY

3.1 INTRODUCTION

Linear emission factor equations are the accepted method in emissions inventory modeling to project fleet-average emissions rates in future years. Because emission levels are known to increase with vehicle age, some form of deterioration must be represented in the emission factor equation. The linear form incorporates a deterioration rate (gm/mi increase per 10,000 accumulated miles) that, from a statistical viewpoint, is the simplest assumption satisfying this requirement. There is, however, no engineering basis for the linear increase, and in fact, in-use emissions performance is a far more complex process.

The ARB surveillance data for model year 1975-1979 vehicles are distributed from (approximately) 10,000 to 70,000 miles. Mean accumulated mileage for the 1975-1976 subset is 37,000 miles; for the 1977-1979 subset, 27,000 miles. The data are sparsely populated beyond 50,000 miles, however, and virtually no FTP test data are available on vehicles with more than 100,000 accumulated miles. The accuracy of linear emission factors at high mileage is thus an open question. Since mobile source models track vehicles through 20 years of operation (although with declining travel levels), a modest error in prediction of emissions beyond 50,000 miles could lead to substantial errors in projected inventories.

The engineering viewpoint is that average emission levels are likely to exhibit a declining rate of increase and may effectively plateau at very high mileages. As a physical process, emissions levels cannot increase without bound; engineering analysis can establish upper limits for exhaust emissions, although these are very high compared to average emissions from high-mileage vehicles (e.g., 200+ versus 25 gm/mi CO). An understanding of the causes of deterioration in well-maintained vehicles supports the perception that emissions deterioration between 50,000 and 100,000 miles should be less than in the first 50,000 miles. Data on malperformances of in-use vehicles also suggest that post-50,000 mile deterioration rates should be lower. These considerations also indicate, however, that much of the non-linearity occurs at low mileage and that high-mileage emission rates may deviate substantially from linear emission factors only in the post-100,000 mile interval.

At present, some 150 California vehicles of models years 1975-1979 with 50,000 or more accumulated miles have been tested in the ARB surveillance programs. In comparison to linear emission factors, the current data suggest possible non-linearity but do not provide the conclusive statistical evidence needed to overturn conventional emission factor techniques. A bounding analysis indicates that the non-linearity probably represents no more than a 10 percent reduction in predicted lifetime (100,000 miles) emissions. Data requirements to substantiate these findings are sufficiently high that special testing of high-mileage vehicles in the ARB surveillance programs appears not to be a cost-effective avenue for future research.

The study of high-mileage issues was conducted in two parts. Section 3.2 reviews the engineering basis for expecting non-linear emissions behavior and subjects the surveillance data to statistical analyses. Since much of the results are inconclusive, a second analysis was conducted to establish statistical bounds on the possible magnitude of non-linearities. This bounding analysis is reported in Section 3.3.

3.2 HIGH-MILEAGE EMISSIONS

The ARB in-use surveillance data from projects 1 through 6 have been examined to determine if linear emission factors are appropriate and/or adequate to characterize emissions from high-mileage vehicles. The implication of appropriate is whether deterioration rates are expected to be constant as assumed in the conventional emission factor. Even if inappropriate, linear factors may nevertheless be adequate should the departure from linearity introduce a small error of tolerable magnitude.

The engineering expectation, supported by data discussed here, is that deterioration rates should decline with mileage, so that linear factors are likely to overstate emissions from high-mileage vehicles. Automotive catalysts experience a rigid decline in conversion efficiency during the first several thousand miles of operation, leading to a low-mileage emissions ramp. Idle-mixture maladjustment rates for the California fleet similarly increase rapidly at low-mileage, but then display little trends with further mileage accumulation. Taken together, these suggest that high-mileage deterioration rates should be lower than predicted by linear emission factors. If this position can be demonstrated conclusively with in-use data, alternative emission factor representations could be employed in emission inventory modeling. For example, a deterioration rate proportional to the logarithm of mileage could be an acceptable functional form from a modeling viewpoint, provided that the declining deterioration rate is first established. Both the engineering expectations and the current statistical evidence on non-linear emissions behavior are examined in this section.

3.2.1 Causes of Emissions Deterioration

Well-Maintained Vehicles

The most representative data on emissions deterioration in well-maintained vehicles are provided in a carefully designed study[1/] sponsored by the Coordinating Research Council (CRC). Although the study's purpose was to determine the emissions effects of an anti-knock fuel additive MMT, baseline data covering 50,000 miles accumulated on MMT-free fuel provide estimates of normal deterioration in oxidation and (single-bed) three-way catalyst systems.

[1/] Coordinating Research Council. June 1979. MMT Field Test Program, Report No. 503.

Included in the study were production vehicles built for sale in the California market under the 1977-1979 exhaust emission standards. Three individual vehicles were selected for testing with MMT-free fuel for each of seven carlines -- five oxidation catalyst and two three-way catalysts. They were given maintenance at intervals specified by the manufacturer and driven on fuels representative of commercial gasolines to accumulate mileage according to a schedule similar to the EPA durability cycle. A major innovation of the CRC study was modification of CVS testing methods to simultaneously monitor engine-out and tailpipe emissions over the FTP. Therefore, the data establish a basis for separating increases in engine emissions from deterioration in catalyst efficiency.

The test results indicate clearly that progressive reduction in catalyst efficiency is the major source of tailpipe emissions increases. Engine-out emissions for HC, CO, and NO_x showed no appreciable increase with mileage for MMT-free fuels. Engine-out NO_x emissions exhibit, in fact, a small decline with accumulated mileage that is to be expected from a loss of engine compression through progressive wear on rings and valves. The data on conversion efficiency for oxidation catalyst systems are shown in Figure 3-1. Approximately half the 50,000 mile loss in efficiency occurs during the initial 5,000 to 10,000 miles. Deterioration rates slow markedly beyond this point and appear to follow a linear decline through at least 50,000 miles.

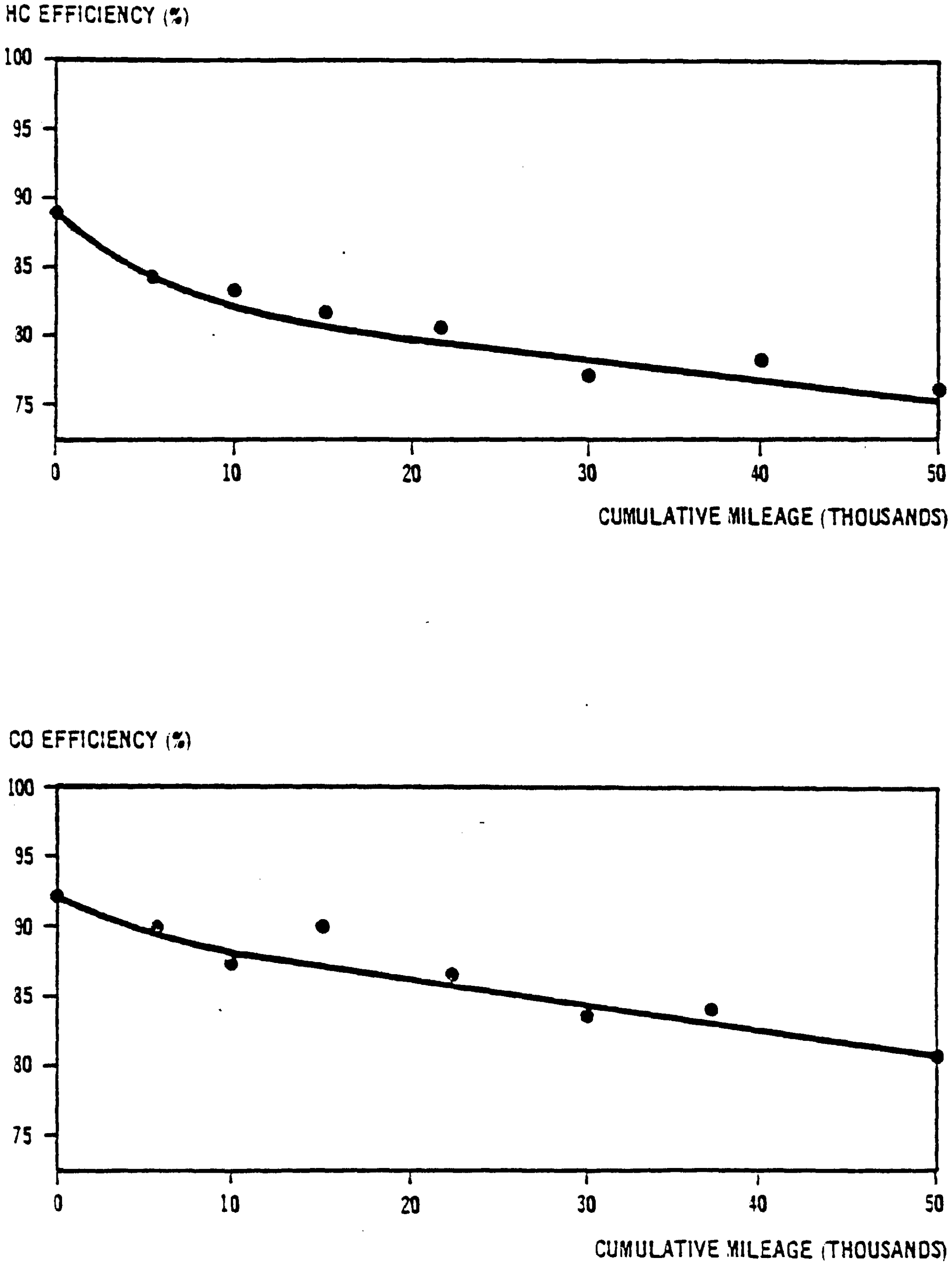
The strong low-mileage non-linearity is confirmed by a second study[2/] that examined the sensitivity of peak catalyst efficiencies to trace levels of lead. The data shown in Figure 3-2 were generated by artificially aging three-way catalysts to the equivalent of 25,000 miles in laboratory testing. The measured efficiencies at 500°C (932°F) indicate little or no deterioration through 15,000 miles on lead-free fuels. Addition of trace levels of

---

[2/] Williamson, W.B., Gandhi, H.S., et al (Ford Motor Company). "Deactivations of Three-Way Catalysts by Fuel Contaminents: Lead, Phosphorus, and Sulfur." SAE Technical Paper Series, #790942.

FIGURE 3-1

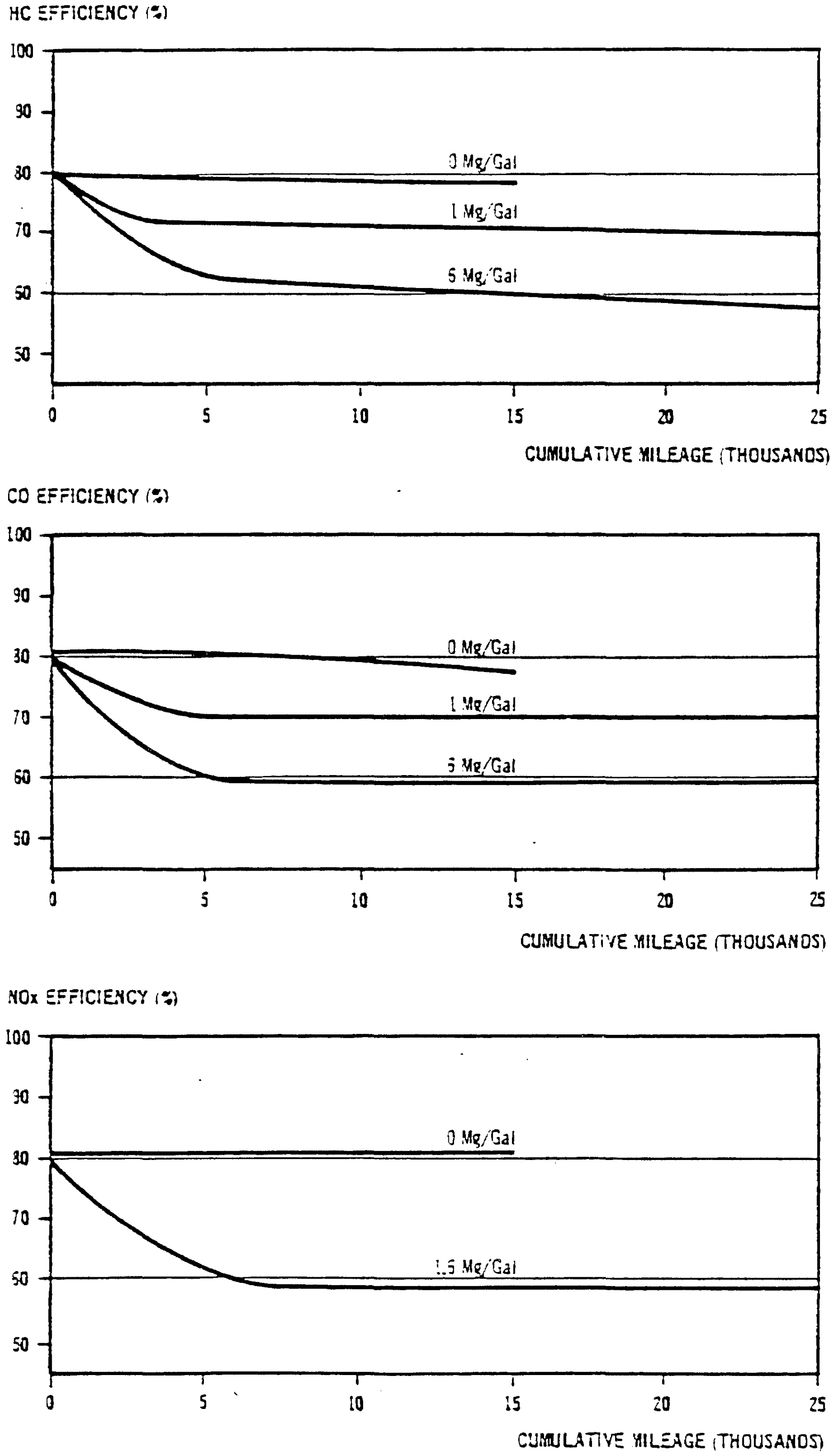
OXIDATION CATALYST EFFICIENCIES



Source: CRC MMT Field Test Program, Report No. 503

FIGURE 3-2

EFFECT OF RESIDUAL LEAD ON PEAK CATALYST EFFICIENCIES



Source: SAE Technical Paper Series, No. 730942 lead (lower curves) caused significant reductions in activity, resulting from the build-up of lead and lead oxide deposits on the active catalystic surface area.

Catalyst efficiency over a transient driving cycle is influenced by the time delay until "light-off," the onset on conversion reactions. Aged catalysts tend to light-off later, leading to decreasing conversion efficiencies as seen on the CRC data. This phenomenon is not reflected in the steady-state data of the second study.

Taken as a whole, the engineering data indicate strong non-linearities in emissions deterioration for well-maintained vehicles. These occur at low mileages (below 10,000 miles) as catalysts are exposed to commercial fuels with trace lead contaminants. Emissions deterioration slows at high mileages and, at least through 50,000 miles, appears to follow a linear trend. However, no comparable data exist to validate a continued linear deterioration in the 50,000 - 100,000 mile interval.

Average In-Use Vehicles

The average emissions deterioration rate observed for in-use vehicles is strongly influenced by trends in the rates of occurrence for emissions malperformances. These tend to increase engine-out emissions with minimal effect on catalyst efficiencies, although air pump failure/tampering clearly impacts catalyst performance. For an individual vehicle, a malperformance has an immediate impact on tailpipe emission levels and, depending on its effect on catalyst durability, may change the vehicle's deterioration rate. Averaged over the vehicle fleet, malperformances will increase the in-use emission factor -- both intercept and slope -- by amounts that depend on the emissions impacts and the prevalence in the fleet. If the prevalence increases with mileage, malperformances will increase average deterioration rates even though they may have no effect on catalyst durability.

While failure of emission control hardware contributes to excess in-use emissions, the most frequently occurring malperformances are closely related to maintenance practices or to tampering. While accumulated mileage is the conventional proxy variable, the frequency probably is tied more directly to the vehicle life-cycle -- i.e., malperformance frequencies may change when vehicles move out of the warranty period or into the hands of second owners, regardless of mileage. There is no direct relationship to accumulated miles (as there is for physical deterioration such as engine wear) that justifies an assumption that malperformance frequencies increase in proportion to fleet-average mileage.

Maladjustments of the idle air-fuel ratio have been identified by several studies as the major cause of excess emissions of HC and CO for pre-1980 vehicles.[3,4/] Mileage trends in the rate of occurrence for this malperformance should therefore be reflected in the in-use emission factor. Idle CO concentrations may be used as a surrogate variable for the presence and extent of maladjustment. The ARB surveillance data also provide independent diagnostic comments on whether the mixture adjustment was within manufacturer specification.

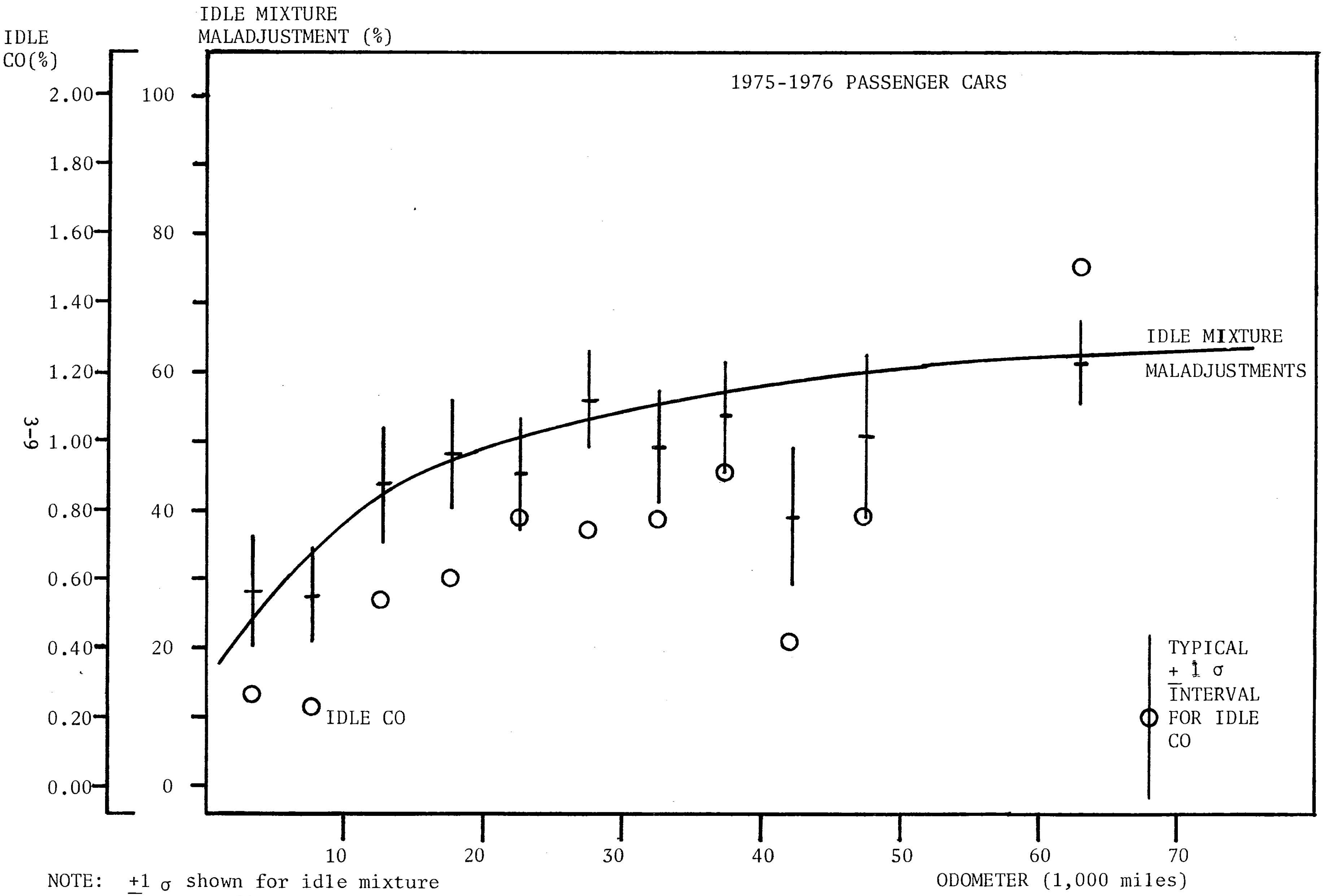
Figure 3-3 shows the trend with accumulated miles in mean idle CO levels and the frequency of mixture maladjustments indicated by diagnostic comments. These data for 1975-1976 vehicles were compiled by grouping vehicles into 5,000 mile intervals below 50,000 miles, while aggregating higher-mileage vehicles into a single group. Although not shown, the data for 1977- 1979 vehicles are substantially the same. As indicated by the smooth curve, the frequency of mixture maladjustments follows a decidely non-linear increase with mileage that is in agreement with the observed trends in idle CO levels. Both measures increase sharply at 10,000 to 15,000 miles and display little increase above this point. (Counterbalancing this is the suggestion that the extent of maladjustment, as indicated by mean idle CO

[3/] Becher, and Rutherford, "Analysis of Oregon's Inspection and Maintenance Program," APCA Paper No. 79-7.3, June 1979.

[4/] Cackette, Lorang, and Hughes, "The Need for Inspection and Maintenance for Current and Future Motor Vehicles. SAE Technical Paper Series, #790782.

FIGURE 3-3

TRENDS IN IDLE MIXTURE MALADJUSTMENTS AND IDLE CO LEVELS



NOTE: ±1 σ shown for idle mixture maladjustments levels, may be relatively greater for post-50,000 mile vehicles, although the limited high-mileage sample precludes a firm determination.)

In combination, the engineering data on well-maintained vehicles and the surveillance data for the dominant malperformance of in-use cars support the expectation that emissions deterioration occurs at a rapid rate over the first 10,000 to 15,000 miles and at a slower rate thereafter. While little data is available on vehicles above 50,000 miles, it appears that emissions deterioration may continue at the lower rate established over 20,000 to 50,000 miles or, as seen for idle mixture maladjustments, may effectively plateau.

3.2.2 Emissions Characterization

Based on the previous discussion, it is likely that linear factors are inappropriate statistical models for in-use emissions. Yet, to be established, however, is whether FTP emissions data support this finding and whether the errors committed are of substantial magnitude. The usual statistical test for non-linear behavior is to estimate linear models augmented with a quadratic term in the independent variable. Significance of the quadratic term at an acceptable level of confidence is taken to be conclusive evidence of the behavior. Subsequent analysis may then use other non-linear forms (e.g., logarithmic functions) to avoid the parabolic curve (high-mileage downturn) forced by the quadratic equation.

For the ARB surveillance data on FTP emissions, linear and quadratic equations have been estimated for each exhaust pollutant (HC, CO, NO_x) in the model year groups 1975-1976 and 1977-1979. These emission factor equations are:

$$\text{Linear} \qquad FTP = a + bx$$

$$\text{Quadratic} \qquad FTP = a + bx + cx^2$$

where x = accumulated mileage

Negative values for coefficients c will indicate that deterioration rates decline with mileage.

The results of this analysis are summarized in Tables 3-1 and 3-2. The confidence (alpha) level, tabulated in parentheses below each quadratic coefficient estimate, is the probability that the observed quadratic effect could arise by random fluctuation in samples of the given size, even though the population under study was strictly linear. For example, the alpha level of 0.33 (HC equation for 1975-1976 vehicles) indicates that 1 of 3 samples from a linear population could evidence the observed quadratic effect. Acceptable alpha levels are generally taken to be 0.05 or 0.01.

In each instance, the quadrated term is found to have a negative coefficient, supporting the expectation of declining deterioration rates. The uniformity of the signs for the six quadratic coefficients is unlikely to arise by chance if there is in fact no non-linearity in emissions behavior. However, in no instance is the quadratic term statistically significant at an acceptable level of confidence. The alpha levels (0.25 and 0.33) for quadratic terms in the HC and CO equation of 1975-1976 vehicles are too high for the quadratic terms to be considered significant. The NO_x term for 1975-1976 (alpha = 0.12) should be considered marginally significant. Alpha levels are higher in the 1977-1979 equations.

On this basis, the results for the 1975-1976 subset may be taken as suggestive, but inconclusive, of non-linear behavior; the data for model years 1977-1979 (with fewer post-50,000 mile vehicles) yield no results pro or con on this issue. What has not been addressed is the statistical power of the samples, or the probability of detecting nonlinear behavior when actually present. The uniformity of quadratic signs suggests that non-linear behavior is present, but is too small to be detected with confidence in the existing data. The bounding analysis described in the following section returns to this point.

TABLE 3-1

NON-LINEAR EMISSION FACTORS FOR MY75-76 CARS

Pollutant		Emission Factor Eqn*	Lifetime Emissions (kg)**	Reduction From Linear Eqn
HC	Linear	$0.63 + 0.22X$	173	-
	Quadratic	$0.39 + 0.40X - 0.021X^2$ (0.33)	169	2.3%
CO	Linear	$7.34 + 2.90X$	2184	-
	Quadratic	$4.92 + 4.64 - 0.218X^2$ (0.25)	2085	4.5%
NO_x	Linear	$2.14 + 0.05X$	239	-
	Quadratic	$1.97 + 0.17X - 0.016X^2$ (0.12)	229	4.3%

---

* Significance (alpha level) of quadratic term in parentheses

**Assessed for 100,000 miles

TABLE 3-2

NON-LINEAR EMISSION FACTORS FOR MY77-79 CARS

Pollutant		Emission Factor Eqn*	Lifetime Emissions (kg)**	Reduction From Linear Eqn
HC	Linear	$0.39 + 0.17X$	124	-
	Quadratic	$0.31 + 0.23X - 0.008X^2$ (0.49)	119	3.8%
CO	Linear	$4.57 + 2.33X$	1622	-
	Quadratic	$3.39 + 3.22X - 0.121X^2$ (0.48)	1546	4.7%
NO_x	Linear	$1.47 + 0.08X$	187	-
	Quadratic	$1.46 + 0.08X - 0.001X^2$ (0.94)	183	2.3%

---

* Significance (alpha level) of quadratic term in parentheses

**Assessed for 100,000 miles

To evaluate the substantive difference that non-linear emission factors would make in lifetime emissions predictions, the total emissions (kg) of each pollutant over 100,000 miles have been calculated from both linear and quadratic equations. As seen in the tables, the quadratic factors represent very small (2-5 percent) reductions in predicted emissions compared to the linear equations. This results from both the small magnitudes of the coefficients and the greater linear slopes of the quadratic forms; while over-estimating high-mileage emissions, linear factors will also under-estimate in a low- to intermediate-mileage range if non-linearities are present. An illustration of this phenomenon may be seen in Figure 3-4 of Section 3.3.

A key comparison to draw in assessing the adequacy of linear factors is whether they result in substantial over-prediction of emissions for the existing high-mileage data. Through project 6, the ARB in-use surveillance program has tested 98 model year 1975-1976 and 53 model year 1977-1979 vehicles with more than 50,000 accumulated miles. These high-mileage subsets have mean odometers of 64,900 and 63,000 miles, respectively, and contain only handfuls of vehicles above 75,000 to 80,000 miles. They provide little information near 100,000 miles and no data beyond this point.

The comparison with predictions of the linear model is made in Table 3-3 by evaluating the linear emission factors at the mean odometers of the high-mileage subsets. As can be seen, mean emissions from the high-mileage subsets are within the standard errors of prediction for the linear equations. Thus the high-mileage surveillance data are indistinguishable from the emission factor lines. There also exists no consistent trend toward over-prediction by the linear equation. These null results indicate that, at least for the mileage range in existing surveillance data, linear factors appear to be fully adequate to characterize average in-use emissions.

3.2.3 Interpretation

A primary conclusion to be drawn from the foregoing discussion is that the divergence of in-use emissions from linear emission factor equations may

TABLE 3-3

COMPARISON OF PREDICTED AND OBSERVED EMISSION LEVELS

	Predicted Emissions (gm/mi) (Linear Eqn.) High Mileage Subset*		Observed Mean Emissions (gm/mi) (High-Mileage Subset)	
1975-1976				
HC	2.12	(0.22)	1.83	(0.15)
CO	26.15	(1.91)	25.04	(2.53)
NO_x	2.46	(0.10)	2.54	(0.13)
1977-1979				
HC	1.47	(0.11)	1.51	(0.21)
CO	19.70	(1.78)	18.59	(3.03)
NO_x	1.95	(0.10)	2.07	(0.17)

---

*Mean odometers (vehicles above 50,000 miles): 64,900 miles for 1975-1976; 63,000 miles for 1977-1979.

NOTE: Standard errors shown in parentheses largely be concentrated at very low mileages (below 10,000 miles) and at very high mileages (above 100,000 miles). Since most of the surveillance data lies between these extremes, emission factors are heavily weighted by data in an essentially linear region. Except for model years in which existing data are at very low-mileages, the predictions of linear factors may thus be adequate estimates of in-use emissions through approximately 70,000 miles. The absence of FTP data at or above 100,000 miles renders it impossible to formulate similar conclusions for higher mileages.

In spite of limitations, the surveillance data do support the engineering expectation that in-use emissions behavior is a non-linear process. However, the evidence is not sufficiently conclusive to overturn the conventional approach to emission factor analysis and introduce non-linear forms. It should be noted that the inconclusive aspect of the analysis implies (in the strictest sense) only that the data are inadequate to detect the non-linear behavior actually present. It does refute the expectation that non-linearities are present that could be substantial at higher mileages. However, the evidence suggests that the emissions reductions might be no more than 5 percent (through 100,000 miles) for non-linear modeling techniques.

3.3 BOUNDING ANALYSIS

Additional high-mileage FTP data will improve the ability of surveillance data sets to characterize in-use emissions. While these data will be added in future surveillance programs, it is unclear at present the priority to be assigned high-mileage issues in ARB's research and testing program. Assessment of the current data's adequacy to characterize high-mileage emissions requires further evaluation of the converse question to statistical significance: how large could the non-linear behavior be and yet remain undetected in existing data? The answer to this question establishes a bound on the potential benefits of additional research.

3.3.1 Conceptual Approach

Stated formally, the bounding analysis evaluates the risk of committing a Type II error in basing the test for significance of non-linear behavior on the existing surveillance data. The Type II error is that of failing to detect an effect of given magnitude that is thought to be present in the population under study. The associated risk is a function of: the sample size, the variance in the data, and the effect's expected magnitude. The risk is frequently considered in the design of sampling experiments, but often neglected in the analysis of experimental data. For cases involving differences in mean values, relatively straightforward calculations can be employed to estimate the required sample sizes (for experimental design) or to bound an effect's magnitude. The methods for evaluating Type II errors in regression analysis are more complicated because the statistical power of the existing sample also depends upon its distribution with the independent variable and the interval in which the non-linear behavior is most strongly apparent. Surveillance data sets concentrated in limited mileage intervals will have little power to resolve trends of any kind with mileage, as may data sets which fail to represent critical portions of the odometer range.

A simulation approach has been employed to estimate bounds on the non-linearity that may be present in the ARB surveillance data. In general terms, the simulation hypothesizes an equation obeyed by the population under study and draws from the modeled population repeated samples of pseudo-data. Variance that is characteristic of the actual population is superimposed on the samples so that each represents a possible outcome of a sampling experiment. Each sample is then analyzed for the non-linear phenomenon of interest and the probability of detecting the effect is tabulated across the ensemble of samples. In the long-run of many such repeated samples, the frequency with which the effect is undetected at a specified level of confidence will estimate the risk due to a Type II error. If the population equation is specified with a variable magnitude for the effect of interest, further simulations can establish an upper bound. This bound is the largest effect that may be present and yet likely remain undetected in a single sample.

The mechanics of the simulation analysis are as follows. Once the population equation and variance are determined, one hundred independent samples are drawn using the sample sizes and odometer distributions of the surveillance data. Each sample is therefore a possible outcome of testing in-use vehicles whose emissions are described by the population equation. Modeled CO emissions from each sample are analyzed using linear and quadratic regression models. The significance tests of the quadratic term are then tabulated across the 100 samples to estimate the likelihood of detecting non-linear behavior.

CO emissions have been chosen as a test case for the simulation study, since the largest non-linearities were estimated in the quadratic regressions for this pollutant. After review of alternative candidates, an inverse exponential growth function:

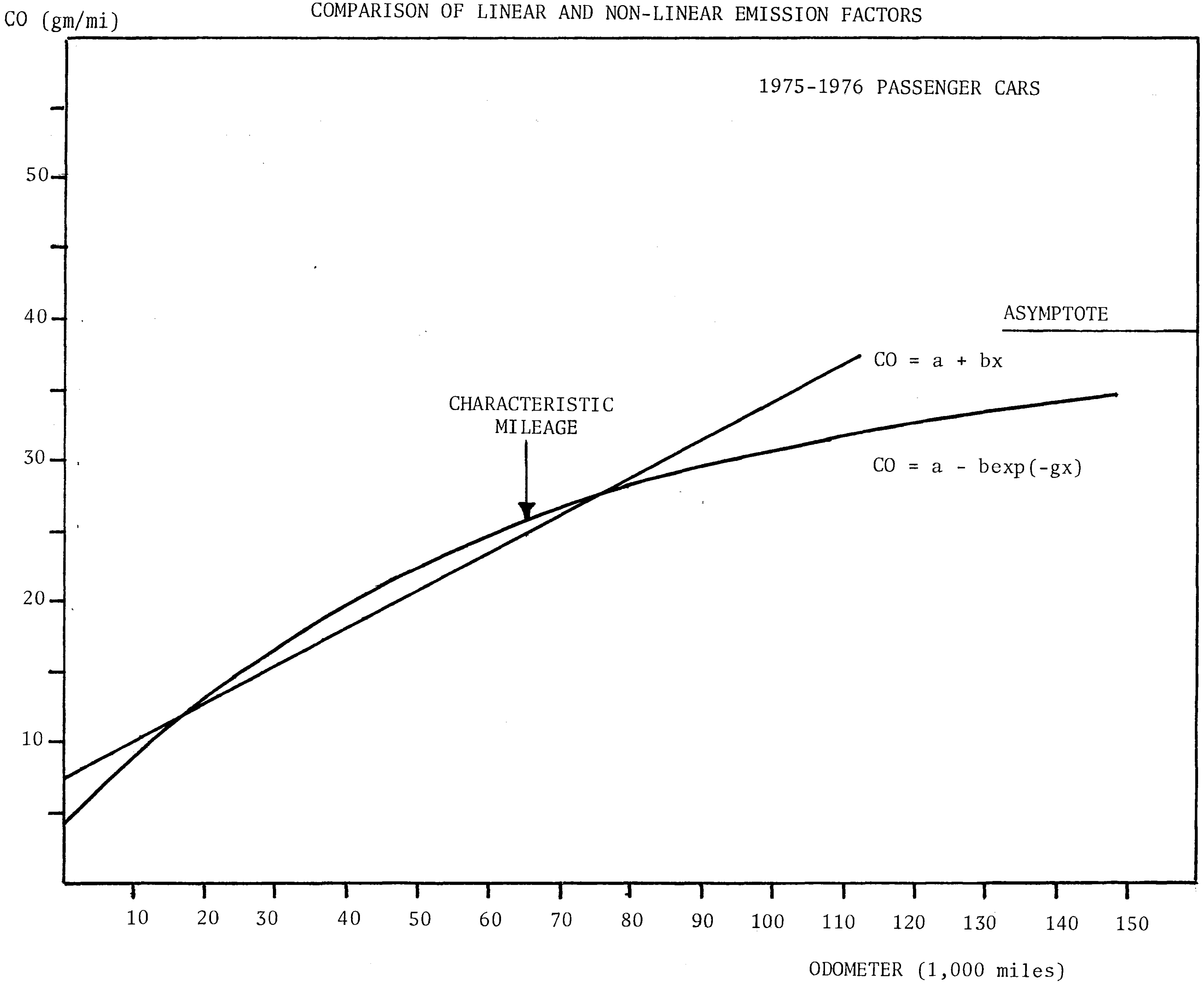
$$CO = a - b \exp(-gx)$$

was chosen for the population equation. As seen in Figure 3-4, this functional form produces a smooth curve with strong non-linearity at low mileage and a shallow slope at higher mileage that is qualitatively similar to the malperformance data for idle mixture maladjustments. Although alternate choices could be made, the hypothesized model for the population is consistent with the established behavior of emission factors and, as can be shown by expansion into a power series, is consistent with a linear approximation (with weak quadratic effect) over limited mileage ranges.

The curves illustrated in the figure are the result of applying non-linear regression techniques to the ARB surveillance data for 1975-1976 vehicles. In comparison to the linear emission factor, the exponential growth model predicts substantially the same emissions over the range of the existing surveillance data. It is only above 90,000 to 100,000 miles that the curves diverge; such behavior is highly likely to go undetected in the existing data. The exponential model has an intercept of $(a-b) \sim 4$ gm/mi and a high-mileage asymptote of $a \sim 38$ gm/mi. The rate at which the curve rises to its asymptote is controlled by the exponential parameter g; the value 1/g

FIGURE 3-4

COMPARISON OF LINEAR AND NON-LINEAR EMISSION FACTORS



defines a mileage parameter that describes the strength of the curvature -- the smaller the value 1/g, the stronger the curvature. Defined precisely, 1/e of the emissions increase between zero-miles and the asymptote has occurred when the average vehicle accumulates 1/g miles. Although the curvature is continuous, the 1/g parameter can be viewed as a "characteristic" mileage locating the "knee" of the growth curve.

To assure that samples of psuedo-data drawn from this population simulate closely the existing ARB surveillance data, several constraints are applied:

(1) The odometer distribution of the samples must exactly match the frequency of ARB data in each 1,000 mile odometer interval.

(2) Mean CO levels averaged across repeated samples must converge to the values observed in the ARB data -- i.e., the ensemble average of CO emissions equals the CO levels in the actual ARB data, although the means in individual data sets will vary randomly.

(3) The slopes of linear emission factor equations estimated for the repeated samples are centered around the values found in the ARB data. Because of the random variation, individual samples will show varying emission factor lines. But, as for mean emissions, the ensemble average must reproduce the results of the ARB data.

(4) The curvature of the equation represents a fixed decrement in predicted emissions over 100,000 miles compared to the linear emission factor estimated from the surveillance data. This is a controlled parameter (RHO) specified by the ratio of predicted emissions in the non-linear form to that of the linear factors. A value of RHO = 0.95 implies a 5% reduction in 100,000 mile emissions.

For a specified odometer distribution (as per constraint 1), constraints 2, 3, and 4 may be used to uniquely specify the population equation -- i.e., determine a single set of coefficients a, b, and g. As a result of the constraints, an ensemble of repeated samples will closely resemble the ARB surveillance data in terms of mean emission levels and the slope of linear emission factors, although due to the data variance each sample will differ.

The variance superimposed on the population was modeled after the variation observed in emissions data. The requirements for the variance distribution are that individual emission levels must be positive and the dispersion of the data should increase with mileage. These are met by a log-normal distribution (characteristic of emissions data) of the ratio of individual values to the mean emissions predicted by the exponential model. Since the distribution is taken to be log-normal, the majority of individual data points lie below the mean (median less than mean) and the samples exhibit low-frequency tails that extend to relatively high emission levels, as shown in Figure 3-5. As mean emissions increase with mileage, the dispersion in FTP emissions also increases. The parameters specifying the variance were initially estimated from the ARB data. They were then adjusted so that over repeated samples the total variance in the psuedo-data closely approximated the actual variance in the surveillance data.

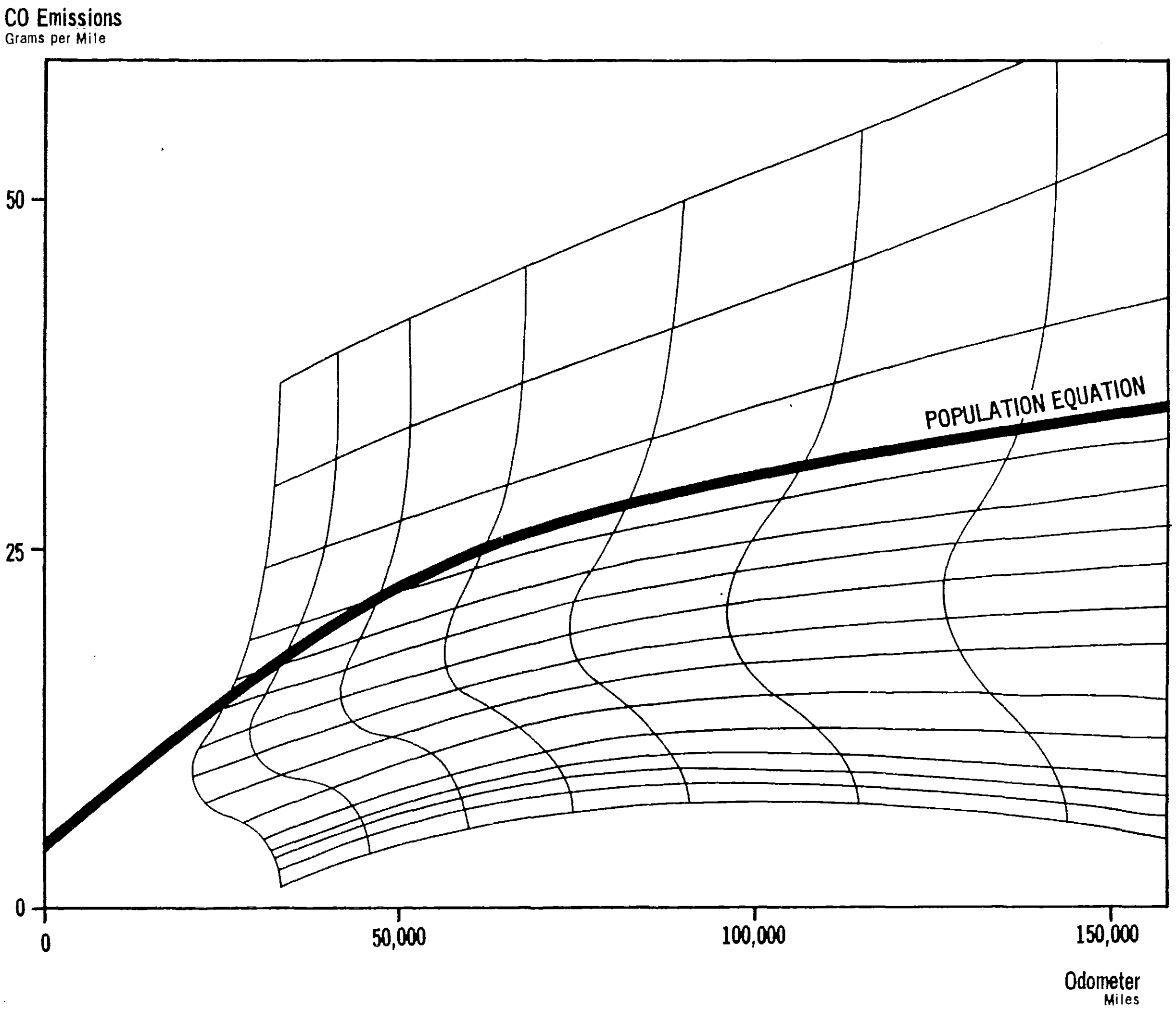
The mathematical manipulations to implement the simulation are relatively complex in comparison to the concepts they represent, and have been omitted in favor of a fuller description of the conceptual approach. Key points to note are that the hypothesized population equation is chosen to be consistent with the qualitative features of emissions behavior and emissions results from the surveillance data. A realistic representation of emissions variation is then superimposed on the population to generate repeated samples of psuedo-data that might result from emissions testing of individual vehicles. While necessarily idealized and simplified in some respects, the final model of CO emissions closely simulates the actual ARB surveillance data.

3.3.2 Simulation Results

Once the simulation model is programmed and validated, the conduct of simulation analyses is relatively straightforward. For the baseline simulations, 100 independent samples were drawn from populations with curvature parameters RHO equal to the reductions in 100,000 emissions (relative to linear factors) estimated from the quadratic regressions. The aggregate

FIGURE 3-5

Illustration of Psuedo-Data for Simulation Analysis



characteristics of these 100 samples are compared in Table 3-4 to the surveillance data sets on 1975-1976 and 1977-1979 vehicles to demonstrate the closeness of the simulation.

The analysis of Type II risks is performed by testing the significance of quadratic regressions in each sample. The beta level for the Type II error is estimated by the fraction of samples in which the quadratic term is found to be insignificant at a 0.05 confidence level. The beta level estimates the probability of failing to detect the non-linearity known to be present in the psuedo-data, and is a function primarily of how large the curvature is (parameter RHO), the data variance, the sample size, and the sample's odometer distribution. Large beta values indicate limited statistical power, while values on the order of 0.10 indicate satisfactory power for the analysis at hand.

As shown in Table 3-5, the baseline simulations indicate that vehicle samples of the size present in the surveillance data have little power to detect non-linear behavior of this magnitude. In 70 to 80 percent of the independent samples, the quadratic term was not significant at the 0.05 level. Thus, basing the analysis of non-linear behavior on existing data runs a very high risk of missing non-linearities on the order of 5 percent over 100,000 miles. For comparison with the results from the surveillance data, the quadratic equation forms estimated on the psuedo-data also are tabulated; averaged over the 100 samples, the estimated coefficients are close approximations to the non-linear behavior actually observed.

The bounding analysis is conducted by increasing the magnitude of curvature (parameter RHO) until the estimated beta level is reduced to an acceptable value or until non-physical results occur. Beta levels are conventionally taken to be slightly higher than the associated alpha level so that if errors are made in the statistical inference, they tend toward failure to detect an effect, rather than the erroneous acceptance of an effect's presence. For an alpha level of 0.05, an associated beta level of 0.10 might therefore be appropriate; in such a design 9 samples in 10 would be expected

TABLE 3-4

COMPARISON OF BASELINE SIMULATIONS TO ARB SURVEILLANCE DATA

	N	Mean CO (gm/mi)	Standard Deviation in CO	Emission Factor Slope	Characteristic Mileage
1975-1976					
ARB Surveillance Data	475	16.7	23.3	2.90	-
Simulation: RHO = 0.957 (Avg. of 100 samples)	475	16.6	23.2	2.91	58,300
1977-1979					
ARB Surveillance Data	505	10.8	16.6	2.33	-
Simulation: RHO = 0.952 (Avg. of 100 samples)	505	10.9	16.5	2.34	85,000

TABLE 3-5

SIGNIFICANCE OF QUADRATIC TESTS
(Baseline Simulations)

	Quadratic Emission Factor	Alpha Level	Beta Level*
1975-1976			
ARB Surveillance Data	$CO = 4.92 + 4.64 - 0.218X^2$	0.25	-
Simulation: RHO = 0.957 (Avg. of 100 samples)	$CO = 4.79 + 4.67X - 0.221X^2$	(0.05)	0.77
1977-1979			
ARB Surveillance Data	$CO = 3.39 + 3.22X - 0.121X^2$	0.48	-
Simulation: RHO = 0.952 (Avg. of 100 samples)	$CO = 3.60 + 3.14X - 0.108X^2$	(0.05)	0.70

---

*Probability of failing to defect quadratic term at specified alpha level to detect the effect's presence. Non-physical results occur when the curvature specified by RHO is too strong to be compatible with all constraints and the requirement that zero mile emissions remain positive. This situation indicates an implausibly large curvature. In either case, the RHO value provides a bounding estimate for the largest emissions curvature likely to be present.

The results of the bounding analysis are summarized in Table 3-6. For 1975-1976 vehicles, the curvature is unlikely to be greater than 15 percent (RHO = 0.85) and probably not greater than 10 percent, unless the existing surveillance data is an unusual random sample (among the 13 percent of such samples) that happens not show statistically significant evidence of a quadratic term. For 1977-1979 vehicles, the curvature is unlikely as great as 20 percent, but could be as large as 15 percent and yet remain undetected in as many as 34 percent of samples. Overall, the non-linearity for 1975-1976 vehicles is likely to fall in the range of 5-10 percent reductions in 100,000 mile emissions, but could be as large as 10-15 percent for model years 1977-1979. Since these are upper bounds, a single figure of 10 percent reduction in 100,000 mile emissions may be used as a guideline for the emissions benefits of quantifying and applying non-linear emission factor modeling techniques.

3.3.3 Additional Data Requirements

One reason for conducting the simulation is that it allows an analysis of how much additional FTP data on high-mileage vehicles would be required to develop conclusive statistical evidence of emission factor non-linearity. Such data could be collected in special surveillance programs targeted to high-mileage vehicles, although the costs of vehicle recruitment and testing would necessarily be weighed against the possible benefits of reduced emission factors.

Several considerations influence the statistical power added by additional high-mileage data: the characteristic mileage (or odometer interval) of the non-linearity, the numbers of new vehicles, the odometer range over

TABLE 3-6

SUMMARY OF BOUNDING ANALYSIS

	1975-1976		1977-1979	
RHO	Beta Level	Characteristic Mileage	Beta Level	Characteristic Mileage
0.95	0.71	48,800	0.69	81,200
0.90	0.13	19,100	0.53	35,600
0.85	0.01*	8,200	0.34	20,000
0.80	-	-	0.12*	12,200

---

*Non-physical results occur (negative intercept) which new vehicles are recruited, and the resulting balance of the augmented data set in covering a broad range of mileages. Since the actual design of a special testing program would require more careful consideration of these alternatives than is warranted here, attention has been limited to the hypothetical testing of vehicles in a 50,000 mile interval from 75,000 to 125,000 miles. To maximize the added power of the new data, it is further assumed that the vehicles are selected to uniformly populate this range. Although admittedly difficult to implement in practice, this simplified model testing program is a reasonable prototype for additional data collection.

For curvatures on the order of 5 percent (RHO = 0.95) as indicated by the surveillance data, the non-linearity is both weak and evidenced at relatively high mileages -- characteristic mileages for the exponential population equation are on the order of 50-80K miles as seen in Table 3-5. A fundamental conclusion of the simulations for this case is that no amount of FTP data in the 75-125K range will likely provide conclusive evidence. The limitations are the variance in the data and the relatively small differences between linear and non-linear emission factor equations in this odometer range. While data collection programs could be specified to map out the entire 0-125K range with effectively arbitrary precision, these are beyond the realm of reason for current or future testing programs.

Since the non-linearity may be as great as 10-15 percent for 1977-1979 vehicles, a second simulation was carried out for RHO = 0.90. An additional 50 vehicles in the 75,000 to 125,000 mileage range would raise the probability of detecting non-linear behavior from 53 percent in the current sample to approximately 70 percent in the augmented sample. This estimate presumes, however, that the true degree of emissions non-linearity is as great as 10 percent over 100,000 miles; there also would remain a 30 percent chance that the augmented data set would not produce conclusive results. For reasons of data variance and the odometer range covered, the Type II risk cannot be reduced to the desired 0.10 level by testing of larger numbers of vehicles. (A more modest requirement of vehicles in the 75,000 to

100,000 mileage range produces no substantial improvement in the quantification of non-linear emission factors, regardless of sample size).

Overall, the prospects are poor for establishing a precedent for non-linear emission factor modeling by FTP testing of high-mileage vehicles. The primary restrictions of this approach, other than cost, are that the existing non-linearity appears to be no more than approximately 10 percent over 100,000 miles and to produce substantial deviations from current linear factors only at low mileage and above 100,000 miles. Recruitment of the necessary very high mileage vehicles would be a difficult task that might not persuasively change the present status.

3.4 SUMMARY

A summary assessment of the adequacy of current emission factors to characterize high-mileage emissions is that, within the range of existing surveillance data, there is no firm evidence that linear factors over-predict (or inadequately represent) in-use emissions. At the same time, engineering analysis supported by the data cited, leads to expectations for low-mileage non-linearities and an attenuation of deterioration rates beyond 50,000 miles. That these effects are only suggested by the data, and not confirmed, results from the apparent weakness in the non-linearity between 10,000 and 70,000 miles and the high variance of emissions data. Virtually nothing is known about FTP emissions from California vehicles near or above 100,000 miles.

A bounding analysis sets a limit of approximately 10 percent reduction in predicted 100,000 mile emissions as the result of quantifying the existing non-linear behavior. It also should be noted that emissions reductions derived through EMFAC will weigh the emission factor curve according to the mileage accumulation rate. This means that the high travel-fraction weights will be placed on the mileage range in which linear factors under-estimate emissions, and low weights at high mileage. Overall, smaller emissions inventory reductions may be expected as a result.

However, this is unlikely to be achieved by the straightforward collection of additional FTP data through the surveillance programs, regardless of expenditure on the recruitment of high-mileage vehicles. Should ARB wish to pursue these emissions issues, it is likely that alternative methods for validating high-mileage emissions characteristics will be required. The most feasible candidate is a validation of high-mileage emissions predictions using idle test data collected in the California I/M program. The large sample sizes should make it possible to determine idle emission rates with high confidence through 125,000 to 150,000 miles. Emission factor adjustments may then be made using techniques similar to those described in the next section.

4. AN EMISSION FACTOR ANALYSIS USING MALPERFORMANCE DATA

In the previous sections, regression analysis was presented to explain observed differences in FTP emissions as a function of accumulated vehicle miles traveled (VMT). Emission of each of the regulated pollutants (CO, HC, NO_x) were found to increase at higher mileage; this was expected because of the gradual deterioration, and occasional malperformance, of emission control system components which occur during the life of a vehicle.

The VMT term thus acts as a surrogate for a large number of more basic explanatory variables relating to the operating condition of emission control system components. While the VMT model is simple and convenient, its explanatory power is limited by vehicle-to-vehicle differences in component deterioration rates and in the occurrences of malperformances. Two vehicles with the same accumulated mileage are likely to have quite different emission levels if one is in well-maintained condition, while the other has (for example) a malfunctioning catalyst and a maladjusted carburetor. In fact, the percentage of variance in emissions which can be explained using the VMT term alone is quite low -- typically less than 10 percent for CO, less than 5 percent for HC, and no more than one percent for NO_x.

A more satisfactory accounting for emissions variance can be obtained at the expense of using a more complex model which reflects details of component malperformances. This section describes an approach for developing such a model.

4.1 INCORPORATION OF ICO AND IHC TERMS

Maladjustment of the idle mixture is known to be a frequent cause of excessive HC and CO emissions. A logical first step to expand the model is to add some measure of idle mixture maladjustment as an independent (explanatory) variable in the regression equations. The idle CO emissions reading (ICO) seems suited for this purpose; it is obtained routinely by both ARB and EPA in emissions surveillance work, and is in fact used by EPA to test for idle mixture maladjustment.*

The augmented model then consists of three regression equations, and can be represented compactly as follows:

$$FTP_i = a_i + b_i x + c_i \cdot ICO$$

where x stands for accumulated mileage in units of tens of thousands of miles, and the index i successively represents each of the three pollutants: CO, HC, NO_x. Since vehicles with dissimilar emission control systems are likely to have substantial differences in emission deterioration rates, the regressions initially were run separately for each of four technology groups and for two model year groups (1975-1976 and 1977-1979). Because of low sample sizes for technologies other than catalyst with AIR, later analysis was subsequently restricted to this one dominant technology group.

Regression results for the largest technology groups, catalyst with air, have been provided in Tables 4-1, 4-2, and 4-3 (one for each pollutant). The basic VMT model appears first in each table, followed by the augmented model described above and referred to as the "VMT-ICO" model. Additional regression models reported in the lower half of each table are discussed later in this section. As seen in Tables 4-1 and 4-2, the addition of the ICO term contributes substantially to the explanatory power of the CO and HC models. For the MY1977-79 group, R-squared values increase from .02 to .31 for CO, and from .08 to .23 for HC. However, no significant correlation is found between ICO and FTP NO_x emissions, as expected since the process of NO_x formation is not strongly related to HC and CO control.

The use of a linear ICO term in these regressions is somewhat simplistic because it implies that FTP (HC and CO) emissions increase without limit at

---

*EPA determines the idle mixture to be maladjusted if idle CO emissions are greater than 0.5 percent.

TABLE 4-1

REGRESSIONS OF FTP CO ON VMT, ICO, IHC USING ARB DATA ALONE

(Catalysts with Air Injection)

Basic VMT Model

MY75-76: $FTP_{CO} = 7.07 + 2.10x$ df = 307
(t-stat): (3.37) (3.57) R^2 = .0398

MY77-79: $FTP_{CO} = 4.45 + 1.59x$ df = 329
(t-stat): (2.89) (2.84) R^2 = .0238

VMT-ICO Model

MY75-76: $FTP_{CO} = 6.14 + 0.99x + 10.05 . ICO$ df = 306
(t-stat): (3.63) (2.06) (12.94) R^2 = .3795

MY77-79: $FTP_{CO} = 3.56 + 0.92x + 9.88 . ICO$ df = 328
(t-stat): (2.75) (1.94) (11.77) R^2 = .3137

VMT-IHC Model

MY75-76: $FTP_{CO} = 6.34 + 1.52x + 0.074 . IHC$ df = 306
(t-stat): (3.19) (2.70) (6.18) R^2 = .1463

MY77-79: $FTP_{CO} = 3.13 + 0.33 + 0.202 . IHC$ df = 328
(t-stat): (2.25) (0.63) (8.91) R^2 = .2141

VMT-ICO-IHC Model

MY75-76: $FTP_{CO} = 6.08 + 0.96x + 9.67 . ICO + 0.0099 . IHC$ df = 305
(t-stat): (3.59) (1.98) (10.75) (0.84) R^2 = .3809

MY77-79: $FTP_{CO} = 3.49 + 0.84 + 9.28 . ICO + 0.0193 . IHC$ df = 327
(t-stat): (2.68) (1.69) (6.92) (0.57) R^2 = .3144

Note: "x" stands for VMT in units of tens of thousands of miles.

TABLE 4-2

REGRESSIONS OF FTP HC ON VMT, ICO, IHC USING ARB DATA ALONE

(Catalysts with Air Injection)

Basic VMT Model

MY75-76:	FTP_{HC}	=	0.549	+ 0.225x					df = 307
	(t-stat):		(1.76)	(2.58)					R^2 = .02
MY77-79:	FTP_{HC}	=	0.283	+ 0.166x					df = 329
	(t-stat):		(3.37)	(5.42)					R^2 = .081

VMT-ICO Model

MY75-76:	FTP_{HC}	=	0.507	+ 0.175x	+	0.458 . ICO			df = 306
	(t-stat):		(1.65)	(2.00)		(3.25)			R^2 = .053
MY77-79:	FTP_{HC}	=	0.247	+ 0.139x	+	0.395 . ICO			df = 328
	(t-stat):		(3.21)	(4.91)		(7.92)			R^2 = .229

VMT-IHC Model

MY75-76:	FTP_{HC}	=	0.401	+ 0.108x	+	0.0017. IHC			df = 306
	(t-stat):		(1.77)	(1.36)		(9.02)			R^2 = .226
MY77-79:	FTP_{HC}	=	0.204	+ 0.091x	+	0.0120. IHC			df = 328
	(t-stat):		(2.76)	(3.26)		(9.90)			R^2 = .293

VMT-ICO-IHC Model

MY75-76:	FTP_{HC}	=	0.405	+ 0.117x	-	0.163 . ICO	+	0.0162 . IHC	df = 305
	(t-stat):		(1.46)	(1.47)		(-1.10)		(8.35)	R^2 = .229
MY77-79:	FTP_{HC}	=	0.207	+ 0.095x	+	0.067 . ICO	+	0.0107 . IHC	df = 327
	(t-stat):		(2.79)	(3.36)		(0.87)		(5.52)	R^2 = .2948

Note: "x" stands for VMT in units of tens of thousands of miles.

TABLE 4-3

REGRESSIONS OF FTP NO_x ON VMT, ICO, IHC USING ARB DATA ALONE

(Catalysts with Air Injection)

Basic VMT Model

MY75-76:	FTP_{NO_x}	=	2.07	+	0.073x					df = 307
	(t-stat):		(17.16)		(2.16)					R^2 = .0150
MY77-79:	FTP_{NO_x}	=	1.52	+	0.066x					df = 329
	(t-stat):		(14.69)		(1.74)					R^2 = .0091

VMT-ICO Model

MY75-76:	FTP_{NO_x}	=	2.08	+	0.078x	-	0.040	.	ICO	df = 306
	(t-stat):		(17.16)		(2.25)		(-0.71)			R^2 = .0166
MY77-79:	FTP_{NO_x}	=	1.53	+	0.069x	-	0.051	.	ICO	df = 328
	(t-stat):		(14.70)		(1.82)		(-0.76)			R^2 = .0109

VMT-IHC Model

MY75-76:	FTP_{NO_x}	=	2.07	+	0.071x	+	0.00033	.	IHC	df = 306
	(t-stat):		(17.08)		(2.06)		(0.46)			R^2 = .0157
MY77-79:	FTP_{NO_x}	=	1.51	+	0.054x	+	0.00189	.	IHC	df = 328
	(t-stat):		(14.49)		(1.38)		(1.16)			R^2 = .0129

Note: "x" stands for VMT in units of tens of thousands of miles.

higher levels of idle CO emissions. Since this seems implausible from an engineering viewpoint, a model containing a quadratic as well as a linear ICO term was tested for i = CO, HC:

$$FTP_i = a_i + b_i x + c_i \cdot ICO + d_i \cdot (ICO)^2$$

Table 4-4 provides results for the FTP CO regressions on the catalyst with air injection technology groups. Both the linear and quadratic corrective ICO terms are determined to be significantly different from zero with 95 percent confidence ($|t|$ statistic > 1.96) for MY1977-79 vehicles, and are found nearly significant for MY1975-76 vehicles. The FTP HC regression results are not shown but are similar. This suggests that a nonlinear functional form of ICO may be appropriate. Using the quadratic ICO term presents problems, however, since it implies that increasing ICO beyond a certain level will correspond to decreasing FTP emissions. This problem can be avoided if a nondecreasing function, such as the natural logarithm of ICO, is used instead of the linear and quadratic pair. While the log term is proven significant (Table 4-4), the R- squared values are reduced by more than half, to the .09 to .03 range. Other possible functional forms of ICO could be tested but it was judged inappropriate to pursue this avenue of investigation within the present task.

The idle HC emission reading (IHC) was also examined as a possible explanatory variable, since IHC is known to be affected by malperformances in the ignition system as well as by idle mixture maladjustment. A strong correlation between IHC and FTP HC emissions was expected, and Table 4-2 shows that the VMT-IHC model outperforms the VMT-ICO model in accounting for variance in FTP HC emissions. The VMT-IHC model also provides a variable accounting for FTP CO emissions, although it is not as powerful in this case as the VMT-ICO model. No significant correlation is found between IHC and FTP NO_x emissions.

Finally, a model was tested which includes both ICO and IHC terms as explanatory variables for FTP HC and CO emissions. Because of the high

TABLE 4-4

REGRESSIONS OF FTP CO ON VMT AND NONLINEAR FUNCTIONS OF ICO USING ARB DATA ALONE

(Catalysts with Air Injection)

VMT-Quadratic ICO Model

MY75-76: $FTP_{CO} = 5.99 + 0.886x + 13.26 \ . \ ICO - 0.519(ICO)^2$ df = 305
(t-stat): (3.55) (1.82) (6.12) (-1.59) R^2 = .3846

MY77-79: $FTP_{CO} = 2.16 + 0.698x + 27.89 \ . \ ICO - 3.538(ICO)^2$ df = 327
(t-stat): (1.76) (1.57) (10.43) (-7.05) R^2 = .4042

VMT-Log (ICO) Model

MY75-76: $FTP_{CO} = 19.18 + 0.752x + 2.04 \ . \ \log(ICO)$ df = 306
(t-stat): (6.60) (1.24) (5.74) R^2 = .1330

MY77-79: $FTP_{CO} = 11.45 + 0.961x + 1.87 \ . \ \log(ICO)$ df = 328
(t-stat): (5.70) (1.73) (5.15) R^2 = .0970

degree of correlation between ICO and IHC, these regressions proved to be less satisfactory than some which included either ICO or IHC alone: regression R^2 values are not substantially increased by inclusion of both terms, and the terms generally can not both be estimated with acceptable confidence (See the VMT-ICO-IHC results in Tables 4-1 and 4-2).

In summary, the foregoing investigation of various regressions using VMT, ICO, and IHC as explanatory variables has produced the following composite model:

$$\begin{aligned} FTP_{CO} &= a_1 + b_1 x + c_1 ICO \\ FTP_{HC} &= a_2 + b_2 x + c_2 IHC \\ FTP_{NOx} &= a_3 + b_3 x \end{aligned}$$

For the largest technology group (catalysts with air pump), the above model explains between 31 and 38 percent of variance in FTP CO emissions in the sample data, between 22 and 29 percent of FTP HC emissions variance, and between 0 and 2 percent of FTP NO_x emissions variance.

4.2 INCORPORATION OF MALPERFORMANCE DATA

Both ARB and EPA conducted thorough inspections of test vehicles in the as-received condition, making note of any adjustments or repairs which were needed. The results of these inspections appeared in each data set in the form of diagnostic fields, with each field representing the diagnosis for a single component. This information, when used in conjunction with VMT, ICO, and IHC, clearly provides a more nearly complete description of the condition of the test vehicle than can be obtained from the three aforementioned variables alone. It was hoped that a regression model expanded to include the diagnostic results would have greater explanatory power than the model discussed in the previous section, and also that it would provide an indication of the relative impact on emissions caused by failures of the various components.

However, there are several problems in implementing this approach. Although ARB and EPA each inspected all of the major vehicle systems and subsystems relevant to emission control (e.g., the carburetor), the two sources typically chose a somewhat different set of specific components to inspect within a given subsystem (e.g., choke, float, carburetor wires, etc.). Also, the diagnostic fields are too numerous to include in a single regression; the ARB data contains 58 fields, while the EPA data has over 80. Moreover, ARB and EPA distinguished between eight and ten types of malperformances (maladjusted, disabled, missing, etc.) so that each field conceivably might have to be represented by several variables in a regression.

It was clear that some of the precision afforded by the original data would have to be sacrificed in order to obtain manageable specifications for regression analysis. A reduction in the number of diagnostic fields was achieved through a recoding process which is summarized in Table 4-5. The numerous original diagnostic fields have been replaced by 19 aggregate variables, each representing a vehicle system important in the emission control process. Since both ARB and EPA inspected at least some components within each of the 19 systems (except that EPA omitted the pulse air system), the new variables are common to both data sources.

The aggregate variables have been coded as dummy variables assigned the value of 1 if a malperformance of any kind has been found in any of the components within the system, or 0 if all of the constituent components passed inspection. A "malperformance" has been defined very broadly to include any disorder noted by inspectors: maladjustment, tampering, disablement, component missing, etc. One obvious defect of the new coding system is that no distinctions have been made between varieties or degrees of malperformances; a system with several disabled components receives the same flagged value as one with a single maladjusted component.

Another problem is that EPA and ARB typically inspected a different set of components within any given system, so that the aggregate variables are generally not equivalent across the two data sources. For example, the new

TABLE 4-5

RECODING SYSTEM FOR COMPONENT MALPERFORMANCE DATA

The underscored group headings are the new variables created for this analysis. Listed underneath each heading are the original diagnostic variables in the Emission Factors and Surveillance databases from which the new variables were formed, together with their original code names (Exxx and Dxx, respectively).

EPA EMISSION FACTORS	CALIFORNIA ARB SURVEILLANCE
Idle Mixture	
E203 Idle Mixture Adjustment	D11 Idle CO or Mixture
Idle Speed	
E204 Idle Speed	D12 Idle RPM
Carburetor	
E201 Carburetor Assembly	D17 Carburetor
E211 Choke Adjustment	D18 Carburetor Choke
E212 Vacuum Diaphragm for Choke	D19 Carburetor Float
E213 Electrical Controls for Choke	
E214 Hoses, Lines, Wires for Choke	
E219 Hoses, Lines, Wires (apparent redundancy)	
E215 Exhaust Heat Control Valve Assembly - Choke	
E216 Actuating Diaphragm - Choke	
E217 Coolant Temp. Searing Vacuum Switches for Choke	
E218 (GM) Check Valve - Choke	
E220 Other - Choke	
E914 Fuel Control/Mixture Device	
E915 (Ford) Stepper Motor	
Fuel Injection	
E206 Fuel Injection Components (1980 Study Only)	D15 Fuel distributor or air flow control unit
	D16 Fuel injector(s)
Exhaust Gas Oxygen Sensor	
E912 Oxygen Sensor	D60 EGO Sensor
Closed-Loop Computer	
E911 Electrical Control Unit - 3-way catalyst	D61 Electronic module or computer
E913 Other Sensors or Switches - 3-way catalyst	
E916 Wires, Hoses, Lines, etc. - 3-way catalyst	

TABLE 4-5 (Continued)

EPA EMISSION FACTORS	CALIFORNIA ARB SURVEILLANCE
Other Fuel System	
E202 Limiter Caps	D13 Fuel Pump
E206 Idle Stop Solenoid	D14 Fuel Filter
E207 Dashpot & Other Throttle Modulators	D20 Idle Speed Solenoid
E208 Fuel Filter Element	D21 Throttle Positioner
E209 Hoses, Lines, Wires - Fuel System	D22 Decel Dashpot
E210 Other - Fuel System	
Air Filter	
E105 Air Filter Element	D23 Air Filter Element
Turbocharger	
E108 Turbocharger Components	D28 Turbocharger
	D29 Turbocharger Wastegate
Other Air Induction	
E103 Temperature Sensors, Switches, Modulators - Air Induction System	D24 Hot Air Intake Heat Stove/Duct (TCS)
E104 (Ford) Delay Valve	D25 Hot Air Intake Snorkel (TAC)
E106 Hoses, Tubes, Lines, Wires - Air Induction	D26 Intake Manifold
E107 Other - Air Induction	D27 Heat Riser Valve
Ignition System	
E301 Distributor Assembly	D30 Spark Timing
E302 Initial Timing	D31 Distributor
E303 Spark Plugs & Wires	D32 Distributor Cap
E304 Vacuum Advance Unit - Ignition System	D33 Rotor
	D34 Points
E305 Spark Delay Devices	D35 Condenser
E306 Coolant Temperature Sensing - Ignition	D36 Electronic Breaker
E307 Hoses, Lines, and Wires - Ignition	D37 Ignition Coil
E308 Dwell	D38 Ignition Control Module
E309 Other - Ignition	D39 Ignition Wire(s)
E303 Initial Timing Limiting Device (1980 Study Only)	D40 Spark Plug(s)
E307 Spark Knock Detector (1980 Study Only)	D41 (VAC) Spark Advance/ Retard Unit
	D42 OSAC or Spark Delay Valve
Evaporative Control System	
E801 Evaporation Canister	D43 Vapor Storage Canister
E802 Canister Filter	D44 Fuel Filler Cap
E803 Hoses, Lines - Evap. Ctl System	
E804 Other - Evap. Ctl System	

TABLE 4-5 (Continued)

EPA EMISSION FACTORS	CALIFORNIA ARB SURVEILLANCE
Catalyst/Misfueling	
E702 Catalyst - exhaust system	D45 Fuel Filler Neck or Restrictor
	D59 Catalyst
Exhaust System	
E701 Exhaust Manifold, Tailpipe, Muffler	
E703 Other - exhaust system	
PCV Valve	
E601 PCV Valve Assembly	D46 PCV Valve or Orifice
E602 Filters - PCV System	D47 Oil Filter Cap
E603 Hoses and Lines - PCV	D48 PCV Filter
E604 Other - PCV	
EGR System	
E401 EGR Valve Assembly	D49 EGR Valve
E402 EGR Valve BPT	D50 EGR Vacuum Amplifier
E403 EGR Time Delay Solenoid	D51 EGR Control System
E404 Venturi Vacuum Amplifier - EGR System	
E405 High Speed Modulator - EGR	
E406 Vacuum Reservoir - EGR	
E407 Coolant Temperature Sensing Valve - EGR	
E408 Hoses, Lines, and Wires - EGR	
E409 Other - EGR	
Pulse Air System	
-	D52 Pulse Air Valve
Air Pump or Belt	
E501 Air Pump Assembly	D53 Air Injection Pump
E505 Drive Belt	D54 Air Injection Pump Belt
Other Air Injection System	
E502 Bypass Valve, Dump Valve - Air Pump System	D55 Diverter, Gulp or Flow Control Valve(s)
E503 Air Diverter Valve	
E504 Check Valve	D56 Air Injection Distribution Manifolds
E506 Hoses, Lines, Wires - Air Injection System	
E507 Other - Air Injection	
E504 Electrical PVS - Air Injection (1977 & 1979 Studies Only)	
E505 Solenoid Vacuum Valve - Air Injection (1977 & 1979 Studies Only)	
E506 Floor Pan Switch - Air Injection (1977 & 1979 Studies Only)	
E507 Vacuum Differential - Air Injection (1977 & 1979 Studies Only)	

TABLE 4-5 (Continued)

EPA EMISSION FACTORS	CALIFORNIA ARB SURVEILLANCE
Miscellaneous Engine	
E901 Engine Assembly	
E903 Valve Adjustment	D63 Intake & Exhaust Valve Adjustment

carburetor variable subsumes 12 inspection items from the EPA data, but only three items from the ARB data. Some types of carburetor malperformances likely to have been detected by EPA may have been missed by ARB's less detailed inspection. It would therefore be expected that a higher aggregate rate of carburetor malperformances would be reflected in the EPA data.

A comparison of malperformances rates between the two data sources, using the 19 aggregate variables, is provided in Table 4-6. Greater percentages of malperformances are generally found in the EPA data, particularly for the air induction system, PCV system, EGR system, and the other (miscellaneous) fuel system category. This suggests that the inspection standards used by EPA were more stringent than those used by ARB. Because of the unresolved differences in inspection procedures between ARB and EPA data sets, and the conflicting rates of emission-related malperformances indicated, subsequent analysis has been restricted to the ARB surveillance data sets alone.

4.3 REGRESSION ANALYSIS USING MALPERFORMANCE DATA

Two approaches were considered for expanding the model to incorporate the detailed malperformance data discussed above. The first approach would make use of the malperformance information in conjunction with an analysis of residuals* from the regressions specified at the end of Section 4-1 (the specifications consist of a VMT-ICO equation for FTP CO emissions, a VMT-IHC equation for HC, and a linear VMT equation for NO_x). The regression equations take into account only VMT, ICO, and IHC and therefore are likely to under predict emissions from vehicles with malperformances in key subsystems (such as catalyst). An indication of the relative importance of each subsystem for control of the three pollutants can be obtained by grouping the sample data by type of malperformance (carburetor, air filter,

---

*A residual is the difference between the actual and predicted (fitted) value of the dependent variable for a single observation in a regression.

TABLE 4-6

PERCENT RATES OF MALPERFORMANCES USING RECODED VARIABLES

(Catalysts with Air Injection)

	MY75-76		MY77-79	
	ARB	EPA	ARB	EPA
Sample Size (Vehicles)	311	79	332	30
Idle Mixture	42.1	54.4	35.3	30.5
Idle RPM	37.6	34.2	36.7	30.5
Carburetor	13.2	38.0	8.4	37.9
Other Fuel System	3.5	83.5	2.1	49.0
Air Induction System	7.7	82.3	5.7	47.9
Turbocharger*	0.0	0.0	0.0	0.0
Ignition	30.5	58.2	24.4	46.6
Fuel Injection*	0.0	0.0	0.0	0.0
Exhaust Gas Oxygen Sensor	0.0	0.0	0.0	0.0
Evaporative Control System	1.0	7.6	0.3	0.0
PCV System	1.3	60.8	1.2	43.0
EGR System	33.1	49.4	17.8	40.7
Air Injection System	3.5	36.7	1.5	34.6
Catalyst and/or Misfueled	1.9	1.3	0.9	0.0
Miscellaneous Engine	3.2	13.9	1.8	25.4

---

*Very few vehicles were equipped with turbocharger or fuel injection components.

etc.), and calculating the mean residual for each group. For example, vehicles with catalyst disorders might be expected to have a higher mean residual than those with air filter disorders; vehicles with no malperformances should obtain a negative mean residual. The standard error of the mean can be used to calculate confidence intervals.

The residuals analysis approach has the advantage of retaining relatively simple regression specifications, and thereby largely avoiding some problems typically encountered when using many variables in a single regression. Unfortunately, this approach is unable to account effectively for multiple system malperformances occurring within the same vehicle. Compound classes could be created e.g., one class might represent vehicles having both catalyst and air filter disorders; however, there are many possible combinations of subsystems and the sample size available for any given one is frequently miniscule.

In the second approach, the aggregate malperformance variables discussed in Section 4.2 would be incorporated directly into the regression equations as independent variables. The model is then specified as follows:

$$FTP_{CO} = a_{CO} + b_{CO} + c_{CO} \cdot ICO + \sum_{i=1}^{nco} d_{COi} \cdot M_i$$

$$FTP_{HC} = a_{HC} + b_{HC}x + c_{HC} \cdot IHC + \sum_{i=1}^{nhc} d_{HCi} \cdot M_i$$

$$FTP_{NOx} = a_{NOx} + b_{NO_x} x + \sum_{i=1}^{nnox} d_{NOx} . i \; M_i$$

where (for CO) $M_1, \ldots, M_{nco}$ represent the malperformance dummy (0/1) variables selected for inclusion in the FTP CO equation, nco is the number of malperformance variables, and $d_{CO.i}$ is the regression coefficient for the ith malperformance variable in the CO regression (the variables in the HC and NO_x equations are defined similarly). This model is able to account for multiple subsystem malperformances found in the sample data. Moreover, the fitted regression equations will be able to provide an estimate of FTP emissions for a hypothetical vehicle with an arbitrary pattern of malperformances, provided that all systems involved were included in the regression specification. Because of these advantages, this approach was chosen for the present analysis.

In this approach, the effects on emissions caused by different system malperformances are assumed to be strictly additive and independent of one another and of VMT and ICO (or IHC). In fact, the cumulative increase in emissions caused by multiple system malperformances can actually be greater or less than the sum of the increases caused by the individual malperformances, because of synergistic and saturation effects. The additional complexity required to account for such interactive effects was, however, judged inappropriate for the current task effort. On the whole, the additive model is considered an acceptable approximation from an engineering viewpoint.

Another problem area is the correlation between explanatory variables in the regression equations. Since vehicles tend to be more poorly maintained as they age, there are positive correlations between VMT and the various malperformance variables, and among the malperformance variables themselves. Additional correlations exist between ICO and the idle mixture malperformance variables, and between IHC and the ignition system variable. The inclusion of correlated independent variables in the regression typically results in reduced coefficient estimates and t-statistics, since explanatory power is in effect being shared by several variables. Therefore, it is desirable to limit the number of explanatory variables in the regressions, using engineering and statistical criteria to select only the most important for inclusion.

The first set of regressions tested included all malperformance variables, except those with a zero rate of occurrence in the ARB data as shown in Table 4-6. These regressions were strictly exploratory in purpose and were intended to help identify which malperformance variables were promising for further investigation. As expected, most the variables could not be established as significant at the chosen confidence level (ninety-five percent).

The regression specifications were then pared down to include only the following malperformance variables: idle mixture, idle RPM, carburetor, ignition system, EGR system, air injection, and catalyst. These seven variables were selected to represent the fuel and ignition systems (where a large percentage of emissions-related problems develop) and the major components designed specifically for emission control.

The results of these regressions are provided in Tables 4-7, 4-8, and 4-9 for CO, HC, and NO_x respectively, with the variables and their coefficients listed in vertical format. The only system malperformances determined to be significantly related to increasing CO and HC emissions are idle mixture for MY75-76, and air injection for MY77-79. However, the carburetor malperformance variable is nearly found significant for CO in MY75-76 with a t-statistic of 1.87, while the idle RPM malperformance is found to have a negative impact on HC emissions for MY75-76.

The fitted coefficient of the air injection variable for the MY77-79 regression is suspiciously high (24.83). Table 4-6 reveals that only 1.5 percent of the MY77-79 sample, or five data points, contained a malperformance in the air injection system. Therefore, little confidence can be placed on the large correlations found between the air injection variable and CO and HC emissions. Similarly, the large coefficients assigned to the catalyst variable in the CO regressions are based on a small number of vehicles in the sample with reported catalyst disorders.

One puzzling result was the significance of both ICO and the idle mixture variable in the MY75-76 CO regressions, since both variables were originally thought to represent the same physical attribute, i.e., the maladjustment of the idle mixture. This suggests that the idle CO reading may be affected by other attributes in addition to idle mixture maladjustment.

TABLE 4-7

REGRESSIONS OF FTP CO ON VMT, ICO, AND SELECTED MALPERFORMANCE VARIABLES USING ARB DATA ALONE

(Catalysts with/Air Injection)

	MY75-76 df=299 R^2=.4134		MY77-79 df=321 R^2=.3623	
	Parameter Estimate	T-Statistic	Parameter Estimate	T-Statistic
(Intercept)	4.55	2.52	2.84	2.10
x (VMT in 10^4mi)	0.85	1.72	0.84	1.75
ICO	9.78	12.33	8.86	10.38
Idle Mixture	8.14	2.96	0.45	0.26
Idle RPM	-4.37	-1.54	1.41	0.82
Carburetor	5.30	1.87	2.23	0.90
Ignition	-0.74	-0.34	-1.42	-0.87
EGR System	-0.46	-0.23	1.27	0.72
Air Injection	-4.03	-0.79	24.83	4.39
Catalyst	7.97	1.17	7.22	0.66

TABLE 4-8

REGRESSIONS OF FTP HC ON VMT, IHC, AND SELECTED MALPERFORMANCE VARIABLES USING ARB DATA ALONE (Catalysts with Air Injection)

	MY75-76 df=299 R^2 = .2547		MY77-79 df=321 R^2 = .3279	
	Parameter Estimate	T-Statistic	Parameter Estimate	T-Statistic
(Intercept)	0.344	1.15	0.165	2.10
X (VMT in 10^4 mi)	0.125	1.52	0.090	3.18
IHC	0.015	8.81	0.011	8.74
Idle Mixture	1.289	2.83	-0.104	-1.05
Idle RPM	-1.296	-2.76	0.084	0.85
Carburetor	0.272	0.58	0.135	0.94
Ignition	0.196	0.54	0.136	1.42
EGR System	-0.415	-1.24	0.062	0.62
Air Injection	-0.086	-0.10	1.074	3.30
Catalyst	-0.159	-0.14	0.133	0.32

TABLE 4-9

REGRESSIONS OF FTP NO_x ON VMT, ICO, AND SELECTED MALPERFORMANCE VARIABLES USING ARB DATA ALONE
(Catalysts with Air Injection)

	MY75-76 df = 299 R^2 = .3753		MY77-79 df = 321 R^2 = .4703	
	Parameter Estimate	T-Statistic	Parameter Estimate	T-Statistic
(Intercept)	1.736	16.41	1.296	15.70
x (VMT in 10^4 mi)	0.016	0.54	0.019	0.64
ICO	-0.095	-2.04	-0.036	-0.70
Idle Mixture	-0.011	-0.07	0.016	0.15
Idle RPM	0.016	0.10	-0.048	-0.45
Carburetor	-0.076	-0.46	-0.177	-1.17
Ignition	0.125	0.97	0.316	3.17
EGR System	1.509	12.67	1.705	15.90
Air Injection	0.462	1.55	-0.449	-1.30
Catalyst	0.262	0.65	-0.219	-0.50

For the NO_x regressions, the EGR system malperformance variable is found significant for both model year groups; this was expected, since the EGR system is the principal component used by manufacturers for NO_x control. The ignition malperformance also passes the t-test, but only for MY77-79 vehicles, and its coefficient estimate is modest (0.316).

The addition of the malperformance variables to the CO, HC, and NO_x regressions clearly has displaced some of the explanatory power of the VMT (x) variable. It might be argued that the VMT variable is superfluous because the system malperformances are explicitly represented in the regression. However, the VMT term serves as a surrogate for the normal deterioration of various components (particularly the catalyst), a process which perhaps is not effectively captured by the malperformance variables alone.

Overall, the most difficult analytical aspect is the estimation of an emissions model in data sets containing relatively small numbers of vehicles with major emissions malperformances. After a variety of exploratory specifications, it was concluded that the best estimation procedure for the current data set is to pool model years 1975-1979 in a single malperformance model. Separate intercept terms were allowed for the 1975-1976 and 1977-1979 groups to capture the effect of more stringent standards in the latter period. However, a common VMT slope for both groups was fit for each pollutant based on findings reported in previous tables. As with VMT, common emissions-response coefficients were estimated for the pooled data for the idle emissions and malperformance terms. The specification and parameters estimates for the final model are summarized in Table 4-10.

While substantially different in form than conventional emission factors, the interpretation of malperformance models is actually straightforward. Vehicles without any malperformance (of those included in the model equation) are characterized by linear VMT-ICO or VMT-IHC equations for (VMT only, for NO_x). For example, a MY75-76 vehicle without major malperformance is expected to have a CO factor of 5.33 + 0.73X + 9.65 ICO. It should be noted, however, that such vehicles are not necessarily well-maintained,

TABLE 4-10

FINAL REGRESSIONS OF FTP CO, HC, AND NO_x ON VMT, ICO, IHC, AND SELECTED MALPERFORMANCE VARIABLES USING ARB DATA ALONE (Catalysts with Air injection)

$\underline{CO}$

$$\left.\begin{array}{lr} \text{MY75-76 } FTP_{CO} = & 5.33 \\ \text{(t-statistic)} & (1.32) \\ & \\ \text{MY77-79 } FTP_{CO} = & 2.71 \\ \text{(t-statistic)} & (2.27) \end{array}\right\} \begin{array}{l} + 0.728X + 9.65ICO + 3.45IDLEMIX + \\ \quad (2.13) \quad\; (16.87) \quad\; (2.88) \\ + 4.33\ AIR + 7.76\ CAT \\ \quad (1.16) \quad\; (1.59) \end{array}$$

$\underline{HC}$

$$\left.\begin{array}{lr} \text{MY75-76 } FTP_{HC} = & 0.442 \\ \text{(t-statistic)} & (2.55) \\ & \\ \text{MY77-79 } FTP_{HC} = & 0.134 \\ \text{(t-statistic)} & (1.97) \end{array}\right\} \begin{array}{l} + 0.096X + 0.0147IHC + 0.202\ AIR \\ \quad (2.08) \quad\; (12.92) \quad\;\; (0.41) \end{array}$$

$\underline{NO_x}$

$$\left.\begin{array}{lr} \text{MY75-76 } FTP_{NO_x} = & 1.721 \\ & (22.22) \\ & \\ \text{MY77-79 } FTP_{NO_x} = & 1.451 \\ \text{(t-statistic).} & (17.61) \end{array}\right\} \begin{array}{l} + 0.017X + 1.595\ EGR \\ \quad (0.87) \quad\; (20.19) \end{array}$$

"X" represents VMT in units of tens of thousand of miles.

since minor malperformances (or major, but very infrequent ones not included in the model) may nevertheless be present. The deterioration rates for such vehicles are reduced in comparison to the conventional linear factor because the confounding effect of increasing rates of occurrence with mileage are controlled by the inclusion of the malperformance terms. Given a relationship between ICO and X (accumulated mileage) for the subset of vehicles without major malperformance, the three-term emissions model also can be reduced to the simple linear form.

Continuing the example, vehicles with idle mixture maladjustment are expected to have increased FTP CO emissions by an amount that depends both on the ICO and IDLEMIX terms. If the maladjustment increases ICO from a specification level of 0.2 percent to 2.5 percent, the incremental FTP impact is 9.65 (2.5-0.2) = 22.2 gm/mi. An additional impact of 3.45 gm/mi is due to idle mixture maladjustment that is not accounted for by the emissions effects observed at idle. Corresponding interpretations hold for the other malperformance terms included in the regression models.

A primary motivation for the development of malperformance models is to present a basis for validation or adjustments to linear emission factor equations. This could include estimating the emissions benefits of anti-tampering programs (e.g., air pump inspection), mandatory maintenance (as in the Colorado I/M program), or idle emission inspection in the California I/M program. Malperformance emissions models may also be applied in validating emission factor projections at high-mileages (above 100,000 miles) where no FTP test data exist. In this latter example, the BAR inspection data on idle emission levels for vehicles in the as-received condition (before repair) could be used to develop a profile of ICO and IHC as functions of mileage through the highest mileage intervals present in the California fleet. These profiles will link emission factor estimates developed from the malperformance equations to an independent (I/M) source of emissions data that does not have the sample size restrictions at high-mileage as existing surveillance data.

4.4 CONCLUSION

The foregoing analysis has produced an enhanced emission factors model which explains approximately 38 percent of variance in FTP CO emissions, 24 percent of variance in FTP HC, and 40 percent of variance in FTP NO_x. This compares with the models which relied on VMT alone, which typically account for much less than 10 percent of variance for the three pollutants.

The new model makes use of detailed inspection data providing information on malperformances occurring in the various vehicle components relevant to emission control. An aggregation process was developed and implemented to prepare the data for analysis; this involved the creation of nineteen variables, each representing the malperformance of a single vehicle system. The variables were then selectively incorporated into the regression model, along with variables representing the idle CO and HC emissions readings (ICO and IHC).

The bulk of improvement in explanatory power for the CO and HC regressions was accounted for by the addition of ICO and IHC, respectively. The incorporation of an EGR system variable was responsible for most of the improvement in the NO_x regression.

5. CONCLUSIONS

Based on the foregoing analysis, the following conclusions should be drawn with respect to the specification of emission factors for 1972 to 1979 mobil year cars.

(1) The addition of the sixth surveillance dataset had little influence on the linear emission factors calculated from the ARB data when compared with emission factors published by ARB in November of 1982. The only changes noted were an increase in deterioration rates for HC and NO_x in 1972 to 1974 model year cars.

(2) The representativeness of the emission factor data collected by EPA in California is extremely poor. The bulk of EPA's California data collection effort focused on three-way catalyst vehicles. Earlier EPA emission factor calculations for California vehicles relied on the use of ARB surveillance data to broaden the representativeness of the sample. Conversations with EPA personnel indicated that EPA has no plans to continue the development of emission factors for California vehicles and that MOBILE3 will rely on ARB factors.

(3) The unrepresentativeness of the EPA data makes it unsuitable for use in calculating emission factors for California vehicles. The poor manufacturer representation increases the chance for bias in technology specific emission factors.

(4) Engineering expectations that non-linear deterioration occurs in low mileage vehicles is suggested but not confirmed through data analysis. Linear deterioration was observed over the range of 10,000 through 70,000 miles. Little is known about the FTP emissions from vehicles with higher mileages. A summary assessment within the range of the existing surveillance data indicates that linear emission factors are adequate predictions of in-use emissions and do not over predict values observed in the data.

(5) A bounding analysis indicates that the maximum reduction in lifetime (100,000 miles) emissions that can be expected through quantification of non-linear behavior is on the order of 10 percent. The surveillance data collection effort required to validate nonlinear behavior is prohibative and confirmation is not assumed. Validation of non-linear behavior is possible through the use idle test data collected by BAR from large samples of California vehicles.

(6) The use of VMT as a surrogate for explanatory variables describing the operating condition of the emission control system obscures the influence of specific malperformances. More sophisticated specifications are available to describe the influence of these variables on in-use emissions. The use idle mixture maladjustments and selected variables describing significant system performance can produce increased explanation of the large variance observed in the data. Nevertheless these models do not significantly improve the predictive capability of the linear emission factor forms.

(7) The use of malperformance variables in emission factor specifications requires the prediction/quantifications of the occurrence of these malperformances in future model years. The marginal gain in accuracy associated with these models is outweighted by the complexity of the specification for use in emission modeling. These specifications however are very useful as a research tool to quantify the influence of system performance on pollutants emitted. They also represent a useful tool to evaluate the benefits of anti-tampering and mandatory maintenance or I/M programs for in-use emission reductions.

APPENDIX A

DESCRIPTIVE STATISTICS FOR EPA THREE-WAY CATALYST DATA

```
                          EPA DATA

                    (3 WAY CAT DATA ONLY)

                     TABLE OF MFR BY MY

MFR          MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |75      |76      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+
      35 |      0 |      0 |      5 |      1 |      6
         |   0.00 |   0.00 |   0.83 |   0.17 |   1.00
         |   0.00 |   0.00 |  83.33 |  16.67 |
         |   0.00 |   0.00 |   2.91 |   0.24 |
---------+--------+--------+--------+--------+
AUDI     |      0 |      0 |      0 |      4 |      4
         |   0.00 |   0.00 |   0.00 |   0.67 |   0.67
         |   0.00 |   0.00 |   0.00 | 100.00 |
         |   0.00 |   0.00 |   0.00 |   0.95 |
---------+--------+--------+--------+--------+
CHRYSLER |      0 |      0 |      2 |      5 |      7
         |   0.00 |   0.00 |   0.33 |   0.83 |   1.17
         |   0.00 |   0.00 |  28.57 |  71.43 |
         |   0.00 |   0.00 |   1.16 |   1.18 |
---------+--------+--------+--------+--------+
DATSUN   |      1 |      0 |      2 |      3 |      6
         |   0.17 |   0.00 |   0.33 |   0.50 |   1.00
         |  16.67 |   0.00 |  33.33 |  50.00 |
         |  25.00 |   0.00 |   1.16 |   0.71 |
---------+--------+--------+--------+--------+
FORD     |      2 |      0 |     52 |    162 |    216
         |   0.33 |   0.00 |   8.67 |  27.00 |  36.00
         |   0.93 |   0.00 |  24.07 |  75.00 |
         |  50.00 |   0.00 |  30.23 |  38.30 |
---------+--------+--------+--------+--------+
GM       |      1 |      0 |     49 |     37 |     87
         |   0.17 |   0.00 |   8.17 |   6.17 |  14.50
         |   1.15 |   0.00 |  56.32 |  42.53 |
         |  25.00 |   0.00 |  28.49 |   8.75 |
---------+--------+--------+--------+--------+
TOTAL           4        1      172      423      600
             0.67     0.17    28.67    70.50   100.00
```

```
                         EPA DATA

                  (3 WAY CAT DATA ONLY)

                    TABLE OF MFR BY MY

MFR          MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |75      |76      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+
HONDA    |      0 |      1 |      0 |      0 |      1
         |   0.00 |   0.17 |   0.00 |   0.00 |   0.17
         |   0.00 | 100.00 |   0.00 |   0.00 |
         |   0.00 | 100.00 |   0.00 |   0.00 |
---------+--------+--------+--------+--------+
MAZDA    |      0 |      0 |      0 |    114 |    114
         |   0.00 |   0.00 |   0.00 |  19.00 |  19.00
         |   0.00 |   0.00 |   0.00 | 100.00 |
         |   0.00 |   0.00 |   0.00 |  26.95 |
---------+--------+--------+--------+--------+
SAAB     |      0 |      0 |      2 |      3 |      5
         |   0.00 |   0.00 |   0.33 |   0.50 |   0.83
         |   0.00 |   0.00 |  40.00 |  60.00 |
         |   0.00 |   0.00 |   1.16 |   0.71 |
---------+--------+--------+--------+--------+
TOYOTA   |      0 |      0 |      2 |     27 |     29
         |   0.00 |   0.00 |   0.33 |   4.50 |   4.83
         |   0.00 |   0.00 |   6.90 |  93.10 |
         |   0.00 |   0.00 |   1.16 |   6.38 |
---------+--------+--------+--------+--------+
VOLVO    |      0 |      0 |     57 |     57 |    114
         |   0.00 |   0.00 |   9.50 |   9.50 |  19.00
         |   0.00 |   0.00 |  50.00 |  50.00 |
         |   0.00 |   0.00 |  33.14 |  13.48 |
---------+--------+--------+--------+--------+
VW       |      0 |      0 |      1 |     10 |     11
         |   0.00 |   0.00 |   0.17 |   1.67 |   1.83
         |   0.00 |   0.00 |   9.09 |  90.91 |
         |   0.00 |   0.00 |   0.58 |   2.36 |
---------+--------+--------+--------+--------+
TOTAL           4        1      172      423      600
             0.67     0.17    28.67    70.50   100.00
```

```
                         EPA DATA
                  (3 WAY CAT DATA ONLY)
                    TABLE OF TECH BY MY

TECH          MY

FREQUENCY    |
 PERCENT     |
 ROW PCT     |
 COL PCT     |75      |76      |78      |79      |  TOTAL
-------------+--------+--------+--------+--------+
 3-WAY  CAT  |      4 |      1 |    172 |    423 |    600
             |   0.67 |   0.17 |  28.67 |  70.50 | 100.00
             |   0.67 |   0.17 |  28.67 |  70.50 |
             | 100.00 | 100.00 | 100.00 | 100.00 |
-------------+--------+--------+--------+--------+
TOTAL               4        1      172      423      600
                 0.67     0.17    28.67    70.50   100.00
```

```
                          EPA DATA

                   (3 WAY CAT DATA ONLY)

                     TABLE OF CYL BY MY

CYL           MY

FREQUENCY|
 PERCENT |
 ROW PCT |
 COL PCT |75      |76      |78      |79      |  TOTAL
---------+--------+--------+--------+--------+
 4       |      1 |      1 |    143 |    284 |    429
         |   0.17 |   0.17 |  23.83 |  47.33 |  71.50
         |   0.23 |   0.23 |  33.33 |  66.20 |
         |  25.00 | 100.00 |  83.14 |  67.14 |
---------+--------+--------+--------+--------+
 5       |      0 |      0 |      0 |      4 |      4
         |   0.00 |   0.00 |   0.00 |   0.67 |   0.67
         |   0.00 |   0.00 |   0.00 | 100.00 |
         |   0.00 |   0.00 |   0.00 |   0.95 |
---------+--------+--------+--------+--------+
 6       |      2 |      0 |     24 |     45 |     71
         |   0.33 |   0.00 |   4.00 |   7.50 |  11.83
         |   2.82 |   0.00 |  33.80 |  63.38 |
         |  50.00 |   0.00 |  13.95 |  10.64 |
---------+--------+--------+--------+--------+
 8       |      1 |      0 |      5 |     90 |     96
         |   0.17 |   0.00 |   0.83 |  15.00 |  16.00
         |   1.04 |   0.00 |   5.21 |  93.75 |
         |  25.00 |   0.00 |   2.91 |  21.28 |
---------+--------+--------+--------+--------+
TOTAL           4        1      172      423      600
             0.67     0.17    28.67    70.50   100.00
```

```
                                  EPA DATA
                            (3 WAY CAT DATA ONLY)
                             TABLE OF TRANS BY MY

          TRANS        MY

          FREQUENCY|
           PERCENT |
           ROW PCT |
           COL PCT |75      |76      |78      |79      |  TOTAL
          ---------+--------+--------+--------+--------+
           AUTO    |      3 |      0 |     90 |    210 |    303
                   |   0.50 |   0.00 |  15.00 |  35.00 |  50.50
                   |   0.99 |   0.00 |  29.70 |  69.31 |
                   |  75.00 |   0.00 |  52.33 |  49.65 |
          ---------+--------+--------+--------+--------+
          SEMIAU   |      0 |      1 |      0 |      0 |      1
                   |   0.00 |   0.17 |   0.00 |   0.00 |   0.17
                   |   0.00 | 100.00 |   0.00 |   0.00 |
                   |   0.00 | 100.00 |   0.00 |   0.00 |
          ---------+--------+--------+--------+--------+
           3 SP    |      0 |      0 |      0 |      2 |      2
                   |   0.00 |   0.00 |   0.00 |   0.33 |   0.33
                   |   0.00 |   0.00 |   0.00 | 100.00 |
                   |   0.00 |   0.00 |   0.00 |   0.47 |
          ---------+--------+--------+--------+--------+
           4 SP    |      1 |      0 |     78 |    158 |    237
                   |   0.17 |   0.00 |  13.00 |  26.33 |  39.50
                   |   0.42 |   0.00 |  32.91 |  66.67 |
                   |  25.00 |   0.00 |  45.35 |  37.35 |
          ---------+--------+--------+--------+--------+
           5 SP    |      0 |      0 |      4 |     53 |     57
                   |   0.00 |   0.00 |   0.67 |   8.83 |   9.50
                   |   0.00 |   0.00 |   7.02 |  92.98 |
                   |   0.00 |   0.00 |   2.33 |  12.53 |
          ---------+--------+--------+--------+--------+
          TOTAL           4        1      172      423      600
                       0.67     0.17    28.67    70.50   100.00
```

```
                                                   EPA DATA

                                            (3 WAY CAT DATA ONLY)
                                               CARGRP=MY 77-79

                                             FREQUENCY BAR CHART
MIDPOINT
  VMTX                                                                           FREQ CUM.   PERCENT   CUM.
                                                                                       FREQ          PERCENT
         |
     0   |****************                                                         78    78    13.11    13.11
         |
     1   |********************************************************************** 349   427    58.66    71.76
         |
     2   |***********************                                                 116   543    19.50    91.26
         |
     3   |********                                                                 39   582     6.55    97.82
         |
     4   |**                                                                        8   590     1.34    99.16
         |
     5   |*                                                                         4   594     0.67    99.83
         |
     6   |                                                                          0   594     0.00    99.83
         |
     7   |                                                                          1   595     0.17   100.00
         |
     8   |                                                                          0   595     0.00   100.00
         |
     9   |                                                                          0   595     0.00   100.00
         |
    10   |                                                                          0   595     0.00   100.00
         |
    11   |                                                                          0   595     0.00   100.00
         |
    12   |                                                                          0   595     0.00   100.00
         |
         ----+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+--
            20  40  60  80  100 120 140 160 180 200 220 240 260 280 300 320 340

                                        FREQUENCY
```

ASSESSMENT OF MISFUELING TRENDS FOR CALIFORNIA VEHICLES

Contract No. A2-065-32
Task 1.B

Prepared for:

CALIFORNIA AIR RESOURCES BOARD
El Monte, California

Prepared by:

ENERGY AND ENVIRONMENTAL ANALYSIS, INC.
1655 North Fort Myer Drive
Arlington, Virginia 22209

June 1985

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (cont'd)

1. INTRODUCTION

Misfueling, the use of leaded gasoline in vehicles certified for the use of unleaded gasoline, is a major source of concern for all air quality regulatory agencies. It inflicts two separate penalties on the environment; first it increases ambient lead levels and second, coats the catalyst and degrades its performance and causes a substantial increase in regulated pollutants. Because of concern about the deleterious impacts of lead on human health and the need to reduce emission inventories to comply with SIP regulations both the California Air Resources Board (ARB) and the Environmental Protection Agency (EPA) have conducted numerous surveys of misfueling over the past 7 years.

Two types of surveys have been conducted. The first requires the use of personnel to either covertly or overtly observe consumers fueling their vehicles. Both EPA and ARB have conducted this type of survey. The second requires personnel to inspect vehicles at predetermined locations, take fuel samples and conduct tests to determine whether a vehicle has been misfueled. EPA has supported these surveys as part of its tampering surveys.

Misfueling cannot be represented by a single metric. There are many dimensions to this behavior. Measurements relevant from an air quality perspective include: the percentage of the catalyst fleet that misfuels in a calendar year; the frequency with which these vehicles misfuel; the number of catalyst vehicles with failed or significantly degraded catalysts due to misfueling. From an enforcement viewpoint the following measurements are of interest: the occurrence of misfueling (% of misfueling vehicles) at self serve vs. full serve stations; and the occurrence of misfueling in urban vs. rural areas. From an energy consumption viewpoint the following measurements are valued: the percentage of leaded gasoline consumed by catalyst vehicles and the percentage of unleaded gasoline consumed by lead certified vehicles.

It is important to understand which of the above misfueling measurements can be calculated from existing EPA and ARB surveys. The surveys of fueling behavior produce a one time record of the percentage of vehicles correctly fueling and misfueling. They cannot provide a measure of the percentage of the catalyst vehicles that misfuel across time (e.g., a calendar year), nor can they produce a measure of the frequency with which these vehicles misfuel. The latter statistic is required to determine the extent of catalyst damage (i.e., the number of consecutive misfuelings). The primary measure resulting from the survey of fueling behavior is the percentage of fuel consumed by catalyst vehicles which is leaded. This is a purchase weighted measure.

The EPA tampering surveys focus on a sample of vehicles at a single point in time but, because checks of the vehicle's condition are performed, it also carries a sense of history or time. The results of those surveys fall somewhere in between a single point in time and a survey of fueling habits across an entire year.* It is difficult to precisely define the measure of misfueling provided by these data.

An alternative measure of misfueling can be developed by observing the fuel purchasing behavior of a selected sample of catalyst vehicles over the course of a year. Such a survey requires the fuel purchase records

---

*The tampering survey checks three indications of misfueling: filler neck enlargement, lead in the fuel and a lead plumbtesmo test of the tail pipe. It is possible for an occasional (once or twice a year) misfueler not to be picked up in the survey. Basically chances for detection degrade with time as the lead content of the fuel decreases, providing the filler neck has not been disturbed.

of these vehicles indicating the type of fuel used. With this information, statistics on the percentage of catalyst vehicles that misfueled during the course of the entire year can be developed. This measure can be titled a fleet involvement rate as opposed to the purchase or volume weighted measure discussed above. With the fuel purchase records, data can also be generated on the frequency with which vehicles successively misfuel and the distribution of the catalyst fleet that misfuels based on the percentage of leaded fuel purchases. For example it can be shown that 4.2 percent of that catalyst fleet misfuels at least 50 percent of the time that fuel is purchased (based on 1981 data). On the other hand 2.8 percent of the catalyst fleet misfuels less than 11 percent of the time that fuel is purchased (based again on 1981 data). This information can be used to evaluate the condition of the catalyst. Overall the diary data of fuel purchase records provides a data base that can be used to produce many measures of misfueling

Energy and Environmental Analysis (EEA), Inc. has access to a data base that can provide information on misfueling behavior both for California vehicles and for 49 state vehicles for calendar years 1981 and 1982. This data has been used to support misfueling analysis for EPA, the Department of Energy (DOE), and private sector clients. This information has been used by EPA to evaluate misfueling activity, by DOE to evaluate the impact of misfueling on the distribution of leaded and unleaded gasoline purchases, and by private sector clients to support comments of lead phasedown rulemakings.

In this report EEA provides a detailed analysis of misfueling activity for calendar years 1981 and 1982 using the NPD Petroleum Marketing Survey

Index data for California vehicles. The results of this analysis are then compared with the results of published ARB misfueling surveys.* The report is organized into the following sections:

- **Section 2** provides a brief overview of the data base and methods employed to calculate the alternative rates of misfueling behavior. Also presented is a summary of the statistics produced characterizing both volume weighted and fleet involvement misfueling rates for 1981 and 1982.
- **Section 3** provides a comparison of the NPD California results with aggregate national values determined from the same data set. Also compared are the California NPD statistics and published ARB results from the same time periods.
- **Section 4** provides a summary of results and comparisons.
- **Appendix A** provides a detailed documentation of the NPD data base and methods used to screen the data. Also presented is a comparison of the results (age distribution, travel characteristics) among NPD and independent data sets to establish the validity of the data base.
- **Appendix B** provides an overview of the methods used to calculate estimated errors.
- **Appendix C** contains examples of purchase logs used by NPD survey participants.

---

*Because of the difficulty in identifying the measures of misfueling resulting from EPA's tampering surveys it was decided to forego comparisons with the EPA data.

2. ANALYSIS OF MISFUELING FOR 1981 AND 1982 CALENDAR YEARS FOR CALIFORNIA VEHICLES

2.1 OVERVIEW OF THE NPD DATA BASE

The purpose of this section is to quantify the rate of misfueling by light-duty vehicle (LDV) owners/operators located in the state of California during 1981. Due to data limitations, light-duty trucks are specifically excluded from the analysis. The survey information is derived from the NPD Petroleum Marketing Index (NPD), a diary panel survey of over 5,000 households conducted by NPD Research Inc. The tables and accompanying descriptive notes in this section present findings without any attempt at interpretation.

Misfueling may be measured in a variety of ways; the appropriate method depends upon the questions to be answered. In this study the misfueling rate is measured by fleet involvement, i.e., the proportion of all catalyst vehicles which are misfueled. Vehicles are categorized on the basis of whether or not they are ever misfueled, on the ratio of leaded fuel purchased to total fuel purchased, and on the maximum number of successive leaded purchases made during the survey period. Most of the tables presented here are aggregated across all model years.

Most previous studies of misfueling have sampled a cross-section of the vehicle population at one point in time. A major advantage of using a diary panel survey is that individuals may be followed through time.

Such a survey provides more complete information about the frequency of misfueling and could allow for detailed studies of the demographic characteristics of misfuelers or motivational factors.

The most common reservation about the use of diary surveys is that they depend upon consistent and truthful self-reporting. Despite concerns about respondents' potential unwillingness to incriminate themselves, the panel participants were quite open about their purchasing behavior and freely indicated the purchase of leaded fuel. The participants know their responses are being collected for gasoline brand market share studies and they are accustomed to reporting detailed information about what they have purchased. The participants do not know that government agencies purchase the raw survey data for studies such as this one. Furthermore, the participants are guaranteed anonymity by NPD Research, Inc. when they agree to participate.

2.2 DATA BASE PREPARATION

The NPD data base contains fuel purchase histories for over 12,000 privately operated vehicles. The data, collected during 1981, contain detailed information about fuel purchases, including date, gallons, type of fuel, and total cost. An example of the purchase logs filled in by respondents is shown in Appendix 2. There is body style/engine information as well as household demographic data associated with each vehicle purchase history. This data base has been used extensively by the U.S. Department of Energy (DOE) to examine trends in fuel consumption, on-road fuel economy and vehicle miles of travel. During the course of this previous work the data were cleaned and established as a SAS data set. As part of this work for DOE, the engine description information provided by the survey respondent was verified (and corrected when necessary using information extracted from the vehicle identification number (VIN).* Based on this engine information, each vehicle has been classified as to whether or not it has a catalyst.

---

*The VIN is reported by the owner, along with the engine description information, when a vehicle first enters the survey.

As part of a recent work effort for the U.S. Environmental Protection Agency (EPA), the catalyst information has been re-examined and verified by EEA. Identification of a catalyst equipped vehicle is made on the basis of the VIN-augmented data for make/model, model year, CID, number of cylinders, fuel system, and type of transmission.

The information presented in this section is based on an analysis of 408 catalyst equipped LDVs from NPD that are regularly garaged in the state of California. A small number of LDVs manufactured during the late seventies required unleaded fuel but were not actually equipped with a catalyst. We have assumed that the ultimate use, if any, of this analysis will be for estimating the effect of misfueling on catalyst vehicles. For this purpose, a misfueling rate among catalyst vehicles is sufficient. Therefore, those vehicles not having a catalyst but requiring unleaded were not included in the study. Another group of LDVs excluded from the study were those participating for less than two months. For the most part, respondents with only one month of participation have very poor record-keeping practices and incomplete purchase histories. Frequently only one or two purchases are reported and typically consist mostly of missing information. A total of 12 vehicles were deleted from the survey for participating less than two months.

In any large data collection effort there is a potential for recording or transcribing errors. To avoid over-reporting the incidence of misfueling, only those vehicles recording at least three leaded purchases during the year are counted as misfuelers. If no more than two leaded purchases are reported, a data error is assumed and the fuel type designation is changed to unleaded. A total of 62 vehicles meet this maximum-of-two leaded purchases criteria. A total of 175 "purchases" showing 0.0 gallons of fuel bought were deleted. In general, these records are null entries representing months when a diary was returned but no fuel was purchased.

2.3 RESULTS OF MISFUEING ANALYSIS FOR 1981

The tables in this report highlight misfueling behavior in the catalyst fleet. They also provide information pertaining to the manner in which the data are weighted and to the way in which unknown fuel type purchases are handled. Each table is prefaced with explanatory notes to assist the reader in interpreting the information presented.

Table 2-1 - Highlights of Misfueling Frequency Distributions - 1981

- Data in this table are weighted on the basis of the NPD projection factors. For information on the effect of alternative weighting methods, see Table 3.
- Total fuel purchased includes leaded, unleaded, and type unknown. For more detail on the treatment of purchases with unknown fuel type, and its effect on the findings, see Table 4.
- Leaded fuel as a percent of total fuel purchased by the catalyst fleet measures misfueling on a gallons purchased basis.
- The leaded fuel under 11 percent and 91-100 percent of total fuel purchased categories measure misfueling on a vehicle basis. Each vehicle's degree of involvement is judged on the basis of how much of their purchase volume is leaded. The vehicle is then assigned to an appropriate category. So, for example, 2.8 percent of the vehicles in the catalyst fleet were misfuelers whose leaded purchases amounted to less than 11 percent of the fuel they purchased during the year. By comparison, 2.6 percent of the vehicles in the catalyst fleet purchased 91 to 100 percent leaded fuel by volume.
- Catalyst fleet involvement in misfueling includes any catalyst vehicle that ever purchased leaded fuel, regardless of quantity or percentage of total fuel purchased over the year.
- There is a small number of vehicles (less than 0.1 percent) of the catalyst fleet who purchased leaded fuel, but less than 10 gallons worth. These vehicles are included in the "Leaded Fuel Under 11% of Total" category regardless of actual percentage.

TABLE 2-1

HIGHLIGHTS OF MISFUELING FREQUENCY DISTRIBUTIONS
1981

Number of Catalyst Equipped Vehicles	408
Leaded Fuel as Percent of Total Fuel Purchased by the Catalyst Fleet	3.6 (1.8)*
Leaded Fuel Under 11% of Total Fuel Purchased by Vehicle (percent of catalyst fleet)	2.8 (1.6)
Leaded Fuel 91-100% of Total Fuel Purchased by Vehicle (percent of catalyst fleet)	2.6 (1.5)
Catalyst Fleet Involvement in Misfueling (percent of catalyst fleet)	9.3 (2.8)
Purchased at least 10 gallons of Leaded Fuel (percent of catalyst fleet)	9.3 (2.8)
Purchased leaded fuel at least 3 times from January to June 1981 and not at all from July to December 1981 (percent of catalyst fleet)	1.4 (1.2)
Purchased leaded fuel at least 3 times from January to June 1981 ´percent of catalyst fleet)	6.4 (2.4)

---

*Values in parentheses are estimated errors in the percent of catalyst fleet.

Table 2-2 - Distribution of Misfueling by Degree of Involvement - 1981

- In the table, vehicle involvement in misfueling is measured by the ratio of leaded fuel to total fuel purchased. A vehicle purchasing a total of 400 gallons of fuel (all types) during 1981, of which 30 gallons are leaded, has a ratio of 30 to 400 or 7.5 percent. This vehicle is placed in the under 11 percent leaded category. Had the same vehicle purchased 350 gallons of leaded, out of 400 gallons total, the ratio would be 87.5 percent leaded and the vehicle would be placed in the 81-90 percent leaded category.

- 2.8 percent of the catalyst fleet, or 30.1 percent of the misfuelers, have leaded fuel purchases totaling less than 11 percent of their annual fuel purchases. By comparison, 2.6 percent of the fleet, or 27.9 percent of misfuelers, purchased 91-100 percent leaded fuel by volume.

TABLE 2-2

DISTRIBUTION OF MISFUELING
BY DEGREE OF INVOLVEMENT
Aggregated Across All Model Years
1981

Leaded Fuel Purchased by Vehicle as Percent of Total Fuel Purchased	Percent of Catalyst Fleet In Category	Estimated Error	Percent of Misfueling Vehicles
Under 11	2.8	1.6	30.1
11-20	1.1	1.0	11.8
21-30	0.8	0.9	8.6
31-40	0.2	0.5	2.2
41-50	0.2	0.4	2.2
51-60	0.9	0.9	9.7
61-70	0.0	0.0	0.0
71-80	0.5	0.7	5.4
81-90	0.2	0.4	2.2
91-100	2.6	1.5	27.9

Table 2-3 - Comparison of Weighting Methods Using NPD's National Data Base

1981

- Each vehicle contributes one observation to the misfueling analysis. When calculating the overall misfueling rate it is helpful if an individual vehicle's contribution can be weighted to account for its importance relative to other vehicles in the fleet. This table compares three methods of weighting.
- Sample weighting with each vehicle assigned a weight of one, does not distinguish among vehicles.
- The NPD projection factor weights are assigned to each household on a monthly basis by NPD Inc. The factors are designed to weight the sample, demographically, to the national level based on income, race, region, and the educational level and occupation of the female head of house. As respondents enter and leave the survey, each household projection factor is adjusted to maintain the national level weighting scheme. The weight used in this study is the sum of these factors over each month a vehicle participates in the survey.
- The Months in Survey method assigns a weight to each vehicle solely on the basis of the number of months a vehicle participates in the survey.
- Overall the three weighting methods produce similar results, although on a model year specific basis there are some differences. This is particularly true for model year 1982 where small sample size is a problem.
- Since the NPD projection factors were designed to weight the survey to a national level on the basis of household demographics, these factors are used in reporting all results except those in this table.

TABLE 2-3

COMPARISON OF WEIGHTING METHODS USING NPD'S NATIONAL DATA BASE

1981

	Vehicle Count NPD			Vehicle Involvement Misfueling Rate (percent) NPD			
Model Year	Sample Weight	Projection Factors	Months in Survey	Sample Weight	Estimated Error*	Projections Factors	Months in Survey
1975	266	366,200	2,665	21.1	4.9	20.7	21.4
1976	444	613,373	4,642	17.3	3.5	16.9	17.1
1977	598	835,144	6,056	16.4	3.0	17.1	16.5
1978	631	841,370	6,380	13.9	2.7	13.0	14.0
1979	520	714,709	5,246	12.7	2.9	12.6	13.0
1980	485	663,866	4,848	10.3	2.7	9.0	9.9
1981	387	402,127	2,649	12.9	3.3	12.7	14.4
1982	22	10,011	75	18.2	16.1	16.4	16.0
TOTAL	3,353	4,446,800	32,560	14.6	1.2	14.3	14.8

---

*Estimated errors would all be based on the unweighted vehicle count, hence they would be nearly identical across the weighting methods. In order to simplify comparisons of the fleet involvement percentages the error estimates have been included only for the sample weight calculation.

Table 2-4 - Comparison of Involvement Under Alternate Assumptions About Unknown Fuel Type Using NPD's National Data Base - 1981

- This table compares three methods of treating unknown fuel types. An unknown fuel type purchase is one in which the respondent has failed to check either the leaded or the unleaded column on the monthly diary log.
- If unknown fuel type is assumed to be leaded fuel, the overall fleet involvement in misfueling is 42.0 percent of the catalyst fleet.
- If unknown fuel type is assumed to be unleaded fuel, the overall fleet involvement in misfueling is 14.3 percent of the catalyst fleet.
- If the unknown fuel for each vehicle is allocated between leaded and unleaded, based on the percentages of known leaded and known unleaded bought for that vehicle, the overall fleet involvement in misfueling is 14.3 percent of the catalyst fleet.
- On average, each vehicle in the survey made 1.3 purchases of unknown fuel type during 1981. This fuel, roughly 12.6 gallons per vehicle, typically represents approximately 2.3 percent of the year's total fuel purchases. Including all of these purchases in the leaded category increases fleet involvement in misfueling by 190 percent.
- The assumption that unknown fuel purchases actually represent leaded fuel is made to test the belief that consumers do not wish to implicate themselves in misfueling. While there may be a handful of respondents whose behavior fits this pattern, most individuals appear to be extremely forthcoming about their misfueling habits. Given the wide distribution of unknown fuel type purchases and people's willingness to report buying leaded fuel, it is likely that most, though not all, unknown fuel purchases are the result of recording error rather than of half-hearted deception.
- The assumption that unknown fuel is unleaded is the most conservative method of allocating unknown fuel. As may be seen in the table, the results are almost identical to those obtained by allocating the unknown fuel between leaded and unleaded.
- All of the tables in this report are based on the assumption that unknown fuel may reasonably be allocated between leaded and unleaded on the basis of the percentages of known leaded and known unleaded bought for an individual vehicle.

TABLE 2-4

COMPARISON OF INVOLVEMENT UNDER ALTERNATE ASSUMPTIONS ABOUT UNKNOWN FUEL TYPE USING NPD'S NATIONAL DATA BASE 1981

		Vehicle Involvement Misfueling Rate					
		Unknown is Unleaded		Unknown is Leaded		Unknown Allocated	
Model Year	Number of Vehicles	Rate (% of fleet)	Estimated Error	Rate (% of fleet)	Estimated Error	Rate (% of fleet)	Estimated Error
1975	266	20.7	4.9	40.8	5.9	20.7	4.9
1976	444	16.9	3.5	39.1	4.5	16.9	3.5
1977	598	17.1	3.0	45.4	4.0	17.1	3.0
1978	631	13.0	2.6	45.7	3.9	13.0	2.6
1979	520	12.6	2.9	40.7	4.2	12.6	2.9
1980	485	9.0	2.5	42.5	4.4	9.0	2.6
1981	387	12.7	3.3	34.5	4.7	12.7	3.3
1982	22	16.4	15.5	45.3	20.8	16.4	15.5
OVERALL	3,353	14.3	1.2	42.0	1.7	14.3	1.2

Table 2-5 - Misfueling Fleet Involvement by Model Year - 1981

- Number of vehicles is a count of the actual, unweighted, number of LDVs in each model year.

- The categories reported here are identical to the third, fourth and fifth items in Table 1. For example, for model year 1976, of which there are 54 catalyst equipped LDVs in the survey, 8.4 percent were misfuelers whose leaded purchases amounted to less than 11 percent of the fuel they purchased during the year. At the same time, 4.5 percent of the model year vehicles purchased 91-100 percent leaded fuel by volume. Overall, 17.0 percent of the model year 1976 vehicles misfueled at least part of the time.

- In general it is assumed that misfueling will increase with vehicle age. Moving backwards from model year 1978, the percentage of vehicles in the 91-100 percent leaded category increases steadily with vehicle age. The involvement rate for model year 1982 is suspect due to small sample size.

TABLE 2-5

VEHICLE INVOLVEMENT MISFUELING RATES BY MODEL YEAR
1981

Model Year	Number of Vehicles	Fuel Under 11% Leaded (% of Fleet)	Fuel 91-100% Leaded (% of Fleet)	Overall Vehicle Involvement (% of Fleet)	Estimated Error in Involvement
1975	32	0.0	10.4	10.4	10.6
1976	54	8.4	4.5	17.0	10.0
1977	66	0.0	2.2	2.6	3.8
1978	78	5.2	1.1	8.8	6.3
1979	61	3.0	1.2	9.4	7.3
1980	69	0.0	0.0	6.9	6.0
1981	44	4.0	0.0	15.4	10.7
1982	4	0.0	64.4	72.9	43.6
OVERALL	408	2.8	2.6	9.3	2.8

Table 2-6 - Repeated Misfueling - 1981

- This table displays the incidence of successive misfueling for the catalyst fleet as a whole. Vehicle involvement rates are percents of the catalyst fleet.

- Vehicles having made at least two leaded purchases in a row are assigned to one of five purchasing categories. The assignment is based on the longest string of leaded purchases made by that vehicle during 1981.

- Vehicles making only singleton purchases of leaded gasoline will not appear in this table. Thus, although 9.3 percent of the fleet misfueled at least once (see Table 5), only 8.9 percent (the sum of the five purchasing categories) of the fleet is represented in Table 6. The remaining 0.4 percent of the fleet that misfueled never purchased leaded twice in a row.

- Percent of leaded purchases is calculated over the entire catalyst fleet. For example, 86.1 percent of the leaded purchases made by the fleet were made by vehicles that have purchased leaded at least 6 times in a row.

- Category assignments are exclusive. A vehicle making two leaded purchases in a row on several occasions, and four leaded purchases in a row on one occasion will be assigned only to the category for vehicles having made four successive leaded purchases.

TABLE 2-6

REPEATED MISFUELING
1981

Maximum Number of Successive Leaded Purchases During the Year	Percent of Catalyst Fleet Vehicles Involved	Estimated Error	Percent of Leaded Purchases By Fleet
2	0.6	0.8	1.3
3	2.1	1.4	4.6
4	0.6	0.8	1.6
5	1.0	1.0	5.8
6 or more	4.6	2.0	86.1

Table 2-7 - Length of Survey Participation - 1981

- In order to be included in this misfueling study a vehicle must have provided data for at least two months. Vehicles reporting for only one month generally provide purchase records with much missing or inconsistent information. In order to reduce the effect of missing data, the minimum reporting requirement was adopted. A total of 12 vehicles were eliminated as a result of this requirement.

- The majority of vehicles contributed a full 12 months of data.

TABLE 2-7

LENGTH OF SURVEY PARTICIPATION
1981

Number of Months In Survey During 1981	Number of Vehicles	Percent of Catalyst Fleet
2	12	2.9
3	9	2.2
4	14	3.4
5	20	4.9
6	8	2.0
7	11	2.7
8	17	4.2
9	16	3.9
10	11	2.7
11	20	4.9
12	270	66.2

2.4 RESULTS OF MISFUELING ANALYSIS FOR 1982

The purpose of this section is to quantify the misfueling rate of catalyst vehicles during 1982. The 1981 and 1982 NPD survey data were processed in the same manner. There are 422 catalyst equipped LDV's regularly garaged in the state of California in the 1982 NPD survey. A total of 16 vehicles were deleted from the survey for participating less than 2 months. A total of 44 vehicles met the maximum-of-two leaded purchase criteria and their leaded purchases were treated as unleaded. In addition, a total of 172 "purchases" showing 0.0 gallons of fuel bought were deleted.

The degree of misfueling increased between 1981 and 1982. This is true on both a volume and a vehicle involvement basis. The following tables highlight the 1982 rates and are structured so as to allow for a direct comparison to the tables presented for 1981. To simplify the presentation the tables describing the weighting methods (Table 2-3) and a comparison of assumptions about unknown fuel types (Table 2-4) have not been included for 1982. The information in the 1982 tables is basically the same as for 1981. Given the number of tables already included it was decided to relieve the reader of yet another set of tables.

Table 2-8 - Highlights of Misfueling Frequency Distributions - 1982

- Data in this table are weighted on the basis of the NPD projection factors. For information on the effect of alternative weighting methods, see Table 2-3.

- Total fuel purchased includes leaded, unleaded, and type unknown. For more detail on the treatment of purchases with unknown fuel type, and its effect on the findings, see Table 2-4.

- Leaded fuel as a percent of total fuel purchased by the catalyst fleet measures misfueling on a gallons purchased basis.

- The leaded fuel under 11 percent and 91-100 percent of total fuel purchased categories measure misfueling on a vehicle basis. Each vehicle's degree of involvement is judged on the basis of how much of their purchase volume is leaded. The vehicle is then assigned to an appropriate category. So, for example, 5.7 percent of the vehicles in the catalyst fleet were misfuelers whose leaded purchases amounted to less than 11 percent of the fuel they purchased during the year. By comparison, 2.4 percent of the vehicles in the catalyst fleet purchased 91 to 100 percent leaded fuel by volume.

- Catalyst fleet involvement in misfueling includes any catalyst vehicle that ever purchased leaded fuel, regardless of quantity or percentage of total fuel purchased over the year.

- There is a small number of vehicles (less than 0.1 percent) of the catalyst fleet who purchased leaded fuel, but less than 10 gallons worth. These vehicles are included in the "Leaded Fuel Under 11% of Total" category regardless of actual percentage.

TABLE 2-8

HIGHLIGHTS OF MISFUELING FREQUENCY DISTRIBUTIONS 1982

Number of Catalyst Equipped Vehicles	422
Leaded Fuel as Percent of Total Fuel Purchased by the Catalyst Fleet	4.2 (1.9)*
Leaded Fuel Under 11% of Total Fuel Purchased by Vehicle (percent of catalyst fleet)	4.7 (2.2)
Leaded Fuel 91-100% of Total Fuel Purchased by Vehicle (percent of catalyst fleet)	2.4 (1.5)
Catalyst Fleet Involvement in Misfueling (percent of catalyst fleet)	13.6 (3.3)
Purchased at least 10 gallons of LeadedFuel (percent of catalyst fleet)	13.6 (3.3)
Purchased leaded fuel at least 3 times from January to June 1981 and not at all from July to December 1981 (percent of catalyst fleet)	2.8 (1.6)
Purchased leaded fuel at least 3 times from January to June 1981 (percent of catalyst fleet)	6.4 (2.3)

---

*Values in parentheses are estimated errors in the percent of catalyst fleet.

Table 2-9 - Distribution of Misfueling by Degree of Involvement - 1982

- In the table, vehicle involvement in misfueling is measured by the ratio of leaded fuel to total fuel purchased. A vehicle purchasing a total of 400 gallons of fuel (all types) during 1982, of which 30 gallons are leaded, has a ratio of 30 to 400 or 7.5 percent. This vehicle is placed in the under 11 percent leaded category. Had the same vehicle purchased 350 gallons of leaded, out of 400 gallons total, the ratio would be 87.5 percent leaded and the vehicle would be placed in the 81-90 percent leaded category.

- 5.7 percent of the catalyst fleet, or 41.9 percent of the misfuelers, have leaded fuel purchases totaling less than 11 percent of their annual fuel purchases. By comparison, 2.4 percent of the fleet, or 17.6 percent of misfuelers, purchased 91-100 percent leaded fuel by volume.

TABLE 2-9

DISTRIBUTION OF MISFUELING BY DEGREE OF INVOLVEMENT

Aggregated Across All Model Years
1982

Leaded Fuel Purchased by Vehicle as Percent of Total Fuel Purchased	Percent of Catalyst Fleet In Category	Estimated Error	Percent of Misfueling Vehicles
Under 11	5.7	2.2	41.9
11-20	3.7	1.8	27.2
21-30	0.7	0.8	5.1
31-40	0.0	0.0	0.0
41-50	0.0	0.0	0.0
51-60	0.6	0.8	4.4
61-70	0.0	0.0	0.0
71-80	0.3	0.5	2.2
81-90	0.2	0.4	1.5
91-100	2.4	1.5	17.6

Table 2-10 - Misfueling Fleet Involvement by Model Year - 1982

- Number of vehicles is a count of the actual, unweighted, number of LDVs in each model year.

- The categories reported here are identical to the third, fourth and fifth items in Table 1. For example, for model year 1976, of which there are 53 catalyst equipped LDVs in the survey, 2.6 percent were misfuelers whose leaded purchases amounted to less than 11 percent of the fuel they purchased during the year. At the same time, 4.8 percent of the model year vehicles purchased 91-100 percent leaded fuel by volume. Overall, 17.9 percent of the model year 1976 vehicles misfueled at least part of the time.

- In general it is assumed that misfueling will increase with vehicle age. Moving backwards from model year 1979, the percentage of vehicles in the 91-100 percent leaded category increases steadily with vehicle age. The involvement rate for model year 1983 is suspect due to small sample size

TABLE 2-10

VEHICLE INVOLVEMENT MISFUELING RATES BY MODEL YEAR
1982

Model Year	Number of Vehicles	Fuel Under 11% Leaded (% of Fleet)	Fuel 91-100% Leaded (% of Fleet)	Overall Vehicle Involvement (% of Fleet)	Estimated Error in Involvement
1975	33	0.0	6.4	19.8	13.6
1976	53	2.6	4.8	17.9	10.3
1977	53	5.1	4.6	14.7	9.5
1978	74	7.6	3.0	15.5	8.2
1979	63	4.7	0.9	5.6	5.7
1980	59	8.0	0.0	16.4	9.4
1981	57	4.7	0.0	8.0	7.0
1982	26	18.8	0.0	18.8	15.0
1983	4	0.0	0.0	8.5	27.3
OVERALL	422	5.7	2.4	13.6	3.3

Table 2-11 - Repeated Misfueling - 1982

- This table displays the incidence of successive misfueling for the catalyst fleet as a whole. Vehicle involvement rates are percents of the catalyst fleet.
- Vehicles having made at least two leaded purchases in a row are assigned to one of five purchasing categories. The assignment is based on the longest string of leaded purchases made by that vehicle during 1982.
- Vehicles making only singleton purchases of leaded gasoline will not appear in this table. Thus, although 13.6 percent of the fleet misfueled at least once (see Table 3), only 13.5 percent (the sum of the five purchasing categories) of the fleet is represented in Table 4. The remaining 0.1 percent of the fleet that misfueled never purchased leaded twice in a row.
- Percent of leaded purchases is calculated over the entire catalyst fleet. For example, 82.6 percent of the leaded purchases made by the fleet were made by vehicles that have purchased leaded at least 6 times in a row.
- Category assignments are exclusive. A vehicle making two leaded purchases in a row on several occasions, and four leaded purchases in a row on one occasion will be assigned only to the category for vehicles having made four successive leaded purchases.

TABLE 2-11

REPEATED MISFUELING
1982

Maximum Number of Successive Leaded Purchases During the Year	Percent of Catalyst Fleet Vehicles Involved	Estimated Error	Percent of Leaded Purchases By Fleet
2	0.6	0.7	1.3
3	3.3	1.7	5.7
4	1.9	1.3	6.2
5	1.3	1.1	4.3
6 or more	6.4	2.3	82.6

Table 2-12 - Length of Survey Participation - 1982

- In order to be included in this misfueling study a vehicle must have provided data for at least two months. Vehicles reporting for only one month generally provide purchase records with much missing or inconsistent information. In order to reduce the effect of missing data, the minimum reporting requirement was adopted. A total of 16 vehicles were eliminated as a result of this requirement.

- The majority of vehicles contributed a full 12 months of data.

TABLE 2-12

LENGTH OF SURVEY PARTICIPATION 1982

Number of Months In Survey During 1982	Number of Vehicles	Percent of Catalyst Fleet
2	17	4.0
3	7	1.7
4	19	4.5
5	23	5.5
6	11	2.6
7	16	3.8
8	18	4.3
9	11	2.6
10	14	3.3
11	21	5.0
12	265	62.8

3. COMPARISON OF NPD CALIFORNIA SURVEY RESULTS WITH OTHER STUDIES

To provide a perspective to the misfueling statistics presented in the previous section it is useful to compare them with the results from the national level NPD analysis conducted for EPA and DOE for the same calendar years Also presented is a comparison with recently published ARB survey results.

3.1 COMPARISON BETWEEN CALIFORNIA AND NATIONAL NPD MISFUELING ANALYSES

Under contract to EPA and DOE, EEA conducted a detailed analysis of national misfueling behavior for calendar years 1981 and 1982.* That analysis employed the same data cleansing, vehicle selection and analytical techniques employed to produce the results in the previous section. To provide a perspective on the peformance of California cars versus the average of the nation as a whole (including California vehicles) most of the tables produced in the preceeding section are replicated with values from the EPA/DOE studies. The discussion format employed in previous section is also followed here. To facilitate the discussion the presentation is organized around the tables not the calendar year. For example, a discussion of Table 1 (Highlights of Misfueling Frequency Distributions) is followed by the results for 1981 and 1982.

---

*Analysis of Light-Duty Vehicle Fuel Switching In the NPD Data Base, prepared for EPA, Ann Arbor, Contract No. 68-01-6558, Work Assignment Nos. 30 and 30A, September 1984.

3.2 COMPARISON BETWEEN NPD CALIFORNIA RESULTS AND ARB SURVEYS

A comparsion undoubtedly of interest to ARB is between the findings presented in the previous section and the results of ARB surveys. Before providing such a comparison it is necessary to review the metrics or basis of measurement that each survey affords. As discussed in the introduction there are two basic measures of misfueling: vehicle involvement - the percentage of the fleet involved in misfueling; and purchase or volume weighted measures - basically the proportion of total fuel purchased by the catalyst fleet that is leaded and unleaded. These are fundamentally different measures of the same behavior and cannot be compared with each other.

The diary survey data allows the calculation of both metrics. From this data it is possible to determine whenever a vehicle misfuels, how often and what percent of the total fuel consumed is leaded and unleaded.* The survey data collected by ARB provides a one time or cross section of a vehicle's fueling behavior during the course of a year. It is impossible to determine the proportion of the fleet involved in misfueling other than at the time of the observation. This data is necessarily purchase weighted and the appropriate metric for comparison with the NPD data is the volume weighted value.

Table 3-6 presents a comparison of the aggregate values observed in NPD California vehicles and the results recently published by ARB; they are in close agreement. For 1981 there is little difference at all; for

---

*For a detailed discussion on the rules employed to clean the NPD data assumptions about possible errors in the data the reader is referred to Appendix A.

1982 NPD shows a value approximately double the result observed by ARB. It is unclear whether this difference implies a true disagreement between the reported levels of misfueling, or reflects instead two statistical factors: the relatively small size of the NPD California sample (with potential for larger variation year-to-year); and the specific geographic areas included in the 1981 and 1982 ARB surveys. It should also be noted that the NPD data does not include light trucks, which are generally regarded as having higher misfueling rates than passenger cars.* The inclusion of these vehicles in the ARB surveys may therefore lead to differences in the misfueling rates.

The values shown for NPD do not represent the percentage of vehicles involved in misfueling, only the volume of fuel or purchases that was of leaded fuel. The actual percentage of the catalyst fleet involved is substantially higher as discussed in Sections 2 and 3.

As a further check on the reasonableness of the data, we compared the results of ARB's and NPD's full serve and self serve misfueling rates. Without burdening the reader with yet another table let it suffice that similar levels of behavior were observed in both surveys. That is, a substantially higher level of misfueling occurred in self serve vs. full serve stations.

---

*EPA tampering surveys have indicated that trucks are more likely to misfuel than cars.

Table 3-1 - Highlights of Misfueling Frequency Distributions

- The sample size for California vehicles is substantially lower than for the national analyses. The smaller size of the California sample leads to a reduction in the statistical accuracy and is reflected in the large estimated errors vis-a-vis the national level data.
- Generally speaking the incidence of misfueling in California appears to be substantially lower than the rate observed for the nation as a whole.
- The level of misfueling activity on a national level increased from 1981 to 1982. The same pattern appears to have occured in California albeit at a lower level.
- The California level of leaded fuel as a percent of total fuel purchased by catalyst equipped vehicles appears to be substantially lower than that observed for the nation. On the other hand the differences in the overall catalyst fleet involved in misfueling appear to be less substantial. This would indicate that more people are misfueling occasionally in California than in the nation as a whole.
- Unlike the other categories there appears to be no increase in the level of dedicated (91-100% of fuel purchased) misfueling in California between 1981 and 1982. This is contrary to the behavior observed at the national level.

TABLE 3-1

HIGHLIGHTS OF MISFUELING FREQUENCY DISTRIBUTIONS - 1981

	National	California
Number of Catalyst Equipped Vehicles	3353	408
Leaded Fuel as Percent of Total Fuel Purchased by the Catalyst Fleet	6.4 (0.8)*	3.6 (1.8)
Leaded Fuel Under 11% of Total Fuel Purchased by Vehicle (percent of catalyst fleet)	5.6 (0.8)	2.8 (1.6)
Leaded Fuel 91-100% of Total Fuel Purchased by Vehicle (percent of catalyst fleet)	3.5 (0.6)	2.6 (1.5)
Catalyst Fleet Involvement in Misfueling (percent of catalyst fleet)	14.3 (1.2)	9.3 (2.8)
Purchased at least 10 gallons of Leaded Fuel (percent of catalyst fleet)	14.5 (1.2)	9.3 (2.8)
Purchased leaded fuel at least 3 times from January to June 1981 and not at all from July to December 1981 (percent of catalyst fleet)	3.0 (0.6)	1.4 (1.2)
Purchased leaded fuel at least 3 times from January to June 1981 (percent of catalyst fleet)	10.0 (1.0)	6.4 (2.4)

---

*Values in parentheses are estimated errors in the percent of catalyst fleet.

TABLE 3-1

HIGHLIGHTS OF MISFUELING FREQUENCY DISTRIBUTIONS - 1982

	U.S.	California
Number of Catalyst Equipped Vehicles	3694	422
Leaded Fuel as Percent of Total Fuel Purchased by the Catalyst Fleet	7.7 (0.9)*	4.2 (1.9)
Leaded Fuel Under 11% of Total Fuel Purchased by Vehicle (percent of catalyst fleet)	7.0 (0.8)	5.7 (2.2)
Leaded Fuel 91-100% of Total Fuel Purchased by Vehicle ´percent of catalyst fleet)	4.3 (0.6)	2.4 (1.5)
Catalyst Fleet Involvement in Misfueling (percent of catalyst fleet)	18.0 (1.2)	13.6 (3.3)
Purchased at least 10 gallons of Leaded Fuel (percent of catalyst fleet)	17.9 (1.2)	13.6 (3.3)
Purchased leaded fuel at least 3 times from January to June 1981 and not at all from July to December 1981 (percent of catalyst fleet)	2.0 (0.5)	2.8 (1.6)
Purchased leaded fuel at least 3 times from January to June 1981 (percent of catalyst fleet)	8.0 (0.9)	6.4 (2.3)

---

*Values in parentheses are estimated errors in the percent of catalyst fleet.

Table 3-2 - Distribution of Misfueling by Degree of Involvement

- The most interesting characteristic over both years is the polarity in misfueling behavior--that is people either misfuel very occasionally (less than 20 percent of the time) or they misfuel almost all of the time ´more than 90 percent). A very small percentage of the misfueling vehicles lies in between these two groups. This pattern of behavior is consistent for both California and the nation.

- The shape of the misfueling distribution for the nation appears relatively unchanged from 1981 to 1982. In California there appears to be a shift from 1981 to 1982 towards more occasional misfueling (and commensurately less dedicated misfueling). Because of the small sample size of the California vehicles it is impossible to confirm this shift as statistically significant. If however this trend could be confirmed it would have adverse implications for the environment. Ideally all misfueling (if it is to occur) should occur in the same vehicles so that the percent of the fleet experiencing catalyst damage is minimized. If however more of the fleet is misfueling even occasionally the opportunity for a greater portion of the catalyst fleet to suffer damage is increased.

TABLE 3-2

DISTRIBUTION OF MISFUELING BY DEGREE OF INVOLVEMENT - 1981
Aggregated Across All Model Years

Leaded Fuel Purchased by Vehicle as Percent of Total Fuel Purchased	Percent of Catalyst Fleet In Category		Estimated Error		Percent of Misfueling Vehicles	
	National	CA	National	CA	National	CA
Under 11	5.7	2.8	0.8	1.6	39.6	30.1
11-20	2.2	1.1	0.5	1.0	15.3	11.8
21-30	0.6	0.8	0.3	0.9	4.2	8.6
31-40	0.3	0.2	0.2	0.5	2.1	2.2
41-50	0.3	0.2	0.2	0.4	2.1	2.2
51-60	0.2	0.9	0.2	0.9	1.4	9.7
61-70	0.3	0.0	0.2	0.0	2.1	0
71-80	0.8	0.5	0.3	0.7	5.6	5.4
81-90	0.5	0.2	0.3	0.4	3.5	2.2
91-100	3.5	2.6	0.6	1.5	24.3	27.9
TOTAL	14.4	9.3	--	--	100.2*	100.1*

---

*Error due to round off

TABLE 3-2

DISTRIBUTION OF MISFUELING BY DEGREE OF INVOLVEMENT - 1982
Aggregated Across All Model Years

Leaded Fuel Purchased by Vehicle as Percent of Total Fuel Purchased	Percent of Catalyst Fleet In Category		Estimated Error		Percent of Misfueling Vehicles	
	National	CA	National	CA	National	CA
Under 11	7.0	5.7	0.8	2.2	38.7	41.9
11-20	3.4	3.7	0.6	1.8	18.8	27.2
21-30	0.8	0.7	0.3	0.8	4.4	5.1
31-40	0.4	0.0	0.2	0.0	2.2	0.0
41-50	0.4	0.0	0.2	0.0	2.2	0.0
51-60	0.4	0.6	0.2	0.8	2.2	4.4
61-70	0.1	0.0	0.1	0.0	0.5	0.0
71-80	0.4	0.3	0.2	0.5	2.2	2.2
81-90	0.9	0.2	0.3	0.4	5.0	1.5
91-100	4.3	2.4	0.7	1.5	23.8	17.6
TOTAL	18.1	13.6	--	--	100.0	99.9*

---

*Error due to round off

Table 3-3 - Misfueling Fleet Involvement by Model Year

- There is a trend towards increased misfueling as vehicles age ´overall vehicle involvement) observed in the national data. The relatively high rates of misfueling that occur for late model year vehicles must be discounted for the relatively small sample sizes. There appears to be a similar trend in misfueling as a function of age in the California fleet, however because of the small sample size and commensurate large estimated errors in involvement it is impossible to confirm the trend.

- There is also a trend towards increased dedicated (91-100 percent leaded) misfueling as a vehicle ages observable in the national data. This trend is also observable in the California data. Conversely the expected decline in occasional misfueling (less than 11 percent leaded) is not as obvious in either the national or the California data.

TABLE 3-3

VEHICLE INVOLVEMENT MISFUELING RATES BY MODEL YEAR - 1981

Model Year	Number of Vehicles		Fuel Under 11% Leaded (% of Fleet)		Fuel 91-100% Leaded (% of Fleet)		Overall Vehicle Involvement (% of Fleet)		Estimated Error in Involvement	
	US	CA	US	CA	US	CA	US	CA	US	CA
1975	266	32	5.1	0.0	9.0	10.4	20.7	10.4	4.9	10.6
1976	444	54	4.9	8.4	4.5	4.5	16.9	17.0	3.5	10.0
1977	598	66	7.2	0.0	5.1	2.2	17.1	2.6	3.0	3.8
1978	631	78	5.0	5.2	3.6	1.1	13.0	8.8	2.6	6.3
1979	520	61	5.4	3.0	2.0	1.2	12.6	9.4	2.9	7.3
1980	485	69	4.7	0.0	0.7	0.0	9.0	6.9	2.6	6.0
1981	387	44	7.5	4.0	0.0	0.0	12.7	15.4	3.3	10.7
1982	22	4	0.0	0.0	10.8	64.4	16.4	72.9	15.5	43.6
Overall	3,353	408	5.6	2.8	3.5	2.6	14.3	9.3	1.2	2.8

TABLE 3-3

VEHICLE INVOLVEMENT MISFUELING RATES BY MODEL YEAR - 1982

Model Year	Number of Vehicles		Fuel Under 11% Leaded (% of Fleet)		Fuel 91-100% Leaded (% of Fleet)		Overall Vehicle Involvement (% of Fleet)		Estimated Error in Involvement	
	US	CA	US	CA	US	CA	US	CA	US	CA
1975	280	33	4.5	0.0	9.2	6.4	25.0	19.8	5.1	13.6
1976	465	53	6.3	2.6	7.3	4.8	22.4	17.9	3.8	10.3
1977	586	53	6.8	5.1	8.8	4.6	22.0	14.7	3.4	9.5
1978	632	74	6.5	7.6	4.2	3.0	17.1	15.5	2.9	8.2
1979	590	63	6.6	4.7	1.6	0.9	14.8	5.6	2.9	5.7
1980	450	59	9.7	8.0	1.4	0.0	16.0	16.4	3.4	9.4
1981	485	57	7.7	4.7	0.3	0.0	13.0	8.0	3.0	7.0
1982	185	26	7.8	18.8	0.6	0.0	16.6	18.8	5.4	15.0
1983	21	4	0.0	0.0	0.0	0.0	16.3	8.5	15.8	27.3
Overall	3,694	422	7.0	5.7	4.3	2.4	18.0	13.6	1.2	3.3

Table 3-4 - Repeated Misfueling

- As previously discussed, a major source of concern is the frequency with which vehicles successively misfuel. The higher the number of succeeding misfuelings the greater the chance for permanent catalyst damage. The number of successive misfuelings required to inflict irreparable catalyst damage varies depending on the source consulted. To present a conservative view of this issue we have selected six or more repeated misfuelings as the threshold for this phenomena.

- The national data indicates a consistent trend towards an increased percentage of the model year fleet involved in six or more repeated misfuelings as the fleet ages. Put another way the chances for catalyst deactivation increases as a car gets older. Again this trend also appears in the California fleet but because of the small sample and high estimated errors it cannot be confirmed.

- Overall from both California and the nation it appears that approximately 50 percent or more of those vehicles involved in misfueling have misfueled at least six times in a row This would indicate that at least half of the misfueled vehicles have suffered significant catalyst damage.

TABLE 3-4

REPEATED MISFUELING - 1981

Maximum Number of Successive Leaded Purchases During the Year	Percent of Model Year Catalyst Fleet Vehicles Involved		Estimated Error		Percent of Leaded Purchases By Fleet	
	MY 75 Fleet					
	US	CA	US	CA	US	CA
2	0.8	0.0	1.0	0.0	0.8	0.0
3	0.7	0.0	1.0	0.0	0.5	0.0
4	2.2	0.0	1.8	0.0	2.7	0.0
5	0.4	0.0	0.7	0.0	0.4	0.0
6 or more	16.5	10.4	4.5	10.6	95.4	100.00
	MY 81 Fleet					
2	0.3	0.0	0.5	0.0	1.1	0.0
3	5.2	11.2	2.2	9.3	24.7	74.9
4	3.2	2.5	1.8	4.6	22.4	13.4
5	0.7	1.7	0.8	3.8	4.5	11.7
6 or more	3.4	0.0	1.8	0.0	47.4	0.0
	Overall Fleet					
2	1.0	0.6	0.3	0.8	1.2	1.3
3	2.5	2.1	0.5	1.4	2.5	4.6
4	1.7	0.6	0.4	0.8	3.2	1.6
5	1.0	1.0	0.3	1.0	2.1	5.8
6 or more	7.8	4.6	0.9	2.0	90.4	86.1

TABLE 3-4

REPEATED MISFUELING - 1982

Maximum Number of Successive Leaded Purchases During the Year	Percent of Model Year Catalyst Fleet Vehicles Involved		Estimated Error		Percent of Leaded Purchases By Fleet	
	MY 75 Fleet					
	US	CA	US	CA	US	CA
2	0.9	1.4	1.1	4.0	0.5	1.1
3	1.0	0.0	1.2	0.0	0.6	0.0
4	3.3	5.2	2.1	7.8	2.5	3.9
5	1.0	0.0	1.2	0.0	0.9	0.0
6 or more	18.1	13.2	4.5	11.5	95.2	95.0
	MY 82 Fleet					
2	2.3	0.0	2.2	0.0	11.9	0.0
3	5.1	8.5	3.2	27.3	20.9	15.2
4	2.1	0.0	2.0	0.0	9.7	0.0
5	0.8	0.0	1.3	0.0	4.0	0.0
6 or more	6.4	13.2	3.5	13.0	53.4	84.8
	Overall Fleet					
2	0.9	0.6	0.3	0.7	1.0	1.3
3	3.4	3.3	0.6	1.7	4.0	5.7
4	3.1	2.0	0.6	1.3	4.4	6.2
5	1.5	1.3	0.4	1.1	2.7	4.3
6 or more	8.9	6.4	0.9	2.3	87.6	82.6

Table 3-5 - Misfueling by Type of Service Station

- Overall for both California and the Nation the proportion of vehicles misfueling at self service stations is roughly double the rate observed for full service stations in 1981 and 1982.
- The proportion of leaded gallons purchased by the catalyst fleet between self and full service stations appears to conform to the proportion of vehicles observed misfueling in these stations. Put another way the fueling habits of the misfuelers appear roughly the same between the station types.

TABLE 3-5

MISFUELING BY TYPE OF SERVICE STATION

1981

	Percent of Purchases		Percent of Gallons Purchased	
	U.S.	CA	U.S.	CA
Full Serve	4.7	2.7	4.9 (0.1)	2.6 (0.5)
Self Serve	8.4	4.6	9.3 (0.1)	5.6 (0.3)

1982

	Percent of Purchases		Percent of Gallons Purchased	
	U.S.	CA	U S.	CA
Full Serve	5.0	1.6	5.5 (0.2)	1.8 (*)
Self Serve	10.1	5.6	11.1 (0.2)	6.3 (0.4)

---

*Small sample size precludes estimate of accuracy.

TABLE 3-6

COMPARISON OF MISFUELING RESULTS FROM NPD CALIFORNIA, ARB SURVEYS*, AND NPD NATIONAL VALUES

(Comparison based on Purchase or Volume Weighted Data)

1981			1982		
NPD California	ARB	NPD National	NPD California	ARB	NPD National
3.6	3.3	6.4	4.2	1.9	7.7

---

*Vehicle Misfueling in California, Norman Kayne and Walter Madlock, California Air Resources Board SAE Paper No. 841355, October 1984.

4. CONCLUSIONS

A substantial effort has been devoted towards the evaluation of misfueling behavior in California and a comparison with the results from a similar analysis of the same data base looking at the nation as a whole. An additional comparison with ARB survey results was also performed. The results are as follows:

- In order to compare the results of separate surveys care must be taken to identify the basis of the information collected and the metrics that result. There is a substantial difference between the metrics that result from surveillance data and the metrics that result from diary survey data. One tracks behavior at a point in time, the other tracks behavior through time.
- Two measures of misfueling are available. The first, vehicle incidence, identifies the proportion of the catalyst fleet that has misfueled during the course of a calendar year. The second, purchase or volume weighted, identifies the proportion of leaded to total fuel purchased by catalyst vehicles. These two metrics are separate measures of misfueling and cannot be compared.
- Using the purchase or volume weighted metric there is substantial agreement between the results of the ARB survey's and the NPD survey for 1981 and 1982. The percentage of fuelings that were misfuelings ranged between 2 and 4 percent.
- Using the vehicle incidence metric (unavailable from the ARB surveys) the following was observed.
    - Approximately 9.3 percent of the catalyst fleet was involved in misfueling in 1981 and this increased to 13.6 percent in 1982.
    - The distribution of the misfueling is polar--that is people either misfuel very occasionally (less than 20 percent of the time they purchase fuel) or almost continuously (between 91 and 100 of the time they purchase fuel). Few people have misfueling habits that fall in between these two categories.
    - There appears to be a trend towards increased misfueling as vehicles age. Thus as the catalyst vehicle fleet ages the rate of misfueling (number of vehicles involved) will increase.

- Approximately 50 percent of the vehicles that misfuel do so at least six times in a row. This would indicate substantial catalyst damage for 4.6 percent of the catalyst fleet in 1981 and 6.4 percent of the catalyst fleet in 1982.
- There is a substantially higher rate of misfueling in self serve vs. full serve stations. This trend was consistent in both years studied and is corroborated by the results from the ARB surveys.

APPENDIX A
DETAILED REVIEW OF THE NPD DATA BASE

1. THE NPD DATA BASE

The survey information in the NPD data base is derived from the NPD Petroleum Marketing Index (PMI), a diary panel survey of over 5,000 households conducted by NPD Research, Inc. Panel members are chosen on the basis of demographic characteristics and geographical location. The panel does not include singles or non-family households.

In order to ensure that consistent demographic information is available from all households, NPD requires an adult female be present in all families selected for participation. The rest of Section 1 of this preface discusses important characteristics of the NPD data base which should be kept in mind when examining the results presented in this report.

1.1 Selection of the Panel

Possible respondents for the PMI panel are selected from the American Shoppers Panel (ASP). Each candidate is sent a letter asking about his or her interest in participating in a vehicle use diary panel. Between 60 and 65 percent of those asked respond and return their initial questionnaires. This percentage includes those people who own no vehicle. The questionnaire requests information including the VIN and other engine characteristics for a maximum of five separate vehicles per family. Since the respondents are already in the ASP, no demographic data needs to be collected. NPD selects a subset of these respondents on a demographic basis in order to maintain a balanced sample for the PMI survey. Those selected are sent a monthly diary, a visor holder, and an introductory letter. There is a 75 to 80 percent response to the first monthly diary

Respondents are guaranteed anonymity and at no time are told who will be using the data, although they do know that companies buy the data for gasoline brand market share studies. Respondents are given a hotline number to call if they have any questions, but the people who staff the phones do not know themselves who the clients are. Participants in the ASP are recruited from a variety of mailing lists. The response rate at this stage varies from 2 to 25 percent, depending upon the scope of a particular recruitment effort.

1.2 Coverage of Leased Vehicles

Survey respondents were asked to include "all leased cars whether leased by a company, a business, or privately by any family member."* Unfortunately, it is impossible to determine from the available data which vehicles are leased by a company or business.

1.3 Determination of Fuel Type Requirement

Although the respondent was asked if the vehicle required unleaded fuel, the answer to this question was not used in this report for classifying the vehicle fuel requirements. Rather, the fuel requirements were determined on the basis of make, model, model year, engine data provided by the respondent, and, when available, were confirmed with engine data obtained from the VIN.

1.4 Trucks in NPD

In any calendar year, nearly 2,000 light-duty trucks participate in the NPD survey. Due to difficulties in determining truck fuel requirements, only about 150 of these trucks may be positively identified as having

---

*Statement from the letter mailed to each potential participant.

catalysts. Because of the small sample size for trucks, only cars were included in this analysis. The exclusion of trucks must be kept in mind when considering the results since the tampering rates for trucks have been reported to be substantially different from those for cars.*

1.5 Non-Catalyst Cars With Unleaded Fuel Requirements

Certain vehicles are required to use unleaded fuel even though they do not have a catalyst. Since it has been assumed that the ultimate use, if any, of this analysis will be for estimating the effect of misfueling on catalyst vehicles, a misfueling rate among catalyst vehicles is sufficient. While it can be assumed that the misfueling rates among non-catalyst unleaded cars are either the same or different, this issue is not relevant. Misfueling among non-catalyst unleaded cars does affect estimates of leaded gasoline consumption and lead emissions. However the number of such vehicles and their contribution to leaded gasoline consumption and lead emissions is small and errors will be small if equal misfueling rates are assumed. Furthermore, the data base has few of these cars, so any separate estimates would have great uncertainty.

1.6 Confidence Intervals

Selected tables in this analysis include a statistic termed "Estimated Errors" to denote the reliability of reported misfueling rates. Inasmuch as the NPD data base is derived from a quota sample, it may be argued that no statistic can reflect the "error of estimate" as applied in the strict sense of a random sample. Nevertheless, it is important to realize

---

*"Motor Vehicle Tampering Survey - 1982," U.S. Environmental Protection Agency, Office of Enforcement and Legal Counsel, Publication No. EPA-330/1-83-001, April 1983, p.20.

that estimates derived from a quota sample are subject to variability and that the analyst must consider the variability of derived estimates in interpreting the findings. A more detailed explanation of the derivation of these estimates may be found in Appendix 1 of the analysis.

2. VALIDITY OF THE DATA

Since the NPD data is from a diary panel survey it is important to examine the make-up of the panel and to determine if observed trends in vehicle use behavior are consistent with results from other surveys.

Tables 1 and 2* present some of the demographic distributions observed in the NPD data base and compare them to distributions seen in the Residential Energy Consumption Survey (RECS) and in the National Family Opinion (NFO) gasoline diary survey. Table 3** presents a comparison of Household Income distribution as observed in NPD and as reported by the U.S. Bureau of the Census. Owing in large part to the exclusion of singles from NPD the demographic profile is not strictly representative of the U.S. as a whole. This being the case, it is crucial to compare trends in driving behavior in NPD with those seen in other sources.

A comparison of monthly trends in household Vehicle Miles of Travel (VMT) is presented in Figure 1.+ Notes on the pages following the figure describe the sampling and estimation techniques used in each of the

---

*Fuel Purchasing Patterns and Vehicle Use Trends From the NPD Research Gasoline Diary Data Base: Data Display, Energy and Environmental Analysis, Inc., prepared for the U.S. Department of Energy, September 1982.

**Ibid.

+Ibid.

TABLE 1

COMPOSITION AND LOCATION OF HOUSEHOLDS: NPD VERSUS RECS AND NFO

(Percent)

	Family Size				
	One	Two	Three	Four	Five or More
RECS	15.8	36.6	18.1	16.1	13.4
NFO	17.1	38.5	16.8	16.2	11.4
NPD	00.0	42.7	20.5	23.3	13.5

	Number of Vehicles			
	One	Two	Three	Four or More
RECS	40.2	42.8	12.3	4.7
NFO	51.3	39.2	8.1	1 4
NPD	32.0	40.7	17.1	10.2

	Census Regions			
	Northeast	North Central	South	West
RECS	21.0	27.2	31.9	19.9
NFO	21.3	29.2	31.7	17.8
NPD	20.4	26.8	32.2	20.6

TABLE 2

AGE AND ECONOMIC STATUS OF HOUSEHOLDS: NPD VERSUS RECS AND NFO

(Percent)

	Age of Head of Household			
	Under 30	30-39	40-49	50 and Over
RECS	16.9	20.5	17.7	44.9
NFO	3.2	19.5	18.0	59.3
NPD	7.8	22.8	16.8	52.6

	Own	Rent	Rent Free	Other
RECS	74.9	23.8	1.3	0.0
NFO	84.5	13.4	0.5	1.6
NPD	87.5	11.2	1.3	0.0

TABLE 3

HOUSEHOLD INCOME (1981$)
NPD VERSUS U.S. CENSUS
(Percent)

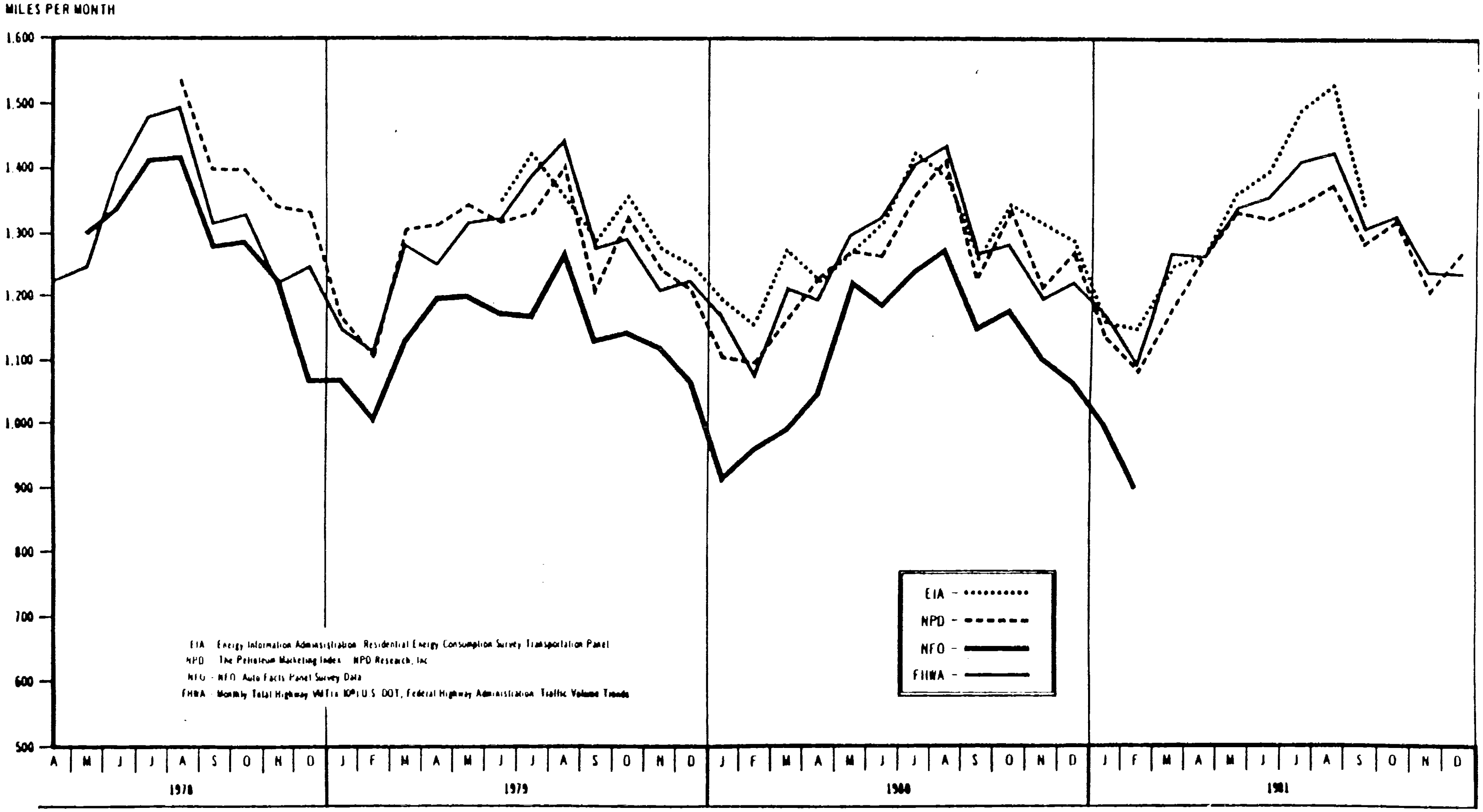
	Under 10,000	10,000 14,999	15,000 19,999	20,000 29,999	30,000 39,999	40,000 49,999	50,000 or More
Census*	25.4	14.4	12.3	21.1	13.1	6.6	7.1
NPD	17.2	18.2	16.2	27.2	13.1	4.7	3.3

---

*U.S. Bureau of the Census, Current Population Reports, Series P-60, No. 134, "Money Income and Poverty Status of Families and Persons in the United States: 1981 (Advance Data from the March 1982 Current Population Survey)," U.S. Government Printing Office, Washington, D.C., 1982.

Note: NPD households include families only, while U.S. Census data uses a broader definition that includes singles.

FIGURE 1
MONTHLY TRENDS IN HOUSEHOLD VMT



EIA - Energy Information Administration, Residential Energy Consumption Survey Transportation Panel
NPD - The Petroleum Marketing Index - NPD Research, Inc
NFO - NFO Auto Facts Panel Survey Data
FHWA - Monthly Total Highway VMT in 10^9 (U.S. DOT, Federal Highway Administration, Traffic Volume Trends

NOTES FOR FIGURE 1

COMPARABILITY OF HOUSEHOLD VMT SOURCES

- Data Sources
  - Energy Information Administration: Residential Energy Consumption Survey Household Transportation Panel (June 1979 to September 1981)
  - Federal Highway Administration (FHWA): Traffic Volume Trends (April 1978 to December 1981)
  - National Family Opinion Poll: NFO/Auto-Facts Gasoline Diary Panel (May 1978 to February 1981)
  - NPD Research, Inc.: Petroleum Marketing Index Diary Panel (August 1978 to December 1981)
- Sampling Techniques
  - RECS: Systematic random sample of households (includes single people and unrelated persons sharing a dwelling)
  - FHWA: Does not sample individual vehicles
  - NFO/Auto-Facts: Quota sample survey of households (includes single people and unrelated persons sharing a dwelling)
  - NPD Research: Quota sample survey of families (single people and unrelated persons sharing a dwelling are not included)
- Estimation Techniques
  - RECS: Odometer readings -- data weighted to national level on the basis of demographic characteristics of household
  - FHWA: City and highway traffic flow counts conducted by State highway departments -- estimate of total travel scaled down by factor of 10^8 for directional trend comparison to household estimates; Census data show approximately 68 million vehicle-operating households in the U.S.
  - NFO/Auto-Facts: Odometer readings -- data are sample-weighted.
  - NPD Research: Odometer readings -- data weighted to national level on the basis of demographic characteristics of household
- Coverage
  - RECS, NFO/Auto-Facts, NPD Research: Report on all vehicles driven (owned/operated) by a household
  - FHWA: Includes trucking and commercial travel

studies. NPD is consistent with these other data sources with respect to monthly trends in household VMT. Further evidence of NPD's consistency may be found in comparisons of annual vehicle miles of travel by vehicle age.

Figures 2 and 3 plot the relationship between annual VMT (Vehicle Miles of Travel) and vehicle age for cars and light-duty trucks. The figures show the well-known trend in decreasing VMT with age as observed in NPD and three other surveys. The other surveys are the Nationwide Personal Transportation Study (NPTS) conducted by the U.S. Department of Transportation, Office of Highway Planning; the Residential Energy Consumption Survey ´RECS) conducted by the Energy Information Administration; and the NFO/Auto-Facts (NFO) national panel diary survey conducted by Auto-Facts, Inc. As can be seen from the two plots, the vehicle-age dependent declines in VMT found in NPD are consistent with those reported by RECS and NFO. The NPTS survey, which consistently reports higher annual VMT than the other data sources, was collected several years earlier than the other studies. Furthermore, NPTS respondents were simply asked to recollect their prior year's mileage accumulation; no effort was made to corroborate the response with odometer records. The RECS survey, which did ask for odometer readings, is a systematic random sample of households, including single people and unrelated persons sharing a dwelling. Since the NPD results are not markedly different from those in RECS, it appears that NPD does not have a serious non-response bias vis-a-vis a random sample with respect to vehicle travel characteristics. In addition, the exclusion of singles seems to have little effect on observed aggregate vehicle use behavior.

The preponderance of evidence suggests that NPD is valid and appropriate for studies of vehicle use behavior in the U.S. Besides having proven itself to be reliable, NPD is also the only currently available source of extensive time series data for the U.S personal transportation fleet.

FIGURE 2

ANNUAL VMT PER CAR BY AGE

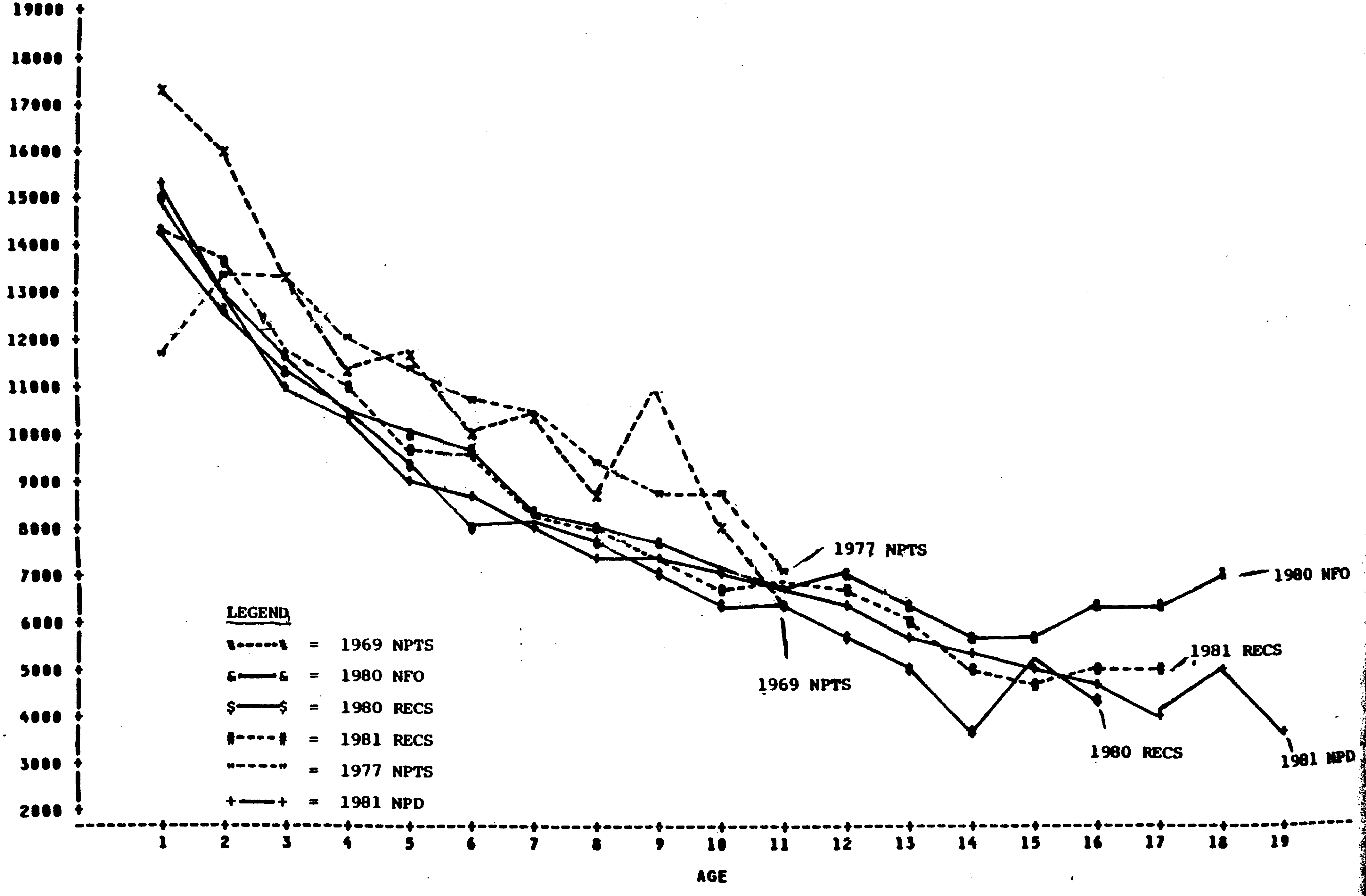
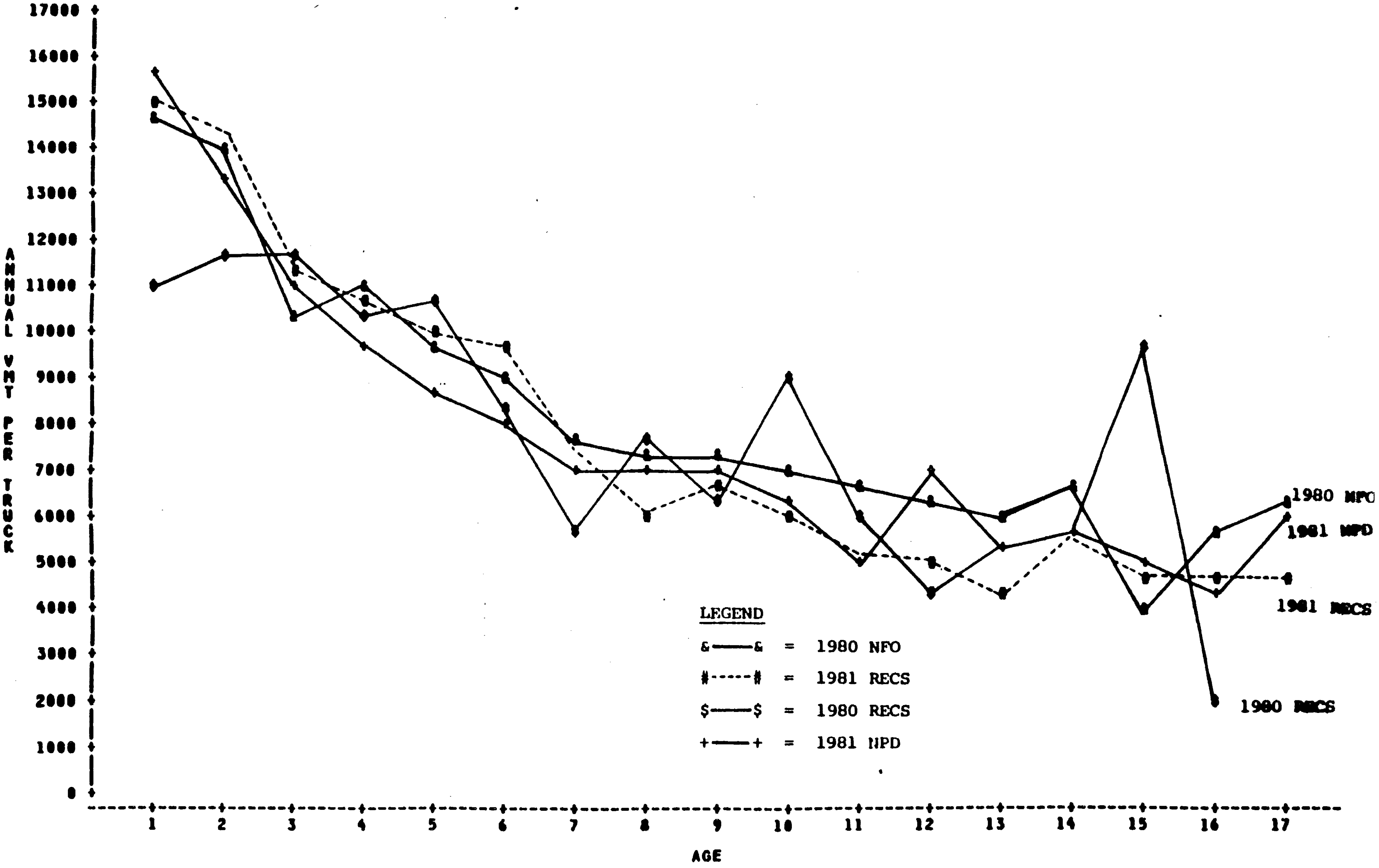


FIGURE 3

ANNUAL VMT PER LIGHT TRUCK BY AGE



Vehicles typically stay in the panel for 10 to 12 months, thus making it possible to take a detailed look at the behavior of many individual vehicle owners over an extended period.

3. ASSUMPTIONS PERTAINING TO THIS ANALYSIS

In processing the NPD data for this analysis, assumptions have been made about unknown fuel type purchases and about reporting errors. These assumptions are discussed in this section.

3.1 Unknown Fuel Purchase Assumption

In approximately 1.6 percent of all purchases reported by the catalyst-equipped cars in NPD, the respondent failed to report whether the fuel purchase was leaded or unleaded. For purposes of assigning misfueling involvement categories the unknown fuel volume is divided between leaded and unleaded fuel on the basis of the ratio between known leaded volume and known unleaded volume for the vehicle making the unknown purchase. Assumptions that unknown is always unleaded or always leaded have also been examined to determine the sensitivity to this approach. The resulting involvement rates are displayed in both the 1981 and 1982 sections of this analysis.*

For purposes of determining the maximum number of successive leaded purchases, unknown purchases are treated as if they were unleaded. It is not practical to randomly assign individual purchases as either leaded or unleaded since on average a vehicle reports only 0.7 unknown purchases during a year. In order to provide bounds for the effect of this

---

*The results may be found in Table 4 "Comparison of Involvement Under Alternate Assumptions About Unknown Fuel Type " This table is included in both the 1981 and 1982 sections.

methodology two alternate cases have been tested. The first case assumes unknown purchases are unleaded; the second case assumes unknown purchases are leaded. Table 4 compares these two cases. Percentages on the diagonal represent vehicles not affected by the manner in which unknown fuel is classified. Percentages to the right of the diagonal represent vehicles that would move into a higher successive purchase category if unknown fuel were assumed to be leaded. When the maximum number of successive leaded purchases is at least three assuming unknown fuel to be leaded has little effect on the distribution of vehicles.

In summary, the unknown fuel type volume is assigned to leaded or unleaded but no attempt is made to correct each individual purchase. If individual purchases were reassigned, the effect on the repeated misfueling statistics is not expected to be large.

3.2 Reporting Errors Assumption

In any large data collection effort there is a potential for recording or transcribing errors.* To avoid overstating the incidence of misfueling, only those vehicles recording at least three leaded purchases during the time they were in the sample are counted as misfuelers. If no more than two leaded purchases are reported, a data error is assumed and the fuel type designator is changed to unleaded.

The rationale for this screening criterion is based on the fact that even with a probability of reporting error as low as 1 percent, there is about one chance in three that one or two misfuelings would be reported

---

*H.T. McAdams, Analysis Memorandum to R. Dulla (EEA), "Reporting Errors in Fuel Purchase Records," under Letter of Agreement No. 026003-1, Contract No. B-F6895-AZ, February 17 and 20, 1984.

TABLE 4

COMPARISON OF ALTERNATE ASSUMPTIONS ABOUT UNKOWN FUEL TYPE PURCHASES USING 1982 NPD DATA

(Using Weighted Data)

Maximum Number of Successive Leaded Purchases Assuming That Unknown is Unleaded	Number of 1982 Vehicles So Classified	Classification Under Alternate Assumptions (Unknown = Leaded) (Percent of Vehicles)						
		0	1	2	3	4	5	6
0	3,050	74.6	11.0	7.6	2.4	1.3	0.8	2.2
1	10		90.4	9.6	0.0	0.0	0.0	0.0
2	34			81.5	4.4	4.3	0.0	9.7
3	106				87.7	3.3	3.0	6.0
4	100					98.4	1.3	0.4
5	55						97.6	2.4
6 or more	339							100.0

during the course of the survey, even though no leaded purchases were actually made. On the other hand, actual misfuelers who misfuel to a significant degree would seldom report as few as two misfuelings during the survey. Thus, the rule is structured to strike a compromise between the two types of errors to provide a refined estimate of the actual vehicle involvement rate in the context of the study.*

Table 5 displays the effect of this methodology on the vehicle involvement rate using data from 1982.** Using the criteria, the overall vehicle involvement rate is 18.0 percent. Without the criteria, taking all leaded designations at face value, the involvement rate is 30.0 percent. Table 6 displays the effect of the methodology on the amount of leaded fuel purchased by the catalyst car fleet Using the criteria, 7.7 percent of the fuel purchased by the catalyst car fleet was leaded. Without that criteria, leaded fuel purchase volume rises to 8.1 percent of the fuel bought by the fleet. Thus, although the methodology reduces the apparent vehicle involvement in misfueling by 40 percent, the change in fuel volume is less than 5 percent.

The criteria results in a conservative lower bound estimate for vehicle involvement and has very little effect on the reported volume of misfueling. An 18 percent vehicle involvement rate, although a lower

---

*H.T. McAdams, Analysis Memorandum to R. Dulla (EEA), "Vehicle Involvement Rate and Its Dependence on Sample Size," under Letter of Agreement No. 026003-1, Contract No. B-F6895-A-Z, March 30, 1984.

**A total of 445 cars meet the maximum-of-two leaded purchase criteria in 1982.

TABLE 5

COMPARISON OF VEHICLE INVOLVEMENT IN MISFUELING IN 1982 RESET MAXIMUM-OF-TWO LEADED PURCHASES TO UNLEADED VERSUS NO RESET

		Vehicle Involvement Misfueling Rate			
		Reset		No Reset	
Model Year	Number of Vehicles	Rate (% of Fleet)	Estimated Error	Rate (% of Fleet)	Estimated Error
1975	280	25.0	5.1	37.2	5.7
1976	465	22.4	3.8	32.1	4.2
1977	586	22.0	3.4	32.3	3.8
1978	632	17.1	2.9	29.1	3.5
1979	590	14.8	2.9	26.9	3.6
1980	450	16.0	3.4	26.3	4.1
1981	485	13.0	3.0	26.9	4.0
1982	185	16.6	5.4	39.6	7.1
1983	21	16.3	15.8	23.2	18.1
Overall	3,694	18.0	1.2	30.0	1.5

TABLE 6

COMPARISON OF VEHICLE INVOLVEMENT IN MISFUELING IN 1982 RESET MAXIMUM-OF-TWO LEADED TO UNLEADED VERSUS NO RESET

		Leaded Fuel Purchased as Percent of Total	
		Reset	No Reset
Model Year	Number of Vehicles	Percent Leaded	Percent Leaded
1975	280	15.3 (4.2)*	15.7 (4.3)
1976	465	11.0 (2.9)	11.3 (2.9)
1977	586	13.4 (2.8)	13.7 (2.8)
1978	632	7.4 (2.0)	7.8 (2.1)
1979	590	4.8 (1.7)	5.2 (1.8)
1980	450	4.2 (1.9)	4.5 (1.9)
1981	485	1.6 (1.1)	2.0 (1.3)
1982	185	2.2 (2.1)	3.0 (2.5)
1983	21	3.5 (7.8)	4.5 (8.9)
Overall	3,694	7.7 (0.9)	8.1 (0.9)

---

*Values in parentheses are estimated errors.

is not insignificant. By comparison, the 1982 EPA tampering survey reports that 10.58 percent of vehicles sampled show at least one positive indication of misfueling.* Since the EPA survey is a random sample and does include trucks, this comparison to NPD is necessarily approximate. However, since EPA reports a higher tampering rate for trucks than for cars a combined sample might be expected to yield a higher rate than for cars alone. A key to the results might be found in the different manner in which the two surveys collected information. The NPD data was collected over a long period of time from individuals who thought they were simply providing marketing information. The EPA data is a compilation of single observations on a random selection of vehicles. One advantage to the EPA method is that classification of a misfueler is based on a direct examination of each vehicle by the survey team. The examination includes a Plumbtesmo test for lead in the exhaust pipe a check of the filler neck restrictor to see if it has been tampered with, and chemical analysis of a gasoline sample to see if lead in the gas tank is above a threshold of 0.05 grams per gallon. There is very little chance that a regular misfueler could escape detection. At the same time, there is some chance that an infrequent misfueler might be overlooked. For example, an individual purchasing leaded gasoline every five or six tankfuls, who uses a funnel to bypass the filler neck restrictor, would show no obvious tampering and might easily have less than 0.05 grams of lead per gallon of fuel in the tank at the time of survey. Furthermore, as noted in the tampering survey, a hastily field-administered Plumbtesmo tailpipe test is unreliable when negative.** Thus, while a positive Plumbtesmo test is reliable evidence of lead in the tailpipe, a

---

*<u>Motor Vehicle Tampering Survey - 1982</u>, U.S. Environmental Protection Agency, Publication No. EPA-330/1-83-001, April 1983, P. 28.

**Ibid.

negative test means only that lead was not detected --the possibility remains that a repeat test under more ideal circumstances would yield a positive result. An additional negative bias is associated with the non-compulsory nature of the survey. Since the EPA survey is openly conducted for a government agency, misfuelers may be extremely hesitant to participate.

In conclusion, while it is difficult to make an exact comparison between NPD and EPA results, each has sources of downward bias and each has strong points. EPA uses a random sample and, through actual examination of the vehicles, has a high probability of identifying regular misfuelers. NPD samples a wider geographic range, including rural areas, and provides demographic information, as well as detailed time-series purchase data. Preference for one type of survey over another is ultimately dependent upon the analysis to be performed and it is the analyst's responsibility to judge the suitability of a particular data base to the task at hand.

APPENDIX B
ESTIMATED ERRORS

Selected tables within the body of this report have included a statistic termed "estimated error" to denote the reliability of key misfueling rates. The purpose of this appendix is to discuss briefly the calculation of this quantity and the considerations that led to its use.

Inasmuch as the NPD data base is derived from a quota sample, it may be justifiably argued that no statistic can reflect the "error of estimate" as applied on the strict sense of a random sample. Nevertheless, it is important to realize that estimates derived from a quota sample are subject to variability and that, as a matter of pragmatism, the issues of bias and variability should be decoupled. The analyst must exercise due caution in selecting a quota sample, considering the purposes of the study, comparison of sample composition and observables (estimates of known quantities) with independent reference sources, and the availability of alternatives to the quota sample's use.

Given that the quota sample is accepted for the purposes at hand the analyst must consider the variability of derived estimates in interpreting the findings. The estimated error statistic is used in this report to reflect the variability of estimates in the sense described above. This calculation follows that of a standard error of estimate derived from a random sample. For sufficiently large samples of size N, the 95 percent confidence limit of an observed proportion p is given by:*

$$CI_{95} = \pm 1.96 \sqrt{p(1-p)/N}$$

In this study, the proportions p are calculated as ratios of vehicles falling within a defined misfueling category to the total number of catalyst vehicles in the survey. Where noted, the proportions are weighted by the NPD projection factors (thereby incorporating both survey participation and control of the sample's demographic balance) In all instances, the sample size is taken to be the (un-weighted) number of catalyst vehicles in the sample.

The resulting estimated error is an approximation to the variability that is present in sample estimates. A more exacting calculation would need to consider the time-series nature of the data (i e extended observations of vehicles across many purchases) and the implications of weighting factors for determining the "effective" sample size. These extended considerations are not germane, however, to the use of the estimated errors as an order-of-magnitude guideline to estimate variability.

---

*Engineering Statistics (Second Edition) by Albert H. Bowker and Gerald Lieberman, Prentice-Hall Inc., 1972 pp. 466-467

APPENDIX C

This appendix contains an example of the purchase logs filled in by NPD panel participants.

AMERICAN SHOPPERS PANEL
P.O. BOX 5401, NEW HYDE PARK, NY 11040

INSTRUCTIONS

1. Slip this diary in the visor holder you attach to the sun visor.
2. Write in ODOMETER (Speedometer) READING when you receive this diary (but no sooner than the date indicated) and again before you return it. There is space for these readings above.
   NOTE: If you replace this vehicle, be sure to write in Odometer readings for BOTH vehicles.
3. Enter all GASOLINE/MOTOR FUEL PURCHASES on the reverse side of this diary.
4. Enter ALL OIL AND ANTIFREEZE PURCHASES on this side of the diary. Make sure purchases at services stations are included.
5. Be sure to tell me all the necessary details if another vehicle is added (see back of return envelope). Tell me about the change in this vehicle (sold, no longer used, disposed of, traded-in, etc.) to the right.
6. All your diaries are important, so please be sure to mail this diary even if no gasoline or oil was purchased this month.

Was the vehicle described on the ID label (top left) SOLD, RETIRED FROM USE, TRADED-IN or otherwise disposed of during the month? ☐ YES ☐ NO

If YES, and vehicle was TRADED-IN or OTHERWISE REPLACED, fill in below for new vehicle:
If NO, please do not fill in below.

MAKE..
(Chevrolet, Ford, Toyota, Dodge, etc.)

MODEL NAME/SERIES..
(Malibu, B 210, Cutlass, Rabbit, TR 7, etc.)

MODEL YEAR 19..........

BODY STYLE..
(2 dr. sedan, 4 dr. sedan, 2 dr. hardtop, hatchback, van, station wagon, pick-up, convertible, camper, etc.)

DOES THIS VEHICLE
REQUIRE UNLEADED FUEL? (✓) One Yes ☐ No ☐
REQUIRE DIESEL FUEL? (✓) One Yes ☐ No ☐
HAVE AIR-CONDITIONING? (✓) One Yes ☐ No ☐

VEHICLE IDENTIFICATION NUMBER (Serial Number.)

☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

ENGINE DISPLACEMENT (See owner's manual. Ex. 180 Cubic Inches, 1.8 Liters.)

.... OR ....
Number of Cubic Inches Numbers of Liters

WHAT TRANSMISSION DOES IT HAVE? (✓) One
☐ 3 Speed Manual ☐ Four Speed Manual ☐ 5 Speed Manual
☐ Automatic ☐ Automatic w/ Overdrive

NUMBER OF CYLINDERS (✓) One
☐ Four ☐ Five ☐ Six ☐ Eight ☐ Rotary

IS IT? (✓) One
☐ Turbocharged
☐ Fuel Injected
☐ Regulator Carburetor

VEHICLE MAINTENANCE—DOES FAMILY MEMBER....
CHANGE THE OIL? (✓) One Yes ☐ No ☐
TUNE-UP THE CAR? (✓) One Yes ☐ No ☐

PRINCIPAL DRIVER:
YEAR OF BIRTH: 19.......... SEX: M ☐ F ☐

WHICH CREDIT CARDS DOES THIS DRIVER CARRY?
....
....
....

GASOLINE/MOTOR FUEL PURCHASES AMERICAN SHOPPERS PANEL **Enter OIL and ANTIFREEZE (Coolant) Purchases on reverse side of card.**

GAS

DATE	BUYER		COPY THIS INFORMATION DIRECTLY FROM GAS PUMP										METHOD OF PAYMENT				DID STATION HAVE..... (✓) all that apply				DID YOU USE... (✓) all that apply				DID YOU FILL UP THE TANK? (✓) one		SPECIAL OFFER			WAS PURCHASE WITHIN 100 MILES OF YOUR HOME? (✓) one	
	Year of birth 19__	Sex	NAME OF BRAND Such as: Shell, Texaco, Exxon, etc.	NAME OF GRADE Such as: Regular, Premium, Super, Diesel, Gasohol, etc.	IS THIS GAS..... (✓) one		NUMBER OF GALLONS/ LITERS PURCHASED		(✓) one		TOTAL AMOUNT PAID (gasoline purchase only)		(✓) one			If credit card was used — NAME OF CREDIT CARD	FULL SERVICE PUMPS	SELF-SERVICE PUMPS	CAR WASH	GROCERY/DAIRY DEPARTMENT	FULL SERVICE PUMPS	SELF-SERVICE PUMPS	CAR WASH	GROCERY/DAIRY DEPARTMENT			Was there a special offer? (✓) one		IF YES, describe offer such as: Price-Off, Free Gift (describe Gift), Discount for Cash, etc.		
					LEADED	UNLEADED	Gal/Lt.	10ths	GALLONS	LITERS	$	¢	CASH/CHECK	CREDIT CARD	STATION CHARGE										YES	NO	YES	NO	write in	YES	NO

GASOLINE/MOTOR FUEL PURCHASES AMERICAN SHOPPERS PANEL **Enter OIL and ANTIFREEZE (Coolant) Purchases on reverse side of card.**

GAS

DATE	BUYER		COPY THIS INFORMATION DIRECTLY FROM GAS PUMP										METHOD OF PAYMENT				DID STATION HAVE..... (✓) all that apply				DID YOU USE... (✓) all that apply				DID YOU FILL UP THE TANK? (✓) one		SPECIAL OFFER			WAS PURCHASE WITHIN 100 MILES OF YOUR HOME? (✓) one	
	Year of birth 19__	Sex	NAME OF BRAND Such as: Shell, Texaco, Exxon, etc.	NAME OF GRADE Such as: Regular, Premium, Super, Diesel, Gasohol, etc.	IS THIS GAS..... (✓) one		NUMBER OF GALLONS/ LITERS PURCHASED		(✓) one		TOTAL AMOUNT PAID (gasoline purchase only)		(✓) one			If credit card was used — NAME OF CREDIT CARD	FULL SERVICE PUMPS	SELF-SERVICE PUMPS	CAR WASH	GROCERY/DAIRY DEPARTMENT	FULL SERVICE PUMPS	SELF-SERVICE PUMPS	CAR WASH	GROCERY/DAIRY DEPARTMENT			Was there a special offer? (✓) one		IF YES, describe offer such as: Price-Off, Free Gift (describe Gift), Discount for Cash, etc.		
					LEADED	UNLEADED	Gal/Lt.	10ths	GALLONS	LITERS	$	¢	CASH/CHECK	CREDIT CARD	STATION CHARGE										YES	NO	YES	NO	write in	YES	NO