

California Inspection and Maintenance Review Committee

An Analysis of the USEPA's 50-Percent Discount for Decentralized I/M Programs

• • • • • • • • • • • •

Prepared by

Joel Schwartz
Special Consultant

Lynn Scarlett, Chair

Detrich Allen
Jerome Aroesty
Joseph Charney
Norm Covell
Elizabeth Deakin
Dennis DeCota
Robert Machon
Joseph Norbeck
Daniel O'Connell
Paul Roberts

Buzz Breedlove, Executive Officer
Joel Schwartz, Special Consultant

900 N Street, Suite 300
Sacramento, CA 95814

February 24, 1995

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
INTRODUCTION.....	5
USEPA'S I/M PERFORMANCE STANDARD.....	5
USEPA'S CASE FOR ITS 50-PERCENT DISCOUNT	6
USEPA'S STATEMENTS ON THE SIZE OF THE DISCOUNT	6
WHAT SOURCES OF DATA DOES THE USEPA CITE TO SUPPORT ITS 50-PERCENT DISCOUNT?	7
WHAT I/M PROGRAM ELEMENTS DOES USEPA CONSIDER TO BE THE MOST IMPORTANT DETERMINANTS OF I/M EFFECTIVENESS?	7
DO THE USEPA'S AUDITS MEASURE I/M EFFECTIVENESS?	8
RESULTS OF USEPA'S TAMPERING SURVEYS OF I/M PROGRAMS.....	24
ASSESSING I/M EFFECTIVENESS USING ON-ROAD AND AMBIENT VEHICLE EMISSIONS DATA	28
CENTRALIZED PROGRAMS	28
DECENTRALIZED PROGRAMS.....	30
AMBIENT OZONE AND CO DATA FROM 1983 AND 1992.....	31
CONCLUSIONS	32
WHERE DO WE GO FROM HERE?	32
REFERENCES.....	34
APPENDIX A	38
APPENDIX B	41
APPENDIX C	42
APPENDIX D	44
APPENDIX E	47
GLOSSARY.....	48

EXECUTIVE SUMMARY

One of the major areas of contention in the vehicle inspection and maintenance (I/M) debate between the state of California and the United States Environmental Protection Agency (USEPA) is the USEPA's 50-percent discount for decentralized I/M programs, relative to centralized programs. The staff of the California Inspection and Maintenance Review Committee (staff) has evaluated the scientific basis for the USEPA's 50-percent discount by evaluating the USEPA's audits and tampering surveys of I/M programs, as well as other researchers' studies of on-road vehicle emissions. We have reached the following conclusions:

- Direct measurements of the variables that I/M is supposed to reduce - emission system tampering and tailpipe emission levels - show little difference between centralized and decentralized I/M programs, and also little or no effect due to current I/M programs. The data do not indicate that decentralized programs are working well. But, neither do they indicate that centralized programs are working any better. We conclude that *whether an I/M program is centralized or decentralized has not been an important factor in determining I/M program effectiveness.*
- The USEPA's audits are not suitable for evaluating I/M programs on two accounts. First, the audits were performed using statistically invalid research designs. Second, since direct measurements of emissions and tampering from on-road vehicles show little or no difference between centralized and decentralized I/M programs, we conclude that the USEPA's audits, at best, evaluate only a subset of the variables affecting I/M performance.
- Combining these results, staff concludes that there is no empirical or scientific basis for a discount for decentralized I/M programs relative to centralized programs.

The following is a summary of our findings for each step in our analysis.

- **On-road and ambient measurements of vehicle emissions indicate that both centralized and decentralized I/M programs have performed poorly**
 - Trends in ambient ozone and CO levels between 1983 and 1992 show that, on average, regions with decentralized I/M and regions with centralized I/M had about the same reduction in the two pollutants.
 - On-road pull-over and on-road remote-sensing studies show that, in both centralized and decentralized I/M programs, there is little or no correlation between vehicle emissions and time since the last I/M test. In other words, cars that recently had an inspection had similar emissions to cars that hadn't had an inspection in a long time.
 - These same studies indicate that cars not subject to an I/M program (i.e., cars registered in nearby non-I/M counties) had emissions similar to the emissions from cars that were subject to I/M.
 - A study of ambient CO levels in Minnesota's annual centralized I/M program showed that 2.5 years after implementation of the I/M program, CO levels declined an additional 1.4% over the decline that would have been expected based on the pre-I/M trend. USEPA's model predicted about a 25% reduction in CO due to the I/M program.

- **USEPA's tampering surveys show little difference in tampering rates between centralized and decentralized I/M programs**

- Centralized programs had slightly lower tampering rates than decentralized programs. However, the small advantage for centralized programs may be due to selection bias. In centralized programs, cars were inspected for tampering as they pulled into centralized test lanes, while in decentralized programs, cars were pulled over at random on the road. Motorists are probably less likely to bring a tampered car to the emission test than they are to drive a tampered car on the road.

I/M programs are supposed to reduce on-road emissions and tampering. However, real-world data on vehicle emissions and tampering do not show a difference between centralized and decentralized I/M programs.

USEPA bases its 50-percent discount, not on measurements of vehicle emissions, but on its audits and tampering surveys of I/M programs. In particular, USEPA places great weight on its covert audits of improper testing rates in I/M programs. In these audits, USEPA set up cars to fail the visual test (by removing catalytic converters for instance), and then covertly took the cars for an inspection. In some cases the cars were set up to fail the emission test as well. When we looked at the audit data, we found a number of methodological and other data problems.

- **USEPA analyzed data that included errors in favor of centralized I/M programs**

USEPA has produced two documents that summarize its various audit reports and analyze its audit data. These two summary documents sometimes do not reflect the actual results obtained by USEPA in its original audits. For example:

- USEPA reports zero-percent improper test rates for both visual and emissions testing in Maryland's centralized program. However, USEPA's Maryland covert audit found a 20 percent improper visual test rate and a 28 percent improper emission test rate. In addition, in four of 10 cases, inspectors failed to verify vehicle identification information. Overall, the improper test rate was 40 percent in Maryland's centralized I/M program.
- USEPA reports a 50 percent improper visual test rate in New Jersey's centralized I/M program. Its audit of New Jersey's program actually found a 66 percent improper test rate.
- USEPA reports a 46 percent improper emission test rate for New York's decentralized program. However, USEPA's New York audit found a 37 percent improper emission test rate.

- **USEPA's audits included structural biases against decentralized I/M programs**

- USEPA cites Arizona's centralized I/M program as having a low rate of improper testing based on its Arizona covert audit. However, in its audit of Arizona's program, USEPA sent a single car to nine different test stations. When a car fails in Arizona's I/M program, future tests of that car are marked on the Arizona I/M computer network as retests. Every inspector after the first one that failed the car would know that the car had previously been tested and had failed. Since inspectors knew they were testing a car that previously failed, USEPA gave Arizona's I/M program an easier test than it gave to decentralized programs.
- USEPA states that is used "subtle deception" in the way it tampered the car used in its covert audit of Georgia's decentralized I/M program. USEPA replaced the car's catalytic converter with something that looked a bit like a catalytic converter. However, in covert audits of centralized I/M programs, USEPA removed the catalytic converters from covert

cars and replaced them with straight pipes. In other words, USEPA gave Georgia's decentralized program a tougher test than it gave centralized I/M programs.

- In Missouri's decentralized I/M program audit, USEPA appears to have targeted its audits to stations that it expected would perform poorly, rather than by random selection of stations. This would unfairly increase the rate of improper testing observed in the audits. In addition, the cars used in this covert audit were 4 and 5 years old. USEPA's tampering surveys show that the actual catalyst tampering rate for cars of this age is virtually zero. Mechanics might be less likely to look for tampered catalysts on newer cars, based on their experience of lower tampering rates in newer cars.

The USEPA's audit data contain a number of methodological biases and idiosyncrasies. Biases in the conduct of the audits make the quantitative results of the audits unreliable. Idiosyncrasies in the methodologies between audits of different I/M programs make the data unsuitable for inter-program comparisons. Finally, spotty documentation of the audit methods and results makes it difficult to systematically assess the meaning of the audit data.

Although the USEPA's audit methodologies may not be scientifically valid, could the audits, in principle, be used to assess the effectiveness of an I/M program in reducing on-road vehicle emissions? We find that the USEPA's audits cannot be used for such an assessment for the following reasons:

- **USEPA did not collect audit data that could be used to assess I/M emission reductions**
 - USEPA does not show a link between proper testing and emission reductions. USEPA's audits generally do not include emission data on cars used in its covert audits. Covert audits did not include repairs and retests that could be used to measure emission reductions.
 - In cases where emission data are provided for cars used in covert audits, the cars have emissions near the failure cutpoints. However, most emission reductions in an I/M program come from detecting and repairing high emitters.
- **USEPA assumes, but does not demonstrate, that its audit measures reflect on-road emission reductions:** USEPA asserts, but does not demonstrate, that the variables it measured relate directly to the effectiveness of I/M programs in reducing on-road emissions. However, USEPA does not present data or analyses that show a relationship between its audit measurements and on-road emission-reduction performance of an I/M program.
- **USEPA considers only a subset of the variables that affect I/M effectiveness:** USEPA overlooks feedbacks that occur when different portions of a program are changed. For instance, if improper testing is curtailed, then motorists might be more likely to seek waivers. If program officials clamp down on waivers, motorists might look for mechanics who can superficially adjust cars to pass the test without fixing defects that might cause high emissions between tests. Motorists might also be more likely to register their cars in non-I/M counties, or not register their cars at all. A whole range of behaviors might occur that are not accounted for in the USEPA's analysis. I/M is a system of interacting variables that includes the motivations of motorists and mechanics. The USEPA's audit measures encompass only a subset of the factors that affect the operation of an I/M system. They are at best an incomplete description of the relevant variables.

- **USEPA does not present a methodology for converting the audit data into a discount:**
The USEPA presents no mathematical or statistical analysis that starts with its audit data and ends with a 50-percent discount. The USEPA presents numerical data on the audit variables, and then makes a number of statements about what variables it believes are important in determining I/M effectiveness. But the path from there to the 50-percent discount (and its numerical precision), or any other posited discount, is left uncharted. Because the USEPA does not show a quantitative link between the audit data and emission reductions due to I/M, it is difficult to understand how the USEPA could have arrived at any particular number for the discount by using the audit data.

The USEPA states that *"the most critical aspect in evaluating an I/M program is the emission reduction benefit it achieves."* We wholeheartedly agree with this statement. It underlies an approach that is essential to determining whether or not any program is achieving real benefits - namely, measuring pollution levels in ambient air and emissions from vehicles as they are driven on the road (or at the very least measuring surrogates that have a demonstrated relationship to on-road emissions). Future measurements of I/M effectiveness should be based on this approach.

While I/M effectiveness has fallen short of expectations, repair of gross-emitting vehicles has been shown to generate significant emission reductions . The California I/M Review Committee's charge is to help point California down the road to an effective I/M program. California's new I/M laws have created many tools for finding gross polluters, sanctioning non-complying motorists, and removing dishonest or incompetent mechanics from the test-and-repair industry. We hope that the USEPA and California will work jointly in developing an I/M program that will capture the potential emission reduction benefits available from effective I/M.

INTRODUCTION

One of the major areas of contention in the vehicle inspection and maintenance (I/M) debate between the state of California and the United States Environmental Protection Agency (USEPA) is the USEPA's treatment of decentralized¹ I/M programs. The USEPA grants half the emission reduction credit to decentralized I/M programs relative to centralized I/M programs. This credit system is referred to as the "50-percent discount." The USEPA states that its 50-percent discount is backed up by data from its audits and surveys of I/M programs around the country. However, after reviewing the USEPA's studies, as well as other relevant information, the staff of the California Inspection and Maintenance Review Committee (staff) concludes that the USEPA's 50-percent discount is not justified. This paper presents our evaluation of the data and analyses behind the USEPA's 50-percent discount.

We evaluate the USEPA's 50-percent discount by asking the following questions:

- Is the USEPA's 50-percent discount based on its audits and surveys?
- Are the USEPA's audits and surveys conducted in a scientifically and statistically valid manner?
- Do better scores on the USEPA's audits translate into greater emission reduction effectiveness for a given I/M program?
- Do the USEPA's tampering surveys show a difference between centralized and decentralized I/M programs?
- Do on-road and ambient measurements of vehicle emissions show a difference between centralized and decentralized I/M programs?

USEPA'S I/M PERFORMANCE STANDARD

The USEPA has developed a performance standard for I/M programs around the country. The performance standard is the maximum fleet average emissions level that a motor vehicle fleet is expected to attain after implementation of an I/M program. The USEPA bases its performance standard on its prediction that a centralized I/M program using IM240 test equipment will achieve the greatest possible emissions benefit from I/M.² This program is termed the "model program" by the USEPA. The USEPA compares the emission reductions achievable from a given I/M program against its model program by running its vehicle emissions model, MOBILE5a, for the model program and for other potential I/M program designs.

In its I/M Rule (USEPA, 1992a) and supporting materials (Tierney, 1993a), the USEPA describes how it will compare a given I/M program to its performance standard. To compare I/M programs to its model program, the USEPA:

- Estimates current average fleet emissions, in grams per mile (gpm), using MOBILE5a. The USEPA inputs local data to MOBILE5a for vehicle registration mix, VMT distribution, ambient maximum and minimum temperatures, average vehicle speeds, federal or California certification standards, etc.

¹ See the glossary for an explanation of I/M program terminology.

² The USEPA's model I/M program also includes pressure and purge testing of the evaporative emissions control system, a \$450 repair cost limit, as well as other features.

- Establishes the performance standard - a target average gpm emission level for cars in a given I/M program region - by running MOBILE5a with the same local inputs, and with the assumption that the USEPA's model I/M program will be implemented.
- Projects the performance standard for years between now and 2010, using assumed future values for the local inputs described above.³
- Runs MOBILE5a for other I/M program designs that a region may wish to implement to see if they meet the performance standard for all years between now and 2010.

The USEPA assumes that when states deviate from its model program, the emission reductions will be less than for the model program. In particular, the USEPA gives half the emission reduction credit to decentralized programs, even if they differ in no other way from the USEPA's model program.⁴

Since 1987, a number of independent studies have called the predictive capability of the MOBILE model series into question.⁵ The fact that vehicle emissions models do not come close to predicting measured on-road emissions is a matter of consensus in the scientific community (Ingalls, 1989; Pierson et al., 1990; National Research Council, 1991; Lurmann and Main, 1992; Fujita et al., 1992; Pierson et al., 1992; Kirchstetter and Harley, 1994). Scientists, economists, and policy analysts have pointed out numerous flaws in the data inputs, assumptions, and methodologies used in MOBILE (Stedman, 1995; Harrington and McConnell, 1994; McConnell and Harrington, 1992). Furthermore, MOBILE does not explicitly include any means of assessing the potential benefits of many types of program enhancements, such as on-road testing, or on-line data acquisition from test-and-repair shops. The problems associated with using MOBILE for determining I/M credit, or for predicting I/M program performance, will be addressed in a future analysis by the I/M Review Committee.

USEPA'S CASE FOR ITS 50-PERCENT DISCOUNT

The USEPA has issued two reports that it says provide the data and analyses that support its 50-percent discount (USEPA, 1991a; USEPA, 1993a). The data in these reports come from overt and covert audits of I/M programs conducted by both the USEPA and state agencies. We will evaluate whether the data and analyses presented in these two reports provide justification for discounting decentralized I/M programs. We will also evaluate the USEPA's audits of individual I/M programs. Our evaluation will determine whether the audits meet scientific standards for data collection and research design that would justify using these audits to make quantitative I/M program comparisons.

USEPA's Statements on the Size of the Discount

The following is a chronology of the USEPA's major policy statements regarding discounting of decentralized I/M programs:

³ The overall *percentage* emission reduction that an I/M program must achieve is usually quoted as the difference between the average gpm emissions from a theoretical fleet that has never been subject to I/M and the average gpm emissions required by the performance standard. In practice, the USEPA identifies the actual reductions necessary from a given fleet as the difference between its current average emissions, as predicted from MOBILE5a, and the USEPA's performance standard level, also as determined by MOBILE5a.

⁴ The 50-percent discount is an input to MOBILE5a. Appendix A shows how the 50-percent discount is hardwired into the model.

⁵ MOBILE5a is the latest in the MOBILE model series. Previous versions include MOBILE3, MOBILE4, and MOBILE4.1. The series of models will be referred to as "MOBILE."

In January, 1991 the USEPA published, "I/M Network Type: Effects On Emission Reductions, Cost, and Convenience" (USEPA, 1991a). This is the first of the two documents that the USEPA says present its data and analyses on the relative effectiveness of centralized and decentralized I/M programs. In this document, the USEPA states:

"The magnitude of the differences [between centralized and decentralized I/M programs] is difficult to accurately quantify, but evidence indicates that decentralized programs may be 20-40% less effective than centralized" .

On November 5, 1992, the USEPA promulgated its I/M rule (USEPA, 1992a). The rule states:

"credits for test-and-repair networks...are assumed to be 50% less than those for a test-only network..."

In November, 1993, the USEPA published "Quantitative Assessments of Test-Only and Test-and-Repair I/M Programs" (USEPA, 1993a). This is the second of the two documents that the USEPA says present its data and analyses on the relative effectiveness of centralized and decentralized I/M programs. In this document, the USEPA states:

"The results [of the USEPA's I/M program audits] along with the tampering survey data, form the basis for EPA's 50% effectiveness discount for test-and-repair programs."

Despite its apparent ambiguity over the size of the discount, the USEPA has affirmed that the discount is 50-percent in a number of recent public statements (see for example, Nichols, 1994, and Tierney, 1995).

What Sources of Data Does the USEPA Cite to Support its 50-Percent Discount?

The USEPA cites two groups of studies for the data and analyses underlying its 50-percent discount:

- USEPA and state covert and overt audits of I/M programs.
- USEPA surveys of tampering rates of vehicles in different I/M programs.

We will describe these sources of data and evaluate what they can or cannot tell us about the relative effectiveness of different I/M programs in reducing on-road emissions of vehicles.

What I/M Program Elements Does USEPA Consider to be the Most Important Determinants of I/M Effectiveness?

The USEPA states the following regarding the factors it considered in generating its 50-percent discount (USEPA, 1993a):

"For the enhanced I/M rulemaking, EPA used data from over 10,000 covert audits to assess the effectiveness of I/M programs. These results, along with the tampering survey data, form the basis for EPA's 50% effectiveness discount for test-and-repair programs."⁶

"The data EPA has used in making decisions about I/M programs comes from several sources, including national tampering surveys, EPA and state audits, and special studies like

⁶ When the USEPA states that it performed "over 10,000 covert audits," it means that it has sent undercover cars to over 10,000 individual covert inspections. Of the 10,413 individual covert audits of I/M stations summarized in USEPA (1993a), 97% were performed by state agencies and 3% were performed by USEPA; 57% were performed by the state of California alone. Another 27% were performed in by the states of Colorado, Michigan and Utah.

the one conducted by the California I/M Review Committee and EPA's study of the Portland, Oregon I/M program. These studies gathered quantitative data on the testing of well over 10,000 vehicles in programs across the country."

"...EPA found in audits of I/M programs, that emission testing was done objectively in test-only I/M programs...On the other hand, the data shows that inspectors in test-and-repair programs routinely attempted to get failing cars to pass the initial test...These data led EPA to reduce the emission test credits by 50% in MOBILE5a for test-and-repair programs."

While USEPA discusses other factors that affect I/M performance, as shown above, it stresses that it believes improper testing, as indicated by covert audits and tampering surveys, is the most important factor that reduces the effectiveness of I/M programs. In a recent presentation to the California I/M Review Committee (Committee),⁷ USEPA staff continued to focus on improper testing as the basis for the 50-percent discount.

Below we evaluate the USEPA's case for its 50-percent discount. We focus, not only on the data sources the USEPA used in its analyses, but also on ambient and on-road studies of vehicle emissions that were not used by the USEPA.

Do the USEPA's Audits Measure I/M Effectiveness?

The USEPA has audited a number of aspects of I/M programs. In Table 1, we summarize the audit data the USEPA obtained for centralized and decentralized programs, as well as the audit data specific to California's I/M program. Each audit measure in Table 1 will be discussed in more detail below. In Appendix D, we summarize the USEPA's data for each I/M program audited.

The USEPA began I/M program audits in 1984 and summarizes its audit procedures as follows (USEPA, 1991a):

- "1) overt visits to test stations to check for measurement instrument accuracy, to observe testing, to assess quality control and quality assurance procedures instituted by the program, and to review records kept in the station;*
- 2) covert visits to test stations to obtain objective data on inspector performance;*
- 3) a review of records kept by the program, including the history of station and inspector performance..., enforcement actions taken against stations and inspectors found to be violating regulations, and similar documents;*
- 4) analysis of program operating statistics, including enforcement rates, failure rates, waiver rates, and similar information; and,*
- 5) entrance and exit interviews with I/M program officials and a written report describing the audit findings and EPA recommendations on correcting any problems found."*

⁷ Presentation by Mr. Gene Tierney and Phil Lorang of USEPA at the February 1, 1995 I/M Review Committee Meeting.

TABLE 1

USEPA Audit Measures of Centralized, Decentralized, and California I/M Programs

Audit Measure		Audit Results ^a			Number of Programs Audited	
		Centralized	California	Decentralized	Cent.	Decent.
% of failures passed after preconditioning		30	37	33	1	2
% of gas analyzers failing calibration test	CO	6	14	23	11	6
	HC	15	9	23		
Ratio of actual emission test failure rate to "expected" failure rate		0.85	0.96	comp. ^b - 0.75 manual - 0.35	12	6 9
Waiver rates in % and cost limits for 1989 ^c	pre-81	11 / \$66	29 / \$50	8 / \$76	11	18
	post-80	10 / \$112	9 / \$175-\$300	6 / \$117		
Trends in percent of cars that fail the emissions test from 1985 to 1988		falling	not included	comp. ^b - falling add comp. - big rise manual - rising	10	2 2 4
% of USEPA covert audits finding improper tests		20	not included	81	3	12
% of state covert audits finding improper tests		none included	19	48	0	13
% of cars switched between initial test and retest ⁸		1	not included	2.4	1	2

Notes: a. A total of 16 centralized and 20 decentralized programs were evaluated on one or more of the audit measures.
 b. comp. = computerized analyzers; add comp. = recently switched from manual to computerized analyzers; manual = manual analyzers.
 c. Waiver rates are given as percent of cars initially failed that received waivers.

Source: Data above are summarized from USEPA, 1991a and USEPA, 1993a.

⁸ Refers to the percent of cases in which an inspector was found to have input a different model year than the actual model year of the vehicle in, possibly in order to improperly raise the failure cutpoints for the vehicle.

Below we will evaluate both the design and conduct of the audit studies themselves, and the USEPA's analysis and interpretation of the audit data. In our evaluation, we will ask the following questions:

- What, if any, are the limitations of the audits in determining the emission reduction effectiveness of an I/M program, and in comparing I/M programs?
- Were there any selection biases in the choice of stations to audit, the choice of undercover cars, etc.?
- Were the sample sizes statistically valid?
- Were standardized methodologies followed among the different audits?
- Does the USEPA define the methodology and criteria by which it takes the raw audit data, processes and analyzes the data, and arrives at its 50-percent discount?

The Committee requested that the USEPA provide it with the data underlying its 50-percent discount. In mid-January, the USEPA provided the Committee with copies of the audit reports from 17 of its I/M program audits.⁹ We focus our analysis on these reports, as well as on the USEPA's analyses of the audit data (USEPA, 1991a; USEPA, 1993a).¹⁰

USEPA Audit Data Include Errors in Favor of Centralized I/M Programs

The USEPA analyzed data that are inconsistent with its original audit reports.

- **USEPA reports a zero percent improper test rate for Maryland even though its audit found several improper testing problems:** The USEPA reports (USEPA, 1993a) that Maryland's centralized I/M program has an improper test rate of *zero percent* on both the visual and emissions tests. However, the actual results of the USEPA's audit of Maryland's I/M program (MD1991a; MD1991b; MD1991c) do show improper testing. For example, Maryland's centralized test stations:
 - Incorrectly passed two of ten covert audit cars with missing catalysts.
 - Did not verify the license and vehicle identification numbers on four of ten cars.
 - Incorrectly passed one of seven cars that were pretested for HC emissions (an error of omission) and improperly failed another car (an error of commission).

The USEPA normally notes a test as improper if one or more portions of the test are performed improperly. Based on the above results, the improper emissions test rate is 28 percent, while the improper visual test rate is 20 percent. When the vehicle identification check is included, 40 percent of the covert vehicles were

⁹ The Committee already had copies of USEPA (1991a) and USEPA (1993a), the two documents in which the USEPA summarizes and analyzes its audit data. The USEPA also provided the Committee with reports on its special study in Portland, Oregon, in 1979, and data from its tampering surveys.

¹⁰ As necessary, we will cite the USEPA's individual audit reports for each state I/M program by the name of the state and the year of publication of the audit. Thus, MD1991 would refer to an audit of Maryland's I/M program published in 1991.

improperly tested for one or more of the emissions test, catalyst check, and vehicle identification information.

In its audit of Maryland's I/M program, the USEPA states (USEPA, 1991d):

"EPA auditors observed many instances where the catalyst could not be seen in the floor mirror. However, in these instances, the inspector did not use the hand held mirrors. They just passed the vehicle...Many of the inspectors are not well trained in identifying the catalytic converter."

In addition, the USEPA's audit data show potential quality control problems in the emission readings at the test lane when compared to the confirmatory pre-test.

- The average CO emissions for eight covert vehicles at the test lanes were 21 percent lower than on the pre-audit emission test.
 - The average HC emissions for seven covert vehicles at the test lanes were 60 percent higher than on the pre-audit emission test.
- **USEPA reports a 50-percent improper test rate for New Jersey, even though its audit found a 66-percent improper test rate:** The USEPA reports a 50-percent improper visual test rate for New Jersey's centralized I/M program (USEPA, 1993a). However, in its audit report on New Jersey (NJ1991), the USEPA states, *"Four of six stations did not reject the undercover vehicles for missing the catalytic converter."* This indicates an improper visual test rate of 66 percent.
- **USEPA incorrectly reports that improper emission testing does not occur in centralized I/M programs:** The USEPA states (USEPA, 1993a), *"audits of 13 test-only programs found no instances of improper emission testing on either initial tests or retests."* (emphasis in original). Only 3 of the 13 audits were covert.¹¹ Of these, the Arizona and New Jersey covert audits provide data for only the catalyst-tampering check (AZ1991, NJ1991). The Maryland covert audit *did* find improper emission testing (MD1991a, MD1991b, MD1991c). Although the USEPA claims that improper emission testing does not occur on retests, none of the three covert audits included repairs and retests of failing vehicles.

In its audit of Maryland's centralized I/M program, the USEPA states (MD1991c):

"The EPA auditors observed many inconsistencies in the way tests were conducted. The SCI [Systems Control, Inc.] attendants were using the last two digits of the VIN [Vehicle Identification Number] on the inspection notice instead of getting the VIN of the vehicle. This is a problem because it does not ensure that the vehicle that was required to be tested is actually the one that was brought to the station."

"Many inconsistencies were also observed during the preconditioning phase..."

¹¹ The USEPA presents data for 3 covert audits of centralized I/M programs in USEPA, 1993a. The audit reports provided to us by the USEPA include only these three covert audits. Thus, we assume these are the only three covert audits.

- **USEPA reports a higher rate of improper testing than it actually found in New York's decentralized program:** The USEPA states (USEPA, 1993a), *"Audits of 49 test-and-repair I/M stations in Missouri and New York...found improper emission testing 34% and 46% of the time on the initial test"* (emphasis in original).¹²

In Figure 2 of the same document, the USEPA reports New York's improper test rate as 42 percent.

However, the USEPA's audit of New York's I/M program found an improper emission testing rate of 37 percent, not 46 or 42 percent (NY1990).¹³

In summary, the data-set that the USEPA used in its analyses of I/M effectiveness does not consistently reflect the actual results of its own audits.

Audits Include Structural Biases Against Decentralized I/M Programs

There are a number of ways in which the USEPA's audit studies included structural biases against decentralized I/M programs.

- **Centralized I/M programs received fewer covert audits than decentralized I/M programs:** The USEPA states, *"Covert audits have not typically been done in centralized programs"* (USEPA, 1991a), and *"EPA has done covert auditing less frequently in test-only programs, mainly because overt audits have never indicated any improper testing problems with emission testing"* (USEPA, 1993a). Overall, the USEPA has covertly audited only 3 centralized I/M programs. On the other hand, the USEPA has performed mainly covert audits of decentralized stations. By performing mainly covert audits on decentralized programs, and mainly overt audits on centralized programs, the USEPA has biased its results against decentralized programs. This is because covert audits might detect more test errors and improprieties than overt audits.
- **USEPA's "covert" audit of Arizona's centralized I/M program may not actually have been covert:** The USEPA audited Arizona's centralized I/M program by taking one car to nine different test stations (AZ1991). One station improperly passed the vehicle even though the catalyst was *"removed and replaced with a section of rusty straight pipe"* (AZ1991). Eight stations properly failed the vehicle. The USEPA concluded that the improper testing rate in Arizona's I/M program is 11 percent.

However, after the first failing test, the "covert" vehicle would have been listed on the Arizona I/M computer system as a "retest." The computer would also list the components that failed. Thus, every inspector after the first one that failed the vehicle would know that the car had previously failed for a missing catalyst.¹⁴

¹²The USEPA actually conducted covert audits of 57 I/M stations, not 49 (38 in Missouri, 19 in New York).

¹³The 34% number for Missouri is the same as the number listed in the Missouri audit report.

¹⁴Telephone conversation with Frone Mahafeey, February 6, 1995. Ms. Mahafeey is director of operations at Gordon-Darby, which administered Arizona's I/M program at the time of the USEPA's covert audit.

Since inspectors knew they were inspecting a failing vehicle, the USEPA gave Arizona's I/M program an easier test than it gave to decentralized I/M programs.¹⁵

- **USEPA may have targeted some decentralized stations for covert audits based on clues that they were performing poorly:** Regarding its covert audits, the USEPA states (USEPA, 1991a),

"The stations visited are randomly selected..."

However, in its Missouri audit report, the USEPA states (MO1993),

"Review of specific station files that had suspiciously low failure rates showed no recent covert audits. EPA targeted these inspection stations for records review through data analysis prior to the time of the audit."

It appears from this statement that the USEPA attempted, in their audits of Missouri's I/M program, to target stations they thought would be most likely to perform poorly.

We looked at several other audit reports for decentralized I/M programs (NH1991, NY1990, MA1989, GA1989, PA1989, KY1989). These other studies do not document how stations were selected for covert auditing. Thus, it is not possible to determine whether selection of stations for covert audits was typically random, or typically biased.

- **Selection of covert cars may have biased audit results:** In the Missouri audit, the covert cars were 4 and 5 years old (MO1993). The 1990 USEPA tampering survey (USEPA, 1993b) shows that the catalyst tampering rate for 4 and 5 year old vehicles is *zero percent*. Whether honest or corrupt, mechanics very likely might be aware that newer cars virtually always have untampered catalysts, and might therefore not bother checking. Furthermore, since catalyst tampering is rarely, if ever, found in newer on-road vehicles, there might not be any significant emission reduction impact from inspecting newer cars for catalyst tampering. If the USEPA had employed older cars in its covert audits, mechanics might have been more diligent in checking for missing catalysts.
- **Some decentralized audits were more difficult to pass than centralized audits:** The USEPA makes the following statement regarding one of the covert cars in its audit of Georgia's decentralized I/M program (GA1988), *"The [1984 Dodge Ram] van had been set up with subtle deception in mind. The catalyst had been removed and replaced with a very small pre-converter."* In contrast, the USEPA simply removed the catalyst and replaced it with a section of pipe in covert audits of centralized I/M programs. Referring to its Arizona audit, for example, the USEPA states (AZ1991), *"The vehicle...had the catalytic converter removed and replaced with a section of rusty straight pipe."* Referring to its Maryland audit, the USEPA states (MD1991c), *"the vehicle had been intentionally tampered by removing the catalyst and installing a straight pipe in*

¹⁵ We might also surmise that the one improper test occurred on the first inspection, since all inspectors who tested the vehicle after its first failure would have known that the vehicle was being *retested*. However, the audit report does not provide information on this issue.

its place." By using "*subtle deception*" in the Georgia audit, and obvious tampering in the centralized audits, the USEPA created a more difficult test for Georgia's decentralized I/M program than it did for the centralized programs.

The Georgia audit data indicate that the USEPA's audit method might have biased the results. The USEPA covertly audited five stations with the 1984 van described above, and all five stations improperly passed the vehicle. Another five covert audits were conducted with a 1983 Ford Escort with the air pump belt removed. In this case, *all five stations properly failed the vehicle for tampering* (GA1988). Based on these data, the USEPA reports an improper test rate for Georgia of 50 percent (USEPA, 1993a). However, the fact that opposite results were achieved with two different undercover cars is an indication of possible test bias. Had the USEPA's undercover 1984 van been tampered in a way comparable to the tampering used in audits of centralized I/M programs, the Georgia I/M stations might have properly failed the vehicle.

- **Decentralized programs may have more ways to fail a covert audit:** In its audit of northern Kentucky's decentralized anti-tampering program, the USEPA found an improper test rate of 71 percent (KY1989). However, the covert audit counted a test as improper if the mechanic failed to check any one of nine different vehicle emission control components. In contrast, covert audits of centralized I/M programs in Arizona and Maryland required only the catalyst (Arizona audit) or the catalyst and fuel inlet (Maryland audit), to be properly checked. In Kentucky, the improper test rate on the catalyst alone was only 14 percent. Kentucky's decentralized program therefore performed better than both Maryland's and New Jersey's centralized I/M programs on the catalyst check.¹⁶

Many decentralized I/M programs include an anti-tampering inspection that covers several components of the emission control system.¹⁷ On the other hand, centralized I/M programs either don't include an anti-tampering inspection, or have an inspection of only the catalytic converter and the fuel inlet. Oregon and Arizona are the only centralized I/M programs that inspect other components as well (USEPA, 1991a). As a result, decentralized I/M stations are more likely than centralized I/M stations to be cited for improper testing on a covert audit, *even though decentralized I/M programs may properly inspect more emission control components overall than a centralized I/M program.*

For example, in 14 covert audits of Kentucky's decentralized I/M program, technicians properly inspected a total of 11 exhaust gas recirculation (EGR) systems, 9 gas caps, 13 fuel inlet restrictors, 12 catalytic converters, 10 positive crankcase ventilation (PCV) systems, five evaporative canisters, 11 air injection

¹⁶ As shown above, improper catalyst inspection rates in centralized programs were AZ - 11%, MD - 20%, NJ - 66%. Recall also that the Arizona audit contained biases that would tend to minimize the measured rate of improper testing.

¹⁷ Decentralized anti-tampering inspections generally include the catalytic converter, fuel inlet, air pump, positive crankcase ventilation (PCV) system, and evaporative canister (USEPA, 1991a, Table 3-4). Some decentralized programs also include other components, such as the exhaust gas recirculation (EGR) system, gas cap, and oxygen sensor.

systems, and 12 oxygen sensors, and properly performed 13 lead tests. Therefore, inspectors in Kentucky's I/M program properly inspected an average of 6.9 key emission control components per inspection (out of 9 possible). In other words, 77 percent of the components that were supposed to be inspected were inspected properly. In nine covert audits of Arizona's centralized I/M program, technicians properly inspected eight catalysts. In 10 covert audits of Maryland's centralized I/M program, technicians properly inspected eight catalysts and 10 fuel inlet restrictors.¹⁸ Thus, in the two centralized I/M programs, inspectors properly inspected 1.4 emission control components per inspection. Thus, although the overall improper test rate for Kentucky was 71 percent, the Kentucky I/M program performed proper inspections on 4.9 times more key emission control components than in the two centralized I/M programs, on a per car basis. Table 2 summarizes these results.

Table 2
Higher Improper Test Rate on USEPA Audits Does Not Necessarily Mean Less Thorough Inspections

I/M Program	Overall Improper Visual/Functional Inspection Rate in USEPA Covert Audits	Improper Catalyst Inspection Rate in USEPA Covert Audits	Number of Emission Control Components Properly Inspected Per Car
Kentucky (Decent.)	71%	14.3%	6.9
Arizona + Maryland (Cent.)	15.5%	15.5%	1.4

Source: USEPA I/M Program Audits - AZ1991, MD1991a, MD1991b, MD1991c, KY1989.

We would like to perform the same type of analysis for other decentralized I/M program audits. Unfortunately, the audit reports often do not provide specific information on what counted as an improper test. However, some of the audit reports indicate that the auditor was watching for improprieties on more than just an emission test and catalyst check. For example:

- In its 1990 audit of New York's decentralized I/M program (NY1990), the USEPA found that five tampering checks were performed improperly out of six covert audits. The New York I/M program requires inspection of 7

¹⁸In its covert audit of Maryland's I/M program, the USEPA watched for proper inspection of the catalyst and fuel inlet restrictor only. In its audit of Arizona's I/M program, the USEPA watched for proper inspection of the catalyst only.

different emission control components. The audit report does not include information on which component checks were incorrectly performed. It is possible, for example, that the catalytic converter check alone might have been performed properly on more than one of the cars.

- In its 1989 audit of Massachusetts' decentralized I/M program, the USEPA states (MA1989), "*A total of 18 covert investigations were conducted during the course of the audit. Eleven of the 18 stations failed to conduct proper tampering inspections. All of the stations committed at least one inspection infraction.*" Based on this, the USEPA reports an improper test rate of 100 percent in the Massachusetts I/M program. The USEPA's audit report does not include any other information on the nature of the test infractions.

The above analysis shows that reports of improper test rates alone are misleading. As we just showed, it is possible for a program that inspects more components, but has higher rates of improper testing, to find more emission control defects than a program that inspects fewer components, but has a lower rate of improper testing.

- **I/M program managers knew that USEPA was conducting audits:**

Undercover cars for covert audits were supplied by the state agency that administers the I/M program. Thus, the agencies knew that the USEPA was auditing their programs, and could potentially have forewarned test stations. However, the potential for such an occurrence might be greater in a centralized program. This might be the case because, in a centralized I/M program, the state's program management officials have a direct day-to-day working relationship with a single contractor. Because the contractor is a direct extension of the government, program officials might also have an incentive to have the contractor perform well.

In contrast, in decentralized programs, hundreds of independent small businesses are licensed by the state and do not have a direct relationship with I/M program managers. Thus, forewarning of decentralized stations might require that I/M program officials contact a multitude of stations. Forewarning a centralized contractor might require only a casual mention of a coming audit during what may be frequent meetings of the contractor and the regulators.

In summary, the USEPA's audits were idiosyncratic, and appear to have been biased against decentralized I/M programs. Biases in the conduct of the audits make the quantitative results of the audits unreliable. Idiosyncrasies in the methodologies between audits of different I/M programs make the data unsuitable for inter-program comparisons. Finally, spotty documentation of the audit methods and audit results makes it difficult to assess systematically the audit data.

What do USEPA's audits measure?

We have seen that the USEPA's methodologies for acquiring and comparing their audit data are not scientifically and statistically sound. However, we must also ask whether the audits, if properly conducted, would indeed measure the emission reduction effectiveness of I/M programs. The USEPA states (USEPA, 1991a),

"The most critical aspect in evaluating an I/M program is the emission reduction benefit it achieves."

However, none of the USEPA's audits measure the emission reductions achieved by I/M programs. The USEPA's audits measure the following I/M program elements:

- Rates of improper testing.
- Effect of preconditioning on emission test failure rates.
- Percent of gas analyzers failing a calibration test.
- Ratio of actual to expected failure rates.
- Waiver rates.
- Trends in percent of cars failing the emission test.
- Percent of cars switched between initial test and retest.

The USEPA did not assess the relationship between scores on these audits and the emission reduction performance of the I/M programs that it studied.

The USEPA assumes, but does not demonstrate, that proper testing results in effective I/M: The USEPA assumes that the combination of a low improper test rate and a low waiver rate indicates that cars are receiving repairs, and therefore that the I/M program is achieving emission reductions. For example, the USEPA states:

"...EPA found in audits of I/M programs, that emission testing was done objectively in test-only I/M programs...On the other hand, the data shows that inspectors in test-and-repair programs routinely attempted to get failing cars to pass the initial test...These data led EPA to reduce the emission test credits by 50% in MOBILE5a for test-and-repair programs." (USEPA, 1993a)

"A waiver represents lost emission reductions to the I/M program, so high waiver rates mean substantial numbers of vehicles that are high emitters are not getting adequate repair. Conversely, a truly low waiver rate indicates that maximum emission reduction benefits are being obtained. However, improper testing is, in a sense, a form of unofficial waiver that may be an alternative to the legitimate waiver system employed by programs. A low reported waiver rate by itself is therefore ambiguous with respect to program success." (USEPA, 1991a)¹⁹

However, the USEPA's audits did not generate appropriate data for assessing an I/M program's effect on emissions. For example:

- **Cars in the USEPA's covert audits might not have been emission tested prior to covert audits:** A number of the USEPA's audits do not give any indication that the covert cars were emission tested prior to the audit (for example, NY1990, AZ1991, NJ1991, MA1989). The audits therefore do not include data that could be used to assess the impact of inspections on emissions.

¹⁹Note that this statement is self-contradictory, asserting that a low waiver rate does, and may not, mean that an I/M program is effective.

- **When emissions were tested, the results were not analyzed:** The Maryland audit was the only one we saw that included emissions data for the undercover cars. However, the USEPA did not include the emissions test data in its report of the audit results (MD1991a). The emissions data for the 10 covert cars appear only in 10 individual inspection reports attached to memos that accompany the audit report (MD1991b, MD1991c).

The Missouri audit report indicates that the covert cars were pre-tested, but does not provide data on emission test results either from the pre-test or the audits (MO1993). In both audits, the USEPA does not present any quantitative analysis of I/M program effectiveness based on the emissions results.

- **Failing cars were not taken through a repair and retest cycle:** In the USEPA's audits, failing cars were not taken through the repair and retest cycle. The USEPA, therefore did not collect data on repair effectiveness.
- **USEPA used marginal emitters or clean cars in its covert audits:** The ten covert cars in the audit of Maryland's centralized I/M program had emissions near the emission standards:
 - The average of the CO emissions of the ten cars was 19 percent below the failure cutpoint of 1.2 percent CO. Nine out of the 10 cars passed the CO emissions test.
 - The average of the HC emissions of the ten cars was 66 percent above the failure cutpoint of 220ppm HC. Five out of the 10 cars passed the HC emissions test.

The cars used in the audit were either marginal emitters or clean cars. However, most of the potential benefit of I/M comes from repairing high emitters (Lawson, 1995a; Lawson, 1995b). The cars that USEPA used in its audit are not the type of cars from which large emission reductions can be achieved. Thus, even if the USEPA collected the data necessary to assess emission reduction impacts, without looking at a representative sample of cars, it would not be possible to assess actual emission impacts.

Use of marginal emitters could also bias the audit results in another way. Inspectors might be more (or less) likely to perform proper tests on cars with higher emissions. If this is the case, then the USEPA's audit results would not reflect the behavior of inspectors on the cars with the greatest potential for emission reductions.

As we showed above, the USEPA believes that ensuring proper testing results in effective I/M. However, the USEPA does not consider the possibility that motorists and/or mechanics might avoid meaningful repairs of high emitters, even if tests are performed properly. If there are other ways to avoid meaningful repairs, then a high rate of proper testing would not be a sufficient condition for an I/M program to be effective. For example:

- **High emitters might obtain certificates of compliance even if tests are performed properly:** Even if an I/M program's improper-test rate and waiver

rate are low, high emitting vehicles might still be able to obtain certificates of compliance. Mechanics can, in many cases, adjust vehicles to pass an I/M test without making the repairs that would be necessary to ensure that the car will remain clean after the inspection. Such adjustments are straightforward for idle tests. Recent evidence indicates they are possible for loaded-mode tests as well (Beebe, 1994; Schwartz, 1994; I/M Review Committee, 1995). Mechanics might also readjust cars after they pass an I/M test. This type of I/M avoidance is possible in both test-and-repair and test-only programs.

Studies show that the behavior described above already occurs. For example, in Arizona's centralized I/M program, a program held up by the USEPA as an exemplary centralized I/M program, the Arizona Auditor General found the following (Arizona Auditor General, 1988):

"Ninety-three percent of all mechanics indicate they have been asked by customers to simply adjust their vehicle to pass the emission test, rather than conduct the appropriate and needed emissions-related maintenance and repairs."

"Ninety-four percent of all mechanics also indicate that they have been asked by customers to re-adjust their vehicles after it has passed the emissions test so that it will run better."

"It is clear from this data that a significant segment of the driving public attempts to circumvent the emissions testing program."

- **Motorists might avoid vehicle registration:** In the absence of other means of avoiding I/M, some motorists may stop registering their cars, or may register them in nearby non-I/M counties.

Thus, even if the USEPA's assertion that testing is performed properly in centralized I/M programs is correct, converting to centralized inspections may simply shift repair avoidance into other areas.

Because the USEPA believes the rate of improper testing is an important measure of I/M effectiveness, we have so far focused on audits of improper testing. We will now look at some of the other audit measures that the USEPA discusses in both its audit reports and in the documents it uses to support its 50-percent discount.

Percent of failures passed after preconditioning:²⁰ When vehicles are not warm, I/M tests can improperly fail them. The USEPA concludes that preconditioning of cars can be properly performed in both centralized and decentralized networks, and with both idle and loaded tests (USEPA, 1991a). The USEPA does not, therefore, appear to consider preconditioning to be a factor in the relative effectiveness of centralized or decentralized I/M programs.

However, the USEPA does conclude that preconditioning cars on a dynamometer is more cost effective than preconditioning them at idle (USEPA, 1991a). The USEPA does not provide any cost analysis to support this conclusion. In addition, the USEPA also

²⁰See the glossary for an explanation of preconditioning.

provides no data, for example, on the capital, training, time or other costs that would determine the relative cost-effectiveness of alternative preconditioning methods.

Percent of gas analyzers failing the calibration test: The USEPA used this audit to estimate the percentage of hydrocarbon (HC) and carbon monoxide (CO) analyzers that were miscalibrated in a given I/M program.

The USEPA found that the average rate of analyzers failing the calibration test was 15 and 23 percent for HC analyzers and six and 23 percent for CO analyzers, for centralized and decentralized programs, respectively. The USEPA also found that California's decentralized I/M program performed better than seven of 11 centralized I/M programs on the HC analyzer audit, and better than three of 11 centralized I/M programs on the CO analyzer audit (nine percent for HC and 14 percent for CO). However, there appear to have been a number of defects in the USEPA's presentation and analysis of these data:

- **Weighting of Audit Data is Biased Against Decentralized Programs:** The USEPA calculates a *weighted average* for the performance of centralized and decentralized programs. It weights each average by the number of analyzers audited in each program (USEPA, 1991a, Table 3-5). Thus, programs in which more analyzers were tested have a larger impact on the average. Because the USEPA tested more analyzers in centralized I/M programs that performed better on the audit, the USEPA's averaging process creates a bias in favor of centralized I/M programs.²¹

If we are interested in how centralized and decentralized programs perform as a group, then we should weight each decentralized and each centralized program equally. Even if we did want to weight one program more than another, we would not want to do it arbitrarily by the number of analyzers we happened to test in each program, but by some process that reflects which programs are most representative.

If we remove the inappropriate weighting that the USEPA employed, we find a marked change in the audit results. For decentralized programs, the average percent of analyzers failing the HC calibration test increases from 23 percent to 24 percent when the weighting is removed. But for centralized programs, the average failure rate increases from 15 percent to 23 percent when the weighting is removed. In other words, the average centralized program is actually similar to the average decentralized program on the HC analyzer calibration audit. On the CO analyzer audit, removal of the weighting brings centralized and decentralized programs closer together, but centralized programs still performed better.

- **USEPA does not analyze degree of calibration error:** The USEPA states (USEPA, 1991a), "*...small errors do not critically affect emission reductions, any significant error affects quality and public confidence in the testing process.*" Nevertheless, the USEPA's analysis of the audit data (USEPA, 1991a) does not include information on the degree of miscalibration. We went back to the audit

²¹ The gas analyzer audit data are presented in Appendix D. As can be seen from the audit data, the USEPA tested many more analyzers in centralized programs that had low calibration failure rates (Connecticut and Illinois).

reports for states that the USEPA includes in its gas analyzer analysis (MD1991a, NH1991, GA1989, PA1989, CT1985), and found only one case in which the degree of miscalibration was reported. In this case, only the *average* level of miscalibration was reported, not the individual data for each analyzer. That case was Georgia, in which the analyzer readings were low by an average of 3.7 percent. The Georgia audit did not measure what effect such miscalibration would have on improper passing of vehicles with excess emissions.

- **USEPA fails to show a link between analyzer calibration and the ability of I/M programs to reduce emissions:** The USEPA's analysis of the gas analyzer audit data does not include any analysis of how analyzer calibration affects emission reductions (USEPA, 1991a). Without knowing both the degree of miscalibration, and the relationship of a given level of miscalibration to false passes of high emitters, it is not possible to determine such a relationship.

It is conceivable that even large calibration errors might not affect the test results of the highest emitters. For example, in the El Monte pilot study, we have found that 75 percent of the HC emissions in excess of the failure cutpoint come from only 17 percent of the vehicles that fail the HC emissions test (Lawson, 1995b).²² These vehicles have emissions that are several times higher than typical failure cutpoints in I/M programs. The USEPA uses a calibration tolerance of +5%/-7%²³ in its analyzer calibration audits. However, an analyzer with readings that are low by even a factor of two is unlikely to miss the highest emitters because their emissions are more than a factor of two above the failure cutpoints. Thus, even large calibration errors might not affect the pass/fail results of the cars that contribute most of the potential emission reductions in an I/M program.

In sum, it is impossible to use the audit data on miscalibration for a quantitative analysis of I/M program effectiveness, because the USEPA does not report or analyze the degree or direction of miscalibration in different I/M programs. Furthermore, the USEPA did not collect data that could show a quantitative relationship between degree of calibration inaccuracy and the percent of total excess emissions that are missed due to false passes of high emitters.

Early emission test failure rates: The USEPA reports the ratio of the actual failure rate to the "expected" failure rate with data it collected from 1983 to 1985. The USEPA derived expected failure rates based on data from Louisville's I/M program for 1988. The USEPA discusses these data at length, but concludes (USEPA, 1991a), "*As with the comparison of decentralized and centralized programs, the failure rates in [Table 3-6, USEPA, 1991a] should not be taken to be reliable indications of which centralized programs were working the best.*" This audit does not, therefore, appear to have played a

²² The El Monte Pilot Study is one of two pilot projects undertaken as a result of a Memorandum of Agreement between the USEPA and the state of California. These pilot projects were designed to gather data on prospective I/M tests, and on the uses of remote sensing.

²³ That is, an analyzer fails the audit if it reads a known concentration of gas more than 5 percent too high, or more than 7 percent too low.

role in the USEPA's assessment of the relative merits of centralized and decentralized I/M programs.

Waiver rates: The USEPA reports waiver rates and repair cost limits for 1989. A waiver indicates that a car that failed an inspection was given a certificate even though it was not completely repaired. Thus, a waiver is a sign that a car's emissions might not have been reduced as much as possible. However, the waiver rate alone is not sufficient to determine the impact of waivers on I/M program effectiveness. One also must know the amount of excess emissions from the waived vehicles that would otherwise have been reduced if a waiver was not available.²⁴ For instance, one I/M program might tend to waive *higher emitting* cars on average than another I/M program, yet both programs might waive the *same percentage* of cars. Without information on the emissions before and after repair of waived vehicles, it is not possible to determine how a given waiver rate affects the emission-reduction performance of an I/M program.

Failure Rate Trends: The USEPA found that failure rates decreased over time in both centralized programs, and decentralized programs with computerized analyzers. The USEPA states (USEPA, 1991a), "*This is expected as more new technology vehicles, which fail less often, enter the fleet and as the program effectively repairs existing vehicles...*" However, the meaning of a trend in failure rates is actually ambiguous for the following reason:

- **Decreasing failure rates may indicate that motorists and mechanics have learned other ways to avoid compliance:** The USEPA provides no evidence that the reduction in failure rates seen in some I/M programs is indeed due to repair effectiveness and/or introduction of new cars to the fleet with more robust emission control systems. A competing hypothesis is that motorists and mechanics learn how to adjust some cars to pass the test without making meaningful repairs. On the other hand, falling failure rates might mean that marginal cars are being repaired, while high emitters are not. There is no way to know without more data. Even if we accept the USEPA's explanation for dropping failure rates, decentralized I/M programs with computerized analyzers performed as well as centralized programs on this audit measure (USEPA, 1991a, Figure 3-2).

²⁴ See the glossary for a definition of excess emissions.

USEPA's Audits Look at Only a Subset of the Relevant Variables

Looking at the audit studies as a whole,

- **USEPA assumes its audit measures reflect on-road emission reductions:** The USEPA asserts, but does not demonstrate, a relationship of its audit measures to the effectiveness of I/M programs in reducing on-road emissions. The USEPA does not present any data or analysis that shows a relationship between its audit measurements and on-road emission-reduction performance of an I/M program. As we will see below, on-road and ambient data, as well as the USEPA's tampering data, are not consistent with the USEPA's interpretation of its audit results.
- **USEPA considers only a subset of the variables that affect I/M effectiveness:** The USEPA overlooks feedbacks that occur when different portions of a program are changed. For instance, if improper testing is curtailed, then motorists might be more likely to seek waivers. If program officials clamp down on waivers, motorists might look for mechanics who can superficially adjust cars to pass the test without fixing defects that may cause high emissions between tests. Motorists may also be more liable to register their cars in non-I/M counties, or not register their cars at all. A whole range of behavioral responses may occur that are not accounted for in the USEPA's analysis. I/M is a system of interacting variables that includes the motivations of motorists and mechanics. The USEPA's audit measures look at only a subset of the factors that affect the operation of an I/M system. These measures are not capable of providing any information on the size and direction of possible feedback effects when I/M program features are altered. The USEPA's audits are, at best, an incomplete description of the variables relevant to I/M effectiveness.
- **USEPA does not present a methodology for converting the audit data into a discount:** The USEPA presents no mathematical or statistical analysis that starts with its audit data and ends with a 50-percent discount. The USEPA presents numerical data on the audit variables, and then makes a number of statements about what variables it believes are important in determining I/M effectiveness. But the path from there to the 50-percent discount (and its numerical precision), or any other posited discount, is left uncharted. Indeed, it is difficult to understand how the USEPA could have arrived at any particular number for the discount by using the audit data.

Based on the foregoing analysis, we conclude that the USEPA's audit measures are not suitable for assessing the relative emission reduction effectiveness of I/M programs. Furthermore, the audit measures are not appropriate for assessing the likely effects of changing various I/M program features.²⁵

²⁵ The USEPA states in its policy documents that its I/M program audits form the foundation of its 50-percent discount (see quotes from USEPA above). However, when the Committee presented some of the above critiques of the USEPA's audits at the February 1, 1995 I/M Review Committee meeting, Mr. Gene Tierney of the USEPA responded with the following statement: *"The audit program was not designed as a way to strictly evaluate how well a program was getting the emission reductions that the MOBILE model was predicting. That was not even close to the intent."*

California's I/M Program Performs Well on Many USEPA Audits

The USEPA's audits do not measure I/M program performance. Nevertheless, California scored well on the USEPA's audit measures when compared to both centralized and decentralized I/M programs. Analyzing the data on the USEPA's terms, we find:

- **Analyzer Calibration (USEPA, 1991a):** California's program scored higher than seven of 11 centralized programs on HC analyzer calibration, and scored higher than three of 11 centralized programs on CO analyzer calibration.
- **Actual vs. "Expected" Failure Rate (USEPA, 1991a):** California's program (ratio of actual to expected failure rate equal to 0.96) performed better than the average centralized program (0.85).
- **Waiver Rates (USEPA, 1991a):** The waiver rate for post-1980 vehicles (the ones we're most concerned with for the future) in California's program was the same as, or lower than, the waiver rates in eight of 11 centralized programs studied.
- **Improper Test Rates (USEPA, 1993a):** California's program (19% improper test rate) performed similar to the average centralized (20% improper test rate) and better than the average decentralized (improper test rates of 49% on state audits, and 76% on EPA audits) I/M program.

Thus, if we accept the USEPA's audit analysis, the audits do not support a discount for California's I/M program relative to centralized programs.

Results of USEPA's Tampering Surveys of I/M Programs

The USEPA's tampering surveys and various on-road and ambient emission studies provide us with a more direct measure of I/M performance. We review these studies below. The USEPA states (USEPA, 1993),

"For the enhanced I/M rulemaking, EPA used data from over 10,000 covert audits to assess the effectiveness of I/M programs. These results, along with the tampering survey data, form the basis for EPA's 50% effectiveness discount for test-and-repair programs."

Between 1985 and 1990, the USEPA and several states performed tampering inspections on about 44,000 vehicles around the country as one means of evaluating I/M effectiveness.²⁶ The USEPA has presented charts and graphs based on these data that indicate that regions with centralized I/M programs have much lower tampering rates than regions with decentralized programs (Tierney, 1993b; USEPA, 1991a; USEPA, 1993a).

Walsh et al. (1994), however, analyzed the USEPA data and found that tampering rates did not vary significantly between program types. They found that:

²⁶USEPA also performed tampering surveys in 1991 and 1992, but has since ended the survey program. The results of the 1991 and 1992 tampering surveys have not yet been analyzed.

- Regions with centralized I/M programs have slightly lower tampering rates than regions with decentralized I/M programs.
- Regions with I/M programs have slightly lower tampering rates than regions without I/M programs.
- The most important factors determining tampering rates were age and mileage of the fleet, regardless of I/M program type.

Figure 1 summarizes these results.²⁷

Walsh et al. reached different conclusions from the USEPA for the following reasons: The USEPA tends to show only "good" centralized programs and "bad" decentralized programs in its charts and graphs of tampering rates. In addition, the USEPA does not adjust its data for age and mileage differences between the automobile fleets of different regions. When all USEPA tampering data are stratified according to age and mileage, differences in tampering rates are accounted for mainly by differences in vehicle age and mileage, and not by differences in I/M program network type.

The slightly lower tampering rates found by Walsh et al., for centralized programs might not mean that tampering rates were actually lower in centralized programs. This is because the USEPA examined cars in centralized programs for tampering *as they arrived at the centralized test lane for their scheduled inspection*, but always examined cars in decentralized programs by pulling them over on the road (USEPA, 1990b). Motorists are probably less likely to submit tampered vehicles for testing at test lanes than they would be to drive tampered vehicles on the road. The USEPA's tampering surveys are thus likely to be biased toward showing lower tampering rates for regions with centralized I/M programs than might actually be the case.

The USEPA cites its observation of low tampering rates in Oregon (about 6 percent) as partial evidence that centralized programs are more effective than decentralized programs. However, evidence indicates that motorists refused to participate in the USEPA's surveys more often in some states. For example, in the 1987 tampering survey, the refusal rate in Portland, Oregon was nine percent, while the average refusal rate nationwide was 4 percent (USEPA, 1988). Studies have shown that motorists who refuse a voluntary on-road inspection drive higher emitting cars on average than those who submit to the on-road inspection. In one remote-sensing study, for example, average CO and HC emissions of "refuseniks" were 2.4 and 2.8 times higher than the emissions from cars that were inspected (ARB, 1994).

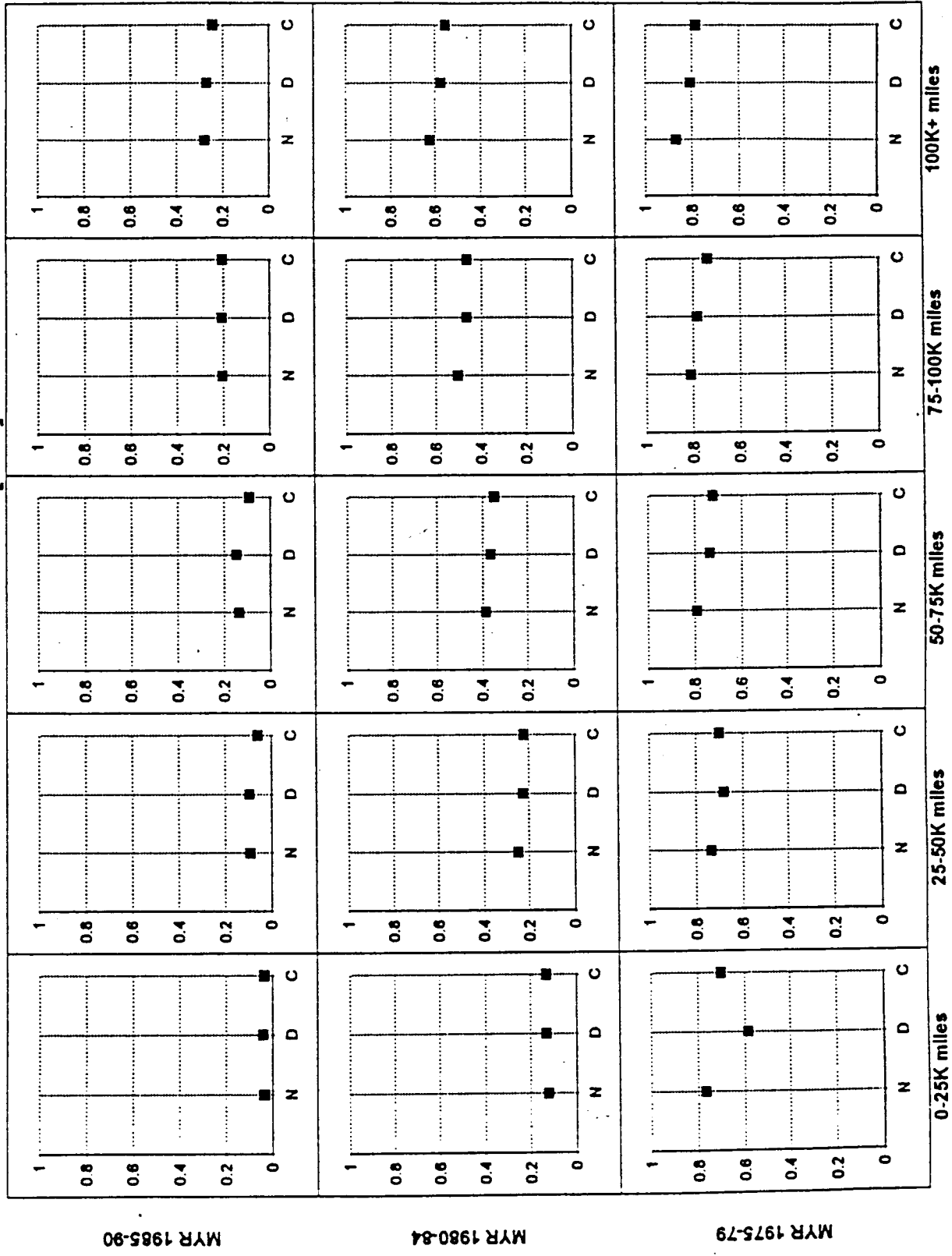
Tampering surveys may be a reasonable surrogate for on-road emissions. The USEPA included an idle emissions test along with its tampering inspections and found that tampered vehicles have idle emissions that are 3 to 6 times higher than the idle test emissions from untampered vehicles (USEPA, 1993b).

The USEPA's data fail to support its claim that its tampering surveys show centralized programs to be superior to decentralized programs. The USEPA's data indicate that

²⁷ USEPA has asked Doug Lawson to look at tampering rates of individual components, rather than overall tampering rates. This analysis will be presented in the near future.

Figure 1

TAMPERING RATES [EPA]



Caption for Figure 1

Figure 1 summarizes data from national tampering surveys. On each of the 15 graphs in the figure, the vertical axis gives the tampering rate measured. The horizontal axis gives the I/M program type (N = no program, D = decentralized, C = centralized). The outer vertical axis divides rows by model year range. The outer horizontal axis divides columns by odometer mileage range. Thus, looking across the top row gives tampering rates for 1985-1990 model year cars as mileage increases. Looking down the middle column gives tampering rates for 50,000 to 75,000 mile cars over the three model year ranges. Note that the differences in tampering rates with model year or odometer mileage for any type of I/M program are much greater than the differences between types of I/M programs.

overall tampering rates differ by only small amounts between centralized and decentralized programs. As a result, the tampering data do not support the USEPA's 50-percent discount for decentralized programs relative to centralized programs.

We have now completed our evaluation of two of the USEPA's main sources of support for its 50-percent discount. We concluded that the USEPA's audit data are statistically invalid and are not suitable for assessing the effect of I/M on emissions. Furthermore, we have concluded that the USEPA's tampering data do not support a discount for decentralized I/M programs. Below, we evaluate a third source of data on I/M effectiveness.

ASSESSING I/M EFFECTIVENESS USING ON-ROAD AND AMBIENT VEHICLE EMISSIONS DATA

As we noted earlier, the USEPA states (USEPA, 1991a):

"The most critical aspect in evaluating an I/M program is the emission reduction benefit it achieves."

While the USEPA believes that emission reduction benefits are the ultimate test of an I/M program, the USEPA did not measure the emission reduction benefits of existing I/M programs. However, other researchers have performed such measurements, and we review their results below.

Centralized Programs

Tucson

Zhang et al. (1994) measured CO and HC emissions from 14,051 vehicles at a Tucson, Arizona intersection with remote-sensing. Of the vehicles measured, 304 were registered to zip codes not included in the Arizona centralized I/M program. Both before and after adjusting emissions for vehicle age, the data indicate that the non-I/M fleet was cleaner than the I/M fleet (see Table 3). The difference, however, was not statistically significant.²⁸ In other words, the data are consistent with a hypothesis that there is no difference in emissions between the I/M and non-I/M fleets.^{29,30} This suggests that the Arizona I/M program may be having little or no impact on vehicle emissions.

²⁸ See the glossary for an explanation of the meaning of statistical significance.

²⁹ The I/M and non-I/M fleets were corrected for differences in model year distribution. However, it should be noted that there might be other differences between the two fleets in terms of demographics, driving conditions, etc., that are not accounted for, but could have affected the results of the analysis.

³⁰ At the February 1, 1995 I/M Review Committee Meeting, Mr. Phil Lorang disputed the validity of using on-road remote sensing measurements to compare emissions from different groups of vehicles. However, data comparing RSD readings with IM240 and FTP readings indicate that fleet average RSD results are a good indicator of fleet average dynamometer emissions. We analyze this issue in more detail in Appendix C.

TABLE 3
Comparison of Emissions of I/M and Non-I/M Fleets in Arizona

Fleet	Mean %CO	Mean %HC	Average Model Year	Age-Adjusted Mean %CO	Age-Adjusted Mean %HC
I/M	1.06	0.077	1984.7	0.99	0.088
non-I/M	0.81	0.075	1986.7	0.89	0.070

Source: Zhang et al. 1994

In a 1988 study, the Arizona Auditor General reached a similar conclusion (Arizona Auditor General, 1988). The Auditor General's report states, "The time series analysis, covering the fourteen-year period between 1974 through 1987, found that VEIP [Vehicle Emissions Inspection Program] did not have an overall effect on ambient carbon monoxide (CO) levels in Phoenix and Tucson...Using MOBILE3, EPA has consistently credited VEIP with a 25 percent reduction in CO and hydrocarbons."

Minneapolis/St. Paul

The Minneapolis/St. Paul metropolitan area implemented an annual, centralized I/M program in July of 1991. The Minnesota Pollution Control Agency has monitored ambient CO levels since 1986 at three metropolitan sites. Scherrer and Kittelson (1994) analyzed ambient CO data from these three monitoring sites from five years before, through two-and-a-half years after the I/M program was implemented. They found that the monitoring data show an overall 1.3 ± 1.4 percent³¹ reduction in CO due to the I/M program. These data are also consistent with a hypothesis that the Minneapolis/St. Paul centralized I/M program had little or no effect on vehicle emissions. During legislative hearings before the I/M program began, the USEPA and the Minnesota Pollution Control Agency asserted that the program would reduce vehicle CO emissions by 30 percent.³²

Chicago

Stedman et al. (1991) used remote sensing devices to measure HC and CO emissions in Chicago in 1990. Stedman et al. compared the remote sensing readings for 8,971 cars to the emissions from those same cars as measured at their most recent I/M test. They found the following:

- Measured on-road emissions showed no correlation with time since the last I/M test. This means that cars that had not been tested for a long time, and cars that had been tested recently, had similar on-road emission rates. If Chicago's

³¹ The stated range gives the 95% confidence interval. In other words, there is a 95% probability that the actual emission reductions lie within the stated range.

³² USEPA has disputed the findings of Scherrer and Kittelson. Appendix B reproduces Scherrer and Kittelson's original paper, USEPA's critique, and Scherrer and Kittelson's response to USEPA's critique.

centralized I/M program had been effective, cars that were about to be inspected would have had higher emissions than cars that had recently been inspected.

- Because most cars have low emissions, low-emitting cars probably dominated the correlation of emissions with time since the last I/M test. Stedman et al., therefore analyzed the correlation of the 401 cars that had failed to pass the I/M test at least twice before passing their most recent test. Stedman et al. reasoned that this sub-fleet would be composed of higher emitting cars (which turned out to be the case) that one would expect to show the greatest benefit from an I/M program. For this sub-fleet the correlation of emissions vs. time since last I/M was again virtually zero ($r^2 = 0.001$ for CO and 0.005 for HC).³³

Because on-road emissions were found to be unrelated to time since the last I/M test, Stedman et al. concluded that Chicago's centralized I/M program is identifying high emitting vehicles, but not repairing them. Thus, Chicago's I/M program is having little impact on on-road emissions.

Decentralized Programs

California

California performed annual random roadside pullover studies in regions subject to I/M in 1989 and 1991, and in regions never subject to I/M in 1990. In these studies, the Bureau of Automotive Repair and Air Resources Board performed a standard Smog Check (if the vehicle owner was willing) on cars pulled over cars at random by the California Highway Patrol.³⁴ Lawson (1992, 1993) analyzed the roadside survey data for these three years and found that:

- In 1989 and 1991, on-road failure rates for cars that had been inspected within the last few months were about the same as on-road failure rates for cars that were due for an inspection within the next few months. If Smog Check were effective, the cars about to be inspected should have failed at a higher rate than the cars that had been recently inspected.
- There was no correlation of idle emissions at the random roadside test with time since the last inspection.
- The failure rate for cars not subject to I/M was lower than the failure rate for cars subject to I/M, even though the non-I/M cars were older on average.

Lawson et al. (1990) measured emissions in the Lynwood area of Los Angeles with remote sensors. Measured on-road emissions showed no correlation with time since the last I/M test.

These results suggest that California's I/M program has not significantly reduced on-road vehicle emissions.

³³ See the glossary for an explanation of the meaning of " r^2 ".

³⁴ Smog Check is the name of California's I/M program.

Denver

Two remote sensing studies have compared I/M and non-I/M fleets in Colorado (Ostop and Ryder, 1989; Zhang et al., 1993; Radian, 1992). Ostop and Ryder used remote sensors to measure the emissions from 1,200 I/M vehicles and 3,400 non-I/M vehicles at a site near the boundary between a county subject to I/M and a county not subject to I/M. The I/M cars were three percent lower emitting on average, but the difference was not statistically significant.³⁵

Zhang et al. and Radian used remote sensors to measure the emissions from 11,170 vehicles in central Denver. Of these vehicles, 2,043 were not registered in areas subject to an I/M program. There was no statistically significant difference between the HC emissions of the I/M and non-I/M vehicles. However, the non-I/M vehicles had CO emissions 13 percent higher than the I/M vehicles, and the difference was statistically significant.

Ambient Ozone and CO Data from 1983 and 1992

Manhard (1994) compared reductions in ambient ozone and CO in regions with centralized I/M and regions with decentralized I/M from 1983 to 1992, using published USEPA data (USEPA, 1993c). The results are shown in Table 4.

TABLE 4

Average Reduction in Ambient Ozone and Ambient CO between 1983 and 1992 for Non-Attainment Regions Aggregated by Type of I/M Program

Type of I/M Program	Ambient Ozone Reductions	Ambient CO Reductions
Centralized	-24%	-36%
Decentralized	-23%	-35%

Source: USEPA, 1993c; Manhard, 1994

Ozone is created partially as a result of local mobile source emissions. Carbon monoxide results almost totally from local mobile source emissions. As Table 4 shows, over the period from 1983 through 1992, reductions in ambient ozone and CO were about the same for regions that had a centralized I/M program and regions that had a decentralized I/M program.

Of course, this could be considered a fairly crude measure of I/M effectiveness. It is also possible that there are systematic differences between regions with centralized and decentralized programs in terms of socio-economics, climate, culture, and other factors (although no such systematic differences are obvious). However, most of the CO

³⁵ See the glossary for an explanation of the meaning of statistical significance.

inventory comes from cars and light trucks. If centralized I/M programs were really achieving 20 to 30 percent reductions in vehicle CO emissions (over and above the effect of fleet turnover), as predicted by the USEPA's models, and if decentralized programs were really half as effective as centralized programs, then surely there would be an observable difference in air quality improvement based on I/M type. That ambient air quality data do not reveal such differences is strong evidence that I/M network type has not been an important factor in determining emission reductions.

In summary, studies of on-road and ambient vehicle emissions reveal that:

- In both centralized and decentralized programs, tailpipe emissions measurements show little or no difference between I/M and non-I/M fleets.
- In both centralized and decentralized programs, both ambient and tailpipe data show little or no emissions reductions attributable to I/M.
- Nationwide average reductions in ambient CO and ozone levels over the last decade appear to be independent of the type of I/M program in a given region.

Thus, measurements of the most important variable that I/M is supposed to affect - on-road emissions of vehicles - indicate that there is little or no difference between centralized and decentralized I/M programs, and that neither is currently producing a significant benefit.

CONCLUSIONS

The USEPA's audits are not suitable for evaluating I/M programs because they were based on statistically invalid research designs not intended for use in I/M program comparisons. Furthermore, even if the audits had been conducted without research biases, the audits do not measure the emission reduction effectiveness of I/M programs. Finally, if we accept the USEPA's interpretation of its audit results, California's decentralized I/M program scored as well or better than many centralized programs.

Direct measurements of the variables that I/M is supposed to reduce - emission system tampering, and tailpipe emission levels - show little or no difference between centralized and decentralized I/M programs, and also little or no effect due to I/M. The data do not indicate that decentralized programs are working well. But, neither do they indicate that centralized programs are working any better. We conclude that whether an I/M program is centralized or decentralized has not been an important factor in determining historical I/M program effectiveness.

We conclude that there is no empirical or scientific basis for a discount for decentralized I/M programs relative to centralized programs.

WHERE DO WE GO FROM HERE?

The USEPA states that *"the most critical aspect in evaluating an I/M program is the emission reduction benefit it achieves"* (USEPA, 1991a). We wholeheartedly agree with this statement. It underlies an approach that is essential to determining whether or not

any program is achieving real benefits - namely, measuring pollution levels in ambient air and emissions from vehicles as they are driven on the road. Future measurements of I/M effectiveness should be based on this approach.

Proponents of the USEPA's model program may assert that, even if our analyses are valid, their model program will overcome all of the failings of current I/M programs by implementing transient loaded-mode testing and separation of test-and-repair. Given that motorists have so far thwarted our best efforts at ensuring that high emitting vehicles are found and repaired, such a contention is not based on empirical observation or scientific analysis.

While I/M effectiveness has fallen short of expectations, repair of gross emitting vehicles has been shown to generate significant emission reductions (Lawson, 1995; El Monte Pilot Study, 1995; Sunoco, 1994; Stedman, 1993; Lawson, 1993). The Committee's charge is to help point California down the road to effective I/M. California's new I/M laws have created many tools for finding gross polluters, sanctioning non-complying motorists, and removing dishonest or incompetent mechanics from the test-and-repair industry. We hope that the USEPA and California will work jointly in developing an I/M program that will capture the potential emission reduction benefits available from effective I/M.

REFERENCES

- ARB, 1994, "On-Road Remote Sensing of CO and HC Emissions in California," Contract No. A032-093, Report prepared by D. H. Stedman et al., University of Denver, February, 1994.
- Arizona Auditor General, 1988, "Performance Audit: Department of Environmental Quality, Vehicle Emissions Inspection and Maintenance Program," Report to the Arizona Legislature, #88-11, December, 1988.
- Aroesty, J., 1994, "Restructuring Smog Check: A Policy Synthesis," RAND Corp. report # DRR-885-CSTC, October, 1994.
- AZ1991, "Arizona Covert Audit Results," Memo from Mr. Gene Tierney to File, September 4, 1991.
- Beebe, J., 1994, Paper presented at the Mobile Sources/Clean Air Conference, Estes Park Colorado, NCVECS, October, 1994.
- CA I/M Review Committee, 1993, "Evaluation of the California Smog Check Program and Recommendations for Program Improvements - Fourth Report to the Legislature," February 16, 1993.
- California I/M Review Committee, 1995, Interview of El Monte Pilot Study repair mechanics, January 4, 1995.
- CT1985, "Draft Audit Report - Connecticut I/M Program," USEPA Region I, 1985.
- El Monte Pilot Study, 1995, Preliminary analyses of the El Monte Pilot Study data, January, 1995.
- Fujita, E., et al., 1992, "Comparison of Emission Inventory and Ambient Concentration Ratios of CO, NMOG and NO_x in California's South Coast Air Basin," *J. Air Waste Manage. Assoc.*, 42, 264 - 276, 1992.
- GA1989, "1988 Georgia Audit Findings," June, 1989
- Glover, E. and W. Clemmens, "Identifying Excess Emitters with a Remote Sensing Device: A Preliminary Analysis," SAE paper 911672, August 5, 1991.
- Harley, R., et al., 1992, "Respeciation of Organic Gas Emissions and the Detection of Excess Unburned Gasoline in the Atmosphere," *Environ. Sci. Technol.*, 26, 2395-2408, 1992.
- Harrington, W., and V. McConnell, 1994, "Modeling In-Use Vehicle Emissions and the Effects of Inspection and Maintenance Programs," *J. Air Waste Manage. Assoc.*, 44, 791-799, 1994.
- Ingalls, M. N., 1989, "On-Road Vehicle Emission Factors from Measurements in a Los Angeles Area Tunnel," Paper No. 89-137.3, presented at the A&WMA 82nd annual meeting, Anaheim, CA, 1989.

- Kirchstetter, T., and R. Harley, 1994, "Measurements of California light duty vehicle emissions in the Caldecott Tunnel," paper presented at the 4th Annual CRC/APRAC On-Road Emissions Workshop, San Diego, California, March 16-18, 1994.
- KY1989, "1989 Northern Kentucky Anti-Tampering Program Audit Report," USEPA Region IV, 1989.
- Lawson, D. R., 1990, "Emissions from In-Use Motor Vehicles in Los Angeles: A Pilot Study of Remote Sensing and the Inspection and Maintenance Program," *J. Air Waste Manage. Assoc.*, 40, 1096-1104, 1990.
- Lawson, D. R., 1992, Letter to Richard Sommerville, Chair of the California I/M Review Committee, October 16, 1992.
- Lawson, D. R., 1993, " 'Passing the Test' - Human Behavior and California's Smog Check Program," *J. Air Waste Manage. Assoc.*, 43, 1567, 1993.
- Lawson, D. R., 1995a, "The Costs of "M" in I/M -- Reflections on Inspection and Maintenance Programs," *J. Air Waste Manage. Assoc.*, in press.
- Lawson, D. R., 1995b, presentation at the February 1, 1995 I/M Review Committee Meeting.
- Lurmann, F., and H. Main, 1992, "Spatial and Temporal Variations in Nonmethane Organic Compounds and NO_x," paper presented at the SCAQS Data Analysis Conference, Los Angeles, California, July 21-23, 1992.
- MA1989, "Massachusetts I/M Program 1989 Audit Report," USEPA.
- Manhard, R., 1994, Untitled analysis of USEPA emission trend data. Can be obtained from Mr. Manhard at (303)773-2079.
- McConnell, V., and W. Harrington, 1992, "Cost-Effectiveness of Enhanced Motor Vehicle Inspection and Maintenance Programs," Discussion Paper QE-22, Resources for the Future, Washington, DC, August, 1992.
- MD1991a, "Maryland Covert Audit Results," Memo from Mr. David Sosnowski to Mr. Gene Tierney and Mr. Kelly Bunker, July 5, 1991.
- MD1991b, "Maryland Covert Audit Results," Memo from Mr. Kelly Bunker to Mr. David Sosnowski, July 26, 1991.
- MD1991c, "State of Maryland, Motor Vehicle Emissions Inspection and Maintenance, Draft Audit Report," USEPA Region III, January, 1991.
- MO1993, "1993 Audit Report of the St. Louis, Missouri Inspection/Maintenance Program," USEPA Region VII, 1993.
- NADA, 1993, "Statement of Issues," Case #93-1004 filed in the US Court of Appeals, District of Columbia Circuit, February 15, 1993.
- National Research Council, 1991, "Rethinking the Ozone Problem in Urban and Regional Air Pollution," NAS Committee on Tropospheric Ozone Formation and Measurement, National Academy Press, 1991.
- NH1991, "Inspection/Maintenance Program Audit Report, New Hampshire," June, 1991.

- Nichols, M., 1994, Letter from Mary Nichols, Assistant USEPA Administrator for Air and Radiation, to Cal-EPA Secretary James Strock, December 30, 1994.
- NJ1991, "National Air Audit System: 1990 Follow-Up Audit of the New Jersey Inspection and Maintenance Program," January 3, 1991.
- NY1990, "National Air Audit System, New York Inspection and Maintenance Program, Final Report," February, 1990.
- OR1979a, "Portland Study Interim Analysis: Observations on Six Months of Vehicle Operation," USEPA # IMS 003/PS-1, January 1979.
- OR1979b, "Portland Study Interim Analysis: Update with Nine Months of Vehicle Operation," USEPA, May 1979.
- OR1985, "1985 National Air Audit, Portland, Oregon, Inspection and Maintenance Audit," June, 1985.
- Ostop, R. L. and L. T. Ryder, 1989, "Ute Pass Carbon Monoxide Emissions Study," City of Colorado Springs, Department of Utilities, Environmental Services Division, March, 1989.
- PA1989, "Pennsylvania Motor Vehicle Emissions Inspection and Maintenance (I/M) Draft Audit Report," USEPA, Region III, November, 1989.
- Pierson, W. R., N. F. Robinson, and A. W. Gertler, 1992, "Review and Reconciliation of On-Road Emission Factors in the SoCAB," paper delivered at the SCAQS Data Analysis Conference, Los Angeles, California, July 21-23, 1992.
- Pierson, W. R., et al., 1990, "Comparison of the SCAQS Tunnel Study with Other On-Road Vehicle Emission Data," *J. Air Waste Manage. Assoc.*, 40, 1495-1504, 1990.
- Radian, 1992, "Colorado Automobile Inspection and Readjustment Program Performance Audit, Final Report," Prepared for the Colorado State Auditor, June, 1992.
- Scherrer, H. C., and D. B. Kittelson, 1994, "I/M Effectiveness as Directly Measured by Ambient CO Data," SAE Paper #940302. March, 1994.
- Schwartz, J., 1994, Memo from Joel Schwartz to Lynn Scarlett that includes a digest of informal interviews with the mechanics in the El Monte Study, October 12, 1994.
- Stedman, D. H., 1995, "EPA Overestimates the Benefit of Biennial I/M," submitted to *J. Air Waste Manage. Assoc.*
- Stedman, D. H., 1993, "Provo Pollution Prevention Program: A Pilot Study of the Cost-Effectiveness of an On-Road Vehicle Emissions Reduction Program," University of Denver/City of Provo, Utah, January 15, 1993.
- Stedman, D. H., G. A. Bishop, J. E. Peterson, P. L. Guenther, I. F. McVey, and S. P. Beaton, 1991, "On-Road Carbon Monoxide and Hydrocarbon Remote Sensing in the Chicago Area," ILENR/RE-AQ-91/14, Report Prepared for the Illinois Department of Energy and Natural Resources, Office of Research and Planning. October, 1991.

- Sunoco, 1994, "Sunoco Emissions Systems Repair Program," Prepared for Paul Durkin, Sun Co., Inc., Philadelphia, PA, 1994.
- Tiao, G. C., et al., "Statistical Analysis of the Effect of Car Inspection and Maintenance Programs on the Ambient CO Concentrations in Oregon," *Environ. Sci. Technol.*, 16, 328-335, 1982.
- Tierney, E. J., 1995, Handout and statements made at an "I/M Stakeholders Meeting," in Ann Arbor, January 24, 1995.
- Tierney, E. J., 1993a, Memo titled "Performance Standard Calculations," from Mr. Gene Tierney to the California Blue Ribbon Committee on Smog Check, July 2, 1993.
- Tierney, E. J., 1993b, "Presentation to the Smog Check Committee," undated.
- US Court of Appeals, 1994, District of Columbia Circuit, Case # 93-1004, decided May 6, 1994.
- USEPA, 1988, "Motor Vehicle Tampering Survey - 1988," Office of Mobile Sources, August, 1988.
- USEPA, 1990b, "Motor Vehicle Tampering Survey - 1989," Office of Mobile Sources, May, 1990.
- USEPA, 1991a, "I/M Network Type: Effects on Emission Reductions, Cost, and Convenience," USEPA-AA-TSS-I/M-89-2, by Mr. Gene Tierney, Office of Air and Radiation, January, 1991.
- USEPA, 1992a, "Inspection/Maintenance Program Requirements; Final Rule," Federal Register, 40 CFR Part 51, November 5, 1992.
- USEPA, 1993a, "Quantitative Assessments of Test-Only and Test-and-Repair I/M Programs," EPA-AA-EPSP-I/M-93-1 Office of Air and Radiation, November, 1993.
- USEPA, 1993b, "Motor Vehicle Tampering Survey - 1990," EPA 420-R-93-001 Office of Air and Radiation, February, 1993.
- USEPA, 1993c, "1992 National Air Quality and Emissions Trends Report," EPA 454/R-93-031.
- Walsh, P. A., D. R. Lawson, and P. Switzer, 1991, "An Analysis of U. S. Roadside Survey Data," Paper presented at the Fourth CRC On-Road Vehicle Emissions Workshop, March 17, 1994.
- Zhang, Y., D. H. Stedman, G. A. Bishop, P. L. Guenther, S. P. Beaton, and J. E. Peterson, 1993, "On-Road Hydrocarbon Remote Sensing in the Denver Area," *Environ. Sci. Technol.*, 27(9), 1885-1891, 1993.
- Zhang, Y., S. P. Beaton, G. A. Bishop, D. H. Stedman, 1994, "Final Report: Tucson Intersection Study of Automobile Emissions," University of Denver, Department of Chemistry, September, 1994.

APPENDIX A

HOW THE 50-PERCENT DISCOUNT IS HARDWIRED INTO MOBILE5A³⁶

The MOBILE Model

In some public statements, the USEPA has cited the results of running its MOBILE5a model to justify its 50-percent discount for decentralized I/M programs.³⁷ However, an examination of the model shows that the 50-percent discount for test-and-repair programs is an *input* to the model and not an *output* of the model. In other words, the model cuts emission-reduction benefits in half if the I/M network type is a test-and-repair network. A 50-percent discount comes out of the model because the USEPA hard-wired it into the model up front. Below, we present the FORTRAN programming code of the relevant portions of MOBILE5a to demonstrate this.

Putting the 50-Percent Discount into MOBILE5a

MOBILE begins by defining an array called DISCNT. An *array* is a variable that has multiple slots for values. In this case, the variable DISCNT has three slots that are assigned values at the beginning of the program. ("Quotes" of the FORTRAN code from MOBILE 5a are set off in boxes.)

```
DATA DISCNT/1.0, 0.50, 0.50/
```

In the language of FORTRAN, the three slots in the array have the following values:

```
DISCNT(1) = 1.0  
DISCNT(2) = 0.5  
DISCNT(3) = 0.5
```

So DISCNT(2) refers to slot number 2 in the array, and it has a value of 0.5.

Any slot in the array can also be referred to with the name of another variable. So, for instance, if we create the variable INTYP and then define INTYP = 3, we would get the following result:

```
DISCNT(INTYP) = 0.5
```

because DISCNT(3) = 0.5 and INTYP = 3.

³⁶ We assume that many people who read this will not be familiar with the FORTRAN computer programming language. The relevant portions of the FORTRAN code are explained in the text.

³⁷ For example, when questioned about various aspects of the 50-percent discount by the California Blue Ribbon Committee on Smog Check in June, 1993, Mr. Gene Tierney of USEPA responded to a number of queries by recommending that the committee run the MOBILE model.

MOBILE5a defines the variable DISCNT as a function of the variable INTYP. INTYP is given three possible integer values, depending on the I/M program network type.

C DISCNT	R	I/M program efficiency discount as a f(INTYP)
----------	---	---

C INTYP	I	I/M Program type: 1 = Test only,
C		2 = Test and Repair (Computerized),
C		3 = Test and Repair (manual)

So, referring back to the discussion of arrays above, we see, for instance, that the discount for a test-only program is 1.0, i.e., no discount, and the discount for either type of test-and-repair program is 0.5, i.e., a 50-percent discount.

PCLFTN and PCLFTO are defined as "correction" factors. As we will see below, these will be used to modify the emission factors calculated by MOBILE, based on the type of I/M program.

C PCLFTN	R	percentage correction factor for new I/M
C PCLFTO	R	percentage correction factor for old I/M

Credit is determined based on the value of INTYP. In the FORTRAN code on the next page,³⁸ NNOVLP is a variable equal to the number of I/M programs in a given region. This appears to be designed to account for the effects of previous I/M programs, or of overlapping I/M programs when a region has a hybrid program.

In the "IF" statement below you will see a calculation of either PCLFTO or PCLFTN. PCLFTO and PCLFTN start with a value of 1.0 and are then reduced based on a number of factors, including the value of DISCNT. For instance, if INTYP has a value of 1, meaning a test-only network, then DISCNT = 1.0 and there is no discount. If INTYP has a value of 2 or 3, meaning a test-and-repair network, then DISCNT has a value of 0.5, and there is a 50-percent discount.

Thus, it is clear that the 50-percent discount is a matter of telling the program to cut any I/M credit by half if the network is a test-and-repair network, and to leave the calculated credit as is if the network is a test-only network

³⁸ Boldface sections of the FORTRAN code are our emphasis.

```
C
C
C Calculate the correction factor only if there is one I/M
C program.
C
  IF(NNOVLP.EQ.1) THEN
    PCLFTO=1.0-(PCREDO*(1.0-WAIV(1))*ENFORC(CRIM(NNOVLP),1)*
* DISCNT(INTYP(NNOVLP)))
    IF(ICY.EQ.ICYIM(NNOVLP)+1.
* .AND.IFREQ(NNOVLP).EQ.2) PCLFTO=PCLFTO*0.5+0.5
    BCLEFT=PCLFTO
  ENDIF
C
```

APPENDIX B

Research Paper on the Effectiveness of Minnesota's I/M Program

This Appendix includes copies of,

- A 1994 SAE paper by Scherrer and Kittelson on the effectiveness of Minnesota's I/M program.

Reprinted with permission from SAE.

Paper No. 940302 , © 1994 Society of Automotive Engineers, Inc.

- USEPA's critique of their paper.
- Scherrer and Kittelson's response to the USEPA's critique.

I/M Effectiveness as Directly Measured by Ambient CO Data

Huel C. Scherrer and David B. Kittelson
University of Minnesota

Reprinted with permission from SAE.

Paper No. 940302 , © 1994 Society of Automotive Engineers, Inc

Reprinted from: **Global Emission Technology and Analysis**
(SP-1043)

SAE *The Engineering Society
For Advancing Mobility
Land Sea Air and Space®*
I N T E R N A T I O N A L

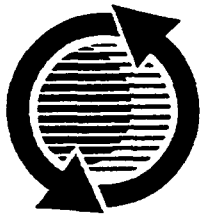
International Congress & Exposition
Detroit, Michigan
February 28-March 3, 1994

The appearance of the ISSN code at the bottom of this page indicates SAE's consent that copies of the paper may be made for personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay a \$5.00 per article copy fee through the Copyright Clearance Center, Inc. Operations Center, 222 Rosewood Drive, Danvers, MA 01923 for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale.

SAE routinely stocks printed papers for a period of three years following date of publication. Direct your orders to SAE Customer Sales and Satisfaction Department.

Quantity reprint rates can be obtained from the Customer Sales and Satisfaction Department.

To request permission to reprint a technical paper or permission to use copyrighted SAE publications in other works, contact the SAE Publications Group.



GLOBAL MOBILITY DATABASE

All SAE papers, standards, and selected books are abstracted and indexed in the Global Mobility Database.

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher.

ISSN 0148-7191

Copyright 1994 Society of Automotive Engineers, Inc.

Positions and opinions advanced in this paper are those of the author(s) and not necessarily those of SAE. The author is solely responsible for the content of the paper. A process is available by which discussions will be printed with the paper if it is published in SAE transactions. For permission to publish this paper in full or in part, contact the SAE Publications Group.

Persons wishing to submit papers to be considered for presentation or publication through SAE should send the manuscript or a 300 word abstract of a proposed manuscript to: Secretary, Engineering Activity Board, SAE.

Printed in USA

90-1203C/PG

I/M Effectiveness as Directly Measured by Ambient CO Data

Huel C. Scherrer and David B. Kittelson
University of Minnesota

ABSTRACT

The effectiveness of a centralized Inspection/Maintenance (I/M) program, implemented in a major U.S. metropolitan area in 1991, is directly measured through ambient CO monitor data. A multi-factoral analysis is developed which quantifies effects due to the interaction of hourly traffic levels with wind vector and ambient temperature conditions, allowing a better measure of I/M effectiveness. The time-trend of the measured CO levels is seen to closely match those predicted by the analysis throughout the 7 year study period. An average ambient CO reduction of 1.3 ± 1.4 percent attributable to I/M is measured, with individual results of +1.5 percent, +5.8 percent, and -3.4 percent obtained for the three monitor locations studied.

INTRODUCTION

The U.S. Environmental Protection Agency has increasingly required the implementation of Inspection/Maintenance (I/M) programs in regions of the United States failing to meet National Ambient Air Quality Standards (NAAQS). As of 1991, 33 states or portions thereof required some form of I/M testing, at a (1989) cost of \$ 570 million (1). Not included in this figure are the costs of vehicle repairs, vehicle mileage, or the value of vehicle owners' time expended in complying with these inspection requirements.

As I/M programs have expanded in the U.S., the basis for projecting that these programs will benefit ambient air quality has become increasingly uncertain. In a review of carbon monoxide (CO) air quality studies through 1981, Jones and Walsh concluded there were only limited comparative or time-series data which supported the projection of I/M benefits (2). Noting the rapid evolution of closed-loop emission control systems, Haskew et. al. found a poor correlation between Federal Test Procedure emissions and I/M test procedure results for 1981-1986 vehicles (3). Results of an EPA / Manufacturer Cooperative I/M Testing Program found poor repeatability in I/M test procedure results (4,5).

However, it may be projected that any I/M selection process, regardless of selection accuracy or repeatability, will

produce ambient air quality benefits. This projection basis assumes that I/M procedures supplement, rather than replace, the periodic maintenance which owners would otherwise provide for their vehicles. Whether such projected I/M benefits are realized in practice is the question addressed by this study.

The 1991 introduction of a comprehensive I/M program in the Minneapolis/St. Paul metropolitan area affords the opportunity to directly measure I/M effectiveness through time-series analysis of ambient CO monitor data. The I/M program studied implements inspection-only, centralized testing with an annual, registration enforced procedure. Idle pass/fail cut points of 220 PPM for hydrocarbons and 1.2% for CO are specified for 1981 and newer vehicles, with less stringent requirements extending back to 1976 model year vehicles. Visual inspections for fuel inlet and catalytic converter tampering are specified. Loaded mode conditioning, followed by a second chance idle re-test, is provided for vehicles failing a first idle test.

Each of these I/M requirements is cited by the U.S. EPA to maximize I/M program effectiveness (1). Virtually all 1.4 million gasoline powered vehicles within the 7,670 square kilometer program region are covered by the I/M requirements. I/M testing began July 1991. During the first annual cycle of testing (July 1991 - June 1992), 9.4% of the vehicles tested failed inspection on their initial visit to one of the 11 inspection facilities.

DESCRIPTION OF THE DATA ANALYZED

Hourly ambient CO data were obtained for monitoring sites operated by the Minnesota Pollution Control Agency. Three sites, centrally located in the metropolitan area, meet the criteria of providing nearly continuous data over a 5 year baseline plus 2 year post-I/M implementation period. Sites 1 and 2 show a history of NAAQS violations, while site 3 is designated as a background site (Table 1).

Meteorological data for the 7 year study period were obtained from the National Weather Service (6). Observations recorded at 3 hour intervals for ambient air temperature, wind direction, and wind velocity are used. The

	NAAQS Monitor #	8 Hour Exceedances > 9.0 PPM						
		1986	1987	1988	1989	1990	1991	1992
Site 1	243300050F05	-	7	2	2	1	2	0
Site 2	243300044F01	8	3	0	2	0	0	0
Site 3	242260056F01	0	0	0	1	0	0	0

Table 1. Ambient CO monitor site identification and violation history.

weather station is centrally located within 10 kilometers of the 3 CO monitoring sites.

Traffic levels at monitor sites 1 and 2 were measured by the St. Paul Traffic Engineering Department at 2-3 year intervals. These traffic counts show levels nearly constant over the past 12 years, averaging 23,420 vehicles per day at site 1 and 22,425 vehicles per day at site 2. Traffic levels associated with site 3 comprise a mix of urban freeway and city boulevard sources in proximity to the site.

ANALYSIS

A simplified empirical relation for the observed ambient CO monitor levels may be expressed as the product of the factors

$$CO = \frac{\text{Source Term}}{\text{per Vehicle}} \times \frac{\text{Vehicle}}{\text{Rate}} \times \frac{\text{Transport}}{\text{and Dilution}} \quad (1).$$

This relation is dimensionally consistent with ambient CO concentration, with the factors expressed in units of

$$[PPM] = \left[\frac{\text{gram}}{\text{vehicle}} \right] \times \left[\frac{\text{vehicle}}{\text{hour}} \right] \times \left[\frac{\text{PPM}}{\text{gram/hour}} \right] \quad (2).$$

Equation 1 represents a one-dimensional simplification of the case where CO deposition and transport from each volume element in proximity to the monitor inlet is considered. Equation 1 may be expanded as the product of a single dimensional factor, and non-dimensional factors which are normalized distribution functions of the independent (known) variables:

$$CO_j = A \exp\{-B t_j\} \times f(\text{hour}_j) \times g(\text{weekday}_j) \times h(\text{month}_j) \times u(\text{temperature}_j) \times v(\text{wind vector}_j) \times [1 - \gamma \times IM(t_j)] \quad (3).$$

CO_j is the predicted hourly CO concentration at time t_j for a given monitor site. The discrete distribution functions $f(\text{hour}_j)$, $g(\text{weekday}_j)$, and $h(\text{month}_j)$ incorporate the periodic variation in vehicle rate. The discrete polar function $v(\text{wind vector}_j)$ is defined by dividing wind speed and direction into discrete levels. This function represents transport and dilution of CO.

The source term per vehicle is represented by 3 factors. The dimensional term $A \exp\{-B t_j\}$ is the in-use emission trend factor for a given monitor site, representing continuous model year turn-over of the vehicle fleet. The discrete

distribution function $u(\text{temperature}_j)$ is defined by dividing ambient temperature into discrete levels.

Finally, the term $[1 - \gamma \times IM(t_j)]$ represents the reduction of in-use emissions in the vehicle fleet due to I/M. $IM(t_j)$ is the proportion of the vehicle fleet completing one or more I/M inspections at time t_j . $IM(t_j)$ is zero prior to I/M implementation, and increases monotonically to 1.0 in the 1 year period following I/M implementation. The factor γ represents I/M effectiveness in reducing ambient CO monitor concentrations. Each of the discrete distribution functions satisfies the normalization relationship

$$\frac{1}{n} \times \sum_{j=1}^n f(\text{hour}_j) = 1, \dots \quad (4)$$

for the n hourly observations recorded over the study period.

Each of the factors A , B , γ , and the distribution functions $f(\cdot)$, $g(\cdot)$, $h(\cdot)$, $u(\cdot)$, $v(\cdot)$ are unique to a particular monitor site. These unknowns may be solved by obtaining the least-squares fit

$$\sigma^2 = \sum_{j=1}^n (CO_j \text{ Observed} - CO_j \text{ predicted})^2 \quad (5)$$

using standard numerical techniques. This analysis differs from a logarithmic regression analysis, which would weight

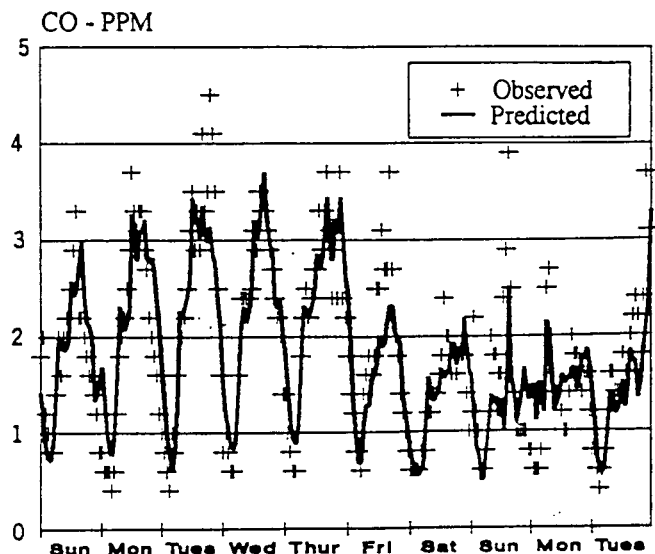


Figure 1. Site 1 hourly observed and predicted ambient CO levels during an arbitrary 10 day interval (April 22 - May 1, 1990).

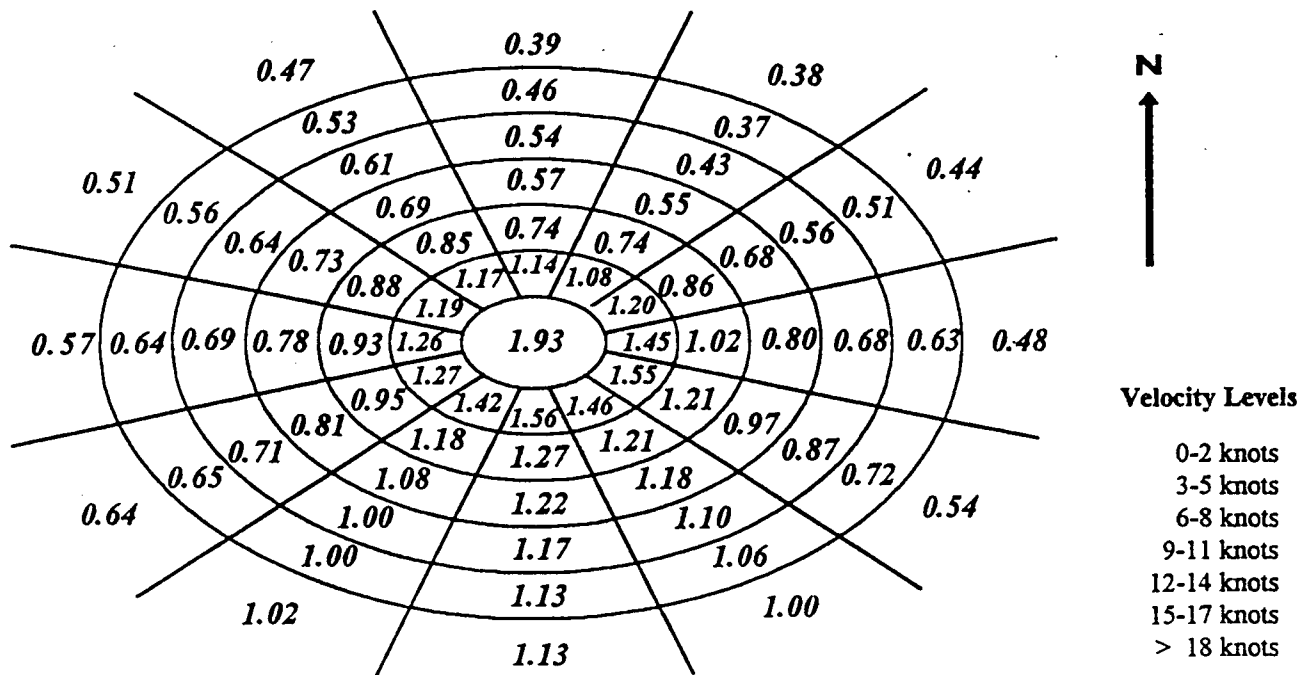


Figure 2. Polar histogram of normalized ambient CO response as a function of wind vector for monitor site 1. Central value is for wind velocities of 0-2 knots, with successive rings in 3 knot increments. CO response values for winds from the south is toward the bottom of the figure.

the difference (observed - predicted) between 20 PPM and 10 PPM the same as the difference between 0.2 PPM and 0.1 PPM. The former difference is considerable in understanding ambient CO monitor levels, while the latter is of minor importance. Thus, linear rather than logarithmic weighting is used due to the associated physical and statutory significance.

RESULTS

Although hourly ambient CO monitor data are very dynamic quantities, a reasonably good fit between observed and predicted values is obtained, as shown in Figure 1. The value of an analysis capable of obtaining such a detailed fit to the observed data is twofold. First, the analysis represents an improved understanding of the time-series data, allowing trends in the data to be interpreted with greater confidence. This is in comparison to exceedance data, which are dominated by extremes in meteorological conditions. Second, the analysis provides insight as to how physical factors and in-use vehicle emission characteristics combine to produce ambient CO concentrations as measured by NAAQS monitors.

Site 1 results for the CO response polar distribution function $v(\text{wind vector})$ are shown in Figure 2. The wind vector is divided into 73 discrete levels, with observed wind velocities of 0 to 2 knots (0 - 3.7 km/hr) comprising the first level. Seventy-two additional levels are defined by dividing wind direction into 12 increments, and wind velocity into 6 increments, as shown by the polar histogram. The maximum CO response at each wind velocity increment is seen to occur

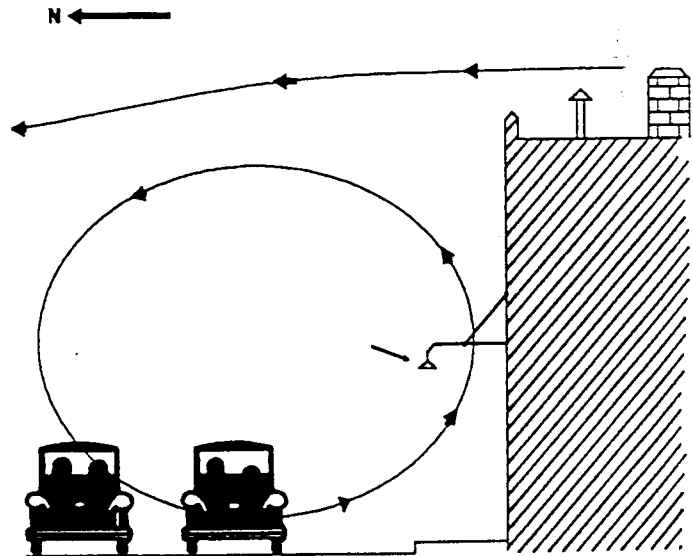


Figure 3. Monitor site 1 bluff body recirculatory flow. The southerly wind vector shown corresponds to enhanced values of the CO monitor response function in Figure 2.

with winds from the due south. A strongly peaked CO response distribution function is also obtained for site 2, with maximum response occurring for winds from the due north.

An interpretation of these results for site 1 is given in Figure 3. Site 1 is located on the south side of a major east-west boulevard. The sampling inlet height is 3.4 m., positioned in front of a building 6 m. in height. Recirculatory flow develops between the boulevard and the north wall of the

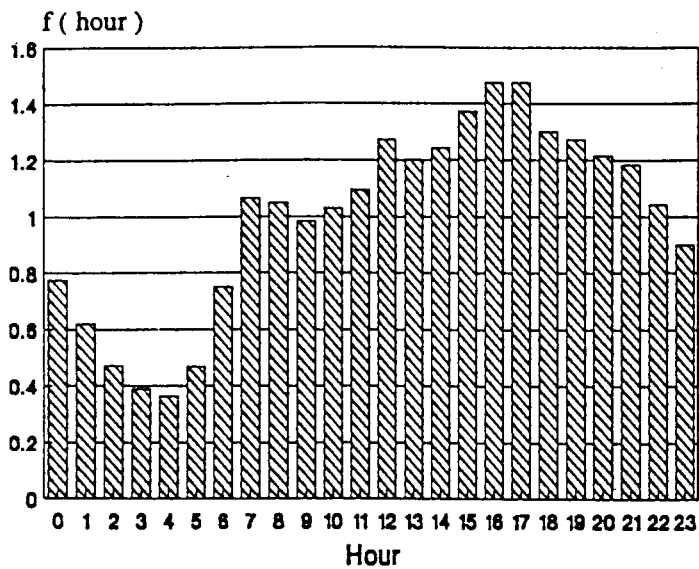


Figure 4. Normalized ambient CO response as a function of hour for monitor site 1, as defined by Equation 3.

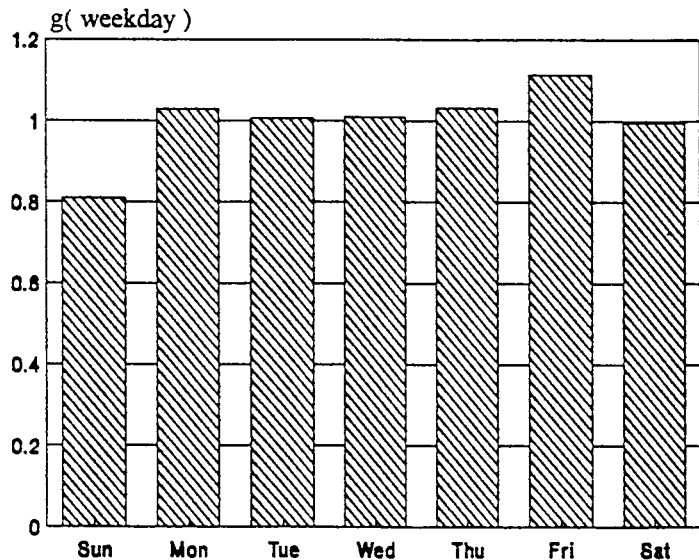


Figure 5. Normalized CO response as a function of weekday for site 1, reflecting weekly traffic patterns.

building on which the monitor inlet is located, maintained by the south wind over the building roof. A symmetric interpretation of the wind vector distribution for site 2 may be given. This monitor inlet is located on the south wall of a building fronting the north side of the same boulevard. Site 2 is located 1 km west of site 1.

Similar results for $f(\text{hour})$ and $g(\text{weekday})$ are obtained for all 3 sites. These results, shown in Figures 4 and 5, follow the patterns seen in hourly and daily traffic count data. Results for the monthly distribution functions are similar between the 3 sites, but counter to the pattern seen in monthly traffic count data. As shown in Figure 6, higher values are obtained for the winter months of October through February. Slightly lower traffic levels (93% of the yearly average) are observed during these months.

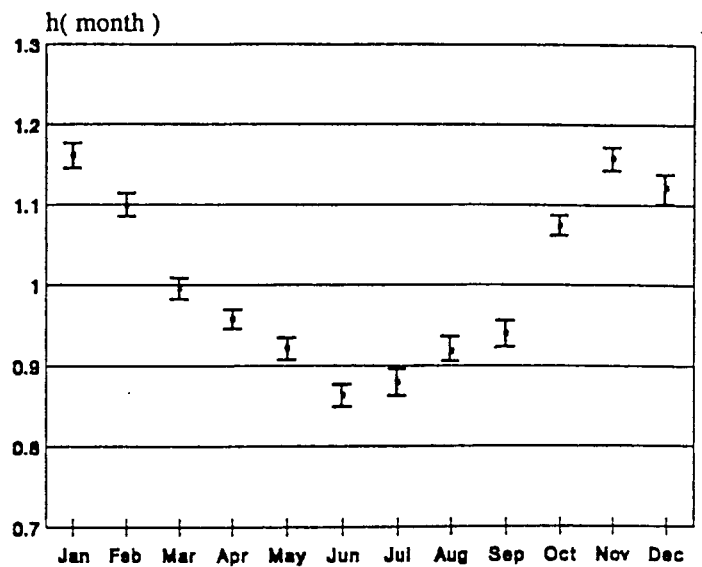


Figure 6. Normalized CO response as a function of month for site 1, with 95% confidence intervals shown.

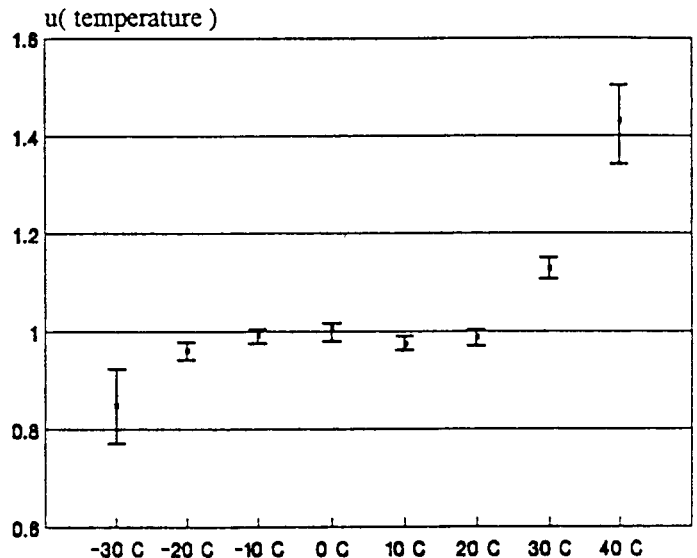


Figure 7. Normalized CO response as a function of ambient temperature for site 1, with 95% confidence intervals. An essentially flat response is shown for the majority of observed ambient temperature conditions.

The distribution function for ambient temperature is defined by dividing observed values into 8 discrete levels of 10°C increments, as shown in Figure 7. The essentially flat response in the range -25°C to $+25^{\circ}\text{C}$ is consistent with a large proportion of the vehicles in proximity to the monitoring sites being in a fully warmed operating mode. Higher response levels in the temperature ranges of 25°C - 35°C and $>35^{\circ}\text{C}$ are observed. It must be emphasized that these results are particular to the ambient CO monitor locations studied. As such, these results do not contradict the well documented relationship between vehicle start-up CO emissions and low ambient temperature, as measured by laboratory testing.

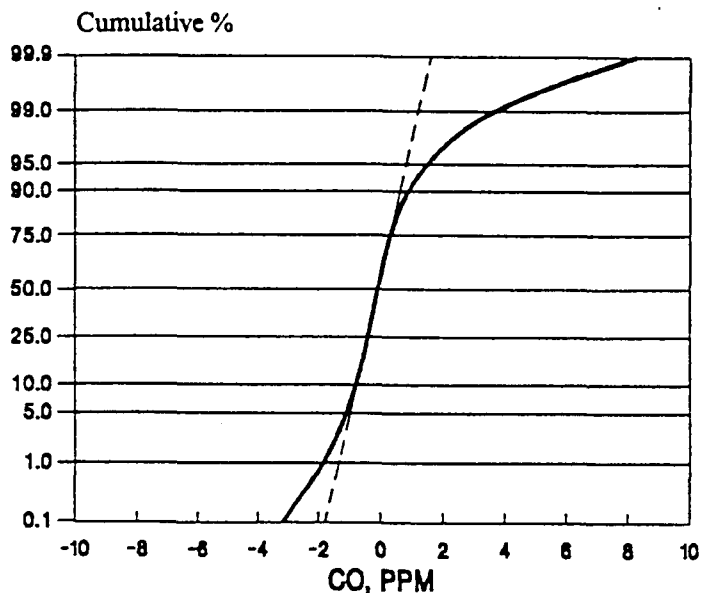


Figure 8. Cumulative CO error distribution (observed - predicted) for site 1, showing a non-gaussian relation. This actual error distribution is used in the confidence interval analysis.

In order to better understand results such as those presented in Figures 6 and 7, a Monte Carlo confidence interval analysis is employed. The set of n sample error values (observed - predicted) is calculated, with the number of observations n more than 50,000 for each monitor site. (The distribution of the sample error values for site 1 is shown in Figure 8). A set of n data values are calculated as the sum of the CO value predicted for the independent variable values (Equation 3) plus a pseudo-random error selected from the set of sample error values. The set of generated data values are re-solved for the parameters A , B , γ , and the discrete distribution functions $f(\)$, $g(\)$, $h(\)$, $u(\)$, $v(\)$ as previously described. This Monte Carlo process is repeated to obtain the confidence interval which brackets 95% of the solution values for each parameter or discrete distribution function level. This Monte Carlo confidence interval technique may be shown to match the theoretical confidence intervals obtained with linear regression analysis (7).

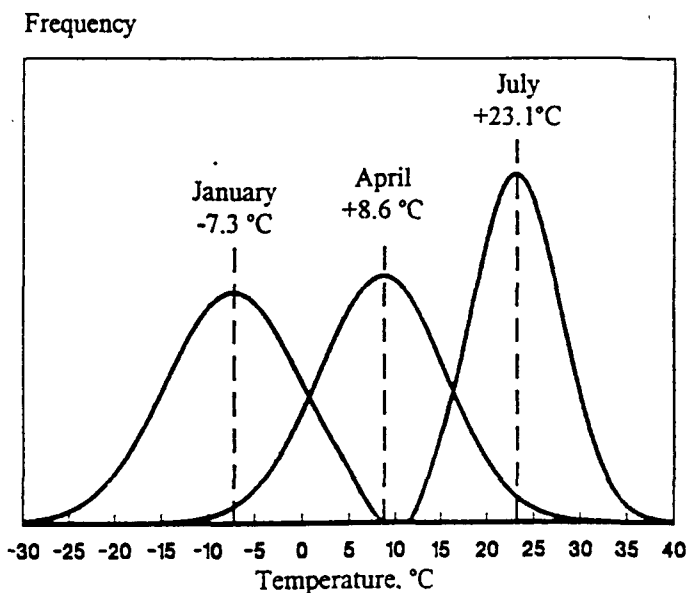


Figure 9. Distribution of hourly ambient temperatures during the study period for January, April, and July, with mean values labeled.

As seen in Figure 7, wide confidence intervals are obtained for conditions with a small sample size, such as ambient temperature ≤ -25 °C (186 observations) or $> +35$ °C (73 observations). For this mid-continent location, a broad distribution of hourly ambient temperatures occurs in a given month, as shown in Figure 9. A significant overlap of ambient temperatures is also seen to occur in successive months. Thus, the underlying temperature conditions allow separation of ambient CO factors related to month (Figure 6) from ambient CO factors related to temperature (Figure 7) with statistical significance. Factors which may be related to higher October through February CO levels include temperature inversion (or vertical mixing) in the atmosphere, seasonal fleet composition, and vapor pressure changes in the fuel supply.

The in-use vehicle emission trend factors A and B , for each of the 3 sites, are presented in Table 2. The pre-exponential term A represents the baseline ambient CO level at time $t=0$, or July 1, 1986. All three sites yield results for B showing ambient CO levels to be continuously

	A , [PPM]	$-B$, [Yr ⁻¹]	$T_{1/2}$, [Yr]
Site 1	2.57 ± 0.03	-0.0616 ± 0.0037	11.3
Site 2	2.63 ± 0.02	-0.0979 ± 0.0036	7.1
Site 3	0.99 ± 0.01	-0.0344 ± 0.0051	20.1

Table 2. Solution values for the in-use emission trend factors (A and $-B$) for each of the 3 sites analyzed. Time in years to obtain a 50% ambient CO reduction is calculated as $T_{1/2} = -\log_e(0.5) / B$.

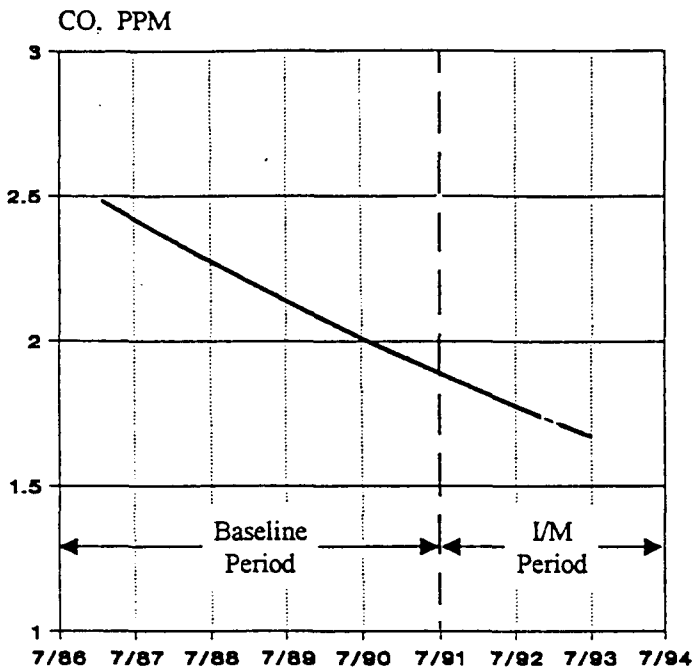


Figure 10. In-use ambient CO emission trend factor, $A \exp\{-Bt\}$, for site 1 over the baseline and I/M implementation periods.

decreasing, but at differing rates. The time for a 50% ambient CO concentration reduction ($T_{1/2}$), calculated from B , is shown for each site.

The in-use vehicle emission trend, $A \exp\{-Bt\}$, is plotted in Figure 10 for site 1. This curve comprises the basic time-series result, with periodic and random influences removed, to which pre and post-I/M implementation data should be referenced. Figure 11 examines a sensitive measure of pre and post-I/M time-series data by plotting the integral

$$\int [CO(t)_{observed} - CO(t)_{predicted}] dt \quad (7)$$

with the predicted values calculated omitting the I/M reduction term $\{1 - \gamma IM(t)\}$. As seen in the figure, observed CO levels very closely follow values predicted omitting the I/M term, both during the I/M implementation and baseline

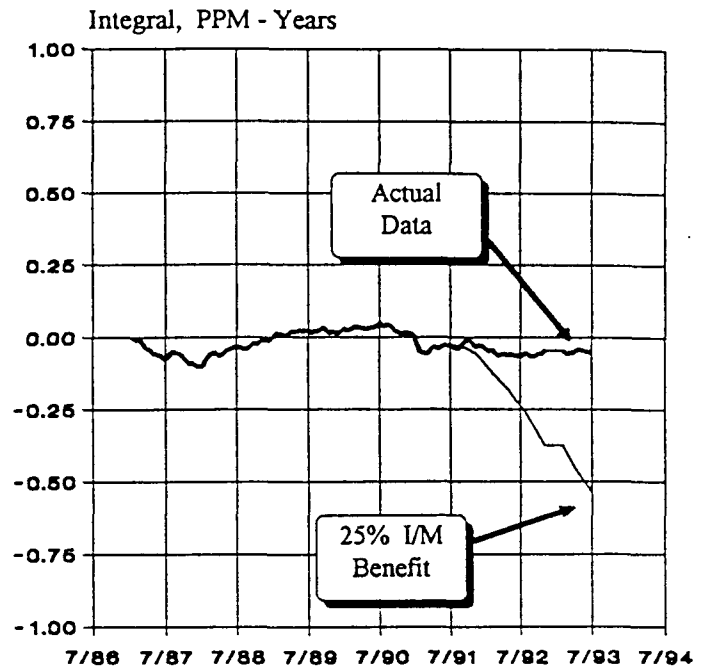


Figure 11. Integral of the difference between observed and predicted ambient CO values for site 1, indicating any systematic change after I/M implementation.

periods. The data fail to demonstrate a systematic reduction in ambient CO monitor levels following I/M implementation, such as the curve calculated for a phased-in 25% I/M benefit reduction.

The I/M benefit factor γ , directly solved for each of the 3 monitor sites, is presented in Table 3. A positive percentage for γ represents a benefit (ambient CO level reduction). The average ambient CO monitor benefit for the three sites is $+1.3 \pm 1.4\%$. The 95% confidence intervals show the tightest intervals for sites 1 and 2, which had the highest CO levels during the I/M implementation period. The widest confidence interval is for site 3, which registered the lowest CO levels throughout the study period.

These results, obtained directly from the same monitors which determine NAAQS compliance, complement recent work comparing emissions from I/M and non-I/M vehicles by remote sensing measurement. No significant difference in hydrocarbon emissions between I/M and non-I/M vehicle populations was found (8).

	Number of Observations	% I/M Benefit $\gamma \times 100$	95 % Confidence Interval
Site 1	52,686	+ 1.5 %	$\pm 2.2\%$
Site 2	53,146	+ 5.8 %	$\pm 2.0\%$
Site 3	51,069	- 3.4 %	$\pm 2.8\%$

Table 3. I/M effectiveness result (γ) for each of the three ambient CO monitor sites analyzed.

Percent Usage of 10% Ethanol Blend

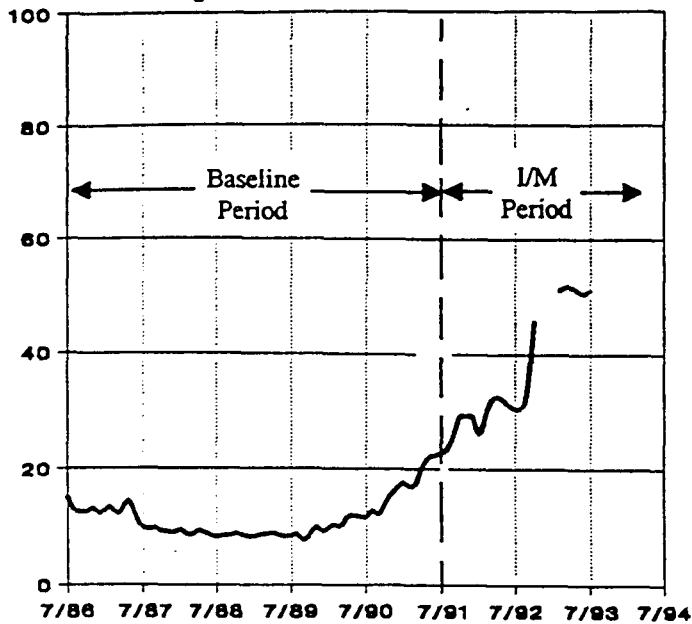


Figure 12. Statewide tax credit data for 10% ethanol blend fuel usage. Increasing ethanol usage occurred simultaneously with I/M program implementation.

DISCUSSION

Two potential sources of systematic error require discussion in examining these results. The choice of fitting the in-use vehicle emission trend to an exponential ($A \exp\{-Bt\}$) rather than linear ($A - Bt$) factor is based on expectations for long-term emission trends. Over the 5 year baseline period of this study, no clear distinction is seen in the data between a linear or exponential trend. The choice of an exponential trend yields higher expected ambient CO levels for the I/M implementation period, resulting in a higher measured I/M effectiveness. This difference is greatest for site 2, with an I/M effectiveness of $-3.7 \pm 2.9\%$ obtained for a linear trend analysis.

A second potential source of systematic error is the use of oxygenated fuel. Statewide, usage of 10% ethanol blend fuel remained low during the baseline period, but increased substantially following I/M implementation (Figure 12). No significant usage of other oxygenates is believed to have occurred during the study period. The metropolitan area studied represents 48% of statewide gasoline consumption, and is thought to closely follow this statewide trend. However, in the metropolitan area only, oxygenated fuel usage was required by law during the period Nov. 1, 1992 through Jan. 31, 1993. In order to reduce potential biasing of I/M effectiveness results, CO data for this period are excluded from the analysis. Any reduction in CO levels due to increasing oxygenated fuel usage during the remainder of the post-I/M implementation period would tend to increase apparent I/M effectiveness.

CONCLUSIONS

An overall I/M effectiveness of $+1.3 \pm 1.4\%$ percent is measured, with individual results of $+1.5\%$, $+5.8\%$, and -3.4% for the three separate monitor sites studied. This one-time benefit is observed to be negligible in comparison to the continuous reductions in ambient CO levels measured at the 3 monitoring sites. Significantly higher ambient CO levels are observed during the winter months, a result not directly attributable to ambient temperature. Ambient CO measurements are seen to be strongly influenced by wind vector / site interactions for the 2 sites studied with a history of NAAQS violations.

While I/M may be projected to supplement normal maintenance practices, it must be recognized that I/M programs have the potential to reduce or eliminate effective, periodic, vehicle maintenance -- offsetting in practice I/M benefits which are speculated in principle to exist. The importance of considering actual data over projections and speculation in formulating sound public policy is illustrated in these findings.

REFERENCES

1. U.S. Environmental Protection Agency, "I/M Network Type: Effects on Emission Reductions, Cost, and Convenience", EPA-AA-TSS-I/M-89-2, January, 1991.
2. K.H. Jones and J.F. Walsh, "The Past and Future Benefits of Automobile Inspection and Maintenance Programs", SAE Paper 831208, August, 1983.
3. H.M. Haskew, J.J. Gumbleton, and D.P. Garrett, "I/M Effectiveness with Today's Closed Loop Systems", SAE Paper 871103, May, 1987.
4. H.M. Haskew, D.P. Garrett, and J.J. Gumbleton, "GM's Results - The EPA/Industry Cooperative Test Program", SAE Paper 890185, March, 1989.
5. U.S. Environmental Protection Agency, "Report on the EPA/Manufacturer Cooperative I/M Testing Program", PB93-208015/GAR, September, 1992.
6. U.S. National Climatic Data Center, "Local Climatological Data Monthly Summary, Minneapolis - St. Paul, MN", ISSN 0198-2745, July 1986 - June 1993.
7. B.A. Berg, "Monte Carlo calculation of Confidence Limits for Realistic Least Squares Fitting", Computer Physics Communications, v.69, pp.65-72, 1992.
8. Y. Zhang, D.H. Stedman et. al., "On-Road Hydrocarbon Remote Sensing in the Denver Area", Environ. Sci. Technol., vol.27, no.9, pp.1885-91, September, 1993.

EPA Response to the Report:

"I/M Effectiveness as Directly Measured by Ambient CO Data"

by Huel Scherrer and David Kittelson
University of Minnesota

Conclusion

For the reasons cited below, EPA does not believe that the study conducted by the University of Minnesota is an accurate analysis of the effectiveness of the inspection and maintenance (I/M) program in the Minneapolis-St. Paul area.

The Study is Built on a False Premise

- The study establishes the baseline emission reductions it attributes to new car technology by linearly extrapolating pre-I/M CO reductions indefinitely into the future.
- This extrapolation suggests that even without I/M, CO emissions will eventually be reduced to zero, due to new car technology and fleet turnover.
- In reality, the reduction in CO due to new car technology will level off as vehicle standards stabilize and fleet turnover to these standards is completed.
- The authors' claim that their conclusions are based on "real-world data" as compared to EPA's computer projections is disingenuous. A "trend" is just another type of projection based upon founded or unfounded assumptions, and data are only "real" to the extent that the formula used is valid and addresses all relevant variables.

The Current Minnesota I/M Program is Mischaracterized

- The Minneapolis-St. Paul area has a basic I/M program, but the authors suggest it is an enhanced program by describing it as a "comprehensive I/M program" that "maximize[s] I/M program effectiveness."
- The authors claim that the program was predicted to "lead to a 25 to 30 percent reduction in carbon monoxide levels," erroneously suggesting that this reduction should be observed at individual monitoring sites.
- Neither EPA nor the Minnesota Pollution Control Agency has ever predicted I/M emission reductions for specific monitoring sites, nor have they predicted that the Minnesota program would achieve a 25% to 30% reduction in overall CO inventory.

APPENDIX C

Can Remote Sensing Data Be Used to Assess Differences in Emissions Between Fleets?

We have presented remote sensing data from Chicago and Tucson that indicate that the centralized I/M programs in these regions may be having little or no impact on the on-road emissions of cars in these regions. In the Tucson study, remote sensing data indicated that the I/M and non-I/M fleets in the Tucson area did not differ in their tailpipe HC or CO emissions.³⁹ In the Chicago study, remote sensing data indicated that there was no correlation of on-road HC or CO emissions with time since the last I/M inspection.

At the February 1, 1995 I/M Review Committee Meeting, Mr. Phil Lorang of USEPA suggested that remote sensing data were too variable to determine differences in emissions between fleets. The Committee's staff responded that, whatever disputes there may be regarding the representativeness of a single RSD reading for a given car, when RSD data are aggregated for a fleet of cars (that has been tested both on a dynamometer and by RSD), the correlation of RSD with the IM240 or FTP tests is quite high. Therefore, RSD data provide a valid way to compare emissions of different fleets of vehicles. On the following pages, we present data showing that fleet average remote sensing readings do indeed show a high correlation with the IM240 and FTP.

Figure C1 shows the results of FTP and RSD measurements for 556 cars in ARB's surveillance fleet (with a regression line drawn through the data). The data we analyzed consist of a single FTP and a single RSD reading for each car.⁴⁰ The data were ranked from lowest to highest FTP emissions and divided into quintiles. The figure is a graph of the average emissions for each quintile on the FTP and RSD. As the figure shows, the averaged remote sensing data provide an accurate reflection of the averaged FTP emissions of the vehicles. The RSD data for this sample were acquired under controlled conditions where cars were driven by the RSD unit at a constant speed.

To see whether RSD is accurate under less controlled conditions, we looked at data for 132 vehicles that were measured by RSD as they drove by on the road, and that were subsequently given an IM240 test at the roadside.⁴¹ These data are plotted in Figure C2 (with a regression line drawn through the data). The median number of RSD measurements per car is two and the average is 2.6. The data thus consist of the average RSD reading paired with the IM240 reading for each car. The data were once again ranked, divided into quintiles, and averaged. As the graph shows, even with this small sample, fleet averages of on-road RSD readings show relatively good correlation with the IM240.

³⁹ The emissions from cars in this study were all measured at the same location with the same instrument, eliminating potential comparison problems that could occur when measurements are compared from different locations and instruments.

⁴⁰ Thanks to Mr. Mark Carlock of ARB for supplying this data set.

⁴¹ Thanks to Mr. Robert Stephens of General Motors for supplying this data set.

The USEPA has also presented research showing that averaged RSD readings achieve good agreement with the averaged IM240 readings (Glover and Clemmens, 1991).

We conclude that fleet RSD measurements are valid indicators of fleet emissions. RSD measurements can therefore be used to measure emission differences between I/M and non-I/M fleets, and also emission differences between fleets of vehicles grouped by time since their last inspection.

APPENDIX D

Summary Of USEPA's I/M Audit Results

Definition of Terms in the Tables Below

Improper Tests: The improper test rate audit

State Failure rate on state audits (in %)

EPA Failure rate on USEPA audits (in %)

Gas Analyzer: The gas analyzer calibration audit.

Number of analyzers tested

HC % of failing hydrocarbon

CO % of failing carbon monoxide analyzers

Early Test Failure Rate: The emission test failure rate earlier in a program's history.

Reported The reported failure rate (in %)

Expected The expected failure rate (in %)

Ratio The ratio of reported to expected failure rates

I/M Program: The I/M program type

Centralized/Contractor Centralized, Contractor-Operated

Centralized/Government Centralized, Government-Operated

Decentralized/Computerized Decentralized, Computerized Analyzers

Decentralized/Manual Decentralized, Manual Analyzers

Notes We have presented the numbers as they appear in USEPA (1991 and 1993a) with the following caveats.

- Where both documents present the same type of audit data, we show the numbers in the more recent publication.
- We have disaggregated the data for programs with manual or computerized analyzers, which the USEPA does not always do.
- In cases where the USEPA presents data for individual cities in the same state, we have averaged the data for each city to arrive at a score for the state (the only exception is Tennessee where there are two different network types, reported separately).
- On the early test failure rate audit, some programs that now use computerized analyzers (and are reported as decentralized/computerized in the table) may have used manual analyzers at the time of the audit.

I/M Program	Improper tests		Gas Analyzer			Early Test Failure Rate			Waiver Rates	
	STATE	EPA	#	HC	CO	Reported	Expected	Ratio	pre-81	post-80

Centralized/Contractor										
AZ		11.0				20.2	36.8	0.55	12	12
CT			32	0	0	17.2	33	0.52	5	4
FL										
IL			47	0	0				11	11
LOU			6	0	0	15.7	16.2	1.00	17	12
MD		0	11	0	0	14.6	14.0	1.00	20	19
MN										
NASH						24.5	25.4	0.97		
WA			12	25	8	19	28.1	0.68	15	15.5
WI			11	18	9	15.3	19.3	0.79	12	9
Average		5.5	19.8	4	2	18.7	26.5	0.7	13.1	11.8

Centralized/Government										
DC			13	23	15	18.4	13.4	1.00		
DE			9	67	22	13.7	7.7	1.00	3	1
IN			24	29	8				10	13
OR			4	50	0	24	38.3	0.63		
MEM			9	44	33	8.1	3.7	1.00	1	2
NJ		50				26.1	27.8	0.94		
Average		50	11.8	37	15	18.1	18.2	0.9	4.7	5.3

Decentralized/Computerized										
AL	53					17.4	23.7	0.69	1	1
CA	19		22	9	14	27.7	28.7	0.96	29	9
CO	10	58	21	29	14				2	1
GA	34	50	21	14	10	6.6	25	0.26	14	12
MA	84	100								
MI	61	96				15.8	12.9	1.00	10	9
MO	58	84				6.7	20.5	0.33	11	14
NH	89	93	14	35	18					
NV	31	67				10.3	29.4	0.35		
NY	42	68				5.1	33.4	0.15		
PA			11	27	18	17.6	19.5	0.90	2	1
VA	60	100				2.3	15.6	0.15		
Average	49.2	79.6	17.8	20	15	12.2	23.2	0.5	9.9	6.7

Decentralized/Manual										
ID						9.8	16.9	0.58	7	26
NC			17	35	65	5.6	21.1	0.27	0	0
UT	11					9.4	21.3	0.44	6.7	4.3
Average	11		17	35	65	8.2	18.8	0.4	4.6	10.1

SUMMARY	Averages for Different Network Types as Noted									
I/M Program	Improper tests		Gas Analyzer			Early Test Failure Rate			Waiver Rates	
	STATE	EPA	#	HC	CO	Reported	Expected	Ratio	pre-81	post-80
Centralized/Contractor		5.5	19.8	4	2	18.7	26.5	0.7	13.1	11.8
Centralized/Government		50	11.8	37	15	18.1	18.2	0.9	4.7	5.3
Decentralized/Computerized	49.2	79.6	17.8	20	15	12.2	23.2	0.5	9.9	6.7
Decentralized/Manual	11		17	35	65	8.2	18.8	0.4	4.6	10.1
Centralized		20.3	16.2	15	6	18.4	22.7	0.8	10.6	9.9
Decentralized	46	79.6	17.7	23	23	11.2	22.3	0.5	8.3	7.7

APPENDIX E

Has the Judiciary Ruled on the Validity of USEPA's 50-Percent Discount?

Recent USEPA documents suggest that a federal court has upheld the USEPA's 50-percent discount for decentralized I/M programs (Nichols, 1994; Tierney, 1995). USEPA refers to a case filed by the National Automobile Dealers Association (NADA) against USEPA (NADA, 1993; US Court of Appeals, 1994).

NADA argued that the USEPA's 50-percent discount is "...arbitrary, capricious, an abuse of discretion or otherwise not in accordance with law" (NADA, 1993). However, NADA did not provide the court with an analysis of the data underlying the 50-percent discount. NADA argued its case on legal and administrative issues, not on the substance of the science underlying the 50-percent discount.

The court ruled that the USEPA was given broad discretion by Congress to determine what constitutes an effective I/M program. Furthermore, the court ruled that "ample evidence in the record supports the EPA's imposition of the 50 percent penalty. The final rule cites several studies resulting from both overt and covert audits of testing facilities that demonstrate considerable levels of improper testing..." The court then goes on to cite specific audit results presented by the USEPA.

The court did not *evaluate* the USEPA's science and analysis. The court merely noted the fact that the USEPA cited studies and data that supported its findings. The court then simply repeated some of the USEPA's findings. In effect, the court ruled on form, not on substance. The relevant section of the court's decision is reproduced on the following pages.

GLOSSARY

- Centralized** Centralized I/M programs require that testing and repair be performed by separate entities. Centralized I/M programs are often also referred to as test-only programs.
- It should be noted that the term "centralized" arose out of the fact that I/M programs with separate testing and repair usually have testing done at a small number of high-volume centers. Nevertheless, it is also possible to have a program with separate testing and repair, but with testing performed at low-volume, neighborhood shops. Such a system would not, strictly speaking, be "centralized."
- Covert Audit** An audit in which a USEPA or state regulator poses as an ordinary motorist bringing a car in for a I/M inspection. The USEPA uses covert audits to determine rates of improper testing in I/M programs.
- Decentralized** Decentralized I/M programs allow participating stations to do both testing and repair. Thus, decentralized programs are often also referred to as test-and-repair programs.
- Excess Emissions** In this case, excess emissions refers to emissions over and above the emission standard that a vehicle must exceed in order to fail an emissions test in a given I/M program.
- IM240** This is the emission test preferred by the USEPA. The IM240 test involves putting a vehicle under variable loads and speeds during a 240 second test.
- Improper Test** The USEPA defines a test as improper if an inspector fails to perform at least one required function during the inspection. For instance, if an inspection is supposed to include a catalyst tampering check, and one is not performed, the test is considered to be improperly performed whether or not the rest of the inspection is performed properly.
- Preconditioning** Preconditioning means running a vehicle for a few minutes before the emissions test in order to ensure that it is properly warmed up. Cars that are not warmed up could fail the emission test even when the car has no reparable defects. In programs using an idle test, preconditioning usually means running the vehicle at a high-speed idle for about two or three minutes. Some centralized programs precondition cars on a dynamometer for 30 seconds to a minute.
- r²** Pronounced "r-squared" and sometimes written this way as well. r^2 is a measure of the correlation between two variables. For instance, if the r^2 between two variables equals 0.7, then 70% of the variation in one variable can be predicted by variation in the other. An r^2 of zero means that the variables are completely uncorrelated, and an r^2 of 1 means that they are perfectly correlated. It is important to remember that a

correlation between variables does not indicate that there is necessarily any causal relationship between them.

**Statistical
Significance**

A result is termed statistically significant when it is determined by a given statistical test to be unlikely to have occurred by chance. For instance, non-I/M cars in Arizona's I/M program were measured to be slightly cleaner than the fleet subject to I/M, but this difference was not statistically significant at a level of 0.05. This means that there is a 95% probability that the two groups of cars have about the same average emissions, and that the measured difference in emissions was due to random fluctuation.

Note that significance, as it is used here, does not necessarily mean "important" or "noteworthy" as it would in common speech. It merely refers to the extent to which measured differences are likely to result from chance fluctuations, or from genuine differences between two populations.

Waiver

A vehicle may in some cases be excused from being repaired in an I/M program. For example, a waiver may be granted if needed repairs would exceed an I/M program's cost limit for that vehicle.