REPORT NO. DOT-TSC-OST-77-32

UNCERTAINTIES IN ESTIMATES OF FLEET AVERAGE FUEL ECONOMY: A STATISTICAL EVALUATION

F.T. Rabe

Environmental Impact Center
55 Chapel Street
Newton MA 02158

JUNE 1977

FINAL REPORT

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
Office of the Secretary
Office of the Assistant Secretary for
Systems Development and Technology
Office of Systems Engineering
Washington DC 20590

Technical Report Documentation Page

1. Report No. DOT-TSC-OST-77-32	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle UNCERTAINTIES IN ESTIMATES OF FLEET AVERAGE FUEL ECONOMY: A STATISTICAL EVALUATION		5. Report Date June 1977
		6. Performing Organization Code
7. Author(s) F.T. Rabe		8. Performing Organization Report No. DOT-TSC-OST-77-32
9. Performing Organization Name and Address Environmental Impact Center* 55 Chapel Street Newton MA 02158		10. Work Unit No. (TRAIS) OS714/R7508
		11. Contract or Grant No. DOT-TSC-1311
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Secretary Office of the Asst. Sec. for System Dev. & Tech. Office of Systems Engineering Washington DC 20590		13. Type of Report and Period Covered Final Report Nov 1975 - Nov 1976
		14. Sponsoring Agency Code

15. Supplementary Notes

*Under Contract to: U.S. Department of Transportation
Transportation Systems Center
Kendall Square
Cambridge MA 02142

16. Abstract

Research was performed to assess the current Federal procedure for estimating the average fuel economy of each automobile manufacturer's new car fleet. Test vehicle selection and fuel economy estimation methods were characterized statistically and sources of uncertainty identified. An empirical evaluation of these methods, based on limited available data, indicated that current estimates cannot determine average fuel economy to within 0.1 mile per gallon, the increment to be used in specifying financial penalties and credits to manufacturers under the Energy Policy and Conservation Act (PL 94-163). Alternative procedures for reducing uncertainties in the estimates were identified, and their potential impact on accuracy was quantified. Results confirm that the accuracy of the estimates could be significantly improved with no increase in sample size.

17. Key Words Fuel Economy, Fleet Average Fuel Economy, Fleet Stratification, Class variance, EPA FE Estimation Procedures	18. Distribution Statement DOCUMENT IS AVAILABLE TO THE U.S. PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	66	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

PREFACE

This report was prepared by the Environmental Impact Center, Newton, Massachusetts, for the U.S. Department of Transportation (DOT), Transportation Systems Center (TSC). The work was done under the Direction of the Environmental and Test Programs Division of TSC. It is a part of a larger examination of methods for reducing transportation fuel consumption being conducted by the Transportation Energy Efficiency Project (TEEP) at TSC under the sponsorship of the Office of the Secretary of Transportation.

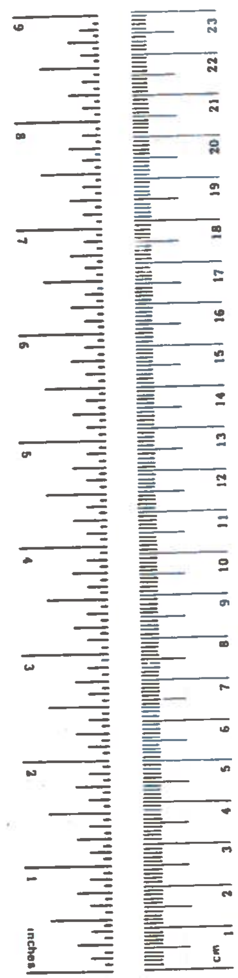
This report is a review, discussion, and evaluation of the procedures used by the U.S. Environmental Protection Agency for calculating the average fuel economy levels of American new car fleets. It concludes that current procedures are not sufficiently accurate to support the administration of the Energy Policy and Conservation Act (PL 94-163) and includes recommendations for increasing the accuracy of fuel economy estimates without the expensive necessity of increasing sample sizes.

Contributions to this study were made by Dr. M. Stephen Huntley, Jr. and Mr. Philip W. Davis of TSC who were technical monitors of this study and Dr. Charles N. Abernethy III, also of TSC.

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F

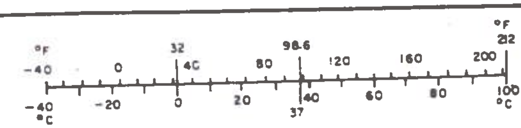


TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF SYMBOLS

C_d = measured fuel consumption of vehicle with drive train d ($d=1,2,\ldots,n_d$)

C_{ij} = measured fuel economy of the j^{th} sample vehicle in class i

$C_{ij\ell}$ = computed average fuel consumption of the j^{th} test vehicle in class i for laboratory cell ℓ

$C_{ij\ell t}$ = measured fuel consumption in the t^{th} test in the ℓ^{th} laboratory cell of the j^{th} vehicle in the i^{th} class

$\overline{F}_f$ = measured average fuel economy of the fleet ($1/\overline{C}_f$)

C_{vt} = measured fuel consumption of the v^{th} vehicle in the t^{th} test ($t=1,2,\ldots,T$)

$\overline{C}_i$ = measured average fuel consumption in the i^{th} class of vehicle ($i=1,2,\ldots,m$)

$\overline{C}_v$ = measured average fuel consumption of the v^{th} vehicle in th t^{th}

e_i = error in fuel consumption estimate for class i

K = constant

n_i = number of sample vehicles from class i

$n_{i\ell}$ = number of laboratory cells used for class i

n_{it} = number of tests per laboratory cell in class i

N = total sample size

P_d = proportion of vehicles in class with drive train type d

P_i = proportion of the fleet in class i

Pr ( ) = the probability of

V ( ) = the variance of

w = width of confidence interval for fleet average fuel economy estimate

w_i = weight of i^{th} item of optional equipment

w_1 = weight of optional equipment sold on more that 33 percent of an engine family

w_2 = weight of optional equipment sold on less than 33 percent of an engine family

$\overline{w}_o$ = average weight of optional equipment for an engine family

z_α = standard normal deviate at the α level of confidence

μc = true mean fuel consumption of a fleet of new cars

μf = true mean fuel economy of the fleet $(1/\mu_c)$

μ_{ij} = true mean fuel consumption in the j^{th} subclass of class i

μ_v = true mean fuel consumption of the v^{th} vehicle in the fleet $(v=1,2,...,M)$

$\underline{\sigma}$ = weighted average of class fuel consumption standard deviations

$\overline{\sigma}_i$ = weighted average variance of fuel consumption in class i

$\overline{\sigma}^2$ = weighted average fuel consumption variance for all classes

$\hat{\sigma}_{\overline{c}}$ $= [V(\overline{C}_f)]^{1/2}$

σ^2_{id} = component of σ^2_i due to class definitions

σ^2_i = variance of fuel consumption in class i

σ^2_{ij} = fuel consumption variance in the jth subclass of class i

$\sigma_{i\ell}^2$ = component of σ_i^2 due to laboratory differences

σ_{ip}^2 = component of σ_i^2 due to production processes

σ_{it}^2 = component of σ_i^2 due to test procedure

σ_v^2 = variance of the v^{th} vehicle

EXECUTIVE SUMMARY

The Energy Policy and Conservation Act of 1975 establishes fuel economy standards for each automobile manufacturer's new car fleet. Starting with the 1978 model year, manufacturers will be subject to financial penalties if the average fuel economy of new cars they sell in the United States fails to meet or exceed mandated levels. The Federal Government has responsibility for estimating fleet average fuel economy levels and assessing penalties and credits. Since penalties vary with 0.1 mile per gallon increments in average fuel economy, small errors in these estimates could cost manufacturers or the Government millions of dollars.

Because fuel economy is a variable rather than fixed quantity, fleet averages can never be determined with complete precision. Statistical sampling and estimation procedures as well as the inherent variability of fuel economy measurements govern the accuracy of fleet average estimates. This study provides a mathematical characterization of sampling and estimation procedures and attempts to quantify potential errors empirically.

Since the Environmental Protection Agency (EPA) will have responsibility for calculating fleet average fuel economy levels, the analysis focuses on their sampling and estimation procedures through the 1975 model year. It should be noted, however, that the primary purpose of these procedures has been exhaust emission certification, rather than fuel economy estimation.

Findings and Conclusions

General

a. Stratified random sampling is a more appropriate approach for estimating fleet average fuel consumption than simple random sampling. By subdividing each manufacturer's new-car fleet into classes containing vehicles with similar engineering and design parameters, greater accuracy can be achieved for any given sample size.

b. There are four major inherent sources of error in fleet average fuel economy estimates using stratified random sampling. Test procedures and laboratory differences cause variation in the measured fuel economy of a single vehicle in repeated tests. Production processes and tolerances cause differences in the fuel economy of even nominall identical vehicles. Finally, the degree of physical differences among vehicles cause further fuel economy variations from vehicle to vehicle within a given class.

c. Insufficient data are available to estimate confidently the magnitude of errors arising from these sources. Limited test results, however, suggest that the sum of test, laboratory, and production variability, measured as the ratio of standard deviation to mean fuel economy, may be in the range of from 5 to 9 percent. The final component of variability, or course, depends on the degree of stratification of the fleet.

d. Inadequate data make quantitative estimates of errors extremely uncertain. Until more extensive tests of fuel economy variability are made, all quantitative assertions concerning the accuracy of fleet average estimates must be viewed with caution.

EPA Fuel Economy Estimates

a. EPA uses a stratified sampling approach. New car fleets are subdivided according to basic engine configurations, engine displacement and emission controls, inertia weight, transmission type, and additional parameters.

b. The major drawback of the procedure is non-random sampling of vehicles within classes. Choices of inertia weight class, transmission, axle ratio, tire size, body style, and engine calibration for test vehicles are not fully representative of all vehicles in each class or for each fleet as a whole. This may create biased fuel economy estimates, although only a detailed empirical investigation could quantify resulting errors.

c. Disregarding these procedural drawbacks, the inherent variability of estimates using the 1975 data sample appears to be significant. Assuming that the coefficient of variation (i.e., the ratio of the standard deviation to the mean) of class fuel consumption is 5 to 10 percent, the width of the statistical confidence interval for the calculated average fuel economy of one major manufacturer was estimated to be 0.8 to 1.8 miles per gallon, depending on the desired level of confidence. While reasonable for many purposes, this level of accuracy obviously would not allow unambiguous determination of penalties and credits under the new requirements.

d. By optimizing the sampling procedure, the above confidence interval widths, under the same assumptions, could be reduced to 0.16 to 0.99 mph with no increase in sample size. Test replications for sample vehicles would further increase accuracy.

e. In order to achieve accuracy to within 0.1 mph, however, sample sizes would have to be drastically increased. The minimum required number of sample vehicles <u>per manufacturer</u> was estimated to be 530. Under less fortuitous assumptions, this number increases to nearly 5,000. Historically, sample sizes have ranged from about 50 to 200 test vehicles for major manufacturers. Such sample sizes would entail major additional expenses to manufacturers and the Government.

Recommendations

a. Representation of transmission types, axle ratios, and other optional equipment within vehicle classes should be randomized to prevent bias.

b. Optimal fleet stratification procedures should be employed to minimize the variance of fuel economy estimates.

c. Additional studies should be made to quantify thoroughly fuel economy variability and its relationship to test and laboratory procedures and equipment, production processes, and vehicle parameters.

d. A thorough cost-benefit analysis should be performed prior to any major increase in the number of vehicles tested per manufacturer.

1.0 INTRODUCTION

The Energy Policy and Conservation Act of 1975[1*] establishes fuel economy standards for new cars sold in the United States starting with the 1978 model year. Manufacturers are subject to financial penalties and credits based on the average fuel economy of their new car fleets. The Federal Government is responsible for preparing estimates of fleet average fuel economy to determine the penalty or credit for each manufacturer.

Since penalties and credits vary with 0.1 mile per gallon increments in average fuel economy, the accuracy of Government estimates clearly is of considerable importance. Slight errors could cost manufacturers millions of dollars in fines. This accuracy depends on random variations in fuel economy measurements as well as vehicle selection and test procedures employed for estimation of fleet average fuel economy.

The research documented here was an attempt to characterize mathematically and quantify empirically the uncertainties in estimates of fleet average fuel economy. Since the Environmental Protection Agency (EPA) has responsibility for preparing fuel economy estimates, their 1975 model year procedures for vehicle selection and testing were assessed to determine needed improvements in accuracy. Finally, alternative measures for improving accuracy were identified and their implications were evaluated.

The following Chapters present findings and conclusions. Chapter 2.0 discusses the mathematical question of how fleet average fuel economy levels can be calculated and characterizes the uncertainty associated with alternative approaches. Chapter 3.0 evaluates fleet classification and vehicle selection criteria and the components of fuel economy variability associated with various classifications and criteria. Chapter 4.0 describes EPA procedures and quantifies the uncertainty in resulting fuel economy estimates. Finally, Chapter 5.0 describes measures for improving the accuracy of current estimates and discusses their implementation.

*Superscripts refer to reference in Appendix A.

2.0 ESTIMATION OF FLEET AVERAGE FUEL ECONOMY

For automobiles in actual use, fuel economy depends not only on vehicle characteristics but also on driving cycles or patterns. Indeed, while engine size, vehicle weight, transmission type, etc., are important, the distances, speed variations, terrain, and climate through which vehicles are driven play an even greater role in determining "on-the-road" fuel economy.

To simplify and standardize the measurement of fuel economy, fixed cycles representing average or typical on-the-road driving patterns are employed. EPA, for example, tests fuel economy over an "urban" and a "highway" test cycle,* subsequently combining the results to obtain a single "composite" value for each vehicle. Such measured values bear no simple relationship to actual in-use fuel economy, but serve as convenient reference points for evaluation of performance and for intervehicular comparisons. The new fuel economy standards are set in terms of EPA's standard driving cycles. The quantity under analysis here is therefore "fuel economy over the Federal driving cycles" rather than in-use fuel economy.

Average fuel economy for a fleet of vehicles is, by definition, equal to the total distance driven by the fleet divided by the total amount of fuel consumed. The arithmetic mean of individually measured fuel economy values will not yield this ratio. If, for example, three vehicles were driven 75, 15, and 100 miles, respectively, the consumed 3, 1, and 5 gallons of gasoline, the true average fuel economy would be 190 miles divided by 9 gallons, or 21.11 miles per gallon. Yet the arithmetic average of measured fuel economy values -- 25, 15, and 20 mpg -- is an even 20 mpg. To avoid such errors, average fuel economy must be calculated by determining the arithmetic mean fuel <u>consumption</u> (in gallons per mile) and then inverting to find average fuel <u>economy</u> (in miles per gallon). The mathematical analysis may be simplified by working primarily with fuel consumption units to avoid consideration

---
*See References 2 and 3 for descriptions of the development of Federal driving cycles.

of inverse functions. Final results may then be translated to the more familiar fuel economy units.

The fuel consumption of every new car is a stochastic (or random) variable with an unknown mean and distribution about the mean. For the fleet as a whole, therefore, the mean fuel consumption will be,

$$\mu_c = 1/M \sum_{v=1}^{M} \mu_v \qquad \text{Eq. (1)}$$

where:

μ_c = the true mean fuel consumption of the fleet

M = the number of vehicles in the fleet

μ_v = the mean fuel consumption of the v^{th} vehicle

The mean fuel economy will be

$$\mu_f = 1/\mu_c \qquad \text{Eq. (2)}$$

where: μ_f = is the mean fleet fuel economy.

If we could test every vehicle in the fleet T times, we could estimate the above parameters as follows:

$$\bar{C}_v = 1/T \sum_{t=1}^{T} C_{vt} \qquad \text{Eq. (3)}$$

$$\bar{C}_f = 1/M \sum_{v=1}^{M} \bar{C}_v \qquad \text{Eq. (4)}$$

$$\bar{F}_f = 1/\bar{C}_f \qquad \text{Eq. (5)}$$

where:

$\bar{C}_v$ = the estimated mean fuel consumption of the v^{th} vehicle

T = number of tests of the v^{th} vehicle

$\bar{C}_f$ = the estimated mean fuel consumption for the fleet

$\bar{F}_f$ = the estimated mean fuel economy for the fleet and M is defined as before. Equations (3) through (5) provide empirical estimates of μ_v, μ_c, and μ_f, respectively. Even though every vehicle has been tested, however, these estimates are uncertain because of the stochastic nature of fuel consumption. Specifically, the fuel consumption of each vehicle has a variance which measures its tendency toward the mean value. Thus, our estimates will also have variance:

$$V(C_{vt}) = \sigma_v^2 \qquad \text{Eq. (6)}$$

$$V(\bar{C}_v) = \sigma_v^2/T \qquad \text{Eq. (7)}$$

$$V(\bar{C}_f) = 1/M^2 \sum_{v=1}^{M} \sigma_v^2/T \qquad \text{Eq. (8)}$$

Obviously, if every vehicle were tested, $V(\bar{C}_f)$ would be very small, since the number of vehicles, M, is very large. The size of the fleet, however, precludes testing every vehicle. We must, therefore, select a relatively small sample of vehicles to represent the fleet. Several sampling procedures are available, the most likely candidates being random sampling and stratified sampling.

In random sampling, each vehicle in the fleet has an equal chance of being selected. The estimating equation then becomes:

$$\bar{C}_f = 1/N \sum_{v=1}^{N} \bar{C}_v \qquad \text{Eq. (9)}$$

Where N, the number of vehicles in the sample, is much smaller than M, and $\bar{C}_v$ is the sample mean fuel consumption of the v^{th} test vehicle. In this case, the variance of the estimates is given by:

$$V(\bar{C}_v) = \sigma_v^2/T \qquad \text{Eq. (10)}$$

$$V(\bar{C}_f) = 1/N^2 \sum_{v=1}^{N} \sigma_v^2/T \qquad \text{Eq. (11)}$$

Where σ_v^2 is the variance of the v^{th} test vehicle. Note that, if every vehicle has the same fuel consumption variance, σ^2, and only one test is made of each vehicle (T=1), Equation (11) simplifies to the more familiar:

$$V(\bar{C}_f) = \sigma^2/N \qquad \text{Eq. (12)}$$

the standard formula for the variance of a sample mean.

Ordinarily, the variance in Equation (11) will be much larger than in Equation (8), since N is a small fraction of M. In random sampling, the variance can be reduced only by increasing the number of vehicles tested or the number of tests per vehicle (T).

Stratified sampling offers a means of decreasing the variance of the estimate without increasing the sample size N. If the manufacturer's fleet of new cars can be grouped into classes more homogeneous with respect to fuel consumption than the fleet as a whole, then the weighted average of fuel consumption among these classes will more accurately estimate the true mean fuel consumptio of the fleet than random sampling.

In stratified sampling, the fleet average is estimated as follows:

$$\bar{C}_f = \sum_{i=1}^{m} p_i \bar{C}_i \qquad \text{Eq. (13)}$$

$$\bar{C}_i = 1/n_i \sum_{j=1}^{n_i} C_{ij} \qquad \text{Eq. (14)}$$

where:

p_i = the proportion of the manufacturer's total production in class i

$\bar{C}_i$ = the average fuel consumption in class i

C_{ij} = the measured fuel consumption of the jth vehicle in class i

n_i = the number of vehicles tested in class i

The n_i vehicles from each class must be selected randomly. The variance of this estimate will be:

$$V(\bar{C}_f) = \sum_{i=1}^{m} p_i^2 \sigma_i^2 / n_i \qquad \text{Eq. (15)}$$

Here, σ_i^2 is the variance of fuel consumption for the i^{th} class of vehicles. Individual vehicle variability is implicit in this variance.

Once the variance of the sample mean has been determined, a confidence interval can be constructed if the statistical distribution of fuel consumption is known. Assuming normal distribution,* the confidence interval is formally defined by:

$$Pr(\bar{C}_f - z_\alpha \sigma_{\bar{C}} \leq \mu_c \leq \bar{C}_f + z_\alpha \sigma_{\bar{C}}) \geq 1 - \alpha \qquad \text{Eq. (16)}$$

where:

P_r = the probability of

z_α = the standard normal deviate at a confidence of α

$\sigma_{\bar{C}} = [V(\bar{C}_f)]^{1/2}$

This confidence interval may be translated into units of fuel economy by inverting:

$$P_r \left[\frac{1}{\bar{C}_f + z_\alpha \sigma_{\bar{C}_f}} \leq \mu_f \leq \frac{1}{\bar{C}_f - z_\alpha \sigma_{\bar{C}_f}} \right] \geq 1 - \alpha \qquad \text{Eq. (17)}$$

Equation (17) holds true no matter which approach is used to determine $\bar{C}_f$ and $V(\bar{C}_f)$.

EPA currently uses a stratified sampling process to estimate the fleet average fuel economy for manufacturers. Automobiles can be grouped conveniently into classes on the basis of engineering parameters. The expected fuel consumption of vehicles within such

---

*Very little is known about the distribution of fuel consumption for single vehicles or groups of vehicles. Normal distribution is assumed for convenience.

classes is much more homogeneous than that expected of vehicles randomly selected from the fleet as a whole. Thus, stratified sampling should be more accurate than random sampling, and we can focus the discussion on the former procedure.

3.0 FLEET STRATIFICATION AND CLASS VARIANCE

3.1 GENERAL

Stratified sampling as applied to estimation of average fuel economy requires that each manufacturer's fleet of new cars be subdivided into strata, or classes, each containing vehicles with similar fuel economy levels. Then, a random sample of vehicles from each class may be selected and tested, and the weighted-average fleet fuel economy may be calculated in accordance with Equations (13) and (14).

3.2 STRATIFICATION PROCEDURE

Classes must be defined prior to actual fuel economy measurements. Therefore, vehicles for which similar fuel economy levels are anticipated a priori should be grouped together. Since the fuel economy of an automobile (over a fixed driving cycle) depends largely on engineering and design parameters, classes may be defined as groups of vehicles with identical or similar vehicle parameters. Engineering and design characteristics provide criteria for fleet stratification.

Virtually every feature of an automobile has some influence on fuel economy. A small fraction of these, however, are dominant.[4] These include:

a. basic engine design (i.e., horsepower, number of cylinders, displacement, method of air aspiration, emission controls, etc.),

b. vehicle weight,

c. transmission configuration,

d. body style, and

e. axle ratio and tire size (N/V).

Vehicles with the same combination of all important parameters may be termed a "configuration," and are nominally identical. Thus, for example, Dodge Darts with the same basic 318 cubic inch

V-8 engine, catalyst, and calibration, in the 4,000 lb inertia weight class, with automatic transmissions, and with standard tires and axle ratio constitute a configuration. All such Darts would be expected to have similar fuel economy.

Configurations form natural classes which might be used in stratified sampling. So many are produced, however, that less detailed class definitions are more practicable. By holding only some of the above parameters constant in class definitions; i.e., including several configurations in each class, the total number of classes for a manufacturer can be reduced. This will permit a smaller sample size, although the accuracy of the resulting estimate of average fuel economy may be reduced.

The choice of which parameters to control in class definitions depends on their relative influence on fuel economy. If parameters can be ranked in terms of their importance, then a hierarchical classification system may be used, as shown in Figure 1. Each manufacturer's fleet is first subdivided according to variations in the most important parameter -- basic engine design in Figure 1. These groups are next disaggregated based on variations in the second most important parameter -- inertia weight. Resulting subgroups are broken up according to the next most important parameter (transmission), and so forth. In principle, this structure can be extended to incorporate every aspect of vehicle design.

In practice, however, the number of parameters controlled in class definitions depends upon accuracy requirements and testing costs. Class definitions can be conceived as horizontal lines cutting across the hierarchy at any level. A class definition line low on the pyramid-shaped hierarchy would result in a large number of classes, each containing very similar vehicles. A line higher on the pyramid would reduce the number of classes, but each class would then contain vehicles with widely varying parameters. Since at least one sample vehicle must be selected from each class, fewer classes yields a smaller minimum sample size and reduced testing costs. Differences among vehicles within each class,

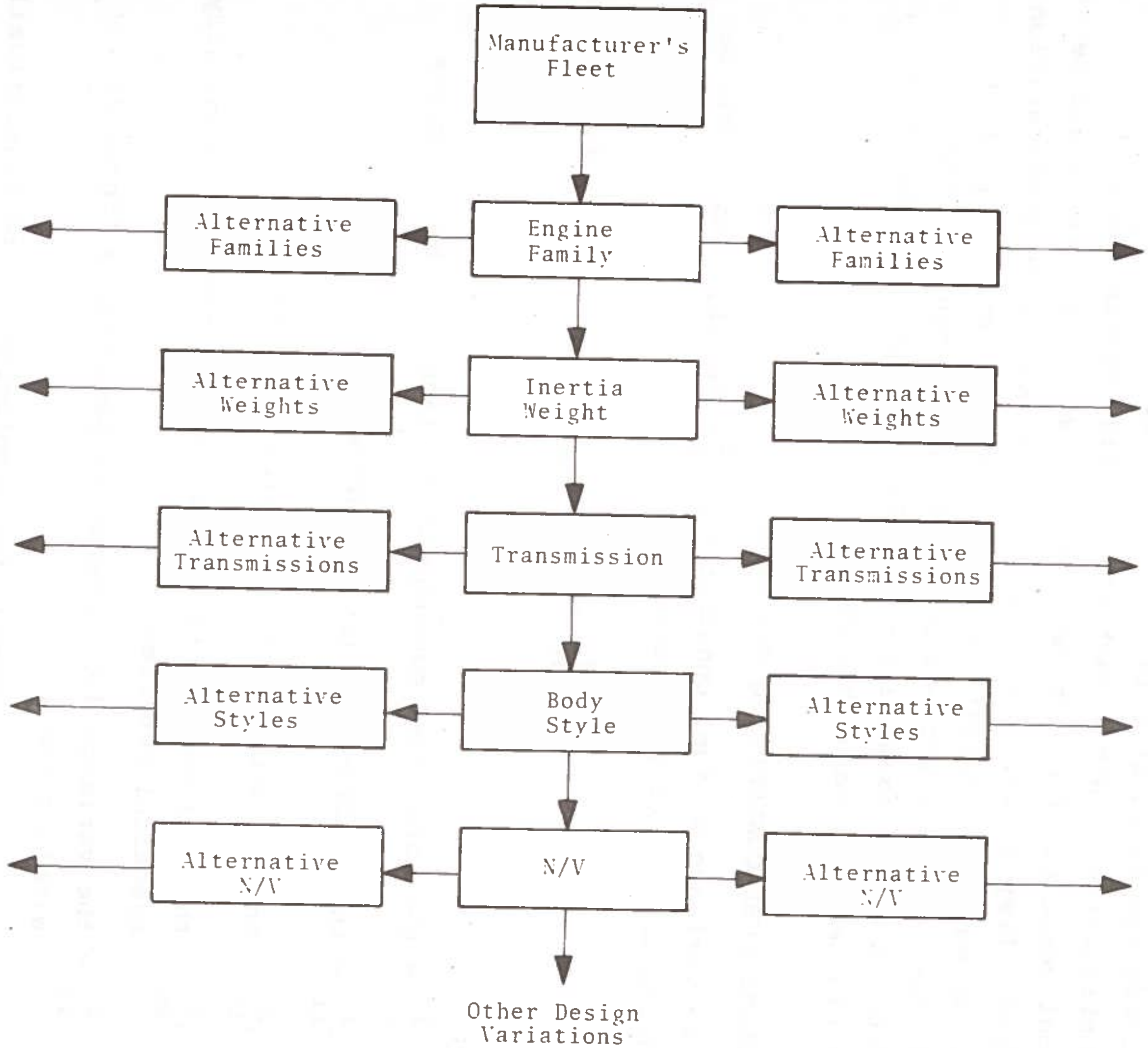


FIGURE 1. FLEET CLASSIFICATION SYSTEM

however, will increase the class fuel economy variance. Yet moving the class definition line down the hierarchy causes a several-fold increase in the number of classes, required sample size, and testing costs.

The bottom level of the hierarchy represents configurations. Classifying the fleet at this level would define a new class for each configuration. Since each class would contain nominally identical vehicles, fuel economy variance within classes would be minimized. Despite this fact, however, defining each configuration as a class would not eliminate variance in the estimate of fleet average fuel economy, nor would it necessarily minimize that variance. A closer examination of fuel consumption variability within classes will make this evident.

3.3 CLASS FUEL CONSUMPTION VARIANCE

The variance of fuel consumption, σ_i^2, for any class i may be considered the sum of four components:

$$\sigma_i^2 = \sigma_{it}^2 + \sigma_{i\ell}^2 + \sigma_{ip}^2 + \sigma_{id}^2 \qquad \text{Eq. (18)}$$

where:

σ_i^2 = the total fuel consumption variance in the i^{th} class

σ_{it}^2 = the variance due to test procedure

$\sigma_{i\ell}^2$ = the variance due to laboratory differences

σ_{ip}^2 = the variance due to production differences in nominally identical vehicles

σ_{id}^2 = the variance due to vehicles parameter differences within class i

and the components are assumed, for simplicity, to be statisticall independent.

The test variance, σ^2_{it}, arises through inaccuracies in measure-
nent and data reduction equipment, differences in ambient conditions
from test to test, and the vehicle operator's inability to follow
perfectly the specified driving cycle. It implies that an individual
vehicle's measured fuel consumption will vary from test to test,
even using the same equipment and personnel.

The laboratory variance ($\sigma^2_{i\ell}$) arises from differences in
equipment, calibration, and personnel from one laboratory to the
next, or from one test cell to the next in a single laboratory.
Thus, testing the same vehicle in different laboratories will
yield somewhat different results. In principle, the test and lab-
oratory variances are independent of the type of vehicle tested,
although they may be related to its fuel consumption level.

Production variance, σ^2_{ip}, is caused by slight differences
among even nominally identical vehicles. Production tolerances,
quality control procedures, and tuning practices cannot be con-
trolled sufficiently to produce identical vehicles. Thus, even
if test and laboratory variances were eliminated, two vehicles
from the same configuration would yield differing fuel consumption
test results.

The final component, σ^2_{id}, is essentially a class definition
variance caused by including vehicles with different engineering
and design characteristics in the same class. Intuitively, σ^2_{id}
depends upon the magnitude of these intervehicular disparities.
Thus, lumping many configurations in one class will greatly in-
crease this component.

Only σ^2_{id} depends on how the fleet is stratified. The test
and laboratory variance are functions of measurement equipment
and procedures, while the production variance depends on engineering
specifications and manufacturing processes. These three components
cannot be affected by the stratification process.

No systematic empirical study of these components of variance
has been performed. EPA and the automobile manufacturers have re-
ported limited results of fuel economy test variability which give

some notion of magnitude. Unfortunately, these reports lack full consistency in the method of calculating variability. Most frequently, results are presented in terms of the sample standard deviation in fuel economy units. Some studies present variation in terms of the sample standard deviation in carbon dioxide emissions (since most of the fuel "consumed" by a vehicle is transformed to CO_2) as an indication of fuel consumption variability. These results may be compared by calculating the coefficient of variation -- the sample standard deviation as a percent of the sample mean -- for each set of results. However, the distinction among test, laboratory, and production variances has not always been clearly made and the values presented below must be viewed with caution.

Several sources[5,6,7] have reported variations in CO_2 emissions for a single vehicle in a single test cell. The reported coefficien of variation of CO_2 emissions was 2.0 percent for EPA, 1.2 percent for General Motors Corporation, and 2.5 percent for Honda.[6]

Laboratory variance can be identified by subtracting test variance from measured values of overall fuel consumption variabilit across test cells. Reports of overall variability from laboratory to laboratory[7] range from 3.8 to 8.8 percent for the ratio of standard deviation to mean fuel economy. This implies that laboratory variability isolated from single-cell test variability is on the order of 2.9 to 8.4* percent (assuming test variability of 2.5 percent). A direct estimate of laboratory variability yielded values of 1.0 to 3.5 percent.

Production variance has been reported at 3.5 percent for the ratio of standard deviation to mean fuel economy in tests of nominally identical vehicles by two manufacturers.[8] It is uncertai whether this value includes test and laboratory variability as well as vehicle variability, but the small magnitude of the reported value suggests that test and laboratory variabilities were factored out of the calculations.

---

*$\sigma_{i\ell} = [(\sigma_{i\ell}^2 + \sigma_{iv}^2)-\sigma_{iv}^2]^{1/2}$; i.e., $2.9\% = (3.8^2-2.5^2)^{1/2}$, etc.

Table 1 summarizes reported values for test, laboratory, and production variabilities. Assuming independence, these components are not simply additive. By taking the square root of the sum of their squared values, however, we can gain an indication of the coefficient of variation for all three components. This value is 4.7 to 9.4 percent, shown in the bottom line of the table. The scarcity of test data make this a very rough approximation.

The above components and their sum apply to classes containing only nominally identical vehicles. The final element of class variance, σ_{id}^2, will arise if vehicles with differing parameters are included in a class. Its magnitude will, in principle, be controlled by the degree of differentiation among vehicles within classes. The statistician, therefore, has the option of reducing the value of σ_{id}^2 to zero by defining each vehicle configuration as a class for his stratified sample.

To summarize, the primary goal of fleet stratification is to define classes with smaller fuel consumption variance than the fleet as a whole. This is most easily achieved by grouping vehicles with similar or identical engineering and design parameters. Class fuel consumption variances depend partially upon the degree of differentiation among vehicles within classes. Even if classes comprise nominally identical vehicles, however, components of variability arising from test procedures, laboratory differences, and production processes will remain. Class definitions, therefore, cannot eliminate uncertainty. Nevertheless, for any given sample size, the classification procedure will determine how nearly the minimum variance is approached for the fleet average fuel consumption estimate.

TABLE 1

REPORTED VALUES FOR COMPONENTS OF FUEL ECONOMY VARIABILITY

Component	Standard Deviation/Mean
Test Variability ($\sigma_{it}/\bar{C}$)	1.2 - 2.5%
Laboratory Variability ($\sigma_{i\ell}/\bar{C}$)	2.9 - 8.4%
Production Variability ($\sigma_{ip}/\bar{C}$)	3.5%
Total Variability Independent of Stratification Process*	4.7 - 9.4%

*The sum is equal to $\frac{1}{\bar{C}} (\sigma_{it}^2 + \sigma_{i\ell}^2 + \sigma_{ip}^2)^{1/2}$

4.0 EPA FUEL ECONOMY ESTIMATION PROCEDURES

4.1 GENERAL

Although no formal program for estimating manufacturer's fleet average fuel economy levels was in effect prior to the Energy Policy and Conservation Act, the Environmental Protection Agency has made informal estimates as a part of their fuel economy labeling effort for the 1975 and 1976 model years. The labeling effort has been based on test vehicles selected for the exhaust emissions certification program. Vehicle selection criteria, originally intended to include all high-emission configurations in the sample, have been modified somewhat to enhance representation of fleet fuel economy. Nevertheless, the original vehicle selection process remains largely intact.

The discussion in the preceding sections provides framework for assessing this process in terms of its accuracy in estimating fleet average fuel economy. The assessment is based on 1975 model year data, the latest available at the time of this study. It should be noted, however, that EPA made several refinements for the 1976 model year and is now completing an internal evaluation of vehicle selection and test procedures for use in the mandatory fuel economy program. Requirements for fleet representation differ substantially for emissions certification and fuel economy estimation; shortcomings of the current procedures for fuel economy estimation are hardly surprising since they were designed for a different purpose.

.2 FLEET STRATIFICATION AND VEHICLE SELECTION

EPA does, in fact, stratify each manufacturer's fleet of new cars based on engineering and design parameters.* The initial grouping is into engine families. To be grouped in the same engine family, vehicles must be identical in:

a. cylinder bore center-to-center dimensions

---

*See 40 CFR, Part 85.

b. dimension from centerline of crankshaft to the top of the cylinder block head face,

c. cylinder block configuration,

d. location and size of intake and exhaust valves,

e. method of air aspiration,

f. combustion cycle, and

g. catalytic converter and/or thermal reactor characteristics.

Further subdivision may (optionally) be based on bore and stroke, surface-to-volume ratio of the cylinder at top dead center, intake and exhaust manifold design, valve sizes, fuel system, and camshaft and ignition timing.

Vehicles in each engine family are next subdivided into engine displacement-exhaust emission control system-evaporative emission control system combinations. These combinations represent vehicle classes in the sense that at least one test vehicle must be selected for each. In practice, however, most engine families contain a single such combination, and there are formal criteria for representing additional variations in parameters within each combination.

In their applications for emissions certification, manufacture list all of the configurations they expect to produce within each engine family, along with sales projections for each. Included in the descriptions are alternative transmissions, body styles, ax ratios, tire sizes, and engine calibrations. In addition, items o optional equipment expected to be included in 33 percent or more of the vehicles sold within each family are listed in their combin weight added to the basic inertia weight of each configuration. Vehicles are grouped according to inertia weight classes containin 250-pound increments up to the 3,000-pound class, and 500-pound increments above 3,000 pounds. These descriptive lists and sales projections provide the information upon which EPA bases selectior of test vehicles. Three groups of vehicles may be selected from each engine family.

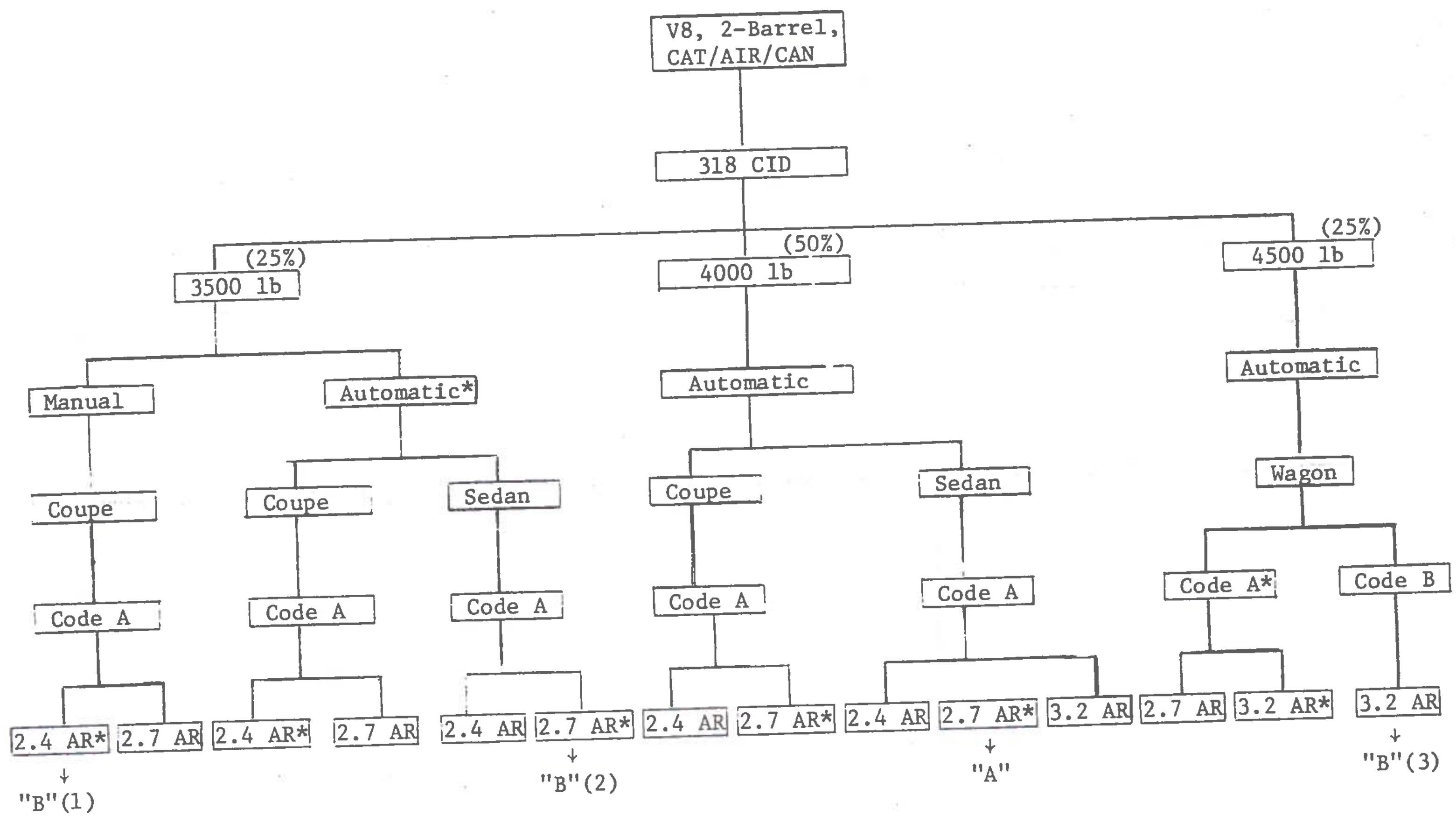
The first group, or "A" vehicles, is selected to represent the projected highest-selling displacement-emission control combinations. One vehicle is selected from each of the projected highest-selling combinations until 70 percent of production for the engine family is represented or until four vehicles have been selected. If 70 percent or more of expected production in an engine family is in a single displacement-emission control combination, two "A" vehicles may be chosen from that combination. Otherwise, only one "A" vehicle may be selected from each high-production combination.

Each "A" vehicle reflects the projected sales-weighted average inertia weight, the projected highest-selling transmission and body style, and the standard axle ratio, tire size, and engine calibration for the displacement-emission control combination it represents.[9]

The second group, comprising "B" vehicles, includes configurations suspected of high emission levels because of their engineering and design parameters. The increasing importance of fuel economy, however, has encouraged the use of "B" selections to represent variations in vehicle characteristics omitted from "A" vehicles.[9] Thus, "B" vehicles might be from the same displacement-emission control combinations as "A" vehicles, but reflect different weights, transmissions, axle ratios, etc. A maximum of four "B" vehicles may be selected.

The third group ("C" vehicles) includes one vehicle from each displacement-emission control combination not represented by an "A" or "B" vehicle. Each "C" vehicle must reflect the projected sales-weighted average inertia weight for its combination, but transmission type, axle ratio, etc., may be selected to represent worst case emissions.[9]

A simplified example will help clarify this procedure. Figure 2 depicts an imaginary engine family of automobiles with a particular V-8 engine, two-barrel carburetor, and emission control system comprising an oxidation catalyst, air injection, and canister to trap cold-start hydrocarbon emission. Only one displacement -- 318 cubic inches -- is produced, but it is used in three weight classes:



*Standard or highest sales

Selections: "A" - 4000 lb, automatic, sedan, Code A, 2.7 AR
"B" - (1) 3500 lb, manual, coupe, Code A, 2.4 axle
(2) 3500 lb, automatic, sedan, Code A, 2.7 axle
(3) 4500 lb, automatic, wagon, Code B, 3.2 axle
"C" - None

FIGURE 2 VEHICLE SELECTION EXAMPLE

3,500; 4,000; and 4,500 pounds. Manual transmission is offered only in the 3,500-pound coupe model; automatic transmissions are standard equipment on the coupe, sedan, and wagon body styles in all weight classes. Only the standard engine calibration (Code A) is offered in coupes and sedans, while the second calibration (Code B) is available for higher performance in the wagon model. Finally, two or three axle ratios are offered for each model, with the standard axle varying from lower to higher weight vehicles.

The percentages in parentheses above each weight class designation indicate the proportion of projected sales for the engine family represented by that inertia weight class. Thus, 50 percent of the vehicles sold are projected to be in the 4,000-pound class, with 25 percent each in the other two classes. Projected sales are displayed in Table 2.

To select vehicles, we note first that the family contains only one displacement-emission control combination, so, a maximum of two "A" vehicles may be selected. The first "A" vehicle is required to reflect the projected sales weighted average inertia weight, most popular transmission and body style, and standard engine calibration and axle ratio for the displacement-emission control combination. We, therefore, select a sedan in the 4,000-pound weight class with automatic transmission, engine Code A, and 2.7 rear axle ratio (plus standard tires, not shown on the figure).

"B" vehicles may now be selected to provide a more thorough representation of vehicular differences within the family. From left to right, we select (1) a 3,500-pound coupe with manual transmission, engine code A, and 2.4 axle ratio; (2) a 3,500-pound sedan with automatic transmission, engine code A and 2.7 axle ratio; and (3) a 4,500-pound wagon with automatic transmission, engine code B, and 3.2 axle ratio. These choices reflect all the available weight classes, body styles, transmission types, engine calibrations, and axle ratios, so, the final "B" and optional "A" vehicle may be omitted. No "C" selection is possible.

TABLE 2 EXAMPLE CALCULATION OF MEAN FUEL ECONOMY

Test Vehicle	Fuel Consumption gpm	Fuel Economy mpg	Projected Sales	Projected % Total Sales
A	0.0690	14.5	100,000	50
B(1)	0.0617	16.2	40,000	20
B(2)	0.0592	16.9	10,000	5
B(3)	0.0820	12.2	50,000	25

$C = (0.5)(0.0690) + (0.2)(0.0617) + (0.05)(0.0592) + (0.25)(0.0820)$

$= 0.0703$ gpm

Average Fuel Economy $= 1/\bar{C} = 14.22$ mpg

After the fuel consumption of each test vehicle has been measured, EPA calculates a projected sales-weighted average for each inertia weight class and each displacement-emission control combination. These results are then summed for all combinations in all engine families to estimate the projected sales-weighted average fuel economy for each manufacturer. In 1976, harmonic averaging replaced the arithmetic averaging used in 1975.

To illustrate this, Table 2 provides data on test vehicles from our imaginary engine family. Fifty percent of sales in the engine family, or 100,000 vehicles, are expected in the 4,000-pound inertia weight class represented by our "A" selection. Of the 50,000 vehicles expected to be sold in the 3,500-pound class, 40,000 will have automatic and 10,000 will have manual transmissions. The remaining 25 percent of engine family sales are expected to occur in the 4,500-pound class represented by the third "B" vehicle. Measured fuel consumption for each test vehicle is weighted by the proportion of engine family sales it represents and these values summed to determine mean fuel consumption for the family. The engine family mean values may then be weighted by the proportions of overall fleet sales they are expected to represent and summed to determine fleet average fuel consumption. Inversion yields fleet average fuel economy.

Although not specifically required to do so, EPA has attempted in the past two years to include one test vehicle from every inertia weight class. They have also attempted to represent all transmission options offered in significant numbers. Thus, while the nominal class definition has specified only basic engine, displacement, and emission controls, most of the classes used in practice have represented specific combinations of basic engine, displacement, weight, and transmission type. In addition, manufacturers are permitted to submit fuel economy data for vehicle configurations not included in the EPA sample. When provided, such supplementary information can also improve the accuracy of the estimates. Because it is optional, however, it cannot be relied upon, and the adequacy of the EPA sample must be considered in isolation.

Several important drawbacks are apparent:

(a) Vehicle selection within classes is not random. The choice of engine calibration, axle ratio, tire sizes, and to some extent transmission type are not random in the statistical sense, so that estimates of class fuel consumption will be biased. One reason for this problem is the lack of clear class definitions; the selection process is oriented around engine families rather than detailed class specifications. As a result, class definitions fall out of the selection process rather than guiding the choice of test vehicles. A more fundamental problem, however, is the current need to choose vehicles with worst case emission features, a need which is difficult to reconcile with random sampling for fuel economy estimates.

(b) Test results of vehicles with different axle ratios and tire sizes are arithmetically averaged. Several axle ratios and tire sizes are typically offered for every class of vehicle. Historically, data on sales by axle ratio and tire size were not available, so, EPA used an arithmetic rather than weighted average. Since the standard axle and tires are usually sold on 90 percent of each model,* the arithmetic average can cause significant errors in estimates of class fuel consumption. In the future, either the selection of axle ratios and tire sizes should be random, or a production weighted average should be used.

(c) Engine family representation is unrelated to production quantities. Not only is the number of sample vehicles arbitrarily limited for engine familes, but the same sampling criteria apply regardless of the differences in production quantities for engine families. Thus, high production families may be represented by the same number of observations as low production families. This increases the variance of the fleet average estimate for a fixed number of total observations.

*No complete data are available on sales by axle ratio or tire size. The major domestic manufacturers, however, have reported their conclusion to TSC personnel that standard tires and axles predominate.

(d) Differences in vehicle weights are not accurately represented. This problem has two components. Most weight classes are defined by 500-pound increments. This allows room for weight-related fuel consumption variations. In addition, the full weight of optional items of equipment sold on more than 33 percent of vehicles in each engine family is added to all test vehicles for that family. Inertia weights may be biased upward as a result, yielding lower measured fuel economy.

These problems had little significance when the EPA test data were used only for fuel economy labeling of models to the nearest mile per gallon. Under the new mandatory fuel economy program, however, very small uncertainties -- 0.1 mpg -- in fleet average estimates have major financial implications. It is useful, therefore, to make rough approximations of the uncertainties now prevailing in the estimates.

4.3 ACCURACY OF THE ESTIMATES

The question of accuracy has two components:

a. What is the error (or bias) in class fuel economy estimates?

b. What is the inherent uncertainty in fleet fuel economy estimates?

Errors in class estimates may not affect the fleet estimate since they will tend to cancel out across classes unless the bias is systematic. Further, they are the result of inappropriate estimation procedures and may be corrected. The inherent uncertainty in fleet estimates, on the contrary, cannot be eliminated even with optimal procedures. These two kinds of problem are assessed separately below.

4.3.1 Errors in Class Fuel Economy Estimates

Ideally, the probability of a specific vehicle configuration (within a class) being selected for testing should be proportional to the fraction of class sales that configuration represents.

With classes defined as engine family-inertia weight-transmission type combinations, the specification of axle ratio, engine calibration, etc., for test vehicles should be based on expected sales of alternative axle ratios, calibrations, etc. Unfortunately, insufficient data are available to address non-random sampling, the most important source of error in class estimates. We can, however, assess the effects of the current representation of inertia weight, transmission type, and axle ratio variations.

<u>Inertia Weight Effects</u>. In determining the inertia weight classes of vehicles selected within a given engine family, the full weight of all items of optional equipment sold on 33 percent or more of that family's vehicles is added. The true average weight of optional equipment, however, is:

$$\bar{w}_o = \sum_i p_i w_i \qquad \text{Eq. (19)}$$

where:

$\bar{w}_o$ = average weight of optional equipment (for an engine family

p_i = market penetration of the i^{th} item of equipment

w_i = weight of the i^{th} item

Assume that the average penetration of items with greater than 33 percent market share is p_1, that the average penetration for those with less than 33 percent is p_2, and that the total weight of the first group of items is w_1, while that of the second group is w_2. Then, the current estimate of the weight of optional equipment is simply w_1, while the actual average weight is:

$$\bar{w}_o = p_1 w_1 + p_2 w_2 \qquad \text{Eq. (20)}$$

and the error is

$$w_1 - \bar{w}_o = (1-p_1)w_1 - p_2 w_2 \qquad \text{Eq. (21)}$$

The error is more likely to be positive than negative, so, there is a good likelihood of bias. However, this will only affect fuel economy in instances where the error causes test vehicles to be reclassified in a higher inertia weight. Otherwise, the dynamometer setting remains the same, and no change in measured fuel economy should occur.

The full range of optional equipment for a given model typically weighs several hundred pounds. Let us assume (arbitrarily) that the error from Equation (21) averages 50 pounds. Then, if the actual inertia weights of vehicles are evenly distributed within the 500-pound range of weight class, some 10 percent of the vehicles will be incorrectly classified at the next higher inertia weight. Those vehicles will be tested at the 500-pound greater dynamometer setting, which typically will cause about a 1-mpg fuel economy penalty. Since 10 percent of the fleet is affected, the net bias in average fuel economy would be about 0.1 mpg.

Clearly, the existence and size of any actual bias depend on the parameters p_1, p_2, w_1, and w_2 for each engine family. Only a detailed study could clarify the issue empirically. It is interesting to note, however, that careful engineering by a manufacturer could cause his vehicles to fall on average at the extreme upper end of each inertia weight class, creating a new bias in the direction of higher fuel economy.

Transmission Effects. As of 1975, EPA selected a vehicle for each transmission type used in 30 percent or more of an engine family. Transmissions used in less than 30 percent of sales were usually ignored. This clearly creates errors. Assume, for example, that two transmissions are offered, and that the more popular version sells in slightly more than 70 percent of the engine family while achieving 10 percent higher fuel consumption than the less popular alternative. Then, the average for the class will be:

$$\bar{C} = (0.7)(1.1X) + (0.3)(X) = 1.07X \qquad \text{Eq. (22)}$$

where X is the fuel consumption of the less popular configuration, while the current estimate would be:

$$\bar{F}' = 1.1X \qquad \text{Eq. (23)}$$

the consumption of the more popular version. The above error of some 4 percent (or 0.6 mpg with a 20 mpg average) is the worst case. Since automatic transmissions are used on almost all new cars in the United State, the fleet average error from inappropriate representation of transmission alternatives should be negligible. Only a detailed empirical study, however, could verify this expectation.

Drive Train Effects. In calculating fuel economy estimates, EPA uses an arithmetic average for vehicles with different axle ratios and tire sizes within a given engine family, displacement, and inertia weight class. This implicitly assumes that vehicles with different drive trains are sold in equal numbers, whereas the standard axle and tire are sold on some 90 percent of all cars. Thus, within a vehicle class, the error to be expected from the current calculation is:

$$e_i = \sum_{d=1}^{n_d} \left(\frac{1}{m} - p_d\right) C_d \qquad \text{Eq. (24)}$$

where:

e_i = error in fuel consumption average for the class

m = number of vehicles tested in the class

P_d = proportion of vehicles in class with drive train type d

C_d = consumption level for vehicle with drive train d

n_d = number of drive trains available in class

It should be noted that e_i depends on the difference in fuel consumption levels for vehicles with different drive trains, because both 1/m and p_i will sum to equal 1 so long as n_i = m. Further, the error may be random and tend to cancel across classes.

To evaluate the magnitude of this error, a subsample of the 1975 certification fuel economy data for one major manufacturer was analyzed. It was assumed that the standard axle ratio was used in 90 percent of the vehicles in each inertia weight class, while the remaining 10 percent was evenly distributed among the alternative axle ratios. In cases where some alternatives were not tested, their fuel economy was estimated for the city (CVS-CH) driving cycle using an equation empirically estimated by EPA. Equation (24) was used to determine the error for each inertia weight, then these errors were summed using a production-weighted average to find the net error for the fleet of cars.

The calculations indicated an error in the fleet-wide fuel economy average for the city cycle of -1.3 mpg. Although a rough approximation, this finding strongly suggests that alternative axle ratios be tested and sufficient data collected to allow averaging on a production weighted basis.

4.3.2 Statistical Uncertainty in Fuel Economy Estimates

The preceding section examined errors from incorrect representation of vehicles in the new car fleet. Now we turn to the more fundamental question of the variability of fuel economy estimates even if the fleet is represented through appropriate sampling and averaging. Of primary concern are the inherent variability of fuel economy tests and the degree of stratification of the fleet.

As noted in Chapter 3.0, the standard deviation of fuel consumption estimates appears to be on the order of 4 to 7 percent of mean values when test, laboratory, and production variabilities are included. The final component of variability is a function of vehicle differentiation within classes; its quantification is beyond the scope of this work. In the following calculations, therefore, the total class standard deviation was assumed to be either 5 percent or 10 percent of the class mean fuel consumption. The former value is optimistic, while the latter value is more realistic for the current degree of vehicle differentiation within classes.

Given the above assumptions, the degree of fleet stratification will determine what proportion of the manufacturer's fleet is represented by each class. Recalling Chapter 2.0,

$$V(\bar{C}_f) = \sum_{i=1}^{m} p_i^2 \sigma_i^2 / n_i \qquad \text{Eq. (15, repeated)}$$

$V(\bar{C}_f)$ will be smaller when the fleet is evenly distributed among classes, so that the values for the p_i are uniformly small. Conversely, a few large values among the p_i can greatly increase $V(\bar{C}_f)$ if the n_i values are held constant.

A general analysis of $V(\bar{C}_f)$ can be made with a few simplifications. Assuming that all classes have the same fuel consumption variance, σ^2, we have:

$$V(C_f) = \sigma^2 \sum_i p_i^2 / n_i \qquad \text{Eq. (25)}$$

Since σ is assumed to be a constant fraction of mean fuel consumption, while Σp_i^2 is a function of the fleet stratification system, we can perform a sensitivity analysis, as shown in Table 3. For any given value of mean fuel consumption, $V(\bar{C}_f)$ will depend on sample size and fleet distribution within classes. The worst case occurs when nearly all the fleet is represented in a single class, but only one test vehicle for that class is selected; Σp_i^2 would then approach unity, and $V(\bar{C}_f)$ would approach σ^2. The best case occurs when the fleet is uniformly distributed among classes, so that:

$$p_1 = p_2 \ldots p_m = 1/N \text{ and } n_1 = n_2 = \ldots n_m = 1$$

so that $$V(\bar{C}_f) = \sigma^2/N \qquad \text{Eq. (26)}$$

where N is the sample size. Thus, given the above assumptions, the variance of the estimate of fleet average fuel consumption will fall somewhere between σ^2/N and σ^2.

Table 3 considers sample sizes of 50, 100, and 200 vehicles and mean fuel economy values of 15, 20, and 25 mpg. To calculate the standard deviation of fleet average fuel economy estimates, the assumed mean values were translated into fuel consumption units

TABLE 3

SENSITIVITY ANALYSIS OF FLEET AVERAGE FUEL ECONOMY VARIABILITY

		N = 50		N = 100		N = 200	
$\bar{F}$ = (mpg)	$\sigma_c/\bar{C}$	Best Case $\sigma_{\bar{F}}$(mpg)	Worst Case $\sigma_{\bar{F}}$(mpg)	Best Case $\sigma_{\bar{F}}$(mpg)	Worst Case $\sigma_{\bar{F}}$(mpg)	Best Case $\sigma_{\bar{F}}$(mpg)	Worst Case $\sigma_{\bar{F}}$(mpg)
15	0.05	0.11	0.75	0.08	0.75	0.05	0.75
	0.10	0.21	1.50	0.15	1.50	0.11	1.50
20	0.05	0.14	1.00	0.10	1.00	0.07	1.00
	0.10	0.28	2.00	0.20	2.00	0.14	2.00
25	0.05	0.18	1.25	0.12	1.25	0.09	1.25
	0.10	0.36	2.50	0.25	2.50	0.18	2.50

Note: N = total sample size

$\bar{F}$ = fleet average fuel economy

σ_c = standard deviation of fleet average fuel consumption

$\bar{C}$ = fleet average fuel consumption ($1/\bar{F}$)

$\sigma_{\bar{F}}$ = standard deviation of estimated fleet average fuel economy, calculated as:

$$\sigma_{\bar{F}} = 1/2 \left[\frac{1}{\bar{C} - \sigma_{\bar{c}}} - \frac{1}{\bar{C} + \sigma_{\bar{c}}} \right]$$

(gallons per mile). Then class fuel consumption standard deviations were calculated as either 5 or 10 percent of mean fuel consumption. The best and worst case fleet average fuel consumption standard deviations were then translated back into the more familiar fuel economy units.

Values in the table are indicative of how the variability of fleet average fuel economy estimates depends upon several factors. The ranges between best and worst case values are particularly noteworthy since they show the importance of the degree of stratification of the fleet. Larger sample sizes will reduce the variability of the estimate only if they allow a more uniform representation of the fleet across classes.

Clearly, the worst case values in Table 3 are overly pessimistic since they assume that one class represents the entire fleet of new cars. Nevertheless, there is substantial concentration of U.S. auto sales in a small number of vehicle configurations. To provide a more realistic approximation of uncertainty, the 1975 model/year fuel economy data for one major manufacturer were analyzed.

In this analysis, the standard deviation of class fuel consumption was again assumed to be either 5 or 10 percent of the measured class mean fuel consumption. To calculate the overall variance of the fleet average fuel economy estimate, the actual proportions of the total fleet falling into each class were determined. For simplicity, it was assumed that vehicle selections within each class were random. Data were supplied by the Automotive Energy Efficiency Program (AEEP) of the TSC.

Results are summarized in Table 4. The standard deviation for the estimate of fleet average fuel consumption was calculated as 0.0009 gallons per mile assuming class standard deviations at 5 percent of mean class fuel consumption and 0.0018 gallons per mile assuming 10 percent. The accuracy of the fleet average fuel economy estimate may be expressed in terms of a confidence interval. Assuming normal distribution for fuel consumption, the widths of 90 to 95 percent confidence intervals are shown in Table 4. They

TABLE 4

CALCULATED VARIABILITY OF A MAJOR MANUFACTURER'S FLEET AVERAGE FUEL ECONOMY ESTIMATE

Assumed Class Variability ($\sigma_i/\bar{C}_i$)	Standard Deviation of Average Fuel Consumption Estimate ($\sigma_{\bar{c}}$) (gal/mile)	Confidence Interval Width* for Fleet Average Fuel Economy Estimate (mpg)	
		90%	95%
0.05	0.0009	0.76	0.90
0.10	0.0018	1.53	1.80

*Assuming normal distribution, the width of the confidence interval is equal to:

$$\left[\frac{1}{\bar{C} - z_\alpha \sigma_{\bar{c}}} - \frac{1}{\bar{C} + z_\alpha \sigma_{\bar{c}}}\right]$$

where $z_\alpha = 1.645$ at the 90 percent confidence level and 1.96 at the 95 percent level. $\bar{C}$ in this case was calculated from the manufacturer's data.

imply that the true fuel economy average for the manufacturer can be established to within about 0.8 to 1.8 miles per gallon depending on confidence level and class variance assumptions. The actual accuracy is probably lower because of non-random vehicle selection and errors in fleet representation.

Even so, this level of accuracy is acceptable for many purposes. However, the new regulations will impose penalties based on 0.1 mpg increments in average fuel economy, so, more precise estimates would be highly desirable. The next section presents alternatives for achieving improved accuracy.

5.0 METHODS FOR INCREASING PRECISION

5.1 INTRODUCTION

In applications of ordinary random sampling, the accuracy of a parameter estimate can be increased only by enlarging the sample. The variance of the estimate in stratified sampling, however, depends on more than just total sample size. It is possible, therefore, to define vehicle selection procedures which will increase the precision of fleet average fuel economy estimates without increasing the size of the sample.

Because the current EPA procedures were originally designed for other purposes, they do not provide optimal fuel economy estimates. Some relatively simple modifications, discussed below, would allow more accurate estimates to be made for any given sample size.

5.2 DISTRIBUTION OF THE SAMPLE AMONG CLASSES

The variance of the fleet average fuel consumption estimate depends on three factors: the proportion of the fleet in each class, the fuel consumption variance in each class, and the number of observations in each class. This was formally expressed in the formula:

$$V(\bar{C}_f) = \sum_{i=1}^{m} p_i^2 \sigma_i^2 / n_i \qquad \text{Eq. (15 repeated)}$$

For the moment, let us consider classes as fixed, with the values of p_i known, and those of σ_i unknown. The question, then, is how the sample of test vehicles should be allocated among classes.

EPA currently does not differentiate among classes in allocating the sample. Since their vehicle selection criteria are the same for every engine family, the result is an approximately equal number of test vehicles for each class. If the total sample size is N and the number of class m, the class sample sizes will be

$$n_1 = n_2 = \cdots = n_m = N/m$$

and $V(\bar{C}_f)$ will become:*

$$V(\bar{C}_f) = m/N \sum_{i=1}^{m} p_i^2 \sigma_i^2 \qquad \text{Eq. (27)}$$

This result may be improved by recognizing that increasing the class sample size, n_i, can help offset large values of p_i:

$$n_i = Np_i \qquad \text{Eq. (28)}$$

$$N = \sum_{i=1}^{m} n_i \qquad \text{Eq. (29)}$$

This approach is known as proportional sampling, and yields a variance for the fleet average of

$$V(\bar{C}_f) = 1/N \sum_{i=1}^{m} p_i \sigma_i^2 \qquad \text{Eq. (30)}$$

Proportional sampling, thus, can decrease the variance by an amount equal to:

$$\frac{1}{N} \sum_{i=1}^{m} \sigma_i^2 p_i (mp_i - 1)$$

In classes with small p_i, the reduced sample size will make little difference. In classes with large p_i, however, the increased value of n_i will significantly reduce variance. When all classes contain approximately equal proportions of the fleet (i.e., $p_i = p_2 = .., = 1/m$), there is little difference in the resulting variances.

Proportional sampling is often used when there is no prior information available on class variances. However, it usually will not provide the theoretical minimum variance for the fleet average estimate. The optimal allocation of the sample may be found by minimizing $V(\bar{C}_f)$ with respect to n_i; that is:

$$dV(\bar{C}_f) = -\sum_{i=1}^{m} (p_i \sigma_i / n_i)^2 \, dn_i \equiv 0 \qquad \text{Eq. (31)}$$

*Actual variance might be higher due to integer rounding-off of the n_i.

solving for n_i gives:

$$n_i = \frac{Np_i\sigma_i}{\sum_{i=1}^{m} p_i\sigma_i}, \quad i=1,2,3\ldots m \qquad \text{Eq. (32)}$$

Here, the class sample sizes are proportional not to the fractions p_i alone but to the products $p_i\sigma_i$. Note that proportional sampling <u>is</u> optimal in the special situation of constant class variances, i.e., $\sigma_1=\sigma_2=\ldots\sigma_m=K$. In this case, Equation (32) yields:

$$n_i = \frac{Np_iK}{K\Sigma p_i} = Np_i \qquad \text{Eq. (33)}$$

In general, of course, the σ_i are not known, but if they can be estimated fairly accurately, the optimal sampling procedure will offer significant improvements over proportional sampling. Indeed, if the optimal approach were followed precisely, we would have:*

$$V(\bar{C}_f) = 1/N\left(\sum_{i=1}^{m} p_i\sigma_i\right)^2$$

$$= 1/N \sum_{i=1}^{m} p_i\sigma_i^2 - 1/N \sum_{i=1}^{m} p_i(\sigma_i-\bar{\sigma})^2 \qquad \text{Eq. (34)}$$

where:

$$\bar{\sigma} = \sum_{i=1}^{m} p_i\sigma_i \qquad \text{Eq. (35)}$$

Comparison of equations (30) and (35) shows that the reduction in variance in going from proportional to optimal sampling is:

$$1/N \sum_{i=1}^{m} p_i(\sigma_i-\bar{\sigma})^2$$

The variance of fuel consumption in each class of vehicles cannot be anticipated with precision. Clearly, however, classes with greater differences among vehicles will have larger fuel con-

*Integer rounding off of the n_i would prevent minimum variance.

sumption variances. It may be possible, therefore, to make rough estimates of class variances in advance of vehicle selection to allow implementation of optimal sampling.

5.3 TEST REPLICATIONS

Optimal sampling within classes as described above will make the most efficient use of a fixed number of test vehicles provided that only one test is performed on each vehicle. The fleet average fuel consumption variance can be further reduced, however, by replicating tests without increasing N, the sampling size. Two components of class variance, test and laboratory variabilities, are susceptible to reduction in this fashion. Production and vehicle differentiation variance components (σ_{ip} and σ_{id}) will remain constant despite replications.

To characterize this effect mathematically, we can consider the effect of testing each sample vehicle in multiple laboratory cells several times. Then:

$$\bar{C}_{ij\ell} = \frac{1}{n_{it}} \sum_{t} C_{ij\ell t} \qquad \text{Eq. (36)}$$

$$\bar{C}_{ij} = \frac{1}{n_{i\ell}} \sum_{\ell} \bar{C}_{ij\ell} \qquad \text{Eq. (37)}$$

$$\bar{C}_{i} = \frac{1}{n_{i}} \sum_{j} \bar{C}_{ij} \qquad \text{Eq. (38)}$$

where:

$C_{ij\ell t}$ = measured fuel consumption in the t^{th} test in the ℓ^{th} lab cell of the j^{th} vehicle in class i

$\bar{C}_{ij\ell}$ = computed average fuel consumption of the j^{th} test vehicle for lab cell ℓ

$\bar{C}_{ij}$ = computed average fuel consumption for the jth vehicle

n_{it} = number of test replications per lab cell in class i

$n_{i\ell}$ = number of lab cells employed for class i

and the other variables are as previously defined.

Each measured fuel consumption value will have random error components arising from the test procedure, laboratory facilities, and, as a representative of class fuel consumption, from production variability and vehicle differentiation. Thus, we may establish the variances associated with the above calculations as:

$$V(C_{ij\ell t}) = \sigma^2_{it} + \sigma^2_{i\ell} + \sigma^2_{ip} + \sigma^2_{id} \qquad \text{Eq. (39)}$$

$$V(\bar{C}_{ij\ell}) = \sigma^2_{it}/n_{it} + \sigma^2_{i\ell} + \sigma^2_{ip} + \sigma^2_{id} \qquad \text{Eq. (40)}$$

$$V(\bar{C}_{ij}) = \sigma^2_{it}/n_{i\ell}n_{it} + \sigma^2_{i\ell}/n_{i\ell} + \sigma^2_{ip} + \sigma^2_{id} \qquad \text{Eq. (41)}$$

$$V(\bar{C}_{i}) = 1/n_i(\sigma^2_{it}/n_{i\ell}n_{it} + \sigma^2_{i\ell}/n_{i\ell} + \sigma^2_{ip} + \sigma^2_{id}) \qquad \text{Eq. (42)}$$

Increasing the number of test replications per laboratory cell will decrease only the test variance. Increasing the number of laboratory cells employed decreases both σ^2_{it} and $\sigma^2_{i\ell}$ since each vehicle is tested more total times over more facilities. Finally, increasing n_i, the number of test vehicles in the class, decreases all components of class variance.

The overall fleet average fuel consumption variance may now be calculated as:

$$V(\bar{C}_f) = \sum_{i=1}^{m} p_i^2(\sigma^2_{it}/n_{i\ell}n_{it} + \sigma^2_{i\ell}/n_{i\ell} + \sigma^2_{ip} + \sigma^2_{id})/n_i \qquad \text{Eq. (43)}$$

and compared with the earlier:

$$V(\bar{C}_f) = \sum_{i=1}^{m} p_i{}^2\sigma_i{}^2/n_i \qquad \text{Eq. (15, repeated)}$$

$$= \sum_{i=1}^{m} p_i^2(\sigma^2_{it} + \sigma^2_{i\ell} + \sigma^2_{ip} + \sigma^2_{id})/n_i \qquad \text{Eq. (44)}$$

with the reduction apparent when n_{it} and $n_{i\ell}$ are greater than one.

If more than one laboratory cell is used for test replicatic the optimal sampling formula from the preceding section must be modified accordingly. Although n_i, $n_{i\ell}$, and n_{it} can be optimized simultaneously to minimize the variance of the fleet average, a full mathematical explication of the process is beyond the scope of this report. The utility of test replications, however, shoul be noted since it is far less expensive to repeat fuel economy tests than to construct and prepare additional test vehicles.

5.4 DEFINITION OF CLASSES

In the preceding discussion, vehicle classes were assumed tc be fixed, so that only the n_i could be altered to improve accurac In fuel economy estimation, however, class definitions are flexib and more detailed classes can be specified, if necessary, until e class contains nominally identical vehicles; i.e., every configur is a class. This can allow a further decrease in the class fuel consumption variances and, thus, a further increase in the accura of the fleet average estimate.

Let us assume that an initial set of classes and class sampl sizes have been specified and class variances anticipated. In th i^{th} class, the variance σ_i^2 is expected to be particularly large. We will, therefore, investigate the benefits of further stratify: the i^{th} class into subclasses. Suppose, for example, that class i contains vehicles with the same basic engine, displacement, inertia weight, and transmission, but several different axle rati and engine calibrations. New subclasses may be defined such that each contains a single axle ratio and engine calibration.

If we sample within these subclasses, our class variance bec

$$V(\bar{C}_i) = \sum_{j=1}^{\ell} p_{ij}^2 \sigma_{ij}^2 / n_{ij} \qquad \text{Eq. (45}$$

where:

$V(\bar{C}_i)$ = variance of estimated average fuel consumption in class i

p_{ij}^2 = proportion of class i vehicles in subclass j

σ_{ij}^2 = variance of fuel consumption in subclass j

n_{ij} = number of test vehicles in subclass j ($\sum_{j=1}^{\ell} n_{ij} = n_i$)

This variance will be smaller than we would have obtained with a random sample of vehicles in class i. Indeed, a random sample would yield:

$$\sigma_i^2 = \sum_{j=1}^{\ell} p_{ij}\sigma_{ij}^2 + \sum_{j=1}^{\ell} p_{ij}(\mu_{ij} - \mu_i)^2 \qquad \text{Eq. (46)}$$

where:

μ_{ij} = the true mean fuel consumption in subclass j of class i

μ_i = the true mean fuel consumption in class i

and the other variables are as defined previously. This is at least as large as the weighted average of subclass variances and will be larger unless all subclasses have the same mean fuel consumption.

The allocation of the n_i observations, of course, should be either proportional or optimal as discussed above. Proportional sampling (when the σ_{ij}^2 are unknown) will provide a variance of:

$$V(\bar{C}_i) = 1/n_i \sum_{j=1}^{\ell} p_{ij}\sigma_{ij}^{\ 2} \qquad \text{Eq. (47)}$$

which is less than the variance of the estimated class mean with random sampling by:

$$1/n_i \sum_{j=1}^{\ell} p_{ij}(\mu_{ij} - \mu_i)^2$$

Optimal sampling, if the σ_{ij}^2 can be anticipated, will provide a further reduction, so that the net decrease from random sampling would be:

$$1/n_i \sum_{j=1}^{\ell} p_{ij}(\mu_{ij} - \mu_i)^2 + 1/n_i \sum_{j=1}^{\ell} p_{ij}(\sigma_{ij} - \bar{\sigma}_i)^2$$

where:

$$\bar{\sigma}_i = \sum_{j=1}^{\ell} p_{ij}\sigma_{ij}$$

This amount is the maximum available reduction in variance from a more detailed stratification of class i vehicles.

Whether or not this maximum reduction can be achieved depends not only on information about the σ_{ij}^2 values, but also on the number of observations, n_i, allocated originally to class i. Clearly, n_i must be equal to or greater than the number of subclasses, ℓ. Even if n_i permits only one observation in each subclass, however, the class variance, $V(\bar{C}_i)$, should be reduced from the random sampling approach.

The central point of this example is that class definitions based on vehicles parameters are not sacrosanct. The same criteria for definition need not, and indeed should not, apply to every class. In some circumstances, an entire engine family might be included in a single class, while another engine family might be stratified completely so that each vehicle configuration represented a unique class. The critical parameters for class definition are p_i and σ_i; physical characteristics of vehicles are important only insofar as they influence σ_i.

5.5 ACCURACY WITH OPTIMAL SAMPLING

We can now explore the potential benefits of optimal stratification and sampling. Of particular interest are: (a) the improvement in precision of the fleet average fuel economy estimate holding sample size constant, and (b) the sample size required to obtain a pre-specified level of accuracy.

Accuracy will be defined in terms of confidence margin width. For a calculated fleet average fuel consumption and standard deviation, we have:

$$Pr\left(\frac{1}{\bar{C}_f + z_\alpha \hat{\sigma}_{\bar{c}}} < \mu_f < \frac{1}{\bar{C}_f - z_\alpha \hat{\sigma}_{\bar{c}}}\right) \geq 1-\alpha \qquad \text{Eq. (17, repeate}$$

where:

$Pr(...)$ = the probability that ...

$\overline{C}_f$ = calculated average fuel consumption

z_α = the standard normal deviate for confidence α

$\hat{\sigma}_{\overline{c}}$ = estimated standard deviation of $\overline{C}_f$

μ_f = true mean fuel economy of the fleet

α = desired level of confidence

The width of the confidence interval is thus:

$$W = \frac{1}{C_f - z_\alpha \hat{\sigma}_{\overline{c}}} - \frac{1}{C_f + z_\alpha \hat{\sigma}_{\overline{c}}} = \frac{2 z_\alpha \hat{\sigma}_{\overline{c}}}{\overline{C}_f^2 - z_\alpha^2 \hat{\sigma}_{\overline{c}}^2} \qquad \text{Eq. (48)}$$

Now $\hat{\sigma}_{\overline{c}}$ is the square root of $V(\overline{C}_f)$, which in optimal sampling is

$$\hat{\sigma}_{\overline{c}} = \left[1/N \sum_{i=1}^{m} p_i \sigma_i \right]^{1/2} \qquad \text{Eq. (49)}$$

The summation represents a weighted average of class fuel consumption standard deviations, which we will call $\underline{\sigma}$. Thus, our confidence interval width may be restated as:

$$W = \frac{2 z_\alpha \underline{\sigma}}{(N)^{1/2} (\overline{C}^2 - z_\alpha^2 \underline{\sigma}^2 N^{-1})} \qquad \text{Eq. (50)}$$

By assuming that $\underline{\sigma}$ is equal to either 5 or 10 percent of mean fuel consumption, we can determine W for several different sample sizes and confidence levels. Results are shown in Table 5, assuming a mean fuel economy of 18 mpg ($\overline{C}_f$ = 0.0556), the mandated standard for 1978.

For current sample sizes of 50 to 200 vehicles per manufacturer, optimal sampling would permit confidence intervals ranging in width from 0.16 to 1.10 mpg, depending on the variance, actual sample size, and confidence level. The sample size of 50 is representative of smaller manufacturers, while 100 to 200 vehicles are typically tested for the major domestic companies.

Even with optimal sampling, it appears that fleet fuel economy averages cannot be estimated to within 0.1 mpg unless sample sizes are increased. We can estimate the necessary sample size for this level of accuracy by simplifying Equation (50):

$$W \sim \frac{2z_\alpha \sigma}{\bar{C}_f^{\,2} (N)^{1/2}} \qquad \text{Eq. (50, modified)}$$

so that:

$$N \sim \left[\frac{2z_\alpha \sigma}{W\bar{C}_f^{\,2}}\right]^2 \qquad \text{Eq. (51)}$$

Using this approximation, and assuming that fleet average fuel economy is near the 18 mpg standard for 1978, required sample sizes at various levels of confidence are shown in Table 6. The required number of test vehicles per manufacturer ranges from 530 to 4980.

5.6 IMPLEMENTATION OF IMPROVED PROCEDURES

The uncertainties in the above values must be stressed. Nevertheless, they have important implications for increasing the precision of fleet average fuel economy estimates. In particular, if accuracy to within 0.1 mpg is required for imposing penalties and credits under the new fuel economy regulations, substantially larger sample sizes will be necessary even with optimal sampling procedures.

Equally important, however, is the need to eliminate errors and bias arising from inappropriate representation of vehicle parameter variations within classes. Problems of optional equipment, transmission, and axle ratio differences as discussed in Chapter 4.0 may be grouped under a single heading: non-random vehicle selection within classes. Without question, this is partially a result of

TABLE 5

CONFIDENCE INTERVAL WIDTHS
WITH OPTIMAL SAMPLING

Sample Size	$\sigma/\bar{C}_f$*	Confidence Interval Width (mpg) for Confidence Level of 80%	90%	95%
50	0.05 0.10	0.32 0.65	0.42 0.83	0.50 0.99
100	0.05 0.10	0.23 0.46	0.30 0.59	0.35 0.71
200	0.05 0.10	0.16 0.33	0.21 0.42	0.25 0.50

*$\bar{C}_f$ = 0.0556

TABLE 6

SAMPLE SIZE REQUIRED
FOR CONFIDENCE INTERVAL WIDTH
OF 0.1 MPG

$\sigma/\bar{C}_f$	Required Sample Size* for Confidence Level of 80%	90%	95%
0.05	530	880	1245
0.10	2120	3510	4980

*Number of vehicles per manufacturer (approximate).

missing data -- the production quantities of different axle ratios within engine families for example. It also arises because of the current need to select high-emission configurations despite their possible irrelevance or non-representativeness for fuel economy calculations.

Indeed, there appears to be a fundamental conflict between emissions certification and fuel economy vehicle selection criteria. Because many emission certification vehicles are low production or non-representative configurations, they "waste" observations from a fuel economy standpoint. Thus, a larger sample will be necessary for a given level of accuracy when such vehicles are included. In addition, care must be taken that supplemental vehicles are tested from the same class as the non-representative emission vehicles, so that biased estimates of class fuel economy are avoided.

Further and more detailed study of the components of fuel economy variance identified in Chapter 3.0 is needed prior to implementation of refined theoretical approaches. Obviously, calculations of confidence intervals with assumed levels of class fuel economy variability are no more than guesses. Any attempt to estimate fleet average fuel economy levels with precision in the absence of better information on class variances will be futile.

Finally, in implementing improved procedures, it should be noted that accuracy to within 0.1 mpg, while convenient, is not required from a statistical point of view. There are several optional approaches to determining penalties and credits, discussion of which is beyond the scope of this report. However, the benefits of extreme accuracy should be weighted against the costs of achieving it before the decision to test thousands of additional vehicles is made.

APPENDIX A

REFERENCES

1. Public Law 94-163.
2. Kruse, R. E., and Huls, T. A., "Development of the Federal Urban Survey Schedule," SAE Paper 740970, October 1974.
3. Austin, T. C., and Hellman, K. H., "Passenger Car Fuel Economy During Non-Urban Driving," SAE Paper 740592, August 1974.
4. Austin, T. C., Michael, R. B., and Service G. R., "Passenger Car Fuel Economy Trends Through 1976," SAE Paper 750957, October 1974.
5. Paulsel, C. D., and Kruse, R. E., "Test Variability of Emission and Fuel Economy Measurements Using the 1975 Federal Test Procedure," SAE Paper 741035, 1974.
6. Matula, R. A. "Emissions and Fuel Economy Test Methods and Procedures," Consultant Report to the Committee on Motor Vehicle Emissions. National Research Council, Washington D.C., September 1974.
7. Leiferman, M. W. and Wilson, G. M., "1975 EPA-USMVA Correlation Study," U.S. Environmental Protection Agency, June 1975.
8. "Report on Honda Emission Control Systems for 1975 and Subsequent Model Years," prepared for the Committee on Motor Vehicle Emissions of the National Academy of Sciences, May 1974.
9. Narren, J. M., "1975 Data Vehicle Selection Criteria," U.S. Environmental Protection Agency, January 1974.

APPENDIX B
REPORT OF INVENTIONS

This report provides a review and evaluation of procedures used for calculating U.S. car fleet average fuel economy, and offers recommendations for increasing the accuracy of these calculations. A diligent review of the work performed under this contract has revealed no innovation, discovery, improvement, or invention.

115 copies

U. S. DEPARTMENT OF TRANSPORTATION
TRANSPORTATION SYSTEMS CENTER
KENDALL SQUARE, CAMBRIDGE, MA. 02142

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300