STATE OF ARIZONA
OFFICE OF THE
AUDITOR GENERAL

A PERFORMANCE AUDIT
OF THE

DEPARTMENT OF HEALTH SERVICES VEHICLE EMISSIONS INSPECTION PROGRAM

FEBRUARY 1983

A REPORT TO THE
ARIZONA STATE LEGISLATURE

REPORT 83-1

DOUGLAS R. NORTON, CPA
AUDITOR GENERAL

STATE OF ARIZONA
OFFICE OF THE
AUDITOR GENERAL

February 4, 1983

Members of the Arizona Legislature
The Honorable Bruce Babbitt, Governor
James Sarn, M.D., M.P.H., Director
Department of Health Services

Transmitted herewith is a report of the Auditor General, A Performance Audit of the Department of Health Services, Vehicle Emissions Inspection Program. This report is in response to Senate Bill 1220 enacted by the Thirty-fourth Legislature in 1980.

The blue pages present a summary of the report; responses from the Department of Health Services and other affected parties are found on the yellow pages preceding the appendices.

My staff and I will be pleased to discuss or clarify items in the report.

Respectfully submitted,

Douglas R. Norton
Auditor General

Staff: William Thomson
Peter Francis
Mark Fleming
Sylvia Forte
William Wright
Richard Booth

Enclosure

OFFICE OF THE AUDITOR GENERAL

A PERFORMANCE AUDIT OF THE DEPARTMENT OF HEALTH SERVICES VEHICLE EMISSIONS INSPECTION PROGRAM

A REPORT TO THE
ARIZONA STATE LEGISLATURE

REPORT 83-1

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The Office of the Auditor General has completed a performance audit of the Department of Health Services, Vehicle Emissions Inspection Program. This audit was conducted in response to Senate Bill 1220 enacted by the Thirty-fourth Legislature in 1980.

The Arizona Vehicle Emissions Inspection (VEI) Program requires that certain motor vehicles pass an annual inspection to ensure that their exhaust emissions meet standards established by the Department of Health Services (DHS). The program goal is to protect public health from the effects of motor vehicle pollution. Although the Arizona Legislature originally initiated the VEI Program in 1974 in response to air quality problems in the Phoenix and Tucson areas, emissions inspection is now a central component of the State's plan for meeting requirements of the Federal Clean Air Act.

We found no evidence that the VEI Program has been effective. Results of a time series analysis covering the eight-year period 1974 through 1981 indicated that the emissions inspection program has not reduced carbon monoxide levels in Phoenix and Tucson. Neither the implementation of the mandatory program in 1977 nor dramatic increases in test failures in 1979 had an impact on carbon monoxide concentrations in either city (see page 9).

Several problems and factors may, either individually or together, explain why the VEI Program has not been effective. The Program may not have significantly changed vehicle maintenance behavior. Further, a substantial proportion of motorists admit to circumventing program requirements by readjusting their engines after the emissions tests. Exclusion of older and out-of-state vehicles from test requirements and problems resulting from the variability of automobile emissions may also contribute to the Program's ineffectiveness. However, we do not know for certain why it is ineffective, and thus the effect of policy or program changes addressing any or all of these factors is also unknown (see page 11).

Federal law requires Arizona to operate an emissions inspection program. Unless additional research identifies feasible alternatives, if any, for an effective program, we recommend that DHS operate the VEI Program only as required to avoid impositions of Federal sanctions. Any program changes should be rigorously and independently evaluated before being implemented. We further recommend that DHS identify alternate strategies for reducing automotive air pollution in the absence of an effective emissions inspection program. In addition, the Legislature should consider petitioning the U.S. Congress to review the appropriateness of the emissions testing requirement in the Clean Air Act.

The VEI Program boundaries in Pima County are not appropriate. Although the boundaries are larger than necessary to address carbon monoxide problems in Tucson, the Program boundaries exclude areas which may be required for effective ozone control. We also found, however, effective control of ozone through the VEI Program may not be possible because 1) the Program has not been effective in reducing carbon monoxide levels; 2) the Program does not address nitrous oxides, one of the two precursors of ozone; and 3) few vehicles fail emissions tests for hydrocarbons, the other component of ozone.

The Pima Association of Governments (PAG) and the Department of Health Services should analyze the nature and sources of automotive pollution in the Tucson area. PAG should adjust the program boundaries in Pima County to include all areas which contribute to Tucson's current and potential automotive air pollution and exclude those areas which do not (see page 17).

The air/fuel waiver provision of Arizona's Vehicle Emissions Inspection Program is not cost effective. Although statutes require the Department of Health Services, Bureau of Vehicular Emissions Inspection (BVEI) to issue air/fuel waivers, less than 2 percent of the motorists use this service. As a result, the cost per waiver is approximately $61. The Bureau could save an estimated $121,700 in State and Federal funds annually by eliminating the air/fuel waiver program. We recommend that the Legislature consider amending A.R.S. §36-1772.E. to allow such an action (see page 23).

Neither the Bureau nor Hamilton Test Systems (HTS) conducted the appropriate number of inspection station field audits in fiscal year 1981-82. BVEI conducted only 56 percent of the audits required by its regulations during the year. HTS conducted only 53 percent of the audits required by company policy. Although BVEI and HTS have increased the frequency of inspection station audits, the Bureau needs to strengthen control by 1) instituting a formal management reporting system to inform the Bureau chief that necessary audits are conducted and 2) requiring HTS by contract to conduct two audits per month on each inspection station lane (see page 32).

In addition, the Bureau needs to ensure that its inspectors conduct timely quality assurance inspections of analyzers used by fleet inspectors and private garages as required by Bureau rules. Results of these inspections should be reported to the Bureau chief at least monthly (see page 34).

INTRODUCTION AND BACKGROUND

The Office of the Auditor General has completed a performance audit of the Department of Health Services, Vehicle Emissions Inspection Program. This audit was conducted in response to Senate Bill 1220 enacted by the Thirty-fourth Legislature in 1980.

Arizona Revised Statutes §36-1772, which establishes the Vehicle Emissions Inspection (VEI) Program, requires that certain motor vehicles pass an annual inspection to ensure that their exhaust emissions meet standards established by the Department of Health Services (DHS). The program goal is to protect public health from the effects of motor vehicle pollution. Although the Arizona Legislature originally enacted the program in 1974 in response to air quality problems in the Phoenix and Tucson areas, emissions inspection is now a central component of the State's plan for meeting requirements of the Federal Clean Air Act.

Program Operations

The Bureau of Vehicular Emissions Inspection (BVEI) within DHS administers the inspection program. To achieve the program goal of protecting public health, the Bureau has established three objectives. These are:

- To inspect all vehicles in the nonattainment areas of Maricopa and Pima Counties as required by law,
- To adopt emissions standards which will identify vehicles which are gross and high polluters, and
- To require repairs to vehicles which are gross and high polluters which will reduce carbon monoxide and hydrocarbon emissions to the level of vehicles initially passing the inspection.

The VEI Program tests vehicles to ensure that carbon monoxide and hydrocarbons in their exhaust emissions meet standards established by DHS. The inspection requirement applies generally to gasoline-powered vehicles which are less than 14 years old and located within designated portions of Pima and Maricopa Counties. The Program includes those parts of the two counties which do not meet the carbon monoxide standards of the Federal Clean Air Act. Motorists whose vehicles fail to meet these standards must repair their vehicles and submit to a retest. Motorists must either pass the emissions test or seek one of the several waivers provided under the law or BVEI regulations.

The Bureau does not conduct the actual inspections. A private company, Hamilton Test Systems (HTS), under contract with DHS operates nine permanent facilities and one mobile station where its employees conduct emissions tests for the State. Hamilton conducts about 1.4 million tests per year. Payments to the contractor during 1981 were approximately $5.8 million.

Importance of the Vehicle Emissions Inspection Program

The Vehicle Emissions Inspection Program is the principal strategy for reducing carbon monoxide pollution in Phoenix and Tucson. The Federal Clean Air Act requires Arizona to implement an emissions inspection program in these two cities because neither met the Federal standards for carbon monoxide by December 1982. Failure to implement such a program and develop a plan for meeting the standards by December 31, 1987, could cause the two cities to lose more than $100 million in Federal aid. However, Arizona has an inspection program in effect and Phoenix and Tucson are making acceptable progress in developing the required plans. Thus, neither city appears to be in danger of losing Federal funds in the forseeable future.

Current plans for meeting the 1987 carbon monoxide standards rely on the VEI Program as the primary strategy.* Although other strategies will also be implemented, neither the Environmental Protection Agency (EPA) nor Arizona's regional planning agencies consider these strategies sufficient

---

* Although carbon monoxide levels in Pima County are substantially less than in Maricopa County and the Federal standards may be easier to attain, the primary strategy will most likely be the VEI Program. Pima County has not yet completed work on its draft plan.

to produce the carbon monoxide reductions needed to meet the Federal standards. For example, the draft plan for Maricopa County projects needed carbon monoxide reductions of 449 tons per day to attain the 1987 goal. The plan estimates that the existing VEI Program combined with Federal emission requirements for new motor vehicles will reduce carbon monoxide by 349 tons per day (78 percent). The Maricopa plan projects additional reductions of 71 tons per day (16 percent) if the VEI Program also includes all vehicles produced since 1969.

In addition to the VEI Program, the Maricopa plan for attaining Federal standards also includes strategies to reduce vehicle use and to improve traffic flow. The major efforts in this direction are to maintain transit ridership at current levels and implement specific transportation control measures (reversible lanes, for example). These efforts, however, account for reductions of only 27 tons per day, 6 percent of the needed reduction. Other strategies, such as increasing transit ridership and vehicle occupancy would provide 38 tons per day of the needed reduction.

Organization and Budget

BVEI is one of four bureaus in the DHS Environmental Health Services Division. The Bureau has 23 FTE positions for the current fiscal year and employed 22 persons in Phoenix and Tucson as of August 1982. Seventeen employees work in four sections located in Phoenix. These sections are:

- Administration--processes requests for exemptions, waivers and renewals of mechanic certificates and provides research and clerical support;
- Licensing and Certification--trains and certifies mechanics to perform procedures required for air/fuel waivers and fleet inspections.
- Waiver and Testing--operates the BVEI waiver and retest facility and inspects vehicles owned by small governments;
- Quality Assurance--inspects State stations operated by HTS to ensure compliance with procedures, reliability of equipment and inspects and certifies analyzers used by fleet inspectors and private repair facilities.

Five persons work in the Bureau's Tucson office and perform essentially all of the functions of the Phoenix office with the exception of inspecting State stations. Personnel from the Phoenix office inspect State stations in both cities.

As shown in Table 1, the total budget for the Vehicle Emissions Inspection Program in fiscal year 1982-83 is approximately $6.7 million. The VEI Program has several sources of funding. General Fund appropriations supplemented by Federal grants support most BVEI personnel and activities. In addition, fees collected for emissions tests flow into a revolving fund. BVEI uses these funds to pay Hamilton Test Systems for the emissions tests it conducts. Payments to the contractor account for 90 percent of the VEI Program budget. The revolving fund also includes fees charged for waivers and exemptions and supports the cost of personnel who provide these services.

TABLE 1

VEHICLE EMISSIONS INSPECTION PROGRAM BUDGET
FOR FISCAL YEAR 1982-83

	All Sources	General Fund	Revolving Fund	Federal
FTEs	23	16	5	2
Personal services	$ 507,100	$359,100	$ 108,300	$39,700
Employee-related expenditures	108,026	75,900	23,826	8,300
Professional and outside services	5,960,974		5,960,974*	
Travel:				
In-State	17,350	14,550	2,800	
Out-of-State	2,000		2,000	
Other operating expenses	91,155	53,555	34,500	3,100
Other payments	10,900			10,900
Totals	$6,697,505	$503,105	$6,132,400	$62,000

* Estimated payment to Hamilton Test Systems for emission tests.

Audit Scope

The scope of our audit focused primarily on the need for the emissions inspection program, its impact and the policies and procedures which govern its operation. The audit addressed the following specific issues:

- The effect of the Vehicle Emissions Inspection Program on the carbon monoxide levels in Phoenix and Tucson;
- The appropriateness of VEI Program boundaries in Pima County;
- The cost effectiveness of the air/fuel waiver program; and
- The adequacy of quality assurance for State emissions testing stations and analyzers used by private and fleet facilities.

The report section, Other Pertinent Information, presents detailed information on vehicle inspection fee payments and collections and the inspection equipment used in the Program and HTS test and billing procedures.

Collection and analysis of data in two areas required use of outside professional assistance. The Auditor General contracted with the Center for Informative Evaluation to conduct a time series analysis to assess the impact of the VEI Program on carbon monoxide levels in Phoenix and Tucson. The results of this analysis form the basis for Finding I. A second contractor, the Behavior Research Center, surveyed 800 persons in Phoenix and Tucson who had experiences with the VEI Program. The survey provided information on public perception of air quality, attitudes toward the VEI Program, and their experiences in complying with its requirements. Survey results supplement Auditor General staff research on program effectiveness.

The Auditor General and staff express appreciation to the Department of Health Services and Hamilton Test Systems for their cooperation and assistance during the course of the audit.

FINDING I

THE VEHICLE EMISSIONS INSPECTION PROGRAM HAS NOT REDUCED CARBON MONOXIDE LEVELS IN PHOENIX AND TUCSON.

The Vehicle Emissions Inspection (VEI) Program has not reduced carbon monoxide (CO) levels in Phoenix and Tucson. Neither the implementation of the mandatory program in 1977 nor dramatic increases in test failures in 1979 had an impact on carbon monoxide concentrations in either city. Thus, the VEI Program does not appear to be an effective strategy for meeting the Federal Air Quality Standards in 1987. Although a number of factors may impair the Program's effectiveness, we do not know for certain why it is ineffective. Thus, the effect of policy or program changes addressing any or all of these factors is unknown.

Program Was Established to Improve Air Quality

The goal of the Vehicle Emissions Inspection Program is to protect public health by reducing potentially harmful automotive emissions. These emissions include carbon monoxide, a health hazard in itself, and hydrocarbons, which are a precursor of the pollutant ozone.* Unlike ozone, carbon monoxide is generated almost entirely from automotive emissions. The Arizona Legislature enacted the program in 1974 to address air quality problems in Phoenix and Tucson. In 1977 the U.S. Congress amended the Clean Air Act to require emissions inspection in areas not meeting Federal carbon monoxide and ozone standards by December 1982. The act provides for withholding Federal aid for highway construction and sewage treatment plant construction in states not complying with this requirement. If Arizona did not have the VEI Program in effect, the State could lose up to $116 million per year in Federal aid.

---

* Ozone is formed by sunlight acting on hydrocarbons and nitrous oxides.

Previous Emissions Inspection Studies

Because carbon monoxide is generated largely by motor vehicles, previous evaluations of emissions inspection programs have focused on how the program has affected carbon monoxide levels in the atmosphere. Some previous studies suggest that carbon monoxide levels have been reduced by inspection programs. Department of Health Services (DHS) assessments in the early years of the Program show improvement in carbon monoxide emissions. Federal Environmental Protection Agency (EPA) studies also indicated that emissions testing programs reduced carbon monoxide levels in Phoenix as well as Portland, Oregon.

However, the Oregon study found a reduction in carbon monoxide attributable to emissions testing at only one of four monitoring stations in Portland. Furthermore, at the site where this effect was found, results were confounded by movements of the monitoring probe and by major traffic disruptions in the area. A New Jersey study using a similar time series methodology also found reductions in carbon monoxide levels. In this study, however, the effects attributable to emissions testing could not be separated from the effects of new car emission standards.

DHS evaluations of the VEI Program have reported that carbon monoxide levels were decreasing. However, DHS analysts concluded in two recent studies that the VEI Program, even when combined with other carbon monoxide control strategies, will not enable Maricopa County to achieve Federal standards for carbon monoxide by 1987.

To evaluate the effectiveness of the Arizona Vehicle Emissions Inspection Program in Phoenix and Tucson, we contracted for an impact assessment of the Program using time series analysis covering the eight-year period 1974 through 1981.* Time series analysis has been used to evaluate air quality data in other states (Oregon and New Jersey) but has not previously been employed in Arizona. In our opinion, this methodology represents the most technically sophisticated and appropriate approach for assessing the impact of the State's Vehicle Emissions Inspection Program.

The VEI Program Has Not Reduced Carbon Monoxide Levels in Phoenix and Tucson

Although the VEI Program is intended to protect public health by reducing carbon monoxide levels, results of the time series analysis indicate that the Program has not achieved its desired effect. As a result, Maricopa and Pima Counties cannot rely on the emissions inspection program to meet the 1987 Federal Air Quality Standards.

Arizona implemented the mandatory VEI Program in January 1977. For the first time, motorists whose vehicles failed the emissions test were required to repair the vehicle and return for a retest. Beginning in that year, the program required proof of compliance for vehicle registration. By the end of 1977, all light-duty vehicles had been inspected and if necessary, maintained to standards. Moreover, in January 1979, Program standards were raised and significantly more cars failed the emissions test. The average number of vehicles failing the inspection each month increased by almost 5,900 in Phoenix and more than 2,000 in Tucson.

* The time series methodology employed in this study is the Autoregressive, Integrated, Moving Average (ARIMA) models and methods developed by Box and Jenkins. Carbon monoxide data analyzed in this study included four different measures of carbon monoxide (8-hour monthly high, mean highest 8-hour average, monthly average and mean 5 p.m. readings) at two monitoring stations in Phoenix (Central Phoenix and North Phoenix) and one measure of carbon monoxide (mean highest 8-hour average) at the downtown monitoring station in Tucson. Highest 8-hour average concentrations of carbon monoxide are collected and reported to the Federal Environmental Protection Agency as the basis for determining compliance with Federal Air Quality Standards. For a detailed, technical discussion of study results, see Appendix I.

Results of the time series analysis revealed that carbon monoxide levels in Phoenix and Tucson were not reduced by the implementation of the mandatory emissions inspection program in 1977 or by more stringent standards which took effect in January 1979. After conducting a substantial number of analyses, the consultants concluded that the Program has not been effective:

> "Our findings are that neither the 1977 intervention (mandatory I/M) nor the 1979 intervention (higher I/M standards) had any statistically significant impact on CO levels. Our analyses gave the I/M program every possible "benefit of the doubt." Yet we found not one iota of evidence to support the hypothesis that the Arizona I/M program had an impact on ambient air quality. Given the remarkable degree of consistency--our findings were consistent across interventions, sites, CO indicators, and models--we have the greatest possible confidence in our major conclusion."

Results of the time series study strongly suggest that Maricopa and Pima Counties cannot rely on the Vehicle Emissions Inspection Program to meet Federal Air Quality Standards by the end of 1987. To meet the deadline, Maricopa Association of Governments has prepared a nonattainment area plan which identifies strategies to reduce air pollution levels in Maricopa County. The Vehicle Emissions Inspection Program, although not the sole control strategy, is the major element identified in the plan.* Because the time series analysis shows that the Vehicle Emissions Inspection Program has not reduced carbon monoxide levels in Phoenix and Tucson, the Program cannot be relied on to meet the 1987 deadline for attaining Federal Air Quality Standards.

* Pima County has not yet drafted a nonattainment area plan for 1987. The plan now under development will identify the VEI Program as a major strategy for meeting Federal standards in 1987.

Several Factors May Account for the Program's Ineffectiveness

Several problems and factors may either individually or together explain why the Vehicle Emissions Inspection Program has not been effective. These factors include unchanged behavior, circumvention, automotive engine variability and vehicle exclusions. However, the effect of policy or program changes addressing any or all of these factors is uncertain.

Unchanged Behavior - As noted by our consultants (see Appendix I), part of the theory behind the VEI Program is that motorists cannot or will not voluntarily inspect and maintain their vehicles. However, most motorists tune their vehicles and would continue to do so in the absence of the Program.* Thus, the VEI Program may not affect vehicle maintenance behavior to as large an extent as was originally thought. This, in turn, reduces the impact of the Program.

Circumvention - Post-test engine adjustments may contribute to the reduced effectiveness of the VEI Program. Approximately one in nine motorists (11 percent) circumvented VEI Program requirements during the year. In addition, 20 percent of the survey respondents admitted to circumventing the requirements at some time in the past. Because many persons may be unwilling to admit to such behavior, the actual proportion of motorists circumventing the VEI Program may be higher. The result is that many vehicles which pass the emissions test subsequently operate at a level which does not comply with program requirements.

Motorists circumvent the program by adjusting their engines to pass the emissions test and readjusting their vehicles after completing the test. These individuals take such action because they feel that their vehicles do not operate properly when tuned for low emissions as illustrated by the following examples:

---

* 93 percent of survey respondents stated that if an emissions program were not required, they would tune or adjust their cars about as often as they do now. For a full report on the public survey results, see Appendix II.

- One survey respondent had a mechanic adjust the carburetor so that the vehicle would pass the emissions test. The same mechanic later readjusted the carburetor so the the car would "run right." The motorist noted that "A lot of places are telling people up front that they'll do that now."

- A respondent noted that a mechanic had to readjust the carburetor after it had passed the emissions test because the vehicle kept stalling.

- Another person reported resetting the idle after the emissions test. Prior to the test, a mechanic had tuned the vehicle which reduced the fuel economy from 31 mpg to 23 mpg. The individual's adjustments increased mileage to 27 mpg.

Since at least 11 percent of vehicle operators circumvent the program by making post-test adjustments, their impact on program effectiveness is likely to be significant.

Engine Variability - The variability of automotive engines may also limit program effectiveness. Emissions from an automobile can vary during the 12 months between inspections. As a result, the Vehicle Emissions Inspection Program may not produce the anticipated reduction in carbon monoxide emissions from a given vehicle.

The Motor Vehicle Manufacturers Association (MVMA) cites three major problems associated with automobiles which affect the reliability of emissions testing:

1. Vehicle emissions will vary over a wide range of weather conditions. Temperature, humidity and barometric pressure can cause emissions to vary substantially. For example, a 10 degree temperature change caused carbon monoxide emissions to increase by 50 percent in one test.

2. Vehicle usage immediately prior to the emissions test can affect test results. Measurements taken after a long drive are quite different from

those taken after short drives as might occur during in-town use. Idling while waiting in line for a test can also be a factor.

3. The variety of fuels used in different vehicles have different characterstics that affect measured emissions levels.

In addition, the variation resulting from the different variables is not consistent from one model to another or from test to test. As a result, MVMA suggests that emissions testing focus on detecting only "gross" polluters.

The problem with the variability of vehicle emissions was demonstrated by automobiles submitted for emissions testing by Auditor General staff. Two vehicles provided and tuned by Pioneer Ford were submitted for five series of tests. Each series consisted of emissions adjustments and a pretest reading by an experienced mechanic on a State-inspected analyzer, two or three emissions tests at State stations in Maricopa County and a post-test reading at Pioneer Ford. In three of the series, the vehicle either passed or failed all tests as expected. In two of the series, one of the three State station tests showed unexpected results. A vehicle set to fail passed one of three inspections and a vehicle set to pass failed one of three inspections. Table 2 presents data for the two series which produced unexpected results.

TABLE 2

CARBON MONOXIDE EMISSIONS TEST SERIES RESULTS

	1980 Ford Fairmont			1981 Ford Fairmont		
Carbon Monoxide Standard	2.20% or Less			1.50% or Less		
	Time	Reading	Result	Time	Reading	Result
Pretest	5:00 pm*	6.80%	Fail	9:05 am	0.01%	Pass
Station tests	9:20 am	8.41	Fail	10:05 am	0.02	Pass
	10:07 am	6.34	Fail	10:39 am	0.02	Pass
	10:41 am	1.54	Pass	10:58 am	3.03	Fail
Post-test	12:45 pm	6.20	Fail	11:40 am	0.01	Pass

* Tested on previous evening

Vehicle Exclusions - The effect of excluding vehicles from the emissions inspection program over the past five years is uncertain. Air quality planners project that excluding older vehicles will increase in significance in future years. In addition, out-of-state vehicles driven by winter residents may contribute to the lack of program effectiveness.

Although older vehicles have not been the source of a significant proportion of carbon monoxide emissions in recent years, planners expect their significance to increase. The VEI Program excludes vehicles more than 13 years old. Because emissions from older vehicles are likely to be more significant in future years, the Maricopa Association of Governments is recommending that the VEI Program include all vehicles manufactured after 1969.

Out-of-state vehicles driven by visitors for extended periods during the winter months may also contribute to carbon monoxide problems in Phoenix and Tucson. Their impact is unknown and no information on their numbers or emissions is available. However, both traffic volume and the number of carbon monoxide violations increase in Maricopa County during the winter months when these visitors are in residence.

Improving VEI Program Effectiveness

The potential for improving VEI Program effectiveness by addressing the above problems is not certain. Bringing older vehicles into the program can be readily achieved although the impact is unclear. Other problems may be impossible to address effectively.

MAG and DHS have taken action to include older vehicles. Expanding the program to include older vehicles would result in a projected drop in emissions. However, the relative impact of this action depends on continuation of Federal new car emissions standards which are under review by the U.S. Congress. No data indicate the extent to which these changes will affect carbon monoxide levels.

Problems such as the variability of automotive emissions, circumvention and the impact of winter visitors will be more difficult, if not impossible, to address. The variability of vehicle emissions means that even vehicles passing the inspection may continue to contribute to air pollution. Effectively reducing circumvention would require stricter control of motorist and mechanic behavior. Violations may be virtually impossible to detect without use of unacceptable, intrusive methods. Finally, controlling emissions from out-of-state vehicles would be impractical.

CONCLUSION

The Arizona Vehicle Emissions Inspection Program has not been effective in reducing carbon monoxide levels in Phoenix and Tucson and may not be a reliable strategy for meeting the 1987 Federal air quality standards. Although several problems and factors associated with the Program may account for its ineffectiveness, we do not know for certain why it is ineffective. Thus, the effect of any policy or program changes in these areas is also unknown. As a result, Arizona needs to develop alternatives to the VEI Program for reducing automotive pollutants.

RECOMMENDATIONS

1. Unless additional research identifies feasible alternatives, if any, for an effective VEI Program, DHS should operate the Program only as necessary to avoid imposition of Federal sanctions. Any program changes should be rigorously and independently evaluated before being implemented.

2. In the absence of an effective VEI Program, DHS and Arizona's regional air quality planning agencies should develop and evaluate alternate strategies for reducing motor vehicle pollution.

3. The Legislature should consider petitioning the U.S. Congress to review the appropriateness of the Federal requirements for emissions testing in light of the findings in this report and the alternate strategies for reducing air pollution developed by DHS.

FINDING II

THE VEHICLE EMISSIONS INSPECTION PROGRAM BOUNDARIES DO NOT EFFECTIVELY ADDRESS AUTOMOTIVE AIR POLLUTION PROBLEMS IN PIMA COUNTY.

The Vehicle Emissions Inspection (VEI) Program boundaries are not appropriate in the Tucson area. The current boundaries are larger than necessary to control the existing carbon monoxide problem. However, the program boundaries do not include all areas which may be required to control future ozone pollution.

Establishment of Nonattainment Area Boundaries

The VEI Program covers the Pima County nonattainment area (NAA) for carbon monoxide (CO). The Clean Air Act requires that each state designate such areas wherever air quality does not meet Federal standards. NAAs must encompass the areas where violations occur and be sufficiently large to adequately address the source of the problem. Tucson exceeds air quality standards for carbon monoxide and, as a result, comes under these requirements.

The Pima Association of Governments (PAG), working in conjunction with the Department of Health Services (DHS), designated the NAA in 1978 based on the concept of an air shed. The high mountain ridges surrounding Tucson define an air shed which is the basis for the Pima County NAA. An air shed is an area in which air pollutants mix but cannot escape in the absence of a major weather front. Thus, the air shed is intended to provide the basis for area-wide air quality management.

The one exception to the air shed approach is the Town of Marana which requested to be excluded from the NAA on the grounds that the town had no air quality violations. Citing reports that Marana showed no ozone violations, PAG approved this exclusion.* DHS concurred and the Environmental Protection Agency (EPA) approved the nonattainment area boundaries in 1979. EPA based its approval on the fact that 1) the carbon monoxide NAA encompassed the areas where carbon monoxide violations occurred and 2) no data supported including Marana in the ozone area.

Pima NAA Boundaries Are Larger Than Needed to Control Carbon Monoxide in Tucson

The Pima NAA is larger than necessary to address carbon monoxide problems in Tucson. Carbon monoxide pollution is relatively isolated in Tucson, caused largely by commuter traffic. However, the NAA encompasses an extensive area, almost half of which does not contribute to Tucson's commuter traffic.

Carbon monoxide pollution is limited to a small area of Tucson. Violation of Federal carbon monoxide standards occurred at only one of the three monitors in the city during 1980. Two monitors recorded violations in 1981. These monitors are located at intersections which experience the heaviest traffic flow. Thus, the carbon monoxide problem in Tucson appears to be largely a matter of emissions from commuter traffic in specific places rather than a general problem throughout the air shed.

Commuter traffic travels into Tucson from only about half of the Pima County NAA--those areas within approximately 10 to 20 miles of the central city (Figure 1). In contrast, the Maricopa NAA consists almost entirely of areas which contribute commuting traffic to Phoenix. Between 72 and 83 percent of the vehicle miles traveled in Pima County occurs within the urbanized portion of the county. Thus, to the extent that any program is effective, control of carbon monoxide pollution in Tucson probably can be achieved with smaller VEI Program boundaries than are presently in effect.

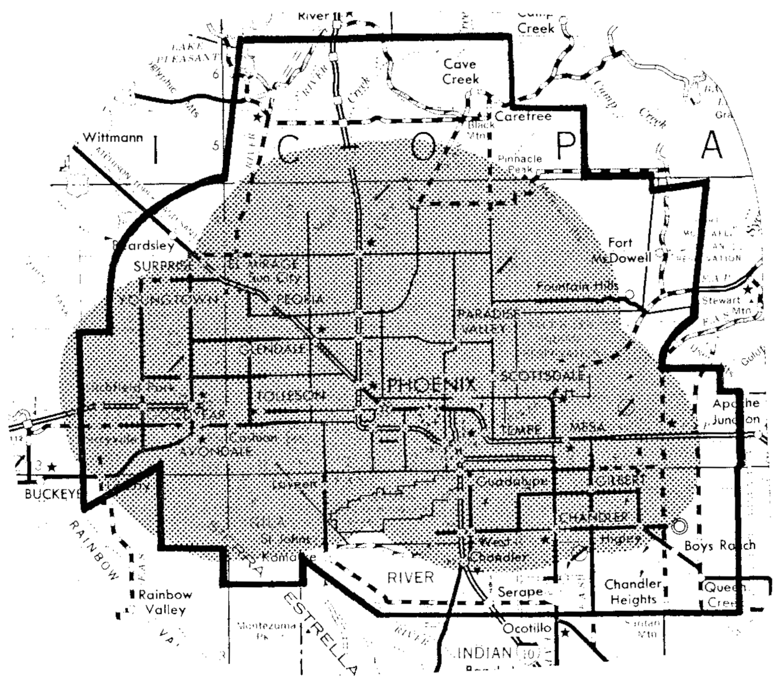
* At the time the nonattainment area was defined, Tucson exceeded standards for both ozone and carbon monoxide.

FIGURE 1

COMMUTING PATTERNS IN ARIZONA CARBON MONOXIDE NONATTAINMENT AREAS

Pima County Nonattainment Area

Maricopa County Nonattainment Area



Areas generating major commuter traffic

Our analysis indicates that outlying areas can be excluded from the NAA without affecting carbon monoxide pollution in Tucson. Thus, excluding the Town of Marana is appropriate. To be consistent, however, other outlying areas should be removed from the carbon monoxide nonattainment area if they do not generate commuter traffic. Two areas, Green Valley and Oro Valley, requested exclusion in 1979. Although PAG directed its staff to investigate such action, the data required by EPA were not available to support exclusion. However, EPA now only requires that an area demonstrate by any means that no problem exists. Our analysis suggests that Green Valley contributes little to the carbon monoxide problem in Tucson. Consequently, excluding Green Valley would probably not adversely affect the problem. Oro Valley, on the other hand, lies within the commuting radius and, thus, contributes to Tucson's carbon monoxide problem.

Larger Boundaries May Be Required to Effectively Control Ozone

If effective control of ozone can be accomplished through the VEI Program, it may require an area-wide VEI Program. Because ozone occurs and spreads throughout the air basin, the more circumscribed area which would effectively control carbon monoxide may not alleviate ozone pollution. If ozone control requires a larger area, then no outlying areas should be excluded from the emissions inspection program. However, effective control of ozone through the existing VEI Program may not be feasible.

Although Pima County is not currently a nonattainment area for ozone, health officials indicate that the problem is growing and may require attention in coming years. Unlike carbon monoxide, ozone is an area-wide problem because it forms in the upper atmosphere and spreads with the prevailing winds. Since the Pima County NAA is a well-defined air shed, this mass of air may stay in the area for several days until a major weather front pushes it out. To the extent that any emissions inspection program is effective, control of ozone may require including all vehicles* in the air shed in the VEI Program.

* All vehicles means those required by law to undergo emissions tests. Current law requires gasoline-powered vehicles less than 14 years old to be tested.

The extent to which effective control of ozone requires an area larger than needed for carbon monoxide pollution control depends on the extent to which nonurban areas contribute the precursors of ozone--hydrocarbons (HC) and nitrous oxides (NO_x)--to the atmosphere. These outlying areas account for between approximately 17 and 28 percent of vehicle miles traveled in the NAA. PAG and DHS need to analyze carefully the nature of ozone pollution in the area to determine if nonurban vehicle travel contributes significantly to the ozone problem in the area. If nonurban vehicles are significant contributors, then the NAA boundary should be changed to include Marana. Although no one area is likely to show significant impact on air quality, a finding that the overall nonurban contribution is significant would provide no basis for excluding any outlying area. If the nonurban contribution is not significant, then PAG should consider reducing the size of the NAA to exclude all outlying areas.

In all decisions on the NAA boundaries, however, DHS and PAG should consider that effective control of ozone may not be possible through the existing VEI Program, regardless of its boundaries, because of the following reasons:

1. Available data indicate the VEI Program has not had an impact on carbon monoxide levels (see Finding I). Thus, factors which limit the Program's effect on carbon monoxide may also affect its impact on ozone.
2. Some health officials indicate that nitrous oxides, which are not controlled by the VEI Program, may have more impact on ozone than hydrocarbons. Since the VEI Program has been in effect hydrocarbon levels have decreased, but nitrous oxides and ozone levels have risen at the rate of 4 to 5 percent annually in Pima County.
3. Few vehicles fail emissions tests for hydrocarbon levels. Only 20 percent of 200,000 test failures involve hydrocarbons. The remaining failures involve carbon monoxide only.

CONCLUSION

The Pima County NAA boundaries do not provide an appropriate basis for addressing current and potential air quality problems caused by motor vehicle emissions in the Tucson area. Depending on the nature of the pollution problem, the boundaries are either too large or too small for an effective VEI Program. However, other factors suggest that the existing VEI Program may not effectively reduce ozone, regardless of its boundaries.

RECOMMENDATIONS

1. The Pima Association of Governments and Department of Health Services should analyze the nature and sources of automotive pollution in the Tucson air shed to determine the most effective area for controlling current and anticipated air quality problems. In evaluating the optimum size for the nonattainment area, PAG and DHS should also consider the ability of the existing VEI Program to effectively control ozone.

2. Based on the results of the analysis recommended above, PAG and DHS should adjust the nonattainment area boundaries to either a) include all areas which contribute to Tucson's current and potential automotive air pollution or b) exclude those areas which do not.

FINDING III

THE BUREAU OF VEHICULAR EMISSIONS INSPECTION CAN SAVE $121,700 IN STATE AND FEDERAL FUNDS ANNUALLY BY ELIMINATING AIR/FUEL WAIVERS.

The air/fuel waiver provision of Arizona's Vehicle Emissions Inspection Program is not cost effective. Although statutes require the Department of Health Services (DHS), Bureau of Vehicular Emissions Inspection (BVEI or Bureau) to issue air/fuel waivers, less than 2 percent of the motorists use this service. The Bureau could save an estimated $89,500 in State funds and $32,300 in Federal funds annually by eliminating the air/fuel waiver program. While the Bureau maintains that such a change would limit its authority to regulate the repair industry, there is no indication that current statutory provisions intended to give the Bureau such broad regulatory powers.

Statutory Provisions for Air/Fuel Waivers

As a service to motorists, State law requires the Bureau to issue air/fuel waivers for vehicles failing the emissions inspection test for carbon monoxide (CO) only. Arizona Revised Statutes (A.R.S.) §36-1772.E. states:

> "Vehicles which fail the curb idle test solely because of air gas mixture are entitled to a certificate of waiver upon correction of the problem, after having furnished satisfactory evidence of correction to the director on a form to be prescribed by the director."

To comply with A.R.S. §36-1772.E., the Bureau requires air/fuel waiver applicants to repair the vehicles at State-approved facilities and to have the waiver form signed by a State-certified mechanic. To participate in the program, repair facilities must register their analyzers with the Bureau and employ at least one automotive technician certified by the Bureau. The Bureau inspects registered analyzers periodically and trains mechanics prior to issuing certification.

Carbon monoxide emissions or faulty air/fuel mixture is the most common cause for failing emissions inspection. The Bureau inspects over 1.2 million vehicles annually of which approximately 160,000 fail the test solely due to carbon monoxide problems.* These vehicle owners are eligible to seek an air/fuel waiver.

Most Eligible Motorists Do Not Use the Air/Fuel Waiver Program

Although most motorists whose vehicles fail the emissions tests are eligible to apply for an air/fuel waiver, the overwhelming majority simply return their vehicles to a State station for a free retest.** Compared to the retest, obtaining an air/fuel waiver is a complicated, lengthy procedure.

Only 2,400 (or 1.5 percent) of the 160,000 motorists eligible to apply for air/fuel waivers utilized the waiver procedure during fiscal year 1981-82. Instead the vast majority (up to 98 percent) of eligible motorists returned to the inspection station for a free retest. According to BVEI, the low utilization in fiscal year 1981-82 is typical of most years.

The air/fuel waiver procedure is more complicated and costly than the free retest. The free retest is simple and convenient. Motorists have repairs made at facilities of their own choice and then take the free retest at the inspection stations. To obtain a waiver, however, the Bureau requires applicants to 1) repair the vehicles at State-approved facilities, 2) have completed forms signed by a State-certified mechanic, 3) pay $1 for the certificate in addition to the $5.44 paid for the initial test, and 4) mail the forms to the Bureau's central office. Motorists may not have time to mail the forms requesting the waiver to the Bureau especially if it is close to the vehicle's re-registration deadline. Thus, providing air/fuel waivers as an alternative to the free retest at the inspection stations has only marginal benefits for motorists.

---

* An additional 40,000 vehicles fail both the CO and hydrocarbon (HC) tests.

** Under the terms of the current contract with Hamilton Test Systems, a motorist whose vehicle failed the first test can have a retest at no additional charge.

Air/Fuel Waiver Program Is Costly

The cost of administering the air/fuel waiver program is substantial and not supported by fees charged. The estimated annual cost for processing 2,400 air/fuel waivers is $147,000, or 19 percent of the Bureau's budget, while the income from the $1 fee charged for each waiver is about $2,400. The waivers do not generate sufficient income to reasonably cover the administrative costs. Thus, $144,600 of the waiver service's administrative expenses are subsidized by other funding sources.* Increasing the fee charged for the waiver, however, is not a feasible solution since the current cost is $61 per waiver.

The air/fuel waiver service's administrative costs include inspection of registered analyzers in repair facilities, yearly certification of mechanics, training, clerical, travel and equipment expenses. Table 3 gives the cost breakdown for work activities related to processing 2,400 air/fuel waivers.

TABLE 3

AIR/FUEL WAIVERS PROCESSING COST,
FISCAL YEAR 1982-83

ITEM	ANNUAL COST
Personnel:	
Inspectors	$ 59,600
Trainers	22,300
Administrative (Asst. Bureau chief)	5,600
Clerical	6,700
Employee related	22,600
Operating overhead	18,900
Vehicles (travel)	6,500
Calibrating gas	4,800
Total	$147,000

* Of the $144,600, $112,400 are General Fund appropriations and $32,200 are Federal funds.

Eliminating the Air/Fuel Waiver Program Would Save $121,700

BVEI could save approximately $89,500 in State funds and $32,200 in Federal funds if State law did not require air/fuel waivers. Without this requirement the Bureau would not need to maintain surveillance over private repair facilities.

Since the waiver program is of marginal benefit to motorists, the Bureau could save approximately $121,700 in State and Federal funds annually by eliminating the program. Currently, the Bureau has 700 registered analyzers in repair facilities and 800 certified mechanics employed by these facilities.* The Bureau inspects analyzers and trains mechanics for certification. The Bureau can eliminate the requirements for registered analyzers and certified mechanics in repair facilities if the statutes did not require air/fuel waivers. As Table 4 indicates, this would save $121,700 in State and Federal funds or 16 percent of the Bureau's fiscal year 1982-83 budget.

---

* The Bureau also has approximately 400 registered analyzers for fleet operations and certifies 765 mechanics who inspect fleet vehicles. The cost savings projection is for air/fuel waivers procedures only and does not reduce inspection and training for fleet operations and for the Hamilton inspection stations.

TABLE 4

PROJECTED ANNUAL SAVINGS
BY ELIMINATING AIR/FUEL WAIVERS*

ITEM	FTEs	AMOUNT OF SAVINGS
State funds:		
Personal Services -		
Inspectors	2.5	$ 50,200
Clerical	.5	5,800
Employee-related expenditures		13,400
Operating overhead expenses		11,200
Vehicles (travel)		6,500
Calibrating gas		4,800
Subtotal State	3.0	91,900
Less current income		2,400
Net State savings		89,500
Federal Funds:		
Personal Services -		
Trainer	1.0	22,300
Employee-related		5,400
Operating overhead		4,500
Subtotal Federal	1.0	32,200
Total: All Funds	4.0	$121,700

* Based on the 1982-83 budget

By eliminating the air/fuel waiver program the Bureau could reduce its staff by four positions: 2.5 inspectors, one trainer and .5 clerical positions. The four FTEs' salaries and related employee and operating costs amounts to a $112,800 annual reduction. In addition, the Bureau would save $11,300 in travel and calibrating gas expenses. The total reduction would be $124,100. Since the Bureau will not receive the $2,400 annual income from the waivers fees, the net savings are $121,700 in State and Federal funds annually or 16 percent of the Bureau's 1982-83 budget.

These savings do not include the cost of staff members who devote only a small portion of their time to the air/fuel waiver function while the $147,000 is based on total processing cost. Thus, the Bureau can implement the savings without jeopardizing other functions. Moreover, the State could save at least $89,500 of General Fund monies even if the federally sponsored training program is not reduced in scope.

Bureau May Be Exceeding
Legislative Intent

Although air/fuel waivers are not cost effective, the Bureau has not requested the Legislature to abolish the statutory requirement. Rather, the Bureau utilizes the requirement as a tool to regulate the automobile repair industry in general. The Bureau chief stated that the purpose of certification is to ensure that mechanics are trained to properly repair all vehicles rather than to sign off on the few air/fuel waivers. However, there is no indication that the provisions of A.R.S. §36-1772.E. were meant to give the Bureau such broad regulatory power over the repair industry.

CONCLUSION

The Bureau could save $89,500 in State funds and $32,200 in Federal funds annually if State law did not require air/fuel waivers. Few motorists use the waiver process, which is costly. Eliminating the air/fuel waivers will not inconvenience the public. Currently over 98 percent of those eligible for air/fuel waivers simply return to the inspection station for a free retest.

RECOMMENDATIONS

1. The Legislature should consider amending A.R.S. §36-1772.E. to delete the requirement for an air/fuel waiver program.

2. If the Legislature deletes the air/fuel waiver requirement, BVEI should:

   a. Delete its administrative rules and regulations relating to air/fuel waivers and

   b. Review its policies, procedures and work activities and eliminate those functions related to air/fuel waivers. The Bureau should revise its budget and reduce personnel to reflect this program change.

FINDING IV

ALTHOUGH THE BUREAU OF VEHICULAR EMISSIONS INSPECTION AND THE CONTRACTOR HAVE INCREASED THE FREQUENCY OF EQUIPMENT CHECKS, ADDITIONAL CONTROLS ARE NEEDED TO ASSURE ACCURATE AND RELIABLE EMISSIONS TESTING.

Both the Bureau of Vehicular Emissions Inspection (BVEI) and Hamilton Test Systems (HTS) have increased the frequency of equipment field audits at inspection stations; however, additional controls are needed to ensure accurate and reliable emissions testing. Specifically, BVEI needs to 1) improve management control over its field audit program and 2) require HTS by contract to conduct inspection station field audits. In addition, BVEI needs to conduct more timely audits of fleet and registered analyzers in private facilities.

Need for Frequent Emissions Equipment Checks

BVEI and Hamilton Test Systems assure the accuracy and reliability of emissions testing equipment by conducting periodic field audits at the vehicle emissions inspection stations. A field audit is an inspection which uses a blend of gases of known proportions to test the ability of equipment to sample and analyze emissions accurately. In effect, a field audit approximates actual emissions test conditions.

HTS equipment is among the best available for emissions testing, but even the most reliable equipment needs to be checked frequently. Administrative Rule R-9-3-1025.A. requires BVEI to conduct equipment checks twice each month at each of the 30 permanent test lanes in Phoenix and Tucson.*

* The mobile inspection station in Green Valley is checked once each month.

In addition, HTS has established its own policies calling for two additional BVEI-type field audits per month. Thus, a total of four field audits should be conducted monthly at each test lane. This frequency of auditing is comparable to the minimum requirements in other states with centralized programs. For example, California, which operates a centralized emissions inspection program using Hamilton equipment, reported that field inspectors check stations three to four times monthly.

Field auditing is not the sole quality assurance procedure. The DHS-HTS contract requires HTS to calibrate analyzers periodically. Calibration differs from field auditing in that separate test gases are used for each component of the system (e.g., carbon monoxide, hydrocarbons). Although this procedure does not test the entire system's ability to sample and analyze the combination of gases in vehicle exhausts, calibration does ensure that each component accurately records gas concentrations. Auditor General review of a sample of HTS records for fiscal year 1981-82 shows the contractor to be in compliance with calibration requirements of the contract.

BVEI and HTS Did Not Conduct Appropriate Number of Audits

BVEI did not conduct the number of inspection station field audits required by regulation during fiscal year 1981-82. Further, HTS did not conduct all field audits required by internal company policy.

BVEI conducted only 56 percent of required field audits in fiscal year 1981-82. The Bureau did not conduct required field audits from July 1981 through March 1982. No field audits were conducted at all during February. These field audits were not conducted because the employee responsible for this function failed to carry out his duties.* BVEI management was unaware of this deficiency until a quality assurance techni-

* The employee has since been terminated from service.

cian, who was not responsible for inspection station field audits, informed the Assistant Bureau Chief of the problem. The Bureau began complying with its regulation in April 1982 and since then has conducted virtually all required field audits of inspection stations.

Hamilton Test Systems also did not conduct all field audits required by internal company policies in fiscal year 1981-82. Until March 1982, HTS informal policies called for one monthly field audit for each test lane. In April 1982, HTS field audit policies were revised to require two such audits per month. HTS failed to conduct the audits because they were uncertain as to the specific procedures to be used. DHS had not supplied HTS with needed information for testing and HTS was in the process of converting from one equipment system to another. In addition, some test results were not being recorded. Like DHS, however, Hamilton has complied with its policies since April 1982.

TABLE 5

BVEI AND HTS STATE STATION AUDITS,
FISCAL YEAR 1981-82

	DHS Lanes Audited		HTS Lanes Audited	
Month	Number	Percent of Required	Number	Percent of Required
1981				
July	32	52%	0	0%
August	33	54	0	0
September	40	66	0	0
October	24	39	7	23
November	18	30	0	0
December	7	11	21	68
1982				
January	42	69	21	68
February	0	0	29	94
March	30	49	19	61
April	62	102	47	77
May	60	98	61	100
June	62	102	38	62
Totals	410	56%	243	53%

Additional Controls Are Needed

Although BVEI and HTS have increased the frequency of their equipment audits at inspection stations, additional controls are needed. BVEI needs to 1) institute a formal management reporting system and 2) require HTS by contract to conduct field audits.

The Bureau has not yet established a formal management reporting system to inform top management of field audit performance. The Bureau Chief does not know on a timely basis 1) whether all required field audits have been conducted and 2) the nature and extent of test lane equipment failure. Thus, there is no assurance that the problem which occurred in fiscal year 1981-82 will not recur.

While Hamilton Test Systems now conducts two field audits per test lane per month, the contractor is not required to do so. The DHS contract with Hamilton does not require regular field audits of inspection stations nor reporting of results to DHS. Requiring HTS by contract to conduct two monthly audits and to report results to DHS would ensure that a total of four field audits are performed on each test lane per month. As noted earlier, this frequency of field auditing is comparable to the minimum requirements in other states which have inspection programs similar to Arizona's.

BVEI Did Not Conduct Timely Audits of Analyzers in Private Facilities

BVEI also did not conduct timely inspections of fleet and registered analyzers in private facilities in fiscal year 1981-82. As a result, faulty analyzers may have remained in service longer than necessary.

Although BVEI rules and policies require the Bureau to inspect fleet and registered analyzers every three months, approximately 37 percent of the inspections exceeded the 90-day requirement in fiscal year 1981-82. As shown in Table 6, fleet and registered analyzer inspections which exceeded 90 days averaged 117 and 124 days, respectively, and ranged as high as 264 days--nearly 9 months. These inspections were not conducted because employees failed to carry out this function properly and management did not take action to address the deficiency.

Because inspections have not been timely, faulty analyzers may have remained in service longer than necessary. One-fourth of the analyzers in our sample failed the quality assurance test in 1981-82. If tested as required, faulty analyzers would be taken out of service and adjusted to meet BVEI accuracy standards. This would provide more assurance that vehicles inspected at fleet stations and private garages are properly tested. At least 10 percent of the vehicles inspected annually pass through these facilities.

TABLE 6

TIMELINESS OF QUALITY ASSURANCE INSPECTIONS
FOR SAMPLE OF FLEET AND REGISTERED ANALYZERS,
FISCAL YEAR 1981-82

	Total	Within 90 Days		Greater Than 90 Days		Average Days In Excess of 90
		Number	Percent	Number	Percent	
Fleet Stations	167	116	69%	51	31%	117
Registered Analyzer Stations	152	84	55	68	45	124
Total	319	200	63%	119	37%	121

CONCLUSION

Both BVEI and Hamilton Test Systems have increased the frequency of equipment field audits at inspection stations. However, additional controls are needed to assure accurate and reliable emissions testing. DHS also needs to conduct more timely inspections of fleet and registered analyzers in private facilities.

RECOMMENDATIONS

1. BVEI should establish a system of twice monthly reports to the Bureau Chief, noting the number of State inspection station lanes field audited by both Bureau and Hamilton Test Systems personnel and the audit results.

2. BVEI should require by contract that Hamilton Test Systems conduct two field audits per month and report the results of the audits to the BVEI Quality Assurance Manager.

3. The BVEI Quality Assurance Manager should ensure that inspectors conduct timely quality assurance inspections of fleet and registered analyzers as required by Bureau regulations and policy. Results of these inspections should be reported at least monthly to the Bureau Chief.

OTHER PERTINENT INFORMATION

During the audit, other pertinent information was developed regarding 1) vehicle inspection fee payments and collections and 2) the emissions test equipment and procedures used in the program.

Vehicle Inspection Fee Payments and Collections

Since January 1981 the Bureau of Vehicular Emissions Inspection (BVEI) has paid Hamilton Test Systems (HTS) $575,000 less than the amount billed by the contractor. This shortfall in payments occurs for two reasons: 1) there is a significant time lag between emissions tests and remittance of funds by the county assessors and the Motor Vehicle Division (MVD), and 2) fees are not collected for all tests performed. Although BVEI has proposed ways to address the cash flow problem, the problem of collecting for all tests conducted has not been addressed.

Collection Process - Current statutes provide for emission test fees to be collected by the county assessors or the Motor Vehicle Division of the Department of Transportation at the time of vehicle registration.* These funds are remitted on a monthly basis to BVEI which then pays HTS for tests performed. Vehicle owners must pay $5.44 for the first inspection and the same amount for the third, fifth and seventh subsequent tests. Each of these paid tests includes a free retest, if needed.

Shortfall in Payments - As shown in Table 7, BVEI has paid HTS approximately $575,000 less than the total amount billed by the contractor since January 1981. This shortfall in payments to HTS represents the difference between the value of emission tests performed by HTS ($11.2 million) and the amount BVEI received from the county assessors and MVD ($10.6 million) for the period January 1981 through September 1982.

---

* MVD has taken over the vehicle registration function in Pima County. In addition, some vehicle owners register their vehicles at the MVD central headquarters in Phoenix.

TABLE 7

BILLINGS AND PAYMENTS FOR EMISSIONS INSPECTIONS, JANUARY 1981 THROUGH SEPTEMBER 1982

Period	Amount Billed	Payments Amount	Payments Percent of Billing	Unpaid Balance Amount	Unpaid Balance Percent of Billing
January - December 1981	$ 6,252,344	$ 5,783,680	92.5%	$468,664	7.5%
January - September 1982	4,915,808	4,809,115	97.8	106,693	2.2
Total	$11,168,152	$10,592,795	94.8%	$575,357	5.2%

The shortfall grew substantially during the first year of the current contract to approximately $469,000. The shortfall grew more slowly in 1982--increasing about $107,000. BVEI paid HTS 98 percent of the amount between January and September 1982 compared to 93 percent in the first year of the new contract. Total payments since January 1981 are 95 percent of total billings.

Causes of Shortfall - The shortfall in payments to HTS appears to occur for two reasons. First, there is a significant time lag between emissions tests and remittance of funds to BVEI. Second, fees are not collected for all tests performed by HTS.

The amount of time that may elapse between an emissions inspection and the receipt of fee for that inspection appears to be a major contributor to the difference between billings and payments. Although HTS bills for inspections at the end of each month, BVEI may not receive payment for as long as 135 days after a vehicle is inspected. As a result, BVEI receipts and payments have been less than HTS billings in 13 of the 21 months since the beginning of the current contract.

The time lag between emissions testing and remittance of funds is permitted by DHS rules and regulations. Regulations require emissions inspections 90 days prior to the registration renewal date. Regulations also stipulate that county assessors or MVD remits inspection fees collected at the time of registration by the 15th of the month following collection. Thus, collection procedures create a "pipeline effect" which delays receipt of inspection fees up to 135 days. If vehicles are registered late, this time lag can be even greater.

Results of a recent DHS audit confirmed that the time lag exists between testing and remittance of funds by county assessors and MVD to BVEI. Fees for 41 percent of emissions tests conducted in October 1981 were not collected until the following or subsequent months. In addition, auditors found that some funds were held for up to six weeks after collection. However, such retention is permitted by BVEI regulations and the audit found no evidence that county assessors and MVD were not properly handling emissions fees or were not complying with applicable regulations.

The shortfall in payments also occurs because fees are not collected for all tests performed. According to BVEI analyses, approximately 2.3 percent of HTS billings since July 1981 include fees for vehicles which do not qualify for registration. These are vehicles which failed the initial test and have not passed or were waivered at the time HTS billed for the initial test. The DHS audit identified 3.9 percent of October 1981 tests which were still uncollected in March 1982.

BVEI also feels that some motorists do not pay as required for all tests performed on their vehicles. These are motorists who must retest their vehicles a third, fifth or even seventh time and should pay fees for these odd-numbered tests in addition to the initial test fee (even-numbered tests are free retests). The BVEI-HTS contract requires BVEI to collect fees for more than one test for a motorist only when HTS notifies the assessor or MVD that additional fees are required.

Although HTS has implemented some procedures in an attempt to ensure that the inspection certificates reflect all charges for multiple tests, motorists can easily circumvent these procedures. Station personnel punch entry documents (renewal certificates or titles for example) to show that the vehicle has been tested. However, motorists can submit vehicles for additional tests and avoid charges by switching documents. For example, Auditor General staff submitted two vehicles for eleven paid inspections. Only one HTS inspector identified prior tests that should have been added to the inspection certificate.

Alternatives for Reducing Shortfall - Although BVEI has proposed ways to address the cash flow problem which contributes to the shortfall in payments, the problem of collecting fees for all tests conducted has not been addressed.

BVEI has identified two solutions to the collections cash flow problem. First, fees could be collected by Hamilton at the time of emissions testing. This method of collections was used prior to January 1981, but was inconvenient for the public. Hamilton accepted only cash payment and personal checks were not accepted. Collecting $5.44 in cash or any other odd amount, at the time of testing, would represent an even greater inconvenience if Hamilton continued to require cash payment.

Second, BVEI has proposed that county assessors and MVD remit emissions testing fees on a daily basis. This would improve cash flow and is consistent with a similar Auditor General recommendation regarding remittance of other vehicle registration fees by the counties (see Auditor General Report number 82-3).

Neither DHS nor HTS have developed procedures to ensure collection of fees for all tests performed. This would require improving controls at inspection stations to identify motorists who receive multiple inspections. Use of single entry documents or unique vehicle identification numbers may be needed to allow HTS to notify county assessors and MVD of the proper amount to be collected.

Emissions Test Equipment and Procedures

An important part of the Vehicle Emissions Inspection (VEI) Program audit was a review of test reliability and accuracy of records. Finding IV addresses the major reliability issue, quality assurance. Our research indicates that the equipment used to inspect vehicle emissions is as good as any available. Moreover, test procedures and data handling are, with two exceptions, adequate to ensure reliable testing and billing.

Equipment - Competitors and a major client alike told Auditor General staff that the Arizona VEI Program contractor, Hamilton Test Systems (HTS), is a leader in the emissions testing field. An official of the California emissions program described HTS as a pioneer in the field. A representative of a company which competed against HTS for the Arizona contract said that he knew of no better equipment and that his company used HTS as a standard when making its bid. Another competitor described its equipment as superior but added that it designed its equipment to meet California standards. HTS designed the original equipment for that state's program. Finally, a representative of one of the major automobile manufacturers described HTS as a leader in both developing testing equipment and test procedures. Thus, to the extent that HTS maintains its equipment, the equipment is as good as any available. As noted in Finding IV, however, the Department of Health Services (DHS) has not adequately inspected HTS equipment in the past to ensure that it functions up to its capability.

Test Procedures - The VEI Program test equipment is fully automated to provide accurate testing and recording of test data. All that an inspector must do is enter and verify the vehicle identification data, insert the tailpipe probe and the system automatically samples the exhaust, analyzes the gas and reports the results. According to HTS personnel, station inspectors cannot affect the conduct of the test. Moreover, the system is programmed to abort testing if major problems occur that would result in erroneous test results.

However, HTS personnel do not always enter the proper engine size information (number of engine cylinders) into the system, resulting in the possibility that vehicles will be measured on the wrong standards. In 6 of 14 instances where Auditor General staff submitted vehicles for emissions inspection, station personnel did not ask the number of cylinders or check under the hood. The entry document did not provide this information. In five of the six instances, station personnel incorrectly assumed the vehicle had six rather than four cylinders. This error reduced allowable emissions from 2.5 to 2.2 percent--a 12 percent reduction.

<u>Billing Procedures</u> - Hamilton Test Systems' control over the billing process is sufficiently adequate to ensure that bills submitted to the Department of Health Services are accurate. Although some weaknesses exhist in the recording and processing of data which results in the HTS bill, the weaknesses do not significantly affect billing accuracy.

Auditor General EDP staff reviewed the process by which HTS records and handles test information. This process begins when a vehicle enters a test station and ends with the preparation of a bill that HTS submits to the Department. The process is reliable and accurate. Direct recording of test information on the station computer, processing of all records with no deletions and separation of programming and systems functions are strong control features in the process.

Although EDP staff noted several weaknesses, only two appeared significant. First, station inspectors sometimes fail to accurately enter the proper number of engine cylinders. As noted above, this problem occurred with vehicles submitted by Auditor General staff for inspection and can result in erroneous results. The second weakness is that HTS can change vehicle information. This produces the potential for false billing. However, a spot check of billings did not indicate that such false billings had occurred.

GLOSSARY

AIR/FUEL WAIVER: Waiver issued to vehicles failing the emissions test for carbon monoxide only. A motorist may obtain an air/fuel waiver upon correcting the problem and furnishing satisfactory evidence of correction to the Bureau of Vehicular Emissions Inspection.

AIR SHED: Area in which air pollutants mix but cannot escape in the absence of a major weather front. An air shed provides the basis for area-wide air quality management.

AMBIENT AIR: The air around us.

BUREAU OF VEHICULAR EMISSIONS INSPECTION: The Bureau within the Department of Health Services that administers the Vehicle Emissions Inspection Program.

CARBON MONOXIDE (CO): Odorless, colorless gas emitted by motor vehicles. Carbon monoxide deprives the blood of oxygen. A large concentration can kill; a smaller amount can cause dizziness, fatigue, headaches and slower driving reaction time.

FIELD AUDIT: An inspection of an emissions analyzer using a blend of test gases in which the relative proportion of each gas is known. Field audits are conducted to ensure accuracy and reliability of emissions analyzers.

HAMILTON TEST SYSTEMS (HTS): A private company which operates the State emissions inspection stations under a contract with the Department of Health Services.

HYDROCARBONS (HC): Atmospheric pollutant caused by unburned gasoline in engine exhaust by evaporation losses from petroleum storage and handling and by evaporation from organic solvents. Hydrocarbons are one component of the pollutant ozone.

MARICOPA ASSOCIATION OF GOVERNMENTS (MAG): Regional Planning Organization responsible for developing plans for reducing automotive air pollution in Maricopa County.

NITROUS OXIDES (NO_x): Atmospheric pollutant produced where fuel is burned at high temperatures, such as in motor vehicles or power plants. This pollutant is a component of ozone.

NONATTAINMENT AREA (NAA): An area where the air quality standards are violated for a given pollutant. In Arizona, NAA boundaries for carbon monoxide are also the boundaries in which the emissions inspection program is in effect.

OZONE: Photochemical oxidants which result from reactions in the atmosphere between hydrocarbons and nitrous oxides in the presence of sunlight. Ozone is popularly known as smog.

PIMA ASSOCIATION OF GOVERNMENTS (PAG): Regional Planning Organization responsible for developing plans to reduce automotive air pollution in Pima County.

TIME SERIES ANALYSIS: A statistical method of studying movements in a set of chronologically ordered observations.

TONS PER DAY: Unit of measurement for carbon monoxide and hydrocarbons.

VEHICLE EMISSIONS INSPECTION (VEI) PROGRAM: A program enacted by the Arizona Legislature requiring certain motor vehicles to pass an annual inspection to ensure that their exhaust emissions meet standards established by the Department of Health Services. The standards specify minimum emissions for carbon monoxide and hydrocarbons.

WRITTEN RESPONSES TO THE AUDITOR GENERAL'S REPORT

ARIZONA DEPARTMENT OF HEALTH SERVICES

PIMA ASSOCIATION OF GOVERNMENTS

HAMILTON TEST SYSTEMS, INC.

CONSULTANT'S REBUTTAL TO WRITTEN RESPONSES

ARIZONA DEPARTMENT OF HEALTH SERVICES

Office of the Director

BRUCE BABBITT, Governor
JAMES E. SARN, M.D., M.P.H., Director

February 1, 1983

AUDITOR GENERAL

Mr. Douglas Norton
Auditor General
111 West Monroe Street
Phoenix, Arizona 85007

Dear Mr. Norton:

This is in response to the performance audit report of the Vehicle Emissions Inspection Program conducted pursuant to Senate Bill 1220 enacted by the 34th Legislature in 1980.

In our meeting on January 11, 1983 and in our memorandum of January 17, 1983, we pointed out serious problems with the time series analysis and the resultant conclusions provided by your consultant, the Center for Informative Evaluation. In spite of our well founded objections which are supported by experts in the field of Air Quality, the performance audit report continues to be based upon a study which lacks understanding of air quality complexities and is misleading.

The analysis is flawed to the point of rendering it invalid. The analysis fails to support the conclusion that "the I/M program has had no salutary impact on ambient air quality in Arizona." A more appropriate conclusion of the study might be that the model, data and consultant utilized were unable to detect any effect either positive or negative. More importantly the contractor did not make a sufficient attempt to reconcile his analysis with the significant empirical evidence which shows the program is effective in reducing emissions. The following experts support our position and their comments can be found as Attachments 1, 2 and 3:

> John Trijonis, Ph.D., an environmental scientist with extensive experience in air pollution studies in California. Dr. Trijonis has also served on several National Academy of Sciences air quality committees.
>
> Robin Dennis, Ph.D., Staff Scientist for the National Center for Atmospheric Research, Boulder, Colorado. Dr. Dennis is involved in the development of stochastic models of ambient CO data in Denver.
>
> Mr. Phil Lorang, Chief, Technical Support Staff, U.S.E.P.A., Ann Arbor, Michigan, and author or co-author of several technical papers relating to the effects of vehicle emissions on air quality.

Terry Woodfield, Ph.D., Assistant Professor of Mathematics at Arizona State University, also reviewed the report to evaluate the application of the methodology. His comments, noting impediments to the use of this Report in

Mr. Douglas Norton

Page Two.

decisionmaking, are included as Attachment 4. Our detailed comments relating to the study are contained in Attachment 6. Additional comments regarding Finding I and comments on the Summary, Findings II, III, IV, and Other Pertinent Information are contained in Attachment 5. Attachment 7 is Hamilton Test Systems' response to Finding IV. *

We are in the process of initiating a more detailed study of carbon monoxide air quality in Arizona. This study will be conducted by recognized professionals in the field of air quality. Because of the considerable expert opinion which is in disagreement with your consultant's report, we recommend that you delay the issuance of your report or appropriately modify it until such time as the detailed study is complete.

We appreciate your concern shown in considering our comments on these drafts. If you have any questions concerning our comments, please contact R. Fred Iacobelli, Chief, Bureau of Vehicle Emissions Inspection at 255-1135.

Sincerely,

James E. Sarn, M.D., M.P.H.
Director

JES:RFI:db

Attachments

* (The Auditor General has chosen to include this as a separate response. Therefore, there is no Attachment 7.)

SANTA FE RESEARCH CORPORATION
ROUTE 7 BOX 124K • SANTA FE, NEW MEXICO 87501 • (505) 983-6568

January 10, 1983

R. Fred Iacobelli, P.E.
Acting Chief, Bureau of Air Quality
Arizona Department of Health Services
1740 West Adams
Phoenix, Arizona 85007

Dear Mr. Iacobelli:

As per your request, I have reviewed the January 1983 report "Time Series Analysis of the Impact of the Vehicle Emissions Inspection Program (I/M) on Ambient Air Quality in Phoenix and Tucson." In my opinion, the analyses and results in this report do not support its principal conclusion -- that "the I/M program has had no salutary impact on ambient air quality in Arizona." More importantly, this report alone certainly cannot be considered adequate justification to discontinue or modify the current I/M program. The report does, however, suggest the need to analyze and resolve possible inconsistencies between ambient air quality trends and calculated emission trends.

I find that the aforementioned report contains several important omissions and defects. Before I discuss the specific deficiencies, however, I would like to raise a more general issue. This issue concerns the general difficulty in interpreting trends of ambient air quality data. I have been involved in numerous air quality trend studies performed for federal, state, and local agencies. In some of these studies, I have found that ambient air quality can yield confusing if not wrong implications. Not only are air quality data continually confounded by meteorological fluctuations, but also unexplicable aberrations often creep in due to undocumented changes in monitoring procedures or other unknown factors. I think that it is very important to analyze air quality data to make sure that ongoing control programs are having the expected effect. When the air quality trends confirm the predictions based on emission analyses, the results are reassuring and are supportive of the control programs. When the ambient data seem to contradict emission trend data, such as <u>may</u> be the case for CO in Arizona, then there is a need to analyze the problem further and resolve the discrepancy. However, considering the vagaries of meteorology and potential errors in ambient air quality data, such a discrepancy should definitely <u>not</u> be immediately interpreted as a fault in the emission data or a failure of the control program. Such an interpretation would be particularly inappropriate with regard to the results of the report in question.

I would now like to address some of the major deficiencies in the report; these are discussed below:

- The main conclusion in the report, "The I/M program has had no salutary impact on ambient CO air quality in Arizona," is overstated. Assuming that the statistical results are correct, a more reasonable conclusion would appear to be that "no evidence can be found that the I/M program has had an impact on CO air quality in Arizona." More importantly, depending on the standard errors of the statistics (see next comment), the conclusion might actually be "no evidence can be found that the I/M program has not had the expected 3-15% improvement in air quality during the two interventions."

- A major omission in the report is the lack of standard errors for the trend statistics. If these standard errors are quite large, the results of the report may not be inconsistent with a 10%, 20%, or 30% improvement in CO air quality due to I/M.

- A major deficiency in the statistical approach is the failure to explicitly control for traffic growth, federal emission standards, and meteorology. All three of these phenomena can potentially produce greater effects than the I/M program. Traffic levels evidently grew by about 50% over the 8-year study period; new car standards produced a large decrease in total emissions over the period; and meteorology leads typically to 10-20% fluctuations in air quality indices on a year-to-year basis. It is a major defect to search for the impact of I/M without explicitly controlling for the other factors, especially when the other factors can be included in an accurate fashion (e.g. gasoline sales adjusted for fuel efficiency to represent traffic, emission models for the new car standards, and Arizona DHS meteorological normalization procedures). The authors of the report say that they have "controlled" for these factors, but this is true only to the extent that meteorological fluctuations are purely seasonal and that traffic growth and federal emission reductions represent a constant trend over the period. Neither of these latter two assumptions is justified. Considering that the three exogenous factors are likely to be stronger than the I/M effect and are likely to be confounded historically with the I/M effect, it is important to treat them explicitly in the statistical model.

- At several points in the report, the authors erroneously assume that vehicular CO emissions depend on fuel efficiency. This is not true. Vehicle emission standards and factors are expressed in gm/mi, and total vehicular emissions depend on these factors and traffic levels (e.g. miles traveled). There is no consistent relationship between fuel efficiency and vehicular emissions. Fortunately, I think that this error does not affect the statistical results in the report. The error does, however, seem to bear on some of the interpretations and qualifications made by the authors.

- The functional form used to represent the I/M program -- stepped increases spread out over 12 months in 1977 and 1979 -- may be inappropriate. As I understand it, the Arizona DHS has expectations and evidence that there was a voluntary response to I/M in 1976 and that there was a continued improvement in 1978 and 1980 as vehicle owners modified their behavior in response to I/M. I think that, in further analyses, the appropriate I/M transition function should be selected in consultation with Arizona DHS.

The above paragraphs summarize my major conclusions concerning the report. If necessary, I would be happy to meet with you to discuss any of these issues in greater detail. Please contact me if you have any questions or if you require any further information.

Sincerely,

John Trijonis

ATTACHMENT 2

NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

P. O. Box 3000 • Boulder, Colorado 80307

Telephone: (303) 494-5151 • TWX: 910-940-3245 • Telex: 45 694 • FTS: 322-5151

14 January 1983

Fred Acabella
Arizona Department of Health Services
State Health Building
Phoenix, AZ 85007

Dear Fred:

This letter is in response to your request for our comments on portions of a report by Richard McCleary on his time series analysis of the impact of IM on ambient CO levels in Phoenix and Tucson. Currently we are involved in an effort to develop stochastic models of ambient CO data in Denver, therefore we are well aware of the complexity of the problem and current literature on the topic.

It is difficult to comment on the analysis because so little actual data is included in the two appendices we received, so our comments must be general. The annual means shown in Table 3a of Appendix I raise some interesting questions, however. The annual means over all of the data sets show generally the same pattern of ups and downs by year: a decline 1974-76, levelling off 1977-81. The random fluctuations about this general trend are to be expected due to differences in meteorology from year to year. Coincidentally, there are <u>increases</u> in 1979 in all data sets, and in 1977 in all data sets except Phoenix Sunnyslope. Such fluctuations are not unusual but they illustrate the difficulty of separating intervention effects from the equally abrupt changes which can be caused by changing weather patterns.

Therefore it is essential to have some understanding of causal relationships and to view the CO data in relation to important explanatory variables. We cannot accept McCleary's contention that a univariate ARIMA model will accurately separate IM effects from all other effects and that it is not necessary to understand the reasons for observed changes.

The accuracy of any estimated parameters requires that the model be correctly formulated. There are several reasons to doubt the correctness of McCleary's ARIMA model: (1) As noted above, random fluctuations due to unexplained sources actually led to increased CO levels in the years the IM effects were expected to occur, hence changes in weather and traffic patterns should be investigated as potential causes. (2) Emissions changes due to IM predicted by the MOBILE2 model differ from the one-year step change used in McCleary's model. (3) It has been shown that strongly seasonal time series based on deterministic physical processes are likely to be incorrectly described by a standard ARIMA model and require, instead, a periodic ARMA model. Points (2) and (3) require further elaboration.

In regard to point (2), EPA tested the effects of IM on emissions in Portland, Oregon, over several successive inspections and, if necessary, readjustments at 3 month intervals. The results, incorporated in MOBILE2, showed that further reductions in mean emissions were obtained on successive inspections (Rutherford, 1982). MOBILE2 predicts an initial step change due to the first inspection and additional smaller step changes due to the second and third inspections. Also, in the Arizona program mandatory annual inspections of motor vehicles began in January 1976 but repair of failed vehicles was voluntary. We would assume that some few failed vehicles were repaired in 1976. Thus it is quite possible that some IM effect would be incorporated in the trend portion of a univariate model over several years and would not appear as a step change.

In regard to point (3), the actual models used are not shown in either of the two appendices that we have seen. We can only assume that they were standard univariate ARIMA models of the type described in McCleary's book (Applied Time Series Analysis for the Social Sciences, 1980). These methods use seasonal differencing to accomodate the cyclic behavior of seasonal data, and assume that the mean and autocovariance function of a differenced series are homogeneous over all seasons. However, in strongly seasonal data the assumption of homogeneity can be quite inappropriate. Tiao and Grupe (1980) have shown that this can lead to a misspecified model which looks acceptable, but is much less accurate than a model which includes periodic means and/or periodic covariances. Cleveland and Tiao (1979) demonstrate the effects of such an inappropriate model choice on monthly ozone data. The standard ARIMA model clearly did not produce as good a fit as the alternative, a periodic ARMA model. This simply means that different parameter estimates were needed for different months of the year to get a good fit. Similar effects can be expected for monthly CO data because, like ozone, it has a strong physically determined seasonal pattern.

Lacking controls that can be shown to account for changes in weather, traffic, and federal emission standards, it is possible that McCleary's model is incapable of isolating the effects of IM. In his model it is likely that some weather effects are included in the step change component which he attributes to IM and that some IM effects are included in the trend component which is supposed to represent only non-IM changes. There is no way of telling, since the model has no explanatory power. Thus McCleary's conclusion that IM had no impact is overstated. It would be more accurate to say that no impact could be detected using this particular model.

It is also possible that the Arizona data base for 1974-81 is incapable of supporting any model of the necessary precision. McCleary's recommendations on development of a future data base are excellent and should be implemented.

We agree with McCleary that the 1979 Arizona DHS study of ambient effects of IM has serious problems which put its conclusions in doubt. McCleary shows a lack of understanding of the empirical studies of emissions effects of IM, however. Meteorology and climate are not relevant in a controlled FTP test of tailpipe emissions. Thus the controlled emissions tests offer empirical evidence in support of IM. Whether emissions reductions due to IM can be discerned in ambient CO measurements remains to be seen.

References

Cleveland, W.P. and G.C. Tiao (1979). Modeling Seasonal Time Series. Revue Economic Appliquee, 32, pp. 107-29.

Rutherford, J. (1982). Derivation of I/M Benefits for Pre-1981 Light Duty Vehicles for Low Altitude, Non-California Areas. EPA-AA-IMS-82-3.

Tiao, G.C. & M.R. Grupe (1980). Hidden Periodic Autoregressive-Moving Average Models in Time Series Data. Biometrika, 67, 2, pp. 365-73.

Very truly yours,

Robin Dennis

Mary Downton

Mary Downton

RD:MD:brm

ATTACHMENT 3

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

ANN ARBOR, MICHIGAN 48105

OFFICE OF
AIR, NOISE AND RADIATION

JAN 19 1983

Fred Iacobelli
Bureau of Vehicle Emission
Inspection
600 No. 40th Street
Phoenix AZ 85008

Dear Mr. Iacobelli:

Thank you for the opportunity to comment on the January 6, 1983 draft report titled "Time Series Analysis of the Impact of the Vehicle Emissions Inspection Program (I/M) on Ambient Air Quality in Phoenix and Tucson." I have already commented by an earlier letter on a previous draft of this report.

My most serious complaint about the previous draft applies to this report also, namely that the authors react too negatively to their failure to reject the null hypothesis that I/M has had an effect on CO levels. The authors seem to go beyond conventional statistical reporting to persuade the reader that the inability to reject the null hypothesis proves that the null hypothesis should be accepted. I am left wondering whether the authors have simply lost sight of the limits of hypothesis testing as a method of positive proof, or whether they have a pre-existing animus towards I/M in particular. Their unexplainable failure to report the standard error of the estimates makes it impossible for the reader to judge how much weight to attribute to the inability to reject the null hypothesis. If as I suspect the standard errors are large, then any number of more optimistic hypotheses about I/M's effect could also be immune from rejection. It is true that the point estimates of I/M's effect are small, but if the standard errors are large, there is a good probability of seeing such results by pure chance even if I/M were as effective as believed.

The authors make much of having used three CO sites and up to four CO measures in the analysis, for a total of eight site-measure combinations. They try to convince the reader that if the I/M program failed to show a significant benefit in eight "quasi-experiments" the I/M program has been positively proven to be ineffective. In fact, however, sites in the same city and different measures of CO are not independent experiments for testing I/M. The study really has only given I/M two opportunities to prove itself, once

(Corrected 2/2/83)

P. Lorang

for Phoenix and once for Tucson. Even these may not be independent, since regional weather patterns could influence both cities.

The report in a number of places says that emissions decrease as fuel economy increases. This reveals a lack of understanding of Federal emission control regulations, which regulate emissions in grams per mile, not per quantity of fuel burned.

The report refuses to acknowledge that results from emissions tests of in-use vehicles in I/M programs are empirical data and can be used as evidence that I/M reduces CO emissions. Only a disbeliever in the conservation of mass could argue that emissions can be reduced without improving air quality. Of course, it is possible to pose arguments that the very large body of evidence from tests of in-use vhicles in Portland, Los Angeles, and Phoenix could be misleading for reasons such as potential recruiting bias, differences between the FTP and real world driving conditions, etc. However, the report for the most part chooses to ignore this body of evidence rather than discuss its merits fairly. This is unfortunate, since I believe the evidence supports at least a tentative conclusion that I/M works, and that it is therefore unfair to prefer a null hypothesis that it has zero effect.

In the main part of the analysis, the authors use a model which ignores the possibility that irregular year-to-year fluctuations in traffic volume and meteorology could be obscuring a positive benefit for I/M. The authors' attempts to rationalize this model are statistically incorrect. My understanding is that the ARIMA model's N_t term controls for linear and seasonal trends but not for year-to-year weather difference. The report does not adequately inform the reader of this shortcoming.

The report includes a short section on results obtained for a model which uses traffic volume estimates and a termperature inversion variable. *In principle*, these two variables should explain some of the variation in CO levels, thus reducing the standard errors and improving the model's power. Since traffic and weather may be correlated with the I/M intervention, including them as variables can change the expected value of the regression estimators also. It turns out that the normalized data are sometimes more variable than before normalization and the I/M effects are still not significantly different from zero. The fact that the data become more variable with the normalization suggests to me that perhaps the particular temperature inversion variable is for some reason a poor indicator of CO dispersion potential, and that other meteorological variables such as wind speed or relative humidity should have been tried. In any case, the fact that even with this particular normalization it is impossible to reject the null hypothesis of no effect from I/M does not prove that I/M has no effect, as the authors attempt to conclude.

If I can be of further assistance, please call.

Sincerely yours,

Phil Lorang

Phil Lorang, Chief
Technical Support Staff

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

ANN ARBOR, MICHIGAN 48105

OFFICE OF
AIR, NOISE AND RADIATION

R. Fred Iacobelli, P.E.
Acting Chief, Bureau of Air Quality
Arizona Department of Health Services
1740 West Adams
Phoenix, Arizona 85007

DEC 1982
RECEIVED
Bureau of Air
Quality Control

Dear Mr. Iacobelli:

Thank you for the opportunity to review the draft Appendix titled "Time Series Analysis of the Impact of the Vehicle Emissions Inspection Program on Ambient Air Quality in Phoenix and Tucson" by Richard McCleary and Barbara C. Nienstedt. Since I am responsible for EPA's activities in evaluation of inspection and maintenance (I/M) programs, I am interested in all such efforts. As you know, my staff supervised the study by G. C. Tiao and his associates of the impact of the I/M program in Portland, Oregon on that city's CO air quality, so I am familiar with the type of functional models which can be used for this purpose. The Tiao study is listed as Reference 22 in the draft Appendix you sent me, which speaks quite favorably of it.

I have organized my comments on the draft Appendix to address in sequence the modeling method, the interpretation of the results of the modeling effort, the recommendations offered at the end of the Appendix, and my own conclusions about the overall value and significance of the study reported in the Appendix.

Modeling Method and Issues

First, the Appendix describes the study by Tiao in Portland as the only good prior work on the subject of estimating I/M benefits from CO air quality data. Unfortunately, the Appendix does not use the same functional form or specification as Tiao did for the model which relates CO levels to the variables that contribute to those levels. Tiao found it best to express CO concentration as the product of a CO emissions/mile emission factor term [$ke^{\beta t}(1 + I/M_t^{d})$, in Tiao's symbols], a traffic volume term [TR_t], and a meteorology term [RH_t]. This makes sense, since CO emissions result from the product of an emission rate per mile times VMT and are diluted depending on meteorology. In contrast, the Appendix uses several different models in which the travel variable and meteorology variable are added, not multiplied. The Appendix does not use any term that is clearly an emission factor term. The Appendix does mention that a "normalized" model was also used (p. 37) which may be more like the more logical Tiao approach, but this "normalized" model is not very well described in the Appendix. All else aside, the use of the additive model in the Appendix should still give meaningful results but they will be less directly indicative of effects on air quality.

Second, the Appendix uses the single highest 8-hour CO average each month to describe CO air quality during that month. Tiao used the average CO value for the month. The single highest value is notoriously variable, since it can be influenced by traffic volume and weather during very short periods and these in turn may flucturate greatly. Since the model in the Appendix uses monthly averge fuel sales as the measure of traffic volume and monthly average relative humidity as the meteorology variable, the CO statistic should clearly be based on the monthly average as well. (If the monthly average is not readily available, the average of each day's highest 8-hour value should be used.) Using the single monthly high value adds noise to the time series and makes it much harder to show that individual factors have a statistically significant effect on CO levels.

Third, as mentioned, the Appendix used fuel sales as the measure of traffic volume for lack of actual traffic counts. The Appendix points out the shortcomings of this approach, but neglects to mention one shortcoming that could have been corrected: average vehicle fuel economy has been increasing, so the real increase in VMT over time has been faster than indicated by fuel sales alone. The Appendix ignores this. As a result, it fails to recognize part of the reduction in CO emissions per mile over time, a reduction that in part may be attributable to I/M. I have enclosed some tables from standard U. S. Department of Transportation reports which give estimates of the fleet fuel economy in each calendar year.

Fourth, the Appendix uses a model which does not explicitly account for the beneficial effects of the new car emission standards--the Federal Motor Vehicle Control Program (FMVCP). Tiao did recognize the FMVCP as a possible effect, and in fact did find from the data that the FMVCP caused a 5% per year reduction in CO emissions and ambient levels. It is inconceivable that the FMVCP has not also had an effect in Phoenix and Tucson. The Appendix accounts for the effect only as part of the "noise" term. I am not familiar with how this noise term is treated in the ARIMA model used by the Appendix, but the failure to recognize the possible effect of the FMVCP may have negative consequences for the results of the analysis.

Fifth, the Appendix uses a very simple variable form to describe the intervention due to I/M. It assumes that I/M has a constant effect on ambient levels (except for an additional impact from the change in failure rate) once all cars have been inspected one time. A better variable would have been the number of vehicles which had been inspected, failed, and then passed the retest after repair. This would capture in one variable the effect of starting the I/M program in 1977 and increasing the failure rate in 1979. It would also reflect any changes in waiver rates over time; presumably cars which receive waivers do not contribute as much to reducing CO levels as do successfully repaired vehicles. Since CO levels are affected by all cars on the road and not just those being inspected each month, it is necessary to use a cumulative vehicle count for this variable. To reflect gradual deterioration in CO emissions back towards non-I/M levels following inspection and repair, I recommend Tiao's method of discounting inspections from previous months by progressively larger percentages. See page 24 of Tiao's final report to EPA for details. Another advantage of this method is that one could then easily estimate the benefit of reducing the waiver rate through a higher cost limit, better training for mechanics, or more rigorous review of waiver applications.

Finally, the Appendix differs from the Tiao method of accounting for seasonality. I am not familiar enough with the two methods to say whether this is likely to have an effect on the results.

<u>Interpretation of Modeling Results</u>

The Appendix estimated the effect of the I/M program in Phoenix on CO levels in several different ways and got results ranging from a reduction of 17% to a reduction of 42%. This range corresponds roughly to what EPA models of I/M would suggest. However, the Appendix reports that none of the estimated benefits are statistically significant. The Appendix concludes that there is therefore no evidence that

I/M has had an effect, and even implies that the evidence shows it does not have an effect. This is an oversimplification of the meaning of the statistical results and will mislead many lay persons who read the Appendix. The modeling results actually need more careful explanation so that the average reader understands the significance of the analysis in the Appendix.

The Appendix is presumably correct that none of the estimates of I/M benefits are "statistically significant." This is a technical term with a very specific meaning. It means that one cannot say with much certainty that the effect of I/M is different from zero. In other words, the effect may be zero, and the positive estimates could be merely random in nature. This is a slanted way to look at the results, and is used so frequently by statisticians only because it is convenient and easy to explain. In fact, according to statistical theory the analysis says the real I/M effect is more likely to be exactly what the analysis estimates it to be (17% to 42% depending on method), than any other value. Admittedly it could be zero, but it could just as likely be higher than the 17% to 42% estimate. Essentially, all the study accomplished was to say that the "noise" in the ambient CO levels is too high to be able to tell what is causing the CO levels to vary from month to month. In effect, we know no more than before the study was made.

This being the case, the statements on page 31 (lines 7 and 8, lines 16 and 17), page 32 (lines 5 and 6), page 33 (lines 15 through 17), page 34 (lines 12 through 18), page 37 (lines 3 through 7), page 40 (lines 7 through 10), and page 41 (lines 5 through 8) are misleading because they imply there is evidence that I/M is not working, or that the failure to find clear evidence that I/M is working is by itself an indication of I/M's failure. The study has in no way proved that I/M is not working. It is a very common result to find that nothing can be proved statistically in this field, and should not be a surprise.

There is substantial evidence from other sources that I/M does reduce CO emissions and therefore ambient CO levels. This evidence includes EPA's study of emissions of cars in Portland, the Tiao study of ambient CO in Portland, the EPA study of the CO emissions of Phoenix cars compared to cars in non-I/M cities* (cited on pages 6 and 7 of the Appendix), and a small EPA study of the I/M program in the South Coast Air Basin of California. This evidence creates a presumption that must be clearly rebutted by a statistical analysis, if the statistical analysis is to be held up as proof that I/M has no effect. In fact, however, the study in the Appendix does not rebut this presumption. If asked, I am sure the authors would concede that the study results could be honestly restated to say, "There is no evidence that I/M does not reduce amibent CO levels by at least 17%."

The inconclusive nature of the study's findings are due in part to the poor choice of the CO statistic to model. As mentioned above, use of the monthly high 8-hour average adds noise to the analysis. The inconclusive results are also due in part to the fact that in Phoenix there has been only one major I/M intervention, in 1977. In contrast, in Portland the biennial, on/off pattern of inspections is a repeated intervention which gives the statistical analysis a better opportunity to distinguish the effect of I/M from random noise. For this reason, it is not surprising that a study in Pheonix would be less conclusive than the Tiao study in Portland.

Of course, the inconclusive and significantly insignificant results may have been otherwise if the Appendix had used the Tiao modeling method exactly.

Recommendations in the Appendix

Any suggestion that the I/M program be terminated based on the results of this study would certainly be an overreaction.

*The Appendix says on page 7 that considering that there was no control for meteorological factors, it is difficult to place much faith in this EPA study. This conclusion is based on a mistaken understanding of the EPA study. The EPA study compared vehicle emissions under the rigid conditions of the standardized FTP emissions test. Meteorological factors are therefore of no concern in comparing Phoenix to other cities. If I/M in Phoenix has lowered CO emissions on the FTP, as this study proves, it must have lowered CO emissions on the streets of Phoenix as well, to at least some degree.

The recommendations regarding the keeping of air quality data cannot be disputed, although I am not in a position to tell what this would require. As you probably know, EPA keeps a computer file of air quality data, called the SAROAD data base. If EPA can be of assistance to Arizona, let me know.

The recommendation regarding traffic counts seems to make sense on the surface. However, since traffic counts will never be available from the pre-I/M period, it may be of little use to begin collecting them now.

The recommendation that the I/M program be evaluated with a proven statistical methodology also makes sense. However, there is so much noise in air quality data that even the best evaluation may still be inconclusive. I object to any assumption that any one air control measure must conclusively prove itself based on an analysis of air quality data. Control measures should be proven effective in reducing emissions, as I/M has been. Common sense then says that air quality will also be improved.

I disagree that a desert community without I/M should be established as a control site for future studies. The desert community would have to have similar weather, traffic concentrations, traffic growth, topology, and new car sales as Phoenix to be a reliable control site. This is very unlikely in a small town. A control site is not strictly necessary for a time series intervention analysis, as the Appendix demonstrates. I would like to point out that the Tiao study did not use Eugene as a control site in the common sense, despite an indication to this effect on page 4 of the Appendix. Portland and Eugene air quality levels were never compared directly. Tiao merely verified that inspecting cars in Portland had no effect on Eugene's air quality. If it had, some unidentified variable would have been suspected as causing the CO reductions in both cities. With respect to evaluating the effect of I/M in Phoenix and Tucson, I am generally not in favor of comparing an I/M city and a non-I/M city and attributing the difference to I/M. If you must do so, it is probably better to use composites of a number of other cities of approximately the same size and population density as Phoenix and Tucson.

Conclusion

The Appendix you sent me contains a less-than-the-best analysis of I/M in Phoenix and Tucson, even though it may be the best that was possible given the authors' resources. The analysis was inconclusive, and is not inconsistent with other studies that have shown that I/M reduces CO emissions and ambient CO levels. The write-up in the Appendix seriously simplifies the explanation of the findings, or lack of findings, and may mislead lay readers.

Sincerely yours,

Phil Lorang, Chief
Technical Support Staff

ATTACHMENT 4

Comments on McCleary - Nienstadt Report

P. 2 "Control" is a little confusing. The models actually account for the effects of exogenous variables through trend and seasonal components, or, such variables are replaced by trend and seasonal components which account for roughly the same amount of explained variation.

P. 17 "ARIMA models must be used" should be restated to say ARIMA models may be used.
There are other valid time series modeling procedures.
"Econometric type" models, multiplicative models, nonlinear models, unobserved components models.

P. 18 Quasi-experimental evidence is subject to interpretation. Lack of control series seriously hinders decision making. Even significant impacts might be misleading as they could be caused by unobserved factors other than the I/M program.
Excellent discussion of how quasi-experimental nature of project handled. Comprehensive inspection of threats to validity.

P. 24 Some question as to the assumption of a uniform intervention, but unlikely that different coding would drastically change results. Worthy of investigation.

P. 38 In discussion of statistical power of models, the statement made is true, but since H_0 was accepted, power is an important issue.
"Controlled meteorological variance in CO by seasonal ARIMA structures" means that meteorological variables were almost entirely explained by a seasonal component which was included in the CO model.

Met. Var. = seasonal + noise
high % low % of variance
CO = trend + seasonal + noise
same as above, accounts for all but small % of variation due to meteorological variables.

i.e., model with seasonal + met. accounts for about same % of variation as model with seasonal alone, so no need for met. in model.

Additional statistical information needed in the report:

*percent of variation in CO explained by model (desire low standardized residual variance, or residual mean square).
*what percent decrease in CO could have been detected: standard error of predictors? Standard error of impact parameter?

Impediments to decision making:

1) No control series
2) Short length of series
3) Quasi-experimental design (validity problems)
4) Power of statistical tests (decision based on accepting H_0).

Possible impediments to statistical conclusions:

1) Nature of residuals, size of residual variance.
2) Distribution theory: statistical test that accepts null hypothesis of white noise may have "low" power, assumption of white noise essential.
3) By nature of statistical testing procedure, some models will produce "significance" merely by chance.

Evaluation:

Excellent analysis using quasi-experimental design.
Comprehensive validity check.
Knowledgeable application of ARIMA modeling to intervention analysis.
Comprehensive modeling of univariate series of interest. Correct interpretation of results, but lack mention of residual variance or nature of residuals.

Statistical analysis seems to warrant accepting of null hypothesis of no effect of intervention, but conclusion should be qualified to mention problems of sample size and power. Otherwise, limitation of statistical procedures are clearly delineated.

I partially endorse conclusion that I/M is ineffective.
I wholly endorse recommendation that data collection/design problems be considered more carefully.

T. J. Woodfield

Terry J. Woodfield

ATTACHMENT 5

Arizona Department of Health Services' Comments on
Performance Audit of the Department of Health Services'
Vehicle Emissions Inspection Program

S U M M A R Y

P. i

To state that no evidence was found that the Vehicle Emissions Inspection program has been effective is to refute the law of conservation of mass. Figures 1, 2 and 3 show the reduction in tail pipe carbon monoxide emission concentrations resulting from repairs to vehicles that failed the emissions test. Figure 1 is for all vehicles inspected. The downward trend indicates the effect of dropping older vehicles from the program and including newer ones. Figure 2 shows the effect of repairs to 1969 model year vehicles, and Figure 3 shows the effect of repairs on 1975 model year vehicles. Figures 2 and 3 represent two different technologies which were included in the program during the years 1976 through 1982. It can be noted that the average emission performance deteriorates between inspections, yielding the net effect of the emissions reductions shown by the shaded area. There is also a positive but less definitive benefit of the inspection program included in the area between the dotted lines and the before-repairs plot due to program induced behavioral changes. These figures also illustrate that there was a profound effect of the program between the 1976 testing and the 1977 testing, and that there were effects of the more stringent test standards prior to the 1979 testing. This phenomenon reveals that the implementation of the mandatory maintenance in 1977 and the increase in test failures in 1979 were blended rather than discrete, and consequently the statistical model was insensitive to them.

The decline in idle tail pipe emissions between the 1976 testing and the 1977 testing, shown in Figures 1, 2 and 3, indicates that the program changed vehicle maintenance behavior, at least for owners of vehicles that are in the program. However, it is suspected that when a vehicle is exempted from the inspection program because of age, maintenance habits will revert to preinspection days and emissions will increase.

The statements regarding circumvention are typical of the uninformed and untrained mechanic. Vehicle manufacturers must meet three criteria in the design of engine performance: (1) Federal emissions certification, (2) corporate average fuel economy, and (3) customer demand driveability. Engine design and tuning specifications optimize all three so that any vehicle in good mechanical condition (including operation of all emission control devices) when tuned to manufacturer's specifications will have low emissions (pass the Arizona emissions test and the Federal test procedure), good fuel economy and optimized performance.

Variability in vehicles does occur. However, emissions factors and the impact of inspection/maintenance programs are based on empirical results from

CO EMISSIONS (Idle)
All LDV Inspected Vehicles

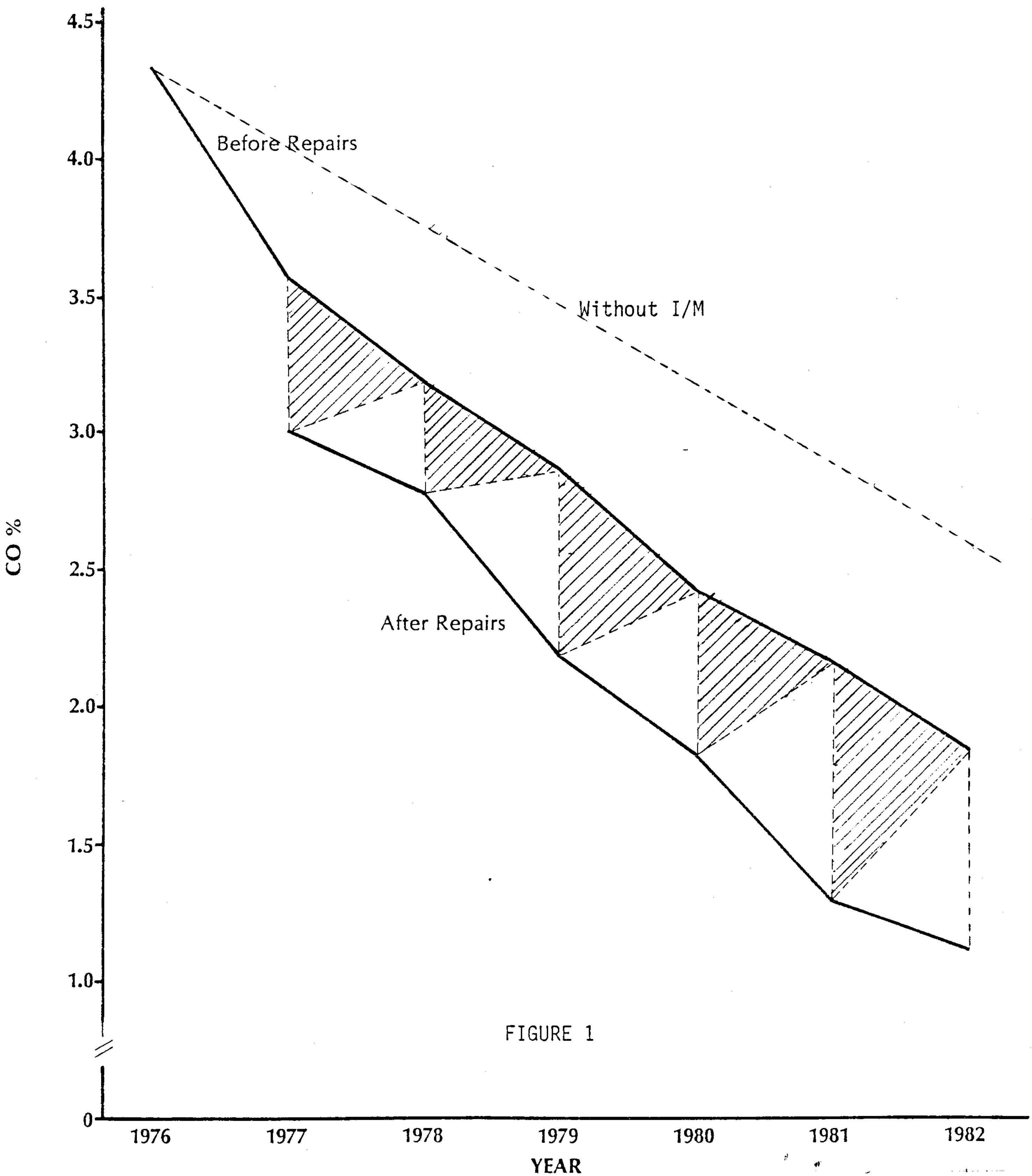


FIGURE 1

CO EMISSIONS (Idle)
1969 Model Year

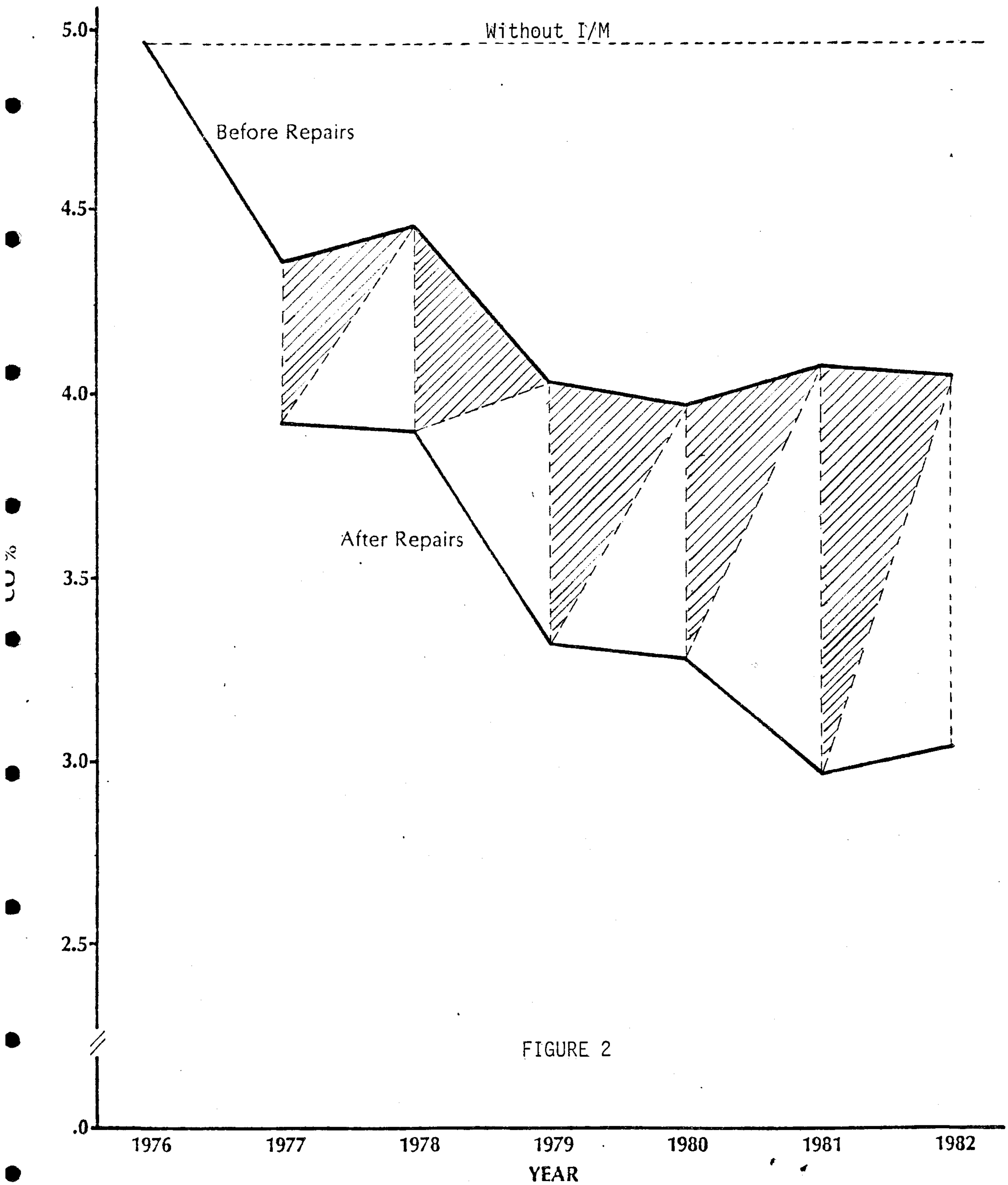


FIGURE 2

CO EMISSIONS (Idle)
1975 Model Year

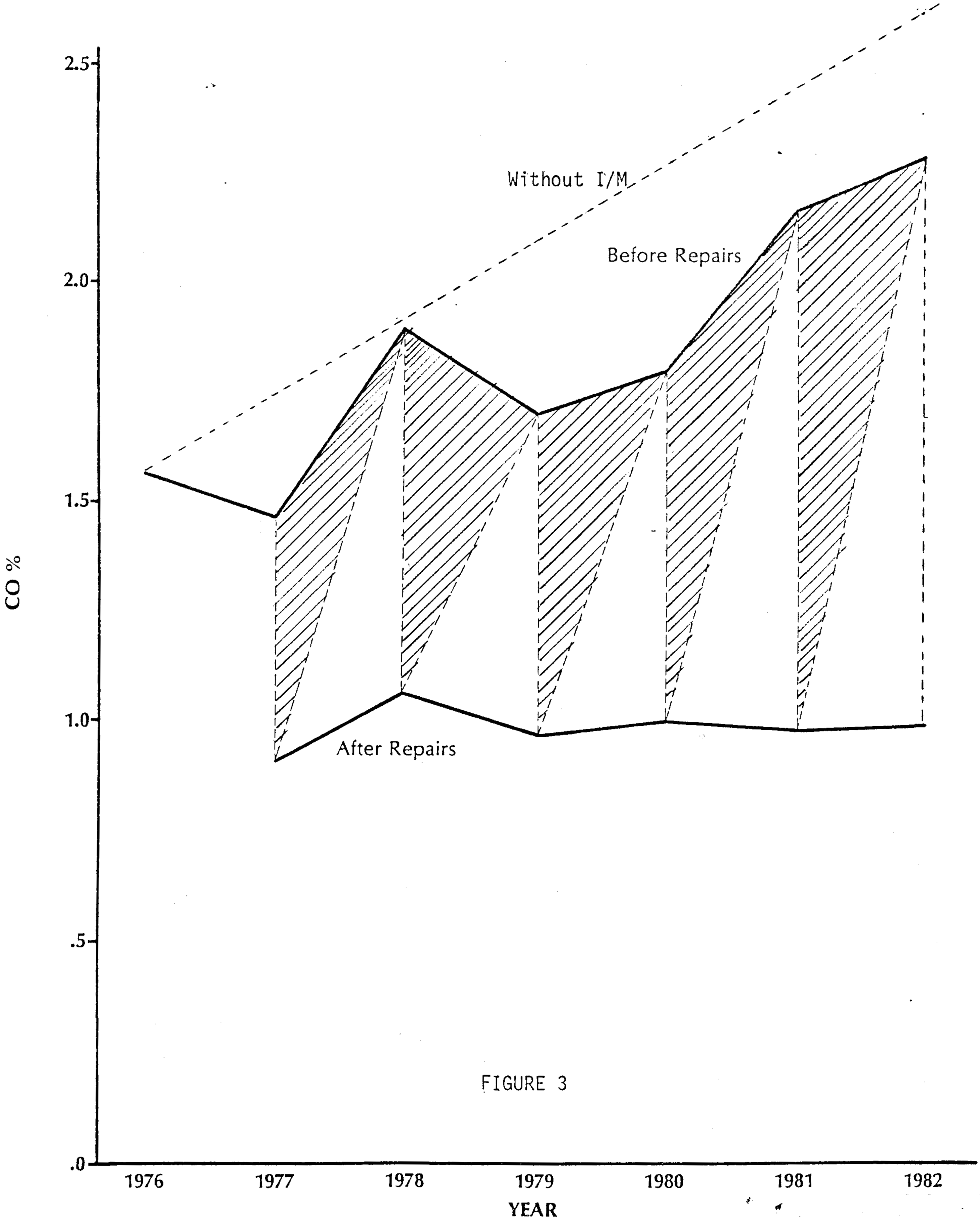


FIGURE 3

the testing of in-use vehicles. Consequently there are provisions for engine variability in the test design.

It is evident from inspection data, which is partially shown in Figures 1, 2 and 3, that emissions reductions are obtained from vehicles entering the program for the first time, and that further emission reductions are gained from the periodic inspection/maintenance process. The inspection program could become more effective by including older and out-of-state vehicles.

The statement that it is not known for certain why the program is ineffective is a result of a conclusion based on a grossly deficient statistical analysis. A more correct statement would be that empirical evidence shows the program is effective in reducing emissions. However, the time series analysis conducted by the Center for Informative Evaluation was unable to determine the effect of these reductions on air quality. The effect of policy or program changes can certainly be evaluated in terms of their effect on emissions.

P. ii

The recommendation to operate the program only as required to avoid imposition of Federal sanctions neglects to recognize the healthful benefits of the emissions reductions and is contrary to the declaration of policy contained in ARS 36.1700. Every effort should be made to make the Vehicle Emissions Inspection Program as effective as possible.

Fifty two potential strategies for reducing the automotive air pollution were identified and evaluated by the Phoenix Air Quality Maintenance Area Task Force prior to formation of the Nonattainment Area Plan for Carbon Monoxide and Photochemical Oxidants in December 1978. Effective alternatives to inspection/maintenance such as vehicle free zones, gasoline rationing and mandatory maintenance were found unacceptable and rejected.

The boundaries of the nonattainment area plan for carbon monoxde and ozone in Pima County have been revised and adopted for proposal by the Pima Association of Governments Regional Council. In order to reflect this change in the inspection/maintenance program, the revision must be submitted to and approved by the Administrator of the United States Environmental Protection Agency.

The statements regarding ozone are questionable and misleading. First, there is no evidence or combustion theory that would relate the perceived ineffectiveness in carbon monoxide control to ineffectiveness in hydrocarbon control. Although fewer vehicles fail the emissions test for hydrocarbons than carbon monoxide, the attendant fleet-wide improvement in tail pipe concentration from repairs to those vehicles that fail is greater for hydrocarbons than for carbon monoxide. Improvements in hydrocarbon idle tail pipe emissions can be depicted in a manner similar to that shown for carbon monoxide in Figures 1, 2 and 3.

To state that the air/fuel waiver provision of Arizona's Vehicle Emissions Inspection Program is not cost effective ignores the fact that the associated provision for registration of emissions analyzers and certification of mechanics has greatly reduced or avoided the expense to program users of multiple trips to repair facilities resulting from second failures due to improper repairs or erroneous emissions measurements during repair.

P. iii

The situation regarding the lack of quality assurance audit has been corrected. However, it should be pointed out that there is no evidence that the quality of the inspection operation suffered because of this deficiency.

The effects of tampering with emissions control devices and misfueling have not been addressed in the report. These two actions on the part of vehicle owners have an obvious and serious impact on the effectiveness of the program when the statutory cost limitation on repairs prohibits the enforcement of repairs. A management report system has been instituted.

F I N D I N G I

P. 7

We totally disagree with the Finding because the basis for the Finding, the analysis prepared by the Center for Informative Evaluation, is poorly conceived, of inferior design, misleading due to the consultant's lack of understanding of air quality complexities, and grossly flawed by the inappropriate use of statistics. As an example, there is one particular problem that we must emphasize because it is critical to the design of the consultant's time series analysis, and his failure to recognize it renders the results meaningless.

The hypothesis being tested is, "Did the ambient air quality improve starting in January, 1977 with the inception of the inspection/maintenance program, and again in 1979 when failure rates increased?" The results are a statement of whether this improvement occurred or not. If the air quality trend due to factors other than inspection/maintenance was stable prior to and after inspection/maintenance implementation, this test would be fair and appropriate. A visual inspection of the air quality trend shown in Figure 4, however, indicates that remarkable changes in the air quality trend were occurring prior to inspection/maintenance implementation and continued through to the present time. These changes in the slope of the trend indicate that the net effect of cleaner new cars and increased traffic has been changing. Apparently the year to year reduction in emissions attributable to cleaner new cars exhibits a decreasing ability to offset the rapid growth in Phoenix traffic. Given these observable facts, it is unreasonable to test whether inspection/maintenance improved the air quality trend.

The fact that the net effect of the uncontrolled factors is a changing air quality trend invalidates the test. Since the test is reported anyway, it should at least be recognized as being biased against showing an air quality benefit from inspection/maintenance, because the net effect of the major control strategy (Federal new car standards) was a decreasing benefit.

The Department of Health Services staff advised the staff of the Auditor General and his consultants of this potential problem early in the audit. The consultant's verbal response and subsequent written discussion in the report are similar. The consultants contend that the effect of cleaner new

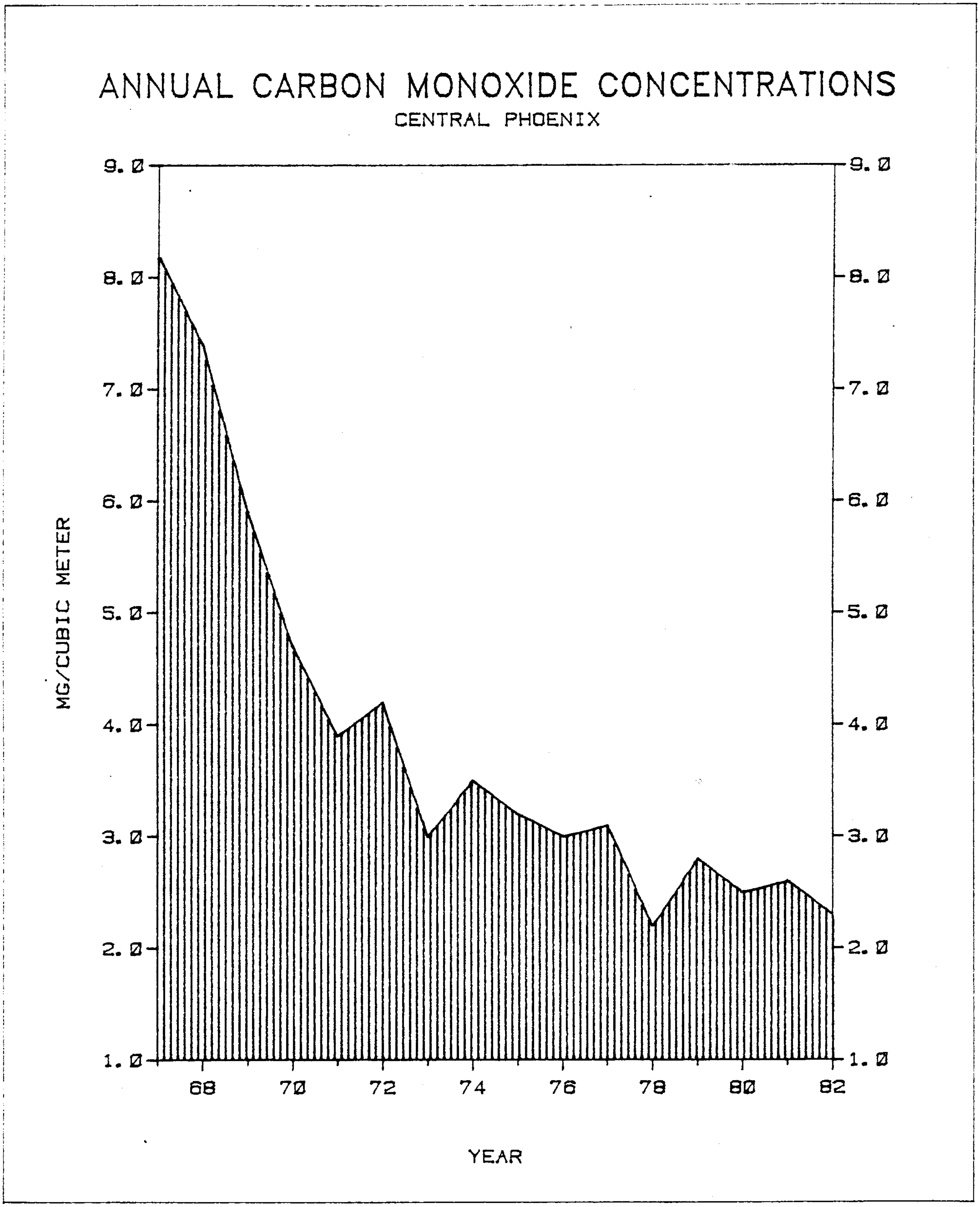

ANNUAL CARBON MONOXIDE CONCENTRATIONS
CENTRAL PHOENIX
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
MG/CUBIC METER
68
70
72
74
76
78
80
82
YEAR

FIGURE 4

cars and the effect of increased traffic "tend" to cancel and therefore the balance between these opposing variables does not need to be dealt with explicitly. The two previous inspection/maintenance time series studies in New Jersey and Portland recognized the problem and dealt with it.

It should be clear that this cancelling assumption is in error, and that the statistical test was fatally flawed in its design.

Further, inspection data indicate that the effects of the mandatory maintenance in 1976 and the change in inspection standards in 1979 were diffused rather than discrete and the statistical test was insensitive to them. Detailed comments on the time series analysis are contained in Attachment 6.

Our comments regarding the certainty of program "effectiveness" are covered in the preceding section.

P. 8

The statement concerning the Oregon study, "Furthermore at the site where this effect was found, results were confounded by movements of the monitoring probe and the major traffic disruptions in the area," is flagrantly misleading inasmuch as the study found an additional annual percentage reduction of approximately 12% at this site with an indication that the inspection/maintenance effect is statistically significant at the 5% level.

P. 10

In our opinion, which is supported by those knowledgeable of air quality complexities, the time series analysis as conducted by the Center for Informative Evaluation is invalid and cannot be used in determining that the Vehicle Emissions Inspection Program has not reduced carbon monoxide levels in Phoenix and Tucson and cannot be relied on to meet the 1987 deadline for attaining the Federal ambient air carbon monoxide standard. Further, inspection/maintenance is not to be relied on as the only strategy to attain ambient air quality standards. Other strategies are required.

P. 11

The statement regarding the unchanged behavior and circumventions is addressed in the preceding section of this attachment.

P. 12

With regard to engine variability, the statement that carbon monoxide emissions may increase by 50% with a 10^{o} temperature change is not confirmed by our test results and cannot be cited as a general example. There is little difference in summer and winter emissions levels. In one analysis of our summer test results, there was scant evidence that suggested higher hydrocarbon readings at temperatures over 105^{o}F with waiting time in excess of five minutes.

Vehicle usage immediately prior to the emissions test may affect test results to the extent that the engine may not be at normal operating temperature if the vehicle has been driven only a short distance prior to testing. The information brochure accompanying Vehicle Registration Renewal Notices advises vehicle owners to have the vehicle at operating temperature when they arrive at the inspection station. The brochure also provides instructions for long periods of idle while waiting for the test. Further, the optional conditioning mode of the test provides additional safeguards.

In our testing we have found no evidence that the grade or type of gasoline will affect emission concentrations. Gasohol, propane and liquid natural gas will.

P. 13

It must be recognized that MVMA's suggestion that emissions testing focus on detecting only gross polluters reflects automobile manufacturers' interests in avoiding the warranty provision of the Federal Clean Air Act. The Federal Clean Air Act requires the manufacturer to repair under warranty any vehicle that fails a legislated emissions inspection for the first two years or 24,000 miles beginning with the 1981 model year.

Variability of a vehicle set to fail the test is to be expected. Only tuning to manufacturer's specifications can assure consistent emission readings, and then, as demonstrated by EPA, some variability is inherent. It is extremely misleading to suggest that, based on the test of one vehicle engine, variability affects program effectiveness. There are some technical difficulties with 1981 Fords if the vehicle has been idling for some time before the emissions test. There is no evidence to demonstrate that the 1981 Ford was adjusted to manufacturer's specifications, and by the admission of the Auditor General's staff, the vehicle was not subjected to the conditioning mode and a second test. The failure is usually rectified by the conditioning mode and retest.

P. 14

We certainly expect some adverse impact on carbon monoxide emissions from vehicles driven by winter visitors for extended periods during the winter months. There is a similar impact from those driven by new residents of Phoenix and Tucson who fail to register at the time of entering children in school or accepting employment. Correction of these problems must come from enforcement of the Motor Vehicle Code. It should be added that during the worst months (November, December and January) for ambient carbon monoxide levels, the vehicle miles traveled are approximately 2%, 11% and 11% above the annual average respectively.

To assume that the impact of the inspection of older vehicles in the Vehicle Emissions Inspection Program is dependent upon the continuation of Federal new car emissions standards is in error. Estimates of the effects of both Federal new car standards and the inclusion of older vehicles in the inspection program have been calculated independently of each other, and are available from MAG and DHS.

P. 15

The conclusions and recommendations have been adequately addressed in our comments on the Summary.

FINDING II

We have addressed Finding II in our comments on the Summary

FINDING III

We reiterate our position regarding the benefit to program users of mechanic certification and quality assurance aspects of the air/fuel waiver program. There is a benefit to program users in higher quality repairs and reduced failure rate on retest. In an analysis of a random sample of over 1700 vehicle inspection reports, it was found that the retest failure rate of "untrained" mechanics is 25.6%, whereas the failure rate of "trained" mechanics is only 13.8%.

We believe there is an intent to provide the inspection program users whose vehicles fail the test with high quality one-time repairs. The adjunct function that has been associated with the air/fuel waiver provision is to provide consistent and accurate emission measurements in the repair industry which agree with emission measurement at the inspection station, a quality of repairs that will insure a high probability of passing retest the first time, reduce engine variability and optimize emission reductions resulting from repairs.

Should the air/fuel waiver be eliminated, we recommend a provision for continued quality assurance surveillance of repair industry analyzers and a continued provision for mechanic training. Eliminating the provision for registered analyzers and mechanic training will inconvenience the public.

FINDING IV

As mentioned in our comments on the Summary, improved management control over the field audit program is in place. The benefits to the State of a contractual change requiring Hamilton Test Systems to perform twice monthly field audits vs. an operational manual requirement will be carefully considered.

Hamilton Test Systems' comments are contained in their letter of January 20, 1983, Attachment 7.

OTHER PERTINENT INFORMATION

The shortfall in payment to the contractor, Hamilton Test Systems, Inc., as of January 31, 1983 is approximately $660,800. Of this, approximately $192,200 has accrued during calendar year 1982. The collection of inspection fees at the time of inspection would prevent further accrual in the shortfall.

CONSULTANT'S REPORT TO THE AUDITOR GENERAL, REVISED JANUARY 20, 1983 (APPENDIX 1)

P. 2 Top

"The questions addressed in this evaluation include:
- What were the effects of the I/M program on ambient air quality?
- What aspects of the program were most or least effective with respect to air quality? And
- How is ambient air quality likely to change in the future?"

The last two questions were not substantively addressed in the report. The first question was dealt with; however, to be specific, the study attempted to find what the effects 1977 and 1979 interventions had on ambient CO concentrations.

"Our models 'control' these exogenous variables nevertheless---through trends, seasonal components, and other 'noise' structures---permitting unbiased estimates of I/M impacts on ambient CO."

The term "control" in experimental and statistical work refers to deliberate, explicit treatment of interferring conditions. The seasonal, trend, and noise treatments in this report are indirect and incomplete. In the referenced New Jersey and Portland studies, efforts were made to control for meteorological and traffic variability. Apparently at times the authors are aware of the correct use of the word "control"---on Page 22..."we will ordinarily report one trend (increase or decrease) figure with the understanding that this trend is the confounded effect of the many uncontrolled forces." (underline added).

P. 3 2nd Paragraph

"There are only three sure methods for reducing mobile source primary air pollutants in the atmosphere. These methods are
1) Reducing vehicle miles traveled by building mass transit systems, taxing motor fuel at a prohibitively high rate, etc;
2) Reducing fuel consumption independent of vehicle miles traveled by making vehicles more fuel efficient."

Reducing fuel consumption does not assure reduced emissions and should not appear in this list. Engine design, fuel, and control devices determine the potential emissions of an engine.
The maintenance of engines and control equipment is a sure way to reduce emissions, which should be a part of any list of "sure" methods of reducing emissions. The authors' opinions on the effectiveness of I/M programs to achieve this is irrelevent.

P. 8 3rd Paragraph

"The research controlled meteorological variance in CO levels with a delta 'temperature' factor..."

The work referenced used delta temperature and wind speed to calculate a daily "met factor." Although simple in derivation and form, the met factor explains over 80 percent of the day to day variability of the peak CO measurements. The authors didn't use this index or present any alternatives.

P. 9 1st Paragraph

"The DHS I/M evaluations used a related (and equally inappropriate) statistical method as well as questionable data (PSI CO readings)."

The questionable data was obtained from the same source as used by the consultant. The authors provide no support for this statement; however, if the data is questionable, then it equally effects their study. The comment on inappropriate statistical methods is also not supported. We are confident that we have not violated a law of statistics by comparing the best-fit curves of CO to a meteorological index for the same month of consecutive years.

P. 9 2nd Paragraph

"In closing this discussion, we can only say that the effectiveness of the Arizona I/M program with respect to ambient air quality remains to be tested empirically."

The emission test report by EPA (Rutherford) and the exhaust measurements by Hamilton Test Systems are empirical evidence that the Arizona I/M program reduces emissions.
It is common sense and an indisputable, scientific conclusion that reduced emissions necessarily result in lower ambient concentrations.

P. 13 Last Paragraph

Hypotheis #1 - "Beginning in January, 1977, ambient CO air quality in Phoenix and Tucson improved."
Hypothesis #2 - "Beginning in January, 1979, ambient CO air quality in Phoenix and Tucson improved again."

These hypothesis statements reflect the lack of thought given to the CO air quality situation. For these hypotheses to be a valid test of the expected effect of the I/M interventions, the following conditions should be demonstrated to exist:

1) The potential increase in emissions due to VMT are exactly counterbalanced each year by the introduction of cleaner new cars.
2) The meteorological conditions were either the same each year or they are explicitly controlled in the statistical procedure.
3) The interventions are discrete and of sufficient magnitude to be founded by the time series model.

Apparently, no preliminary tests were made to check these basic considerations. The stated hypotheses are therefore unsupportable and present an unfair test of I/M effectiveness.

High quality data in the geographic area affecting each monitor, vehicle registration, and emission data (grams/mi) for the vehicle fleet are essential to verify the first condition.
Since meteorological conditions frequently vary substantially from year to year, this variability must be accounted for in addition to the simple adjustment for seasonality, in order to satisfy the second requirement.
As regards the third condition, hypotheis #1 depends on a discrete startup of I/M in January, 1977, despite the fact that the authors are aware of empirical data supporting the conclusion that the program had an effect in 1976. The intervention is therefore probably not discrete, startup being diffused over a two-year period.
Moreover, the initial (1st year) reduction on emissions (according to MOBILE2) is only 10 percent. Under these circumstances, time series analysis may not be expected to find this intervention. Finding the 1979 intervention is an even more dubious undertaking, since MOBILE2 estimates an impact of less than 5 percent in the year following this intervention. The exhaust measurements, which are essentially ignored in the report, may be the only means of detecting these minor year to year reductions in fleet emissions.

P. 18 2nd Paragraph

"To illustrate this logical point, note that the germane threats to internal validity are limited to:

2) Federal new car standards relating to fuel efficiency: In a time series context, we expect CO levels to drop gradually over a period of years due to this factor alone."

Same mistake regarding fuel efficiency as noted earlier. Additionally, the list of threats to internal valdity doesn't mention meteorology this time. This is a particulary bothersome oversight because meteorological variability is probably the most significant of the uncontrolled variables on a year to year basis. One must suspect that the authors didn't want to include meteorology because its trend couldn't be characterized as gradual in any direction.

P. 19 2nd Paragraph

"We expect federal new car standards, improvements in fuel efficiency, changes in vehicle miles traveled, and other fleet changes to be realized incrementally. When Detroit releases a new car model, in other words, we do not expect CO levels in Phoenix or Tucson to respond discretely... I/M impacts, on the other hand, are expected to be discrete. When the I/M interventions occur, that is, we expect a coincident reduction in CO levels distributed over the subsequent year."

Changes in meteorology can obviously be discrete and of sufficient magnitude to totally mask a year to year I/M benefit expected to be in the 3-10 percent range. New car standards also exhibit discrete steps. For example, for the 1975-1979 model years, the emission standard was

15 grams/mile, then starting with the 1980 models, the applicable standard was 7 grams/mile. Obivously as the model year 1980 cars were introduced to the fleet, it would represent a discrete change in the rate of emission reduction due to new cars. Interestingly, this step change in the new car function occurred during the 1979 I/M intervention. Obviously, since the authors didn't address the new car standards, they inadvertently determined that the federal change from 15 to 7 grams per mile, also had no salutary effect.
It is apparent from the pre-I/M air quality trend that the effects of cleaner new cars is discrete.

P. 19 Last Paragraph, Con't on P. 20

"In our analyses, we will ordinarily report one trend (increase or decrease) figure with the understanding that this trend is the confounded effect of many uncontrolled forces. Whatever the source or cause of a trend, however, we can distinguish between incremental effects (trends due to factors other than the I/M interventions) and discrete effects (shifts due to the I/M intervention). The real implication then is that we will have little difficulty finding the discrete impact of the I/M program."

The first sentence quoted above is correct. The verbiage after the first sentence is a continuation of a proposition that the time series analysis is very sensitive to I/M because it alone among the independent variables is discrete (nonincremental). As shown in our previous comment, new car standards operate on air quality in a analogous way to the 1979 I/M intervention. Again, meteorology can be shown to exhibit discrete changes in consecutive years. If the unique (discrete) nature of I/M interventions is as important as the authors seem to be saying, then our comments on this matter should severely limit the strength of the conclusions stated in the report.

P. 20 2nd Paragraph

"In the decade prior to implementation of I/M, ambient annual CO levels decreased by 63 percent. In the first year of the I/M program, the statistic decreased by 23 percent but in the second year of the I/M program, the statistic increased by 14 percent."

The above statement refers to Table 1. The first sentence is correct. The second sentence is not and should read as follows: In the first year of the I/M program, the annual mean CO concentration at the central monitor was 3.3 percent higher than the previous year, but in the second year, the annual mean was 22.6 percent lower. In addition to the mistake, it is also revealing that the second year (1978) was the year analyzed for impact of the first intervention and no significant impact was concluded. 22.6 percent not significant!

P. 20 3rd Paragraph

"These annual figures say nothing about the impact of the I/M program CO air quality because gross annual difference mask a steady downward trend in ambient CO."

The ambient CO trend is not steady, but rather resembles a hockey stick with remarkably different slopes over different time periods.
It is interesting that at least in this context the authors are aware of gross annual differences. These gross differences also necessarily appear in the monthly data used in the study. The causes of the gross year to year differences is predominantly the results of meteorological variability and the changing balance of cleaner new cars versus increasing traffic. Since these gross differences are greater in magnitude than the expected effect of I/M, it is essential that they be understood and controlled. The authors failed to investigate this and did not control for these confounding effects.

"Nevertheless, it appears that increases in vehicle traffic are offset by increases in fleet fuel efficiency. This suggests that vehicle traffic increases will not be a potent threat to internal validity."

A real threat to internal validity is dismissed based on a misconception of the effects of fuel efficiency.
We are aware that "runs" were made to test for I/M commencing in 1976. However, this treatment is not proper either, because only a partial change in public behavior apparently occurred and avaiable data (failure rate) is not a measure of this change. This ambiguous (non-discrete) startup of I/M should have been studied more thoroughly because it may diffuse the initial impact of I/M to the degree that time series might not be an appropriate analytical tool.

P. 20 Bottom Con't on P. 21

"The many diverse forces reflected in these annual statistics (federal new car emissions standards, improved fuel economy, fleet changes, and so forth) do not seem to interact in any specified direction or manner. They appear to cancel each other in some respects, but overall, result in a weak net trend in annual CO levels. The effect of I/M program interventions, in contrast, are focused and easily measured as before/after changes. It is this aspect of the interventions which permit us to test null hypotheses.

Another example of poor (erroneous) analytical perception. The statement that the effect of I/M program interventions in contrast (to the effects of other forces on CO annual concentrations), are focused and easily measured as before/after changes is not supported. As can be seen on Table 1, even a stable statistic like annual mean varies radically from year to year in response to factors other than I/M prior to I/M intervention. For example, the mean statistic for 1971 decreased 17 percent from 1970; subsequently increased by 7 percent in 1972, followed by a 29 percent decrease in 1973, and a 14 percent increase in 1974. In this context, to speak of I/M as relatively discrete effect in comparison to other effects on CO is unwarranted. Furthermore, routine year to year net variability of the uncontrolled independent variables is greater than the year to year benefit of 5-10 percent claimed for I/M. The analysis is clearly and fatally flawed with respect to this item alone.

P. 21 1st Paragraph

"If we assume that the distribution of tests across this 12-month period was uniform, then a theoretically plausible coding for the intervention would be."

$$
\begin{aligned}
It &= 0 \text{ prior to } 1/77 \\
&= 1/12 \text{ in } 1/77 \\
&= 2/12 \text{ in } 2/77 \\
&\ \ \vdots \\
&= 11/12 \text{ in } 11/77 \\
&= 12/12 \text{ in } 12/77 \\
&= 1 \text{ thereafter}
\end{aligned}
$$

This coding for the 1977 intervention is inconsistent with empirical evidence and I/M theory. The following problems exist:

1) The effect of I/M was not 0 prior to 1977.
2) Available data indicates a nine month deterioration period for I/M effects (see work by Tiao).
3) I/M benefits according to theory and exhaust test data are cumulative. The treatment of I/M as a step function spread over one year is incorrect.

P. 28 1st Paragraph

"As the breakdowns in Table 3a and 3b clearly show, however, no threat to internal valdity can have the same effect on all indicators in all sites unless the threats interact in some bizarre manner. Since "Bizarre" means "improbably," we discount this possibility. To the extent that our findings are consistent across indicators and sites then, threats to validity cannot plausibly confound the conclusions drawn from our analyses."

"The breakdowns" show seasonal trends and the relative magnitudes of the four different averaging periods at the two sites. It does not contain information to support the claim that there are no threats to internal validity.

P. 36 Table 4c

In response to the DHS review of the last version of the consultant's report, the authors correct an error in the trend estimate for Phoenix-Sunnyslope mean 5pm readings (pre-1977) in the new Table 4d. Now, however, we see a bewildering set of additional changes in other trend estimates, which do not correspond to the previous Table 4c. The reason for these changes is not explained in the text, so the reader is left to establish for himself which set of values is accurate.

P. 31 Bottom Con't on P. 32

"The I/M program, in our opinion, should not have an incremental impact on CO concentrations. An incremental impact---a post-intervention trend---implies that I/M has a cumulative effect on CO. We can think of no theoretical justification for a cumulative impact and accepted I/M theories (the MOBILE2 model, e.g.) assume a linear relationship between CO concentrations and vehicle emissions."

As mentioned previously, I/M theory and exhaust measurements support the concept of a cumulative benefit. Apparently, the authors failed to familiarize themselves with I/M theory as coded in MOBILE2 and the corroborating data from the Arizona I/M program.

P. 33 2nd Paragraph

"Unless local weather conditions changed abruptly in January, 1977 and then changed again in January, 1979, meteorological effects are also implausible threats to internal validity. Our statistical analyses controlled meteorological variance in CO by seasonal ARIMA structures and, of course, these models give an acceptably accurate description of series variance (Table 4a). Our analyses of meteorological time series suggest further that these models give an acceptable degree of control for weather factors. But this threat, too, is moot because uncontrolled meteorological variance (background noise) in our CO time series could plausibly affect only the statistical power of our models. And if every single parameter estimate in Tables 4b-4d were significant, our conclusions would remain unchanged; null hypotheses of I/M impact would not be rejected."

The discussion on the page preceding the paragraph quoted, is similar to language which has already been criticized. The negation of meteorological variability as a threat to internal validity is unfounded. It is a fact that year to year changes in meteorology occur in Phoenix and Tucson and that these changes are a component of the corresponding differences in measured air quality. For empirical data demonstrating variability, the Table 1* on the next page which had been furnished to the Auditor General, shows the trend of a temperature difference measure, which is probably the best meteorological parameter to associate with ambient pollutant concentrations.

*See "Delta T (So. MTN minus Sky Harbor)" on the next page.

TABLE 1

Delta T (So. MTN minus Sky Harbor)*

Month	Year									
	1974	1975	1976	1977	1978	1979	1980	1981	Marginal	S.D.
January	2.2	6.9	5.5	.6	-.4	-1.6	-.7	3.0	1.94	3.1
February	6.0	5.1	.8	4.8	-2.6	3.3	.9	3.0	2.66	2.8
March	2.5	.3	.9	-.5	-.6	1.4	-2.3	-4.0	-.29	2.1
April	4.5	.9	1.4	.1	.5	3.5	.3	-2.1	1.14	2.1
May	2.5	3.0	.6	-1.4	1.6	-1.5	-1.7	-3.1	0.00	2.2
June	3.0	3.7	2.5	.3	5.6	1.7	5.3	-1.9	2.53	2.5
July	-4.3	-5.4	-4.5	-5.5	-1.4	.1	MSG	-7.1	-4.01	2.5
August	-.9	2.2	.5	-4.5	-.6	-.8	-3.4	-4.7	-1.53	2.5
September	-2.1	.4	-1.7	-2.3	-2.0	1.4	.6	-2.7	-1.05	1.8
October	-1.3	3.5	-.7	.2	2.8	4.9	1.8	-.1	1.39	2.2
November	6.1	7.2	3.2	3.8	.3	6.0	5.1	5.3	4.63	2.2
December	4.9	3.8	4.4	3.3	.2	6.9	5.8	4.5	4.23	2.0
Marginal	1.93	2.63	1.08	-.09	.28	2.11	1.12	-.83		2.3

*5pm to 5am minimum hourly readings averaged across months

The variability of meteorological conditions, year to year, and for the same months across years is large. There is abundant climatological data that refute the contention that significant meteorological variability is restricted to recurring seasonal patterns.

P. 35 Paragraph 1

"Given the remarkable degree of consistency---our findings were consistent across interventions, sites, CO indicators, and models---we have the greatest possible confidence in our major conclision."

Statements of this tone occur in several places in the report. The impression is conveyed that many independent tests were made whose results replicate to high degree. The tests were not as numerous and independent or the results as similar as the authors contend. For example,

1) The air quality trends at the two Phoenix sites are different.
2) The I/M impact estimates for the two Phoenix sites are different.
3) The four air quality measurements were used in Phoenix but only one measurement was used in Tucson.
4) The four air quality measurements in Phoenix are not independent, i.e., they are interrelated because the different averaging periods share common measurements.
5) The 1979 intervention was only tested in Phoenix since the Tucson data was truncated.

P. 38

- "VMT statistics are not available for CO monitor sites..." "These statistics are absolutely essential for any evaluation of the I/M program."
- "Comparing Arizona CO series with data from New Jersey and Oregon, it is clear that Arizona data are more seasonal and, otherwise, less structured. As a practical consequence, this means that relatively long time series will be required for any evaluation of the Arizona I/M program."
- "A real problem with interpreting future evaluations of the I/M program is that there is no status quo control. In an evaluation of the Oregon I/M program, for example, Taio et al, (1981) were able to compare CO levels in an I/M city with CO levels in a non-I/M city. As a principle of research design, comparisons of this sort ensure the validity of evaluation findings. Under ideal circumstances, the comparison increases the statistical power of the analytic model, allowing a precise measure of relatively marginal impacts. For this propose, a CO monitoring site should be established in a desert city not serviced by the I/M program."

Given the absence of these data requirements in the present design, it does not seem thoughtful to conclude "we can therefore state our conclusion with a high degree of confidence," without an accompanying statement of these reservations.

PIMA ASSOCIATION OF GOVERNMENTS
405 TRANSAMERICA BUILDING
TUCSON, ARIZONA 85701
792-1093

January 7, 1983

Mr. Mark Fleming
Auditor General's Office
OEPAD
1700 West Washington
Phoenix, Arizona 85007

Dear Mark:

Please forgive this late response to your inquiry concerning the I/M report. As Mr. Brown discussed with you over the phone, PAG has no direct comments on the report.

In its effort to update the Carbon Monoxide Nonattainment Area, PAG has reviewed the extent of the present nonattainment area and has recommended a reduction in size along the lines of the proposals in your report.

It is PAG's wish to use the minimum resources to achieve the goal of clean air in Pima County.

Sincerely,

Nick C Buchholz

Nick Buchholz
Air Quality Planner

NB/cr

UNITED TECHNOLOGIES
HAMILTON
TEST SYSTEMS

2303 East Thomas Road
Phoenix, Arizona 85016
602/955-9670

AUDITOR GENERAL

CONFIDENTIAL

January 20, 1983

HAND DELIVERED

Mr. Douglas R. Norton
Auditor General
State of Arizona
111 West Monroe Street
Phoenix, AZ 85007

Re: Hamilton Test Systems, Inc. Comments to Auditor General's Report 82-9 on the Vehicle Emissions Inspection Program

Dear Mr. Norton:

This letter is in response to your letter of January 6, 1983 requesting Hamilton to review a revised preliminary report draft of Finding IV of the subject report, which apparently represents your understanding of the changes agreed to during our December 10, 1982 meeting.

We appreciate the changes that have been made, but substantive errors of fact still remain, errors pointed out to you at the December 10 meeting and in our letter to you dated December 13, 1982, which summarized our understanding of the changes to be made. In addition, despite the fact that Hamilton has fully complied with all contractual requirements the bold face paragraph headings and the texts are purposely slanted to imply otherwise.

With respect to Finding IV, our specific comments are as follows:

1st Section (Page 31)

> Your repeated statement that "additional controls are needed to assure accurate and reliable emissions testing" is a conclusion that is not substantiated by any data, and indeed is not supported by the text of Finding IV itself.

2nd Section (Pages 31 and 32)

> Your statement that Hamilton Test Systems assures the accuracy and reliability of emissions testing equipment by conducting periodic field audits at the vehicle emissions inspections stations is not correct.

To ensure and maintain a quality program, Hamilton has established an extensive Quality Assurance Program of which weekly calibration checks are the primary control as outlined on the attachment titled "Functional Explanation and Summary of Weekly Calibration." The description therein and the accompanying tabulated data were again totally ignored, even though submitted in our letter of December 9, 1982.

The monthly Master Technician checks called out in our Maintenance Procedures are not a contractual requirement, but were implemented to duplicate the BVEI calibration audits. While the BVEI audits are an essential part of the State's surveillance of Hamilton's Quality Assurance Program, Hamilton's duplicate checks are not critical to system reliability or integrity and any attempt to depict otherwise is gross misrepresentation.

Your statement that the DHS-HTS contract requires HTS to calibrate analyzers each week is also incorrect. The contract specifies "periodic calibration checks" and "periodic recalibrations" (see Contract Section 6.6 - Maintenance). HTS chose to implement weekly calibration checks as the primary quality control check based on proven equipment design and performance. *

3rd Section (Pages 32 and 33)

The statements that "HTS did not conduct appropriate number of audits" and "further, HTS did not conduct all field audits required by internal company policy," along with "Table 5 representing to show HTS performed only 53% of required audits," appear to be deliberate attempts to imply failure to perform on Hamilton's part, and to discredit system integrity and reliability which is second to none in motor vehicle inspection programs. What are the appropriate or required numbers? "Appropriateness" is dictated by technical factors and "requirements" by contract.

As stated previously, and acknowledged by your report, there is absolutely no requirement for monthly checks by Hamilton.

Curiously your report fails to take full account of the following undisputed facts:

1. Hamilton has complied with all contractual requirements.

2. Hamilton performed 6,271 weekly calibrations during the audit period July 1981 through June 1982.

* Text has been changed to reflect this correction

3. During the audit period July 1981 through June 1982 only 39 calibration failures to meet BVEI specifications were encountered for a failure rate of only 0.62%, which is exceptionally low.

4. In addition to the weekly checks, Hamilton conducted 243 monthly calibrations during this period (although not required by contract) that duplicated the BVEI monthly audits, at its own expense.

5. Since April 1982 at the request of BVEI, Hamilton has voluntarily instituted semimonthly calibrations duplicating the BVEI audits, at no cost to the State.

Of even greater concern to us is your failure to provide HTS with a copy of Findings I, II and III of the report together with the appendices thereto, despite numerous verbal and written requests therefor. Having been given access to the first draft report and invited to the first meeting on December 10, it is difficult to understand why you have steadfastly refused to provide the entire report to us, particularly taking into account HTS' responsibilities under A.R.S. S 36-1775 *et seq*. As the party responsible for conducting the tests for the past seven years, we believe Hamilton is uniquely qualified to comment on the entire report, and we respectfully request one last time that you allow us to review the entire draft report and give us sufficient time to comment on it.

Very truly yours,

HAMILTON TEST SYSTEMS, INC.

Anthony J. Arrigo
Manager - MVI Programs

AJA:jm
attachment

ATTACHMENT

FUNCTIONAL EXPLANATION AND SUMMARY OF WEEKLY CALIBRATION
JULY 1, 1981 THROUGH JUNE 30, 1982

The Hamilton maintenance procedure requires that a calibration check of each analyzer in every lane be performed weekly as a routine. This calibration check is exclusive of and in addition to those performed as a result of some other maintenance function; for example, installation of a new infra-red source. This calibration routine is performed by introducing named calibration gases through each of the four analyzers in each EMS: HC, CO, HHC, and CO_2, and recording the results. The calibration tables in the system software contain the BVEI limits for each gas in the contractual ranges. If a calibration gas reading exceeds the BVEI limit, the software routine locks that EMS and prevents vehicle testing until the problem is corrected. In the event of a "Fail" reading, two "Pass" readings must be obtained after adjustment with the same gas in order to resume testing.

As a foundation for the above calibration procedure, the software also contains all analyzer serial numbers and conversion factors, gas bottle values, and bottle serial numbers. All computations involving the calibration routine are based upon analyzer conversion factors, bottle values, and actual readings from the analyzers. Final calibration results are electronically computed and recorded on the station level floppy disk pending transmission to headquarters.

WEEKLY CALIBRATION CHECKS
JULY 1981 - JUNE 1982

MONTH	NUMBER OF ANALYZER CHECKS PERFORMED	ANALYZERS WHICH FAILED TO MEET BVEI CONTRACT SPECIFICATIONS
July 1981	473	7
August	505	8
September	490	3
October	552	1
November	523	3
December	504	2
Jan. 1982	486	6
February	498	2
March	607	Ø
April	494	Ø
May	496	1
June	643	6
TOTALS	6271	39

CONSULTANT'S REBUTTAL TO WRITTEN RESPONSES

Our report concludes that the I/M program has had no salutary impact on CO concentrations in Arizona. The report is "inaccurate" in only one respect, however; it does not accurately reflect the amount of time and resources devoted to the research. We report the results of only 45 model analyses, for example, but in fact, we analyzed several hundred time series models before we arrived at our conclusion. Many of the comments made by DHS seem reasonable --- although some do not. In any event, each of the points raised in these comments was addressed to some extent or another in our research. While the analyses addressing these points may not have been explicitly described in our report --- although some were and others were implicitly described --- none of the comments were ignored in our broader research project. Our analyses were most comprehensive in this respect and, after reviewing the DHS comments, we must stand by our major conclusion.

The DHS comments raise several consistent objections to our analysis. We will comment on these objections in general, not addressing the specific source or point.

1) Annual weather variation - Many objections focus on annual weather variation that might obscure an I/M impact. First, in our opinion, it is most unlikely that year-to-year changes in weather could coincide exactly so as to make I/M look ineffective; this would be a "coincidence" in the truest sense. Nevertheless, we examined this possibility with in-depth analyses. In addition to the seasonal ARIMA models (which rely on the heavy seasonality of Arizona weather), we used relative humidity, temperature, windspeed, and Delta T inversion measures. None of these analyses generated an I/M impact of the sort claimed by DHS. After this extensive investigation of meteorological variables, our conclusions were unchanged.

2) Annual Traffic Variation - Our opinion with respect to traffic variation is similar to our opinion with respect to meteorology. We examined two relatively distinct measures of traffic but none of these analyses generated an I/M impact of the sort claimed by DHS. Our conclusions remained unchanged even after considering these variables.

The analyses shown in our Tables 5a-5c are typical of the results from analyses where meteorological and traffic variables were explicitly controlled. No other model that we analyzed produced a statistically significant salutary impact, however.

3) Periodic or Spectral Models - One comment raised the possibility that standard ARIMA models may not adequately describe ambient CO processes; it was suggested that some form of "Periodic ARMA" model would give a better description. We examined models of this sort for both Phoenix and Tucson but still found no I/M impact of the sort claimed by DHS. Of course, it is unlikely that one model would fit all CO time series but we addressed this problem in a straightforward manner by examining hundreds of models. If redundancy gives us any protection from logical errors, we are more confident in our conclusions.

4) Functional Forms - We examined at least three functional forms for each time series: linear trend, log-trend, and some "best" trend in between these two extremes. Regardless of functional form, however, we found no salutary impact for I/M.

5) I/M Coding - Several comments question our coding of the I/M impact as an abrupt

change in CO concentrations. We are most sensitive to this comment and we have been since the start of this research. In fact, we examined at least five different types of I/M impact; three of these impact-types are explicitly described in our report. Regardless of impact-type, we found no salutary I/M impact. We cannot address each distinct impact-type raised in the DHS comments due to deadline restrictions. But of course, we examined each of the impact-types suggested in the DHS comments and found no salutary impact of I/M.

6) Emissions Factors - We do not dispute the fact that emissions tests of vehicles are empirical data. Several commentators, especially Mr. Phil Lorang, have raised this point. On the other hand, our research did not deal with this phenomenon and for a good reason. We do dispute the implied statement that tailpipe emission tests can be used as empirical evidence that 1) I/M reduces CO emissions, and 2) that I/M has a salutary impact on abient CO air quality. To evaluate I/M impacts on ambient CO air quality, we must look directly at indicators of ambient CO air quality. Tailpipe emissions tests are not "empirical evidence of I/M effectiveness" in this sense.

7) Fuel Economy - Several commentators have noted our claim that fuel economy and CO emissions are related; they have disputed this claim. First, none of our analyses depend on or are based on this premise and, hence, it has nothing to do with our conclusion. Second, however, there is no strong consensus on this issue in the scientific community as the DHS comments seem to imply. See, e.g., "U.S. Automotive Emissions Controls: How Well Are They Working?" by Professor L.J. White, Am. Econ. Review 72(2), May, 1982). Professor White says, "To the extent that fuel economy improvements involved increases in the pure efficiency of combustion, this would have decreased HC and CO emissions."

In summary, we can only say that none of the comments or objections raised by DHS are surprising. Each was addressed in one way or another by our reseach. None of these comments or objections would cause us to change our conclusions in any substantial way.

Richard McCleary, Ph.D.

Barbara C. Nienstedt

APPENDIX I

CONSULTANTS' REPORT TO THE AUDITOR GENERAL

Revised, January 20, 1983

Time Series Analysis of the Impact of the Vehicle Emissions Inspection Program (I/M) on Ambient Air Quality in Phoenix and Tucson*

Richard McCleary, Ph.D.

Barbara C. Nienstedt

* This report is a revised version of a preliminary draft delivered to the Auditor General on January 6, 1983.

ABSTRACT

In January, 1977, a mandatory vehicle I/M program was implemented in Phoenix and Tucson. Two years later, I/M standards were raised significantly. Both of these policy interventions were aimed at reducing ambient air pollution in the cities. An in-house evaluation of the 1977 intervention found that mandatory I/M had reduced ambient CO concentrations in Phoenix by more than 30 percent. Because this estimate relied on an inappropriate statistical analysis of questionable data, however, this finding cannot be taken seriously.

The research reported here is an evaluation of the 1977 and 1979 interventions on ambient CO concentrations in Phoenix and Tucson. Our research design is a variation of the "time series quasi-experiment." Our data include four conceptually distinct measures of monthly ambient CO concentrations recorded at three distinct sites in two cities. The data were analyzed with ARIMA time series models and methods.

Our findings are, essentially, that neither the 1977 intervention (mandatory I/M) nor the 1979 intervention (higher I/M standards) had any statistically significant impact on CO levels in Phoenix and Tucson. More specifically, we found no evidence whatsoever that the I/M program reduced ambient CO levels in Phoenix and Tucson. This "no-evidence" finding, furthermore, was consistent across a diverse range of CO indicators, monitor sites, cities, and interventions; it was robust to gross changes in statistical models and assumptions. We can therefore state our conclusions with a high degree of confidence.

1.0 Introduction

The U.S. Clean Air Act of 1970, as amended in 1977, requires that states implement car inspection/maintenance (I/M) programs in major cities to reduce hydrocarbon and carbon monoxide (CO) emissions from light-duty vehicles. In May, 1974, the Arizona legislature authorized the Arizona Department of Health Services (DHS) to conduct comprehensive annual emissions inspections in all counties with populations over 350,000. The Arizona I/M program was similar to programs implemented in other states except that it was the first program in the U.S. operated by private contractors. Beginning in January, 1976, the I/M program required emissions inspections of all vehicles applying for registration; maintenance and repair (of vehicles failing inspection) was voluntary, however. Initial I/M regulations required that CO and hydrocarbon tailpipe emissions be tested at high cruise (approximately 50 mph), low cruise (approximately 30 mph), and idle. Compliance with standards for both CO and hydrocarbon pollutants was determined in all three test modes.

Debate over the effectiveness of I/M caused the legislature to reconsider the program in its first year of operation. The decision was finally referred to the voters in the 1976 general election. Prior to the referendum, however, the I/M program was reformed in several ways. Older vehicles were exempted from I/M and compliance with pollution standards for nonexempt vehicles was determined at idle speed only. "Loaded" test modes (high and low cruise) were used for diagnostic purposes only. Finally, vehicles failing retest would be issued a waiver if the repairs required to meet I/M standards cost more than $75. With these modifications, the voters retained the I/M program by a six percent margin.

The research described in this report is an independent evaluation of I/M effectiveness in reducing ambient primary pollutants in the two metropolitan areas served by the I/M program. The questions addressed in this evaluation include:

- What were the effects of the I/M program on ambient air quality?

- What aspects of the program were most or least effective with respect to air quality? And

- How is ambient air quality likely to change in the future?

We address these questions with analyses of monthly CO air quality time series from three sites served by the I/M program. Our research follows the work of Ledolter _et al_. (1979) and Tiao _et al_. (1981) who used time series analysis to evaluate I/M programs in New Jersey and Oregon. Our research is different from this work, however, in that we specify no structural relationships among emissions, traffic, federal interventions, and meteorology. We instead build naive ARIMA "noise" models to test _null hypotheses_ of I/M impacts on ambient CO; our work more closely resembles the work of Box and Tiao (1975) in this respect. Because our models do not include all exogenous variables and relationships, we cannot estimate the impacts of emissions, traffic, meteorology, federal interventions, and so forth on ambient CO. Our models "control" these exogenous variables nevertheless --- through trends, seasonal components, and other "noise" structures --- permitting unbiased estimates of I/M impacts on ambient CO. Our evaluation deals only with these I/M impacts, of course, so this shortcoming of our analysis is not a real concern. To make this limitation of our research explicit, we note that our evaluation will not address a broad range of questions which include:

- What were the effects of federal new car standards on ambient air quality?

- What were the effects of changes in traffic on ambient air quality?

- What were the effects of meteorology on ambient air quality?

In addition to I/M impact estimates, our analyses will estimate the change in ambient air quality not due to I/M. While we can attribute the gross residual change to the joint effect of these variables, we cannot untangle their unique effects. This would require a full structural equation model which, given the limited scope of this research, would not be warranted.

2.0 How the I/M Program Works and Why it Might not "Work"

Since our evaluation tests null hypotheses of I/M impact, the underlying theory of the I/M program is an important issue. There are only three sure methods for reducing mobile source primary air pollutants in the atmosphere. These methods are

1) Reducing vehicle miles traveled by building mass transit systems, taxing motor fuel at a pronibitively high rate, etc;

2) Reducing fuel consumption independent of vehicle miles traveled by making vehicles more fuel efficient; and

3) Making internal combustion engines run "cleaner" independent of fuel efficiency and vehicle miles traveled.

These three methods are components of a broad national policy aimed at improving air quality in U.S. cities. The second and third methods describe (at the risk of oversimplification) federal "new car standards" which, since the late 1960s, have grown increasingly stricter. Although federal new car standards have led inexorably to a cleaner vehicle fleet, they guarantee only that cars will be cleaner and more fuel efficient at the time of manufacture. Thereafter, routine I/M is required. New car standards are best enforced at the federal level. I/M falls neatly into a traditional state jurisdiction, however, and, thus, is best enforced at the state level.

The theory behind I/M programs may be described this way. Vehicle owners cannot or will not voluntarily inspect and maintain their vehicles. Some owners, for example, may not *know* that their vehicles need maintenance or repair. Others owners may know that their vehicles require maintenance but, as a matter of personal taste, will *choose* not to perform the maintenance or repair. I/M must therefore be forced on vehicle owners. As a routine annual procedure, I/M informs owners when and if their vehicles require maintenance; owners who would not then perform the necessary maintenance are forced to do so in order to relicense their vehicles. As a net result of this program, no emission-offending vehicle will go without routine maintenance for more than eleven months. Now summarizing this discussion, we may diagram the theory underlying the I/M program as

```
              Owner Behavior
I/M -----[                      ]----- Cleaner Air
              Vehicle Emission
```

What this diagram means, simply, is that, *in theory*, the I/M program directly impacts owner behavior and vehicle emissions. Its impact on ambient pollution is entirely *in*direct, however. The goal of the I/M program (cleaner air) cannot be realized, in other words, unless it impacts owner behavior and vehicle emissions.

In theory, the I/M program should accomplish these goals in a relatively straightforward manner. Examining the theory's assumptions, however, we see that there are a number of plausible reasons why the I/M program might not "work." Specifically:

- The I/M theory assumes that owners who would voluntarily maintain their vehicles are not aware that their vehicles need maintenance. Since new car warranties often require routine maintenance, this may not be an appropriate assumption for a great number of owners.

- The I/M theory assumes that some owners will not voluntarily maintain their vehicles even when made aware of the need. Since fuel prices have more than quadrupled in the last decade, however, and since routine maintenance may also optimize fuel efficiency, this may not be an appropriate assumption for a great number of owners.

Any I/M program assumes that vehicle owners <u>can</u>not or <u>will</u> not voluntarily perform I/M behavior. If this assumption is unwarranted, the program has not accomplished its primary goal of changing vehicle owner behavior. If vehicle owners were voluntarily performing I/M prior to 1977, in other words, the program would not have an impact on ambient air quality. Furthermore:

- The I/M theory assumes that testing procedures have low false positive ("clean" vehicles incorrectly tested as "dirty") and false negative ("dirty" vehicles tested as "clean") rates. False negative tests will limit the I/M program's effectiveness. False positives, however, will inflate the apparent effectiveness of the I/M program by requiring repair of vehicles that are already "clean." In either case, I/M assumes that field emission tests are reasonably accurate and reliable. The evidence for this assumption is not totally convincing. If emissions offenders in fact cannot be detected with any degree of accuracy, the program cannot possibly work.

- The I/M theory assumes that program-induced behavior is synonymous with emissions-reducing behavior; that any action a vehicle owner takes to pass inspection, in other words, will necessarily result in reduced emissions. But vehicle owners might instead be performing maintenance solely for the purpose of passing the annual test. If this is at all plausible, then the I/M program is not accomplishing its secondary goal. Again, the evidence on this point is is not totally convincing. Public opinion surveys, for example, suggest that some owners repair their vehicles only for the limited purposes of passing the I/M test.

We could expand on the assumptions underlying the I/M program, commenting on the plausibility of each, but this basic outline serves our purposes. We are not arguing that I/M is ineffective or that the theoretical basis for its effectiveness is wholly implausible. Instead, we have tried to demonstrate that there is no basis for believing that the I/M program "works" or that it does not "work". Supporters of the I/M program argue that, given the "commonsense" nature of the program, its effectiveness should be self-evident. But as we have tried to show, commonsense arguments can support both sides of the issue. Lacking empirical evidence, both positions are plausible.

3.0 Empirical Evidence

Supporters of the I/M program argue that it has reduced ambient (CO) pollution (in Phoenix) by 30 percent or more; without a vigorous I/M program, the atmosphere in Phoenix and Tucson would be poisoned in a short time (DHS, 1979). Supporters also cite I/M evaluations in other states as evidence that I/M works in Arizona[1]. Detractors of I/M argue, conversely, that air quality has not improved markedly since 1977, implying that I/M has not been effective; one need only look at Camelback Mountain to see how effective the program has been. Of course, I/M effectiveness cannot be assessed independent of history. Changes coinciding with the I/M program (e.g., population growth) could easily explain why air quality would deteriorate even if the I/M program were spectacularly effective. Both supporters and detractors of I/M are wrong in one respect, however. The effectiveness or ineffectiveness of I/M is not self-evident nor does the empirical evidence support either view.

Computer simulation studies (which are not empirical evidence) support the argument that, without I/M, air quality would have deteriorated more rapidly than it has. The Systems Modeling and Research section of the Arizona Department of Transportation, for example, concluded from a "Monte Carlo" simulation that federal new car standards merely nullified the effect of increased traffic between 1976 and 1979 (Arthur, 1979). Discounting meteorological factors, this implies that CO air quality in 1977-78 would have been virtually identical to 1976 levels without the I/M program. A California computer simulation (Office of the Legislative Analyst, 1980) arrived at a similar set of conclusions. Both studies demonstrated that (simulated) air quality improvements could not be attributed solely to federal new car standards; both attributed air quality improvements to I/M by default.

In another computer simulation, using the EPA emissions model, MOBILE2[2], and linear "rollback" methods, Neuroth and Poynter (1982) projected ambient CO concentrations in Phoenix under several sets of hypothetical circumstances. They found that

> "...compliance with the 8-hour CO standard would only be approached by 1987 if EPA were to relax the ambient standard to allow 5 exceedances of the required CO pollution levels, and if the Arizona I/M program were modified by tightening standards and extending the testing to vehicles older than the present thirteen year old vehicle cutoff. All other combinations of proposed changes would not allow compliance by the year 2000."

The linear rollback method assumes that air quality in a given area varies one-to-one with emissions, so these projections are valid only to the extent that this assumption is valid.[3]

The problem with computer simulation models generally is that, to demonstrate I/M effectiveness, the models must assume that I/M is effective to some extent. Many models assume, for example, that emission tests are accurate and reliable (low false positive and negative rates, i.e.) and that all repairs are directed toward the primary I/M goal (not merely toward passing the test, i.e.). These models assume, in other words, that the I/M theory, as we have outlined it, is wholly correct. The I/M theory is only a theory, however, and simulated evidence in this sense is not evidence.

The body of empirical evidence on the effectiveness of the Arizona I/M program, on the other hand, is inconclusive at best. Two limited evaluations of the program do conclude that I/M is effective in reducing ambient CO (in Phoenix) but both studies are fatally flawed.

The Arizona DHS (1979) analyzed the impact of the I/M program on PSI CO concentrations in Phoenix and, controlling for meteorological and traffic variance, found a substantial I/M impact. The research controlled meteorological variance in CO levels with a "delta temperature" factor, a direct measure of temperature inversion. Variance in CO levels due to traffic was controlled (apparently) by analyzing month-long periods of weekdays. Data were analyzed with ordinary least-squares regression algorithms.

A second evaluation, conducted by Rutherford (1981), compared Phoenix CO emissions with emissions from low-altitude non-California sites without I/M programs. The Emissions Factors (EF) Testing Program had contracted with independent laboratories to perform emission testing in several U.S. cities. Phoenix was an EF test site for several years before and after the inception of its I/M program. Analysis of covariance for Chicago, Houston, St. Louis, Washington, D.C., and Phoenix showed a reduction in emission means due to the I/M program.

Considering that this study analyzed only emissions --- not air quality per se --- its finding says nothing about I/M impacts on air quality. Analysis of covariance in any event ignores temporal relationships among data. The DHS I/M evaluation used a related (and equally inappropriate) statistical method as well as questionable data (PSI CO readings). In both cases, I/M impact estimates are at least biased.[4] Conclusions drawn from either evaluation must be interpreted cautiously, if not wholly discounted.

We have not considered the adequacy of the research designs used in these two evaluations. By "research design," we mean some logical system, couched in terms of pertinent validity issues, which supports scientific inference. Both evaluations used flawed designs or, perhaps, used no design at all. The other weaknesses of these evaluations, particularly the statistical issues, render this point moot. Design is a most crucial issue of this evaluation and and we will now devote considerable effort to explaining our research design. In closing this discussion, we can only say that the effectiveness of the Arizona I/M program with respect to ambient air quality remains to be tested empirically.

4.0 Research Design: Construct Validity

Construct validity, which we define in very general terms as the link between theoretical constructs and their empirical referents (or with their representations, operationalizations, measures, etc.), is the sine qua non of any empirical research. Our theoretical construct is "air quality in areas served by the I/M program." There are infinitely many empirical referents of "air quality in areas served by the I/M program" and the dilemma is that none depicts this theoretical construct perfectly.

In popular conversation, "air quality" means photochemical smog because, probably, people can see smog. Primary pollutants such as CO, hydrocarbons, and ozone, in contrast, are invisible and, hence, do not make for good popular conversation. But the I/M evaluation literature has focussed on ambient CO. Following this tradition, our evaluation of the I/M program will deal only with its impact on "CO air quality."

"CO air quality" is not synonymous with "air quality," of course, but there are two justifications for this narrower definition. As a practical matter, CO data are readily available for both Phoenix and Tucson; measures of other pollutants are not so readily available. More important, however, ambient CO is related to public health.

The critical link between vehicle emissions, CO air quality, and public health rests on the concept of a "dose response." Primary vehicle emission pollutants such as CO constitute a "dose" to which populations are exposed. Impaired lung function, eye irritation, and similar physiological effects are common "responses" to CO doses. CO "dose response" phenomena are not fully understood. People of widely different ages, diets, lifestyles (especially smoking versus non-smoking), and other sensitivities, for example, respond differently to similar CO doses. The responses to high CO doses are fairly well documented. Drowsiness, cardiac changes and pulmonary transformations are common results of CO exposure; even the molecular-level responses to high CO doses are well known (Office of the Legislative Analyst, 1980). However, the responses to relatively low CO doses (ten ppm., e.g.) over long periods of time (years, e.g.), the interaction of CO with other primary and secondary pollutants, and CO dose responses in the "real world" (versus the laboratory) are not well documented. The theoretical link to public health makes ambient CO an important (if not the most important) indicator of air quality. We must remain senstive nevertheless to the limits of our findings. We are evaluating the I/M program only with respect to "CO air quality." Our findings will not, strictly speaking, generalize to "air quality."

Limiting our research to "CO air quality in areas served by the I/M program," still leaves two unresolved construct validity issues: geographical area and CO readings. In both cases, there are no perfect empirical referents but, rather, a plethora of imperfect ones.

Consider geographical area, for example. One "area served by the I/M program" is Maricopa County but no empirical CO measure covers this entire geographical area. All we have are samples of the area taken from discrete CO monitors at two sites (7th Street and Butler and 18th Street and Roosevelt). The germane question, of course, is this: Do CO readings from these sites adequately depict CO air quality across the entire I/M service area? If one were studying real estate prices or demographics, these sites would not adequately depict the phenomena for all of Maricopa County or even for all of Phoenix. These sites probably do give adequate (though imperfect) depictions of CO air quality for the entire I/M service area, however.

In any event, we would be surprised if the I/M program improved (or did not improve) CO air quality at these sites but did not improve (or improved) CO air quality in the rest of Maricopa County. Because we can use two sites to depict the geographic area, our research design has some control over this construct validity issue. If we arrive at the same or very similar findings for both sites, we will have more confidence in our findings than if we arrive at very different findings for both sites.

The issue of construct validity as it relates to area is considerably less problematic than the issue of specific CO measures. Put simply, there are too many possible measures of CO air quality. For example:

- The average monthly CO reading at 5 P.M.
- The highest CO reading in a month
- Some average CO reading in a month

This list of measures is not exhaustive. In the case of monthly average CO readings, for example, there are a great many types of averages (many ways to combine discrete readings into an average, i.e.). Each type of average is related to every other type in some way, of course, so the distinction might appear petty. Nevertheless, it is conceivable that I/M could have an impact on one type of average but not on another.

Now the quality of any empirical referent qua referent is determined by the values of the researcher. Research aimed at assessing compliance with air quality laws, for example, might rely on one empirical referent while research aimed at assessing public health risks might rely on another; and research on health risks to the general population might rely on a different referent than research on the health risks to some subpopulation. Even when the values of the researcher are focussed on a limited context, however, there is no single "best" representation.

Quasi-experimental logic requires that analyses be generalized only to the specific empirical representation. If analysis shows that the I/M program has reduced the average 5 P.M. CO reading at a single monitor in Phoenix, that is, the reduction must be interpreted literally; it says nothing about the impact of I/M on the average 5 P.M. CO reading at other monitor sites (much less for Phoenix as a whole or for Tucson) or about the impact of I/M on some other CO measure. As a rule of logic, we must agree that the I/M program could impact one CO measure or one site or one city but yet have no impact on some other CO measure, site, or city.

But there are practical limits to this philosophical reservation. If we find that the I/M program is effective (or ineffective) on one CO measure, in one site, in one city, our conclusions will be weak. If we arrive at the same finding for another CO measure, in another site, in another city, our findings grow stronger. And if our research leads to similar findings for several CO measures read from multiple sites in two cities, we will have a great deal of confidence in our conclusions. In the end, replication is the only control in our research design for construct validity issues.

The probabilistic notion of validity underlying this dictum appeals to common sense. The scientific procedure associated with this type of logic, empirical "triangulation," requires that one accept this commonsense definition of validity. In terms of pure logic, of course, one can never be ultimately certain that an empirical conclusion is valid. Given a consistent set of findings, however, one can be reasonably certain. Adopting the criterion of consistency a priori, we will be reasonably certain of any conclusions we make. Our research will provide compelling evidence of the I/M impact only to the extent that our findings are consistent.

5.0 Research Design: Internal Validity

Beginning in January, 1977, I/M became mandatory. Due presumably to the putative effectiveness of the I/M program, test emission standards were raised in January, 1979. Beginning that month, significantly more vehicles failed an initial test and were required to be repaired and retested. The hypotheses of our research are deduced from these two interventions.

Hypothesis #1 - Beginning in January, 1977, ambient CO air quality in Phoenix and Tucson improved.

Hypothesis #2 - Beginning in January, 1979, ambient CO air quality in

Phoenix and Tucson improved again.

These hypotheses are transformed into null hypotheses by simple negation. In both cases, the expected improvement in CO air quality will be realized as coincident reductions in CO concentrations distributed over the subsequent years, 1978 and 1980. Reductions due to the second intervention are expected to be relatively smaller than the first.

The most appropriate design for testing these hypotheses is the so-called "time series quasi-experiment." Using the notation of Campbell and Stanley (1966; also, Cook and Campbell, 1979), the time series quasi-experiment may be diagrammed as

...0 0 0 0 0(I)0 0 0 0...

where each "0" is a monthly CO reading and where the "I" is an intervention. In the present case, for example, mandatory I/M was implemented in January, 1977. As a result of this intervention, we would expect the post-intervention CO readings to be statistically different than the pre-intervention readings. Specifically, we would expect them to have a lower level. Within the limits of logical inference, any difference (a reduction, e.g.) could be attributed to the implementation of mandatory I/M.

According to Campbell and Stanley, the time series quasi-experiment is the most powerful nonexperimental design for assessing the impact of planned interventions. Use of time series quasi-experiments to evaluate I/M programs has, nevertheless, been limited. This is probably due to the high level of technical expertise required to interpret time series data, the sophisticated software required for time series analyses, and the lack of readily available time series data. But three important studies have used time series designs and this experience guides our research in several respects.

First, with respect to statistical methods, we cite the work of Box and

Tiao (1975) who used this design in an environmental assessment of photochemical smog in Los Angeles. Recognizing that time series data are serially dependent, nonstationary, and highly seasonal, they state that

> "...the ordinary parametric-nonparametric statistical procedures which rely on dependence or special symmetry in the distribution function are not available nor are the blessings endowed by randomization."

This consideration cannot be overemphasized. Traditional regression methods (including analyses of variance and covariance) are inappropriate for evaluation of I/M programs. Instead, the AutoRegressive Integrated Moving Average (ARIMA) models and methods of Box and Jenkins (1976; also, Box and Tiao, 1965; 1975) must be used. We will use ARIMA models for analysis of all time series data in this research.

In addition to statistical requirements, prior research suggests a number of design issues. New Jersey was one of the first states to initiate an I/M program and, also, one of the first I/M programs evaluated as a time series quasi-experiment. In summarizing their analysis, Ledolter *et al*. (1979) note that

> "...from 1971/1 to 1977/6 CO concentrations at all seven New Jersey air monitoring stations decreased significantly. The average reduction is approximately 28 percent. This reduction can be attributed to the progressively more stringent federal CO emissions standards and to state programs such as the New Jersey car I/M program. Their relative contributions, however, are confounded and are best interpreted jointly."

The New Jersey study illustrates the salient shortcoming of all quasi-experimental designs; quasi-experimental evidence is subject to interpretational confounding by uncontrolled threats to internal validity.

For our purposes, we define a threat to internal validity simply as some factor other than the intervention which may cause a change in CO levels. The germane threat to internal validity in the New Jersey study is, as Ledolter et al. note, the coincident federal new car emission standards which, presumably, also lead to CO reductions.

Campbell and Stanley call the general threat to internal validity raised in this research "history." To illustate the confounding influence of this uncontrolled threat, suppose that we go to two doctors to be treated for an illness. If we are cured, the cure cannot be attributed uniquely to either treatment. Either one or both treatments could have produced the cure. In this specific case, given the coincident implementation of federal new car emission standards, Ledolter et al. could say only that the CO reduction due to I/M ranged between zero and 28 percent. In strictest logical terms (which may be inappropriate, of course, since this excerpt is taken out of context), this is tantamount to saying that the I/M program may or may not have had an impact.

The most appropriate method for controlling "history" as a threat to internal validity, according to Campbell and Stanley, is to incorporate a "control series" into the quasi-experimental design. With this feature, the time series quasi-experiment may be diagrammed as

```
...0   0   0   0   0(I)0   0   0   0...
...0   0   0   0   0   0   0   0   0...
```

The second series has no I/M intervention, so any pre- to post-intervention reduction in CO concentration must be attributed to all factors other than I/M --- to federal new car standards and all other threats to internal validity.

The New Jersey I/M evaluation had no readily available control series but the next major I/M evaluation, conducted in Portland, Oregon by Tiao et al.

(1981), was able to compare CO reductions in Portland with CO reductions in Eugene, where no I/M program had been implemented. Tiao et al. found that

> "At all Portland sites, one can observe a reduction in ambient CO concentrations over the 1970-79 period, with the average reduction ranging from 3.4% to 7.3% per year. The reduction in Eugene is less, the average being 1.9% per year."

This excerpt illustrates quasi-experimental logic. Since the analysis found a reduction in CO concentrations in Eugene (where there was no I/M program), the Portland reductions could not be attributed entirely to I/M. On the other hand, since the Portland reductions were larger than the Eugene reductions, it is unlikely that these effects were due entirely to uncontrolled threats to internal validity such as federal new car standards. The residual Portland reductions can be attributed to the I/M program.

One shortcoming of our research design is that we have no control city for Phoenix or Tucson. CO monitors in Flagstaff and Yuma (which are not serviced by the I/M program) have not operated long enough to provide the desired control data. Lacking a control city, point estimates of the I/M program impact may be confounded with uncontrolled threats to internal validity. Our design has strong controls for threats to internal validity relating to tests of the null hypotheses, however. We will have great confidence in our ability to say whether the I/M program did (or did not) have an impact on ambient CO then but we would have less confidence in our ability to state the size of any nonzero I/M impact.

Let us now examine these features of our design, assessing their ability to control germane threats to internal validity. First, our design is unique in that we have two interventions. As a practical matter then, we are dealing with a weak version of a design that Cook and Campbell (1979: Chapter 5) call the "switching replication" time series quasi-experiment. Each series has an inherent control for certain threats to internal validity.

Second, the I/M program has a specific type of impact which enables us to distinguish it from impacts due to the germane threats to internal validity. To illustrate this logical point, note that the germane threats to internal validity are limited to:

1) Federal new car standards relating directly to emissions: In a time series context, we expect CO levels to drop gradually over a period of years due to this factor alone.

2) Federal new car standards relating to fuel efficiency: In a time series context, we expect CO levels to drop gradually over a period of years due to this factor alone.

3) Increased vehicular traffic in Arizona: In a time series context, we expect CO levels to rise gradually over a period of years due to this factor alone.

4) Floods: Floods in Phoenix during the late 1970s and early 1980s severely disrupted traffic patterns for several months. The effect of this disruption on ambient CO is unknown but empirically evident. In any event, we expect CO levels to change discretely coincident with the floods due to this factor alone.

Now there are two important dimensions of these four threats. First, some of the threats (new car standards for emissions and fuel efficiency) will lead to decreases in CO levels and, thus, will bias null hypothesis tests in favor of the I/M program by exaggerating the I/M impact. Other threats (particularly increased vehicular traffic) will lead to increases in CO levels and, thus, will bias null hypothesis tests against the I/M program. Clearly then, these factors are potent threats to internal validity only in the context of a particular finding. If we find that the I/M program has effectively *reduced* CO concentrations, for example, increased vehicular traffic is not a plausible threat to internal validity; and if we find no effect, it is a plausible threat to internal validity.

With respect to null hypotheses then, these threats present no real obstacle to inference. Second, however, and more important, note that only one of these threats (floods) is expected to have a *discrete* impact on CO levels. We expect federal new car standards, improvements in fuel efficiency, changes in vehicle miles traveled, and other fleet changes to be realized *incrementally*. When Detroit releases a new car model, in other words, we do not expect CO levels in Phoenix or Tucson to respond discretely. Instead, we expect CO levels to respond gradually over a period of many months and years. For example:

```
o
         o
                  o
                           o
                                    o
                                             o
                                                      o
```

I/M impacts, on the other hand, are expected to be *discrete*. When the I/M interventions occur, that is, we expect a coincident reduction in CO levels distributed over the subsequent year. For example,

```
o    o    o    o
               I
               I    o    o
               I         I    o    o    o
               I         I
74   75   76   77   78   79   80   81   82
```

We cannot disentangle the many effects which feed a trend. In our analyses, we will ordinarily report one trend (increase or decrease) figure with the understanding that this trend is the confounded effect of many uncontrolled forces. Whatever the source or cause of a trend, however, we can distinguish between incremental effects (trends due to factors other than the I/M interventions) and discrete effects (shifts due to the I/M intervention). The real implication then is that we will have little difficulty finding the discrete impact of the I/M program.

Table 1 Here

Table 1 shows a set of annual time series data for Phoenix, collected from various sources, which illuminate this discussion. Annual average CO readings from 1967 to 1981 illustrate the potent effect of federal new car standards on CO air quality. In the decade prior to implementation of I/M, ambient annual CO levels decreased by 63 percent. In the first year of the I/M program, the statistic decreased by 23 percent but in the second year of the I/M program, the statistic increased by 14 percent.

These annual figures say nothing about the impact of the I/M program on CO air quality because gross annual differences mask a steady downward trend in ambient CO. Additionally, this is only one limited empirical referent of CO air quality; and these annual figures are apparently influenced by the floods in the late 1970s. Nevertheless, it appears that increases in vehicle traffic are offset by increases in fleet fuel efficiency. This suggests that vehicle traffic increases will not be a potent threat to internal validity. This is an empirical question, of course, which can only be decided in the context of our analyses.

The many diverse forces reflected in these annual statistics (federal new car emissions standards, improved fuel economy, fleet changes, and so forth) do not seem to interact in any specified direction or manner. They appear to cancel each other in some respects but, overall, result in a weak net trend in annual CO levels. The effect of I/M program interventions, in contrast, are focussed and easily measured as before/after changes. It is this aspect of the interventions which permit us to test null hypotheses.

6.0 The I/M Interventions

As a first step in our evaluation of the I/M program, we must decide on an appropriate operationalization of the I/M interventions. In January, 1977, I/M became a mandatory component of annual vehicle registration. By December, 1977, all eligible vehicles had been tested and, if necessary, repaired. If we assume that the distribution of tests across this 12-month period was uniform, then a theoretically plausible coding for the intervention would be

$$
\begin{aligned}
I_t &= 0 \text{ prior to } 1/77 \\
&= 1/12 \text{ in } 1/77 \\
&= 2/12 \text{ in } 2/77 \\
&\;\;\vdots \\
&= 11/12 \text{ in } 11/77 \\
&= 12/12 \text{ in } 12/77 \\
&= 1 \text{ thereafter}
\end{aligned}
$$

The rationale for this coding is fairly obvious. The second intervention, in January, 1979, can be coded analogously and witn the same rationale. For this coding, our time series models are of the form

$$CO_t = w_1 I_{1t} + w_2 I_{2t} + N_t$$

where CO_t is the <u>tth</u> observation of a CO time series

I_{1t} is the <u>tth</u> value of the 1/77 intervention

I_{2t} is the <u>tth</u> value of the 1/79 intervention

N_t is the <u>tth</u> observation of an ARIMA model

The ARIMA model contains all level and trend parameters as well as seasonal and nonseasonal structures. The parameters w_1 and w_2 are scalar weights which give the impacts of the interventions on the CO_t series. A test of the null hypothesis that these interventions had no impact on the CO series is

$$H_0: \ w_1 = w_2 = 0$$

This null hypothesis can be tested against the standard errors of w_1 and w_2 at a nominal level of confidence. If the null hypothesis is rejected, then w_1 and w_2 can be interpreted as estimates of the impact, subject to the pertinent threats to internal validity.

To illustrate these procedures, we will analyze a time series relating to our null hypotheses. Table 2 shows annual retest statistics (in thousands of vehicles) for Maricopa and Pima Counties. From these statistics, we can see that the I/M interventions are not at all minor or trivial. Vehicle retests in Maricopa and Pima Counties increased from zero to 110 and 40 thousand from 1976 to 1977; from 1979 to 1980, retests increased again by 80 and 69 percent respectively. From 1980 to 1981, however, retests dropped in both counties, making it appear as if I/M emissions standards had been dropped.[5] This is an empirical question, of course, which can be answered with analysis of a time series quasi-experiment.

Denoting the number of Maricopa County retests in the <u>tth</u> month by R_t, we posit the model

$$R_t = w_1 I_{1t} + w_2 I_{2t} + N_t$$

where I_{1t} and I_{2t} are step functions changing from zero to one on 1/79 and 1/81 respectively. The iterative identification/estimation/diagnosis model-building strategy of Box and Jenkins (1976; see also, McCleary and Hay, 1980: 2.11) leads to the N_t model

$$N_t = (1 - B^{12})^{-1}(w_0 + a_t)$$

where B is differencing operator, w_0 is a constant, interpreted as an annual

Table 1 - Selected Annual Time Series

Year	CO[a]	MPG[b]	VMT[c]	VMT/MPG
1967	8.2			
1968	7.4			
1969	5.9			
1970	4.7			
1971	3.9			
1972	4.2			
1973	3.0			
1974	3.5	13.4	16.46	1.23
1975	3.2	13.5		
1976	3.0	13.7	19.43	1.42
1977	3.1	13.9		
1978	2.4	14.1	22.08	1.57
1979	2.8	14.3		
1980	2.5	14.9	23.40	1.57
1981	2.6	15.5		

a, Annual Average CO in mg/m^3. Source: Arizona DHS
b, Cars Only. Source: Environmental Protection Agency
c, Total Million VMT. Source: MAG Transportation Planning Office

Table 2 - Annual Retests in Thousands

	1974a	1975a	1976b	1977	1978	1979	1980	1981
Maricopa County				110.2	106.6	189.7	189.1	164.7
Pima County				40.7	41.9	70.8	72.0	60.2

a, No I/M in 1974-75
b, Voluntary I/M in 1976

series trend, and a_t is a white noise random shock. Through a simple algebraic manipulation, we may rewrite this N_t expression as

$$N_t - N_{t-12} = w_0 + a_t$$

And in this form, the nature of the model is made clear. The constant w_0 is an annual increment in retests due, presumably, to all factors other than I/M policy interventions. The stochastic term of this N_t is white noise, indicating that there are no other (exogenous) structures affecting retests other than those captured in the constant. Estimates of the model's three scalar parameters are

w_0 = + 273.0 with standard error = 238.7

w_1 = +5759.7 with standard error = 413.5

w_2 = -2312.5 with standard error = 413.5

Interpreting these estimates, the retest series increases by 273 each year due to unmeasured exogenous structures. In January, 1979, however, the series increased by more than 5700 retests; the increase was permanent and abrupt. In January, 1980, the series <u>de</u>creased, again permanently and abruptly, by more than 2300 retests.

The null hypotheses associated with these two shifts are, in words, the retest series did <u>not</u> change in either 1979 or 1980. To test this null hypothesis, we compare the estimates of w_1 and w_2 with their standard errors. Since both estimates are more than twice the size of their standard errors (roughly equivalent to a t-statistic of $\pm$ 1.96 or the 95 percent confidence level), we reject both null hypotheses. The alternative hypothesis, which we accept, is that the retest series <u>did</u> change in both years.

We have used this exercise to demonstrate the interpretation of ARIMA time series models and the hypothesis testing procedure. We will evaluate the I/M program impact analogously. Substantively, the analysis indicates a shift in program standards from 1980 to 1981. We will accomodate this shift in later analyses by examining all plausible coding schemes for the I/M impact, including using the retest series per se as a measure of the I/M intervention.

7.0 The Data

A practical obstacle to our evaluation of the I/M program is that the state agency charged with operating the I/M program, DHS, does not routinely monitor air quality. Primary responsibility for collecting these data lies instead with the county health departments. And since these agencies have no primary responsibility for running or evaluating the I/M program, air quality data are not collected in an easily analyzed form. No county has a computerized data retrieval system, for example. Our air quality data consequently were collected from county logs by the Office of the Auditor General for the specific purpose of this evaluation. All time series were necessarily hand-calculated at a great cost. This practical factor limited the number of CO time series available for the analysis. We are nevertheless satisfied that the time series collected for the evaluation are sufficient to guarantee a fair test of the I/M null hypothesis.

The impact analyses, reported in subsequent sections, are based on four conceptually distinct CO measures: (1) highest 8-hour reading, (2) monthly mean highest 8-hour reading, (3) monthly mean 5 P.M. reading, and (4) monthly 1-hour mean. We say that these four indicators are "conceptually" distinct because, of course, they are not wholly unrelated. Since each measure is composed of discrete CO readings, each measure is correlated to some extent with every other measure. The degree of correlation is small as a practical matter, however, and this permits us to analyze each indicator as if it were indeed distinct.

One of the four indicators, the mean 5 P.M. CO reading, measures CO at a time of the day when, presumably, concentrations are at their lowest. Peak commuting hours traditionally begin at 5 P.M. and end at approximately 8 P.M. Due to meteorological peculiarities, the highest daily CO concentrations, in Phoenix at least, are thought to occur at 10 P.M. Since this indicator represents a "best case" CO concentration, it gives the I/M program its most conservative test. If I/M impacts this indicator, we would have to agree that it is a spectacularly successful program.

Another of the four indicators, the highest daily 8-hour concentration, is a "worst case" measure of CO air quality. Unlike the other three indicators, the highest 8-hour concentration is not a monthly mean but, rather, a single reading for any given month. This indicator has a much higher level than any of the monthly mean indicators and, presumably, could be more easily impacted by the I/M program. But this indicator also has a higher variance than the other three indicators and, thus, it could easily understate the statistical (though not substantive) significance of an I/M impact. On balance then, we might expect to find impacts in this indicator if the I/M program is at all effective; but we should not be surprised if these impacts are not statistically significant at a nominal level.

The third indicator, mean daily 8-hour high, represents the same "worst case" measure of CO concentrations. Because this indicator is a monthly mean, however, its monthly variance will be much smaller than any highest (single-day) 8-hour CO reading, making the task of statistical detection easier. The level of this indicator will also be lower than the single-day reading, and so, conceptually, it represents a mid-range test between "best" and "worst" case indicators.

The fourth indicator is a mean CO indicator composed of hourly readings from each day of the month. Because it is highly correlated with the total volume of CO in the atmosphere during the month (though not necessarily emitted in that month), this "monthly average" depicts the overall CO air quality in the I/M service area. In a sense, the monthly mean is an "indicator of indicators," and hence, represents a relatively conservative test of the I/M program's effectiveness. I/M could reduce CO concentrations at one time of the day or month, that is, without reducing the monthly mean statistic. The monthly mean statistic could dilute any highly specific I/M impact and we must remain sensitive to this issue.

The four indicators were collected from three monitoring sites: Phoenix Central, Phoenix Sunnyslope, and Tucson Central. Not all four indicators were readily available from each of the four sites, however. The Phoenix Central monitor was the best data source, providing all four CO series from January, 1974 to December, 1981. The Phoenix Sunnyslope monitor had no data prior to January, 1975. This source provided three of the four CO series from that date to December, 1981. Finally, the Tucson monitor provided only one CO series from January, 1974 to December, 1978. For Tucson then, we are unable to evaluate the impact of the 1979 I/M intervention.

Table 3a Here

The Phoenix Central and Sunnyslope series illustrate the strengths of basing an evaluation on several indicators collected at multiple sites. Table 3a, for example, shows annual statistics developed from each of the monthly time series. These annual statistics must be interpreted cautiously because they mask monthly variation. We use them here only for the limited purposes of demonstration. We note first that some of the Sunnyslope indicators show weaker annual trends than the Central indicators. The Sunnyslope mean 5 P.M.

Table 3a - Annual CO Statistics (mg/m^3)

	1974	1975	1976	1977	1978	1979	1980	1981
Highest 8-hr								
PHX Central	15.55	14.98a	11.82b	12.91	10.68	10.99	9.77	10.00
Mean Highest 8-hr								
PHX Central	7.42a	6.56	5.75b	6.05	4.80	5.61	4.71	5.26
PHX Sunnyslope		4.08c	3.52	3.41	3.21	3.54	3.39	3.16
Tucson	3.26b	3.58	3.57	4.18a	3.18b			
Mean 5 P.M.								
PHX Central	1.54a	1.32	1.27b	1.40	1.16	1.19	1.10	.97
PHX Sunnyslope		1.10b	1.10a	.89	.56	1.25	1.17	1.09
Mean 1-hr								
PHX Central	4.06a	3.71	3.29b	3.57	2.77	3.31	2.60	2.98
PHX Sunnyslope		2.52	2.17	1.98	1.70a	2.20	2.18	2.07b

a, highest month; b, lowest month; c, highest and lowest month indicator, for example, shows virtually no trend from 1975 to 1981 while the same indicator for Central show a 26.5 percent reduction. Percent reductions in mean monthly CO, in contrast, are reasonably similar for the two sites (20 and 18 percent respectively) as are the percent reductions in mean highest 8-hour CO (20 and 22 percent respectively). This suggests that the two sites will respond to I/M similarly in terms of two indicators but differently in terms of the third indicator. Overall, the relatively higher CO concentrations read at the Central monitor suggest that it will be the more sensitive (to I/M impacts) of the two sites.

These annual trend statistics also illustrate the need to analyze I/M effectiveness with monthly time series. For some indicators, the lowest CO readings occur prior to either I/M intervention. For others, the opposite is true. Examining only the 1977-78 changes, moreover, it might appear that the initial I/M intervention was effective. But some indicators show greater one-year reductions prior to 1977 while others show stark post-1977 increases. We cannot interpret these annual trends in this way, of course, because they mask monthly changes where one would expect to find I/M impacts. Ultimately, Table 3a shows only that annual trend statistics cannot support inferences about I/M program effectiveness.

Table 3b Here

Table 3b shows another aspect of these series. Broken down by month, we see that some indicators are more seasonal than others. For a given series, furthermore, the Central and Sunnyslope monitors are not always equally prone to variation by seasonal factors. Again, these data should be interpreted cautiously.

The similarities and differences across indicators and sites, as shown in Tables 3a and 3b, demonstrate nevertheless how the multiple CO indicators in this evaluation control germane threats to internal validity. Some of threats outlined earlier operate incrementally and/or seasonally. As the breakdowns in Tables 3a and 3b clearly show, however, no threat to internal validity can have the same effect on all indicators in all sites unless the threats interact in some bizarre manner. Since "bizarre" means "improbable," we discount this possibility. To the extent that our findings are consistent across indicators and sites then, threats to validity cannot plausibly confound the conclusions drawn from our analyses.

8.0 Findings: Impact Analyses

Table 4a Here

As a preliminary step in this task, we constructed univariate ARIMA models for each of the eight CO time series. Statistical descriptions of the ARIMA models, shown in Table 4a, provide crucial insight into these data. We will use the format in Table 4a throughout this report. Models are based on the natural logarithm series and results are expressed in percentages.[6] Confidence bounds associated with each model are set at 95 percent (approximately plus-or-minus two standard errors). Any interval bracketing zero indicates a statistically insignificant parameter.

The "variance explained" statistics shown in Table 4a (computed here as the R^2 coefficient of determination) indicate that the univariate models fit all series except, perhaps, Sunnyslope 5 P.M. CO. The trends for these models indicate modest annual decreases in CO concentrations or, where the parameter is not statistically different than zero, no change in the near future. None of these models accounts for possible I/M impacts, of course, so we reserve judgement on this issue for the time being.

Table 3b - Monthly CO Distributions (mg/m^3)

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Highest 8-hr												
PHX Central	15.0	13.3	11.1	9.9	8.3	9.8	6.0	8.2	10.1	14.2	17.8	18.9
Mean 8-hr High												
PHX Central	9.1	6.9	5.0	5.7	3.3	3.2	2.1	3.3	4.1	7.0	10.2	10.4
PHX Sunnyslope	4.8	4.2	4.8	3.3	2.0	2.6	1.4	2.0	2.4	4.3	5.6	5.9
Tucson	5.0	4.3	3.2	2.6	1.7	2.8	3.2	2.9	2.9	3.3	4.5	5.2
Mean 5 P.M.												
PHX Central	1.6	1.7	1.1	.7	.6	.7	.7	.7	.8	1.3	2.3	2.1
PHX Sunnyslope	1.7	1.4	1.1	.9	.4	.5	.3	.5	.7	.8	1.5	1.9
Mean Monthly												
PHX Central	5.5	4.2	2.8	2.2	1.8	2.1	1.3	1.8	2.3	3.9	6.0	5.8
PHX Sunnyslope	3.1	2.7	2.0	2.0	1.1	1.5	.8	1.0	1.3	2.6	3.3	3.8

Table 4a - Univariate Models

	Variance Explained	Annual Trend	95% Limits
Highest 8-hr			
PHX Central	91.4%	- 5.3%	± 2.4%
Mean Highest 8-hr			
PHX Central	96.6%	- 6.0%	± 3.1%
PHX Sunnyslope	97.5%	- 2.2%*	± 6.4%
Tucson	95.3%	- 1.0%*	± 2.7%
Mean 5 P.M.			
PHX Central	75.9%	- 3.6%*	± 16.9%
PHX Sunnyslope	24.3%	+ 15.3%*	± 37.9%
Mean Montnly			
PHX Central	96.1%	- 4.7%	± 3.1%
PHX Sunnyslope	79.6%	- .2 *	± 8.0%

* Trend estimate not statistically significant at $P_{.05}$

Table 4b - Impact Estimates

Model: $CO_t = w_1 I_{1t} + w_2 I_{2t} + N_t$

	Annual Trends	Impacts 1977	Impacts 1979
Highest 8-hr			
PHX Central	- 13.3%	+ 24.2%*	+ 28.5%*
	(± 9.3%)	(± 40.7%)	(± 34.9%)
Mean Highest 8-hr			
PHX Central	- 14.9%	+ 34.3%	+ 42.3%
	(± 7.6%)	(± 33.7%)	(± 28.3%)
PHX Sunnyslope	- 8.8%*	+ 18.4%*	+ 24.9%*
	(± 12.3%)	(± 37.9%)	(± 39.2%)
Tucson	- 1.2%*	+ 0.0%*	
	(± 3.3%)	(± 0.0%)	
Mean 5 P.M.			
PHX Central	- .5%*	+ 0.0%*	+ 59.4%*
	(± 0.8%)	(± 0.0%)	(±104.6%)
PHX Sunnyslope	- .2%*	- 4.0%*	+207.1%*
	(± 5.4%)	(±168.8%)	(±445.8%)
Mean Monthly			
PHX Central	- 14.6%	+ 42.8%	+ 40.3%
	(± 5.4%)	(± 34.6%)	(± 25.1%)
PHX Sunnyslope	- 6.5%*	+ 4.8%*	+ 40.0%*
	(± 17.3%)	(± 56.6%)	(± 55.9%)

* Parameter estimate not statistically significant at $P_{.05}$

We next incorporated fixed regressor components into the univariate ARIMA models to represent the I/M program. The general model for this analysis was

$$CO_t = w_1 I_{1t} + w_2 I_{2t} + N_t$$

where CO_t is the tth observation of a CO time series and I_{1t} and I_{2t} are dummy variables coded as described earlier; for this model, w_1 and w_2 are estimates of the impacts on the series due to the 1977 and 1979 I/M interventions.

Table 4b Here

The results of these analyses, summarized in Table 4b, are the start of a consistent body of findings we require a priori to support our conclusions. Four of the I/M impacts were statististically different than zero and these represent increases in ambient CO. Based only on these analyses, we cannot reject the null hypothesis but must conclude instead that the I/M program has no salutary impact on ambient CO. We will subsequently reanalyze these data to test the plausibility of several threats to internal validity. Our conclusions remain unchanged, however. The univariate models shown in Table 4a, in that sense, are our final estimates of CO processes in Phoenix and Tucson.

The estimates in Table 4b must be interpreted cautiously. The models used to derive these estimates are nonlinear. Parameters must often be interpreted jointly and statistically insignificant parameters should be deleted from any model before point estimates are interpreted. The "best" trend estimates in this sense are those given in Table 4a. Since each of these series is nonstationary, percent effects furthermore may be misleading. The impact estimates in Table 4b can be used only for testing null hypotheses. In this context, we can entertain questions of statistical power based on these estimates.

Statistical power is a complicated topic in the simplest context but it is especially complicated when dealing with multiple hypothesis tests. Notwithstanding this reservation, statistical power can be addressed in the crudest sense by attributing the largest possible salutary impact to the I/M program. For Central Phoenix, based on 95 percent confidence bounds, the "most optimistic" impacts would be[7]

	1977	1979
Highest 8-hr	- 16.5%	- 6.4%
Mean Highest 8-hr	+ .6%	+ 14.0%
Mean 5 P.M.	- 0.0%	- 45.2%
Mean Monthly	+ 8.2%	+ 15.2%

For Phoenix Sunnyslope, the analogous estimates would be

	1977	1979
Mean Highest 8-hr	- 19.5%	- 14.3%
Mean 5 P.M.	-164.8%	-238.7%
Mean Monthly	- 51.8%	- 15.9%

And finally, for Tucson, the "most optimistic" impact would be a 0.0 percent impact in 1977. One might conclude from these figures that, in Sunnyslope at least, a substantial salutary I/M impact can<u>not</u> be ruled out and this conclusion would be correct. In Phoenix Central and Tucson, on the other hand, one would have to conclude that any salutary I/M impact <u>can</u> be ruled out. And of course, one would also have to conclude that the I/M program actually had a strong <u>negative</u> impact in Phoenix Central. Overall, however, the confidence intervals in Table 4b reinforce our opinion that our findings with respect to I/M impacts are not an artifact of statistical power.

Table 4c Here

Construct validity, which we discussed in the context of CO indicators, applies as well to the manner in which we model the I/M program. A <u>threat</u> to construct validity in this context means simply that we cannot detect an I/M

Table 4c - Impact Estimates

Model: $CO_t = w_1 R_t + N_t$

	Annual Trends	Percent Impact
Highest 8-hr		
PHX Central	- 11.9%	+ .002%*
	(± 9.2%)	(± .006%)
Mean Highest 8-hr		
PHX Central	- 15.7%	+ .005%*
	(± 8.1%)	(± .005%)
PHX Sunnyslope	- 11.8%*	+ .000%*
	(± 16.7%)	(± .005%)
Tucson	- .7%*	+ .003%*
	(± 2.8%)	(± .006%)
Mean 5 P.M.		
PHX Central	- 28.9%	+ .010%*
	(± 20.1%)	(± .007%)
PHX Sunnyslope	- 1.2%*	+ .009%*
	(± 5.7%)	(± .030%)
Mean Monthly		
PHX Central	- 16.2%	+ .005%*
	(± 5.6%)	(± .006%)
PHX Sunnyslope	- 14.4%*	+ .007%*
	(± 16.6%)	(± .051%)

* Parameter estimate not statistically significant at $P_{.05}$

impact because we have fundamentally misrepresented the program in our time series models. To control this threat, we replicated our analyses with five distinct I/M impact models (and literally dozens of variations within each distinct model-type). Since retests may have dropped from 1980 to 1981 in both Maricopa and Pima Counties, for example, (see Table 2), a simple dummy variable coding may give an unfair test of I/M impact. Table 4c summarizes the I/M impacts estimated from the model[8]

$$CO_t = w_1 R_t + N_t$$

where R_t is the actual number of retests in the _tth_ month. These analyses build on our consistent body of findings. All of these estimates are positive (though most are statistically and substantively trivial), so the null hypothesis must stand.

Statistical power seems to be a nonissue in these models too because even the "most optimistic" impacts are extremely trivial. These results are in no way atypical of other results. We analyzed a wide range of models based on the R_t statistics but, _in all cases_, found absolutely no evidence of a relationship between I/M activity (failures, retests, etc.) and CO (except for a few increases which we can ignore).

Table 4d Here

Table 4d illustrates another class of I/M impact models which we examined in the course of this research. The trend estimates summarized in Table 4d are based on the model

$$CO_t = w_0 + w_1 I_{1t} + N_t$$

where w_0 is an estimate of pre-1977 trend and w_1 is an estimate of post-1977 trend. The I/M program, in our opinion, should not have an _incremental_ impact on CO concentrations. An incremental impact --- a post-intervention trend --- implies that I/M has a _cumulative_ effect on CO. We can think of no

Table 4d - Estimated Trend Impacts

	Annual Trends: Pre-1977	Annual Trends: Post-1977
Hignest 8-hr		
PHX Central	- 12.5%*	- 4.4%*
	(± 12.6%)	(± 6.0%)
Mean Highest 8-hr		
PHX Central	- 8.7%*	- 2.1%*
	(± 10.0%)	(± 5.1%)
PHX Sunnyslope	- 14.1%*	+ .6%*
	(± 19.9%)	(± 5.7%)
Tucson	+ 11.9%*	- 13.8%*
	(± 32.7%)	(± 16.1%)
Mean 5 P.M.		
PHX Central	- .1%*	+ .0%*
	(± .2%)	(± .0%)
PHX Sunnyslope	- 45.1%*	+ 15.5%
	(± 63.6%)	(± 13.1%)
Mean Monthly		
PHX Central	- 7.8%	- 1.5%*
	(± 7.7%)	(± 5.9%)
PHX Sunnyslope	- 34.0%	+ 5.9%*
	(± 33.5%)	(± 6.5%)

* Parameter estimate not statistically significant at $P_{.05}$

theoretical justification for a cumulative impact and accepted I/M theories (the MOBILE2 model, e.g.) assume a linear relationship between CO concentrations and vehicle emissions. Although hypothesis tests based on incremental models are thus "generous," none of these analyses reject the null hypothesis in any form. Statistical significance notwithstanding, only one w_1 estimate is more relatively negative than the corresponding w_0 estimate and in that series (Tucson 8-hour mean), the post-intervention trend is estimated from only 24 observations. Otherwise, these estimates indicate that CO concentrations changed at a slower rate after the January, 1977 I/M intervention than before.

The evidence from these analyses actually suggests that the primary I/M impact was a deterioration of CO air quality. While we could think of more than one scenario (or theory) to explain an impact of this sort, there is no need to do so. The sole purpose of these analyses was to test the hypothesis of a null I/M impact. It was tested on two different interventions in two different cities at three different sites using four different CO indicators. In each case, the data could not reject the null hypothesis. So we accept it and because it has survived many fair tests, we must have much confidence in our conclusion that I/M has had no demonstrable impact on CO air quality.

9.0 Analyses of Normalized CO Data

Most of the threats to internal validity discussed earlier cannot, in our opinion, explain away the estimated I/M impacts presented in Table 4b-4d. The putative effects of federal new car emission and fuel economy standards on CO concentrations, for example, would be realized as negative trends. Even when we allow for the possibility of post-intervention trend impacts, however, we find no evidence of an I/M impact on CO (see Table 4d). Vehicle fleet changes due to federal new car standards would tend to bias these hypothesis tests in favor of an I/M impact but since no null hypotheses were rejected, this threat is moot.

Vehicular traffic increases since 1974 also seem irrelevant. The data in Table 1 would suggest that vehicular traffic increases during this period were nullified by fuel economy increases and emission decreases. But in any event, to explain the impact estimates in Tables 4b-4d, vehicular traffic would have had to rise abruptly in 1977 and, again, in 1979, exactly off-setting any CO reductions due to I/M. This is wholly implausible, of course.

Unless local weather conditions changed abruptly in January, 1977 and then changed again in January, 1979, meteorological effects are also implausible threats to internal validity. Our statistical analyses controlled meteorological variance in CO by seasonal ARIMA structures and, of course, these models give an acceptably accurate description of series variance (Table 4a). Our analyses of meteorological time series suggest further that these models give an acceptable degree of control for weather factors. But this threat too is moot because uncontrolled meteorological variance (background noise) in our CO time series could plausibly affect only the statistical power of our models. And if every single parameter estimate in Tables 4b-4d were significant, our conclusions would remain unchanged; null hypotheses of I/M impact would not be rejected.

When considered separately, vehicular traffic and meteorological factors are implausible threats to internal validity. These two factors could operate jointly as a potent threat to internal validity, however. Considering this possibility, we "normalized" each of our CO time series by the formula

$$CONORM_t = CO_t \times (VMT_t \times T_t)^{-1}$$

where VMT_t is a measure of "vehicular miles traveled" and T_t is a measure of "temperature inversion" in the tth month[9]. We then replicated the model analyses shown in Tables 4b-4d using normalized CO as the dependent variable. The results of these analyses are summarized in Tables 5a-5c. Again, the null hypotheses cannot be rejected. One advantage of normalized data, obviously, is that meteorological and traffic variables are explicitly represented in the models and, as a result, the normalized data behave somewhat differently than the raw CO indicators. They are less seasonal, for example, and sometimes more variable. Normalization moreover affects the Central series differently than the Sunnyslope series (the same traffic and temperature inversion series were used for both sites). The results of these analyses nevertheless are consistent with the results of our analyses of the raw CO series. Not only do we find no statistically significant effects but, in almost all cases, the signs of the estimated effects are exactly opposite of what we would expect if the I/M program had any salutary impact on CO.

10. Conclusions and Recommendations

Our analysis began with two related hypotheses. Since 95 percent of the CO in Phoenix and Tucson come from mobile sources, we expected CO levels to drop in 1977, coinciding with implementation of a mandatory I/M program; we did not expect an immediate drop but, rather, a decrease distributed across the year. Similarly, in 1979, emissions standards for the mandatory I/M program were raised. We expected an immediate increase in the number of vehicles failing inspection and, hence, another drop in CO levels; but again, we expected CO levels to be distributed across the subsequent year.

To test these hypotheses, we constructed ARIMA models for eight monthly CO

Table 5a - Impact Estimates

Model: $CONORM_t = w_1 I_{1t} + w_2 I_{2t} + N_t$

	Annual Trends	Impacts 1977	Impacts 1979
Highest 8-hr			
PHX Central	- 1.6%*	+ 3.6%*	+ 1.8%*
	(± 3.2%)	(± 14.4%)	(± 12.1%)
Mean Highest 8-hr			
PHX Central	- .7%	- .1%*	+ 2.8%
	(± .7%)	(± 1.9%)	(± 1.9%)
PHX Sunnyslope	- 7.3%	+ 8.3%*	+ 24.5%
	(± 3.1%)	(± 11.0%)	(± 12.4%)
Mean 5 P.M.			
PHX Central	- .1%*	+ .0%*	- .4%*
	(± .3%)	(± 0%)	(± 18.5%)
PHX Sunnyslope	- 2.4%	+ 2.0%*	+ 7.8%
	(± 1.5%)	(± 4.6%)	(± 4.8%)
Mean Monthly			
PHX Central	- 2.7%*	- 2.4%*	+ 12.0%
	(± 2.7%)	(± 9.7%)	(± 10.2%)
PHX Sunnyslope	- 17.1%	+ 19.9%*	+108.6%
	(± 13.4%)	(±109.0%)	(± 88.3%)

* Parameter estimate not statistically significant at $P_{.05}$

Table 5b - Impact Estimates

Model: $CO_t = w_1 R_t + N_t$

	Annual Trends	Percent Impact
Highest 8-hr		
PHX Central	- 2.5%	+ .001%*
	(± 3.4%)	(± .003%)
Mean Highest 8-hr		
PHX Central	- .6%*	+ .001%*
	(± .7%)	(± .001%)
PHX Sunnyslope	- 9.3%	+ .004%
	(± 3.8%)	(± .002%)
Mean 5 P.M.		
PHX Central	+ .1%*	+ .000%*
	(± .3%)	(± .000%)
PHX Sunnyslope	+ .0%*	+ .001%*
	(± .0%)	(± .002%)
Mean Monthly		
PHX Central	- 1.5%*	+ .000%*
	(± 2.5%)	(± .003%)
PHX Sunnyslope	- 34.5%	+ .014%
	(± 11.3%)	(± .011%)

* Parameter estimate not statistically significant at $P_{.05}$

Table 5c - Estimated Trend Impacts

	Annual Trends Pre-1977	Post-1977
Highest 8-hr PHX Central	- 1.0%* (± 5.0%)	- .8%* (± 1.7%)
Mean Highest 8-hr PHX Central	- 3.7% (± 2.3%)	+ .2%* (± .6%)
PHX Sunnyslope	- 18.6% (± 11.0%)	- .4%* (± 4.3%)
Mean 5 P.M. PHX Central	- .0%* (± .0%)	- .0%* (± .0%)
PHX Sunnyslope	- 7.5% (± 4.6%)	- .0%* (± 1.8%)
Mean Monthly PHX Central	- 15.1% (± 9.0%)	+ 1.0%* (± 2.8%)
PHX Sunnyslope	- 52.7%* (± 54.2%)	- 6.6%* (± 21.6%)

* Parameter estimate not statistically significant at $P_{.05}$

time series collected from three sites in Phoenix and Tucson. Based on these models, we compared pre- and post-intervention time series segments. We found no statistically significant before/after differences. Our findings are that neither the 1977 intervention (mandatory I/M) nor the 1979 intervention (higher I/M standards) had any statistically significant impact on CO levels. Our analyses gave the I/M program every possible "benefit of the doubt." Yet we found not one iota of evidence to support the hypothesis that the Arizona I/M program had an impact on ambient air quality. Given the remarkable degree of consistency --- our findings were consistent across interventions, sites, CO indicators, and models --- we have the greatest possible confidence in our major conclusion.

Conclusion #1: The I/M program has had no salutary impact on ambient air quality in Arizona.

This first conclusion appears to conflict with published reports of the DHS Bureau of Emissions Inspections. A 1979 report of the Bureau claimed a 30 percent decrease in CO levels due to mandatory I/M and improved vehicle mix. This claim is based on questionable assumptions and methods, however. The Bureau used a "Derived Air Quality (DAQ) index," for example, and found that

> "If the DAQ index is a reliable (sic) measure of CO air quality, with the effects of meteorology held constant,..., then we should experience considerable improvement in the DAQ index..."

This assumption should not be accepted without question. The "Derived Air Quality index" must be considered inferior to more grounded, straightforward measures of CO air quality. We cannot understand why the Bureau chose instead to use a speculative, indirect --- albeit perhaps --- reliable indicator of CO levels; nor do we understand what "reliability" means in the DAQ context. To evaluate impacts on CO concentrations, one should look directly at CO concentrations. More important, of course, the statistical analyses used in the Bureau report do not approach the state of the art.

Conclusion #2: The DHS Bureau of Emissions Inspection has no ongoing I/M evaluation program nor even a statistical data system adequate for an external evaluation. Even if the I/M program were marginally effective in reducing ambient CO concentrations, the Bureau of Emissions Inspection could not demonstrate its effectiveness to an impartial, scientific audience.

Based on our analyses, we might recommend that the mandatory I/M program be curtailed or even eliminated. This recommendation would be simplistic, of course, and perhaps even naive. One could argue that no major program should be curtailed or limited on the basis of one retrospective evaluation. On the other hand, ambient CO is a "major problem" and the available evidence leads us to conclude that I/M is not an effective solution.

Recommendation #1: We recommend that Arizona consider alternatives to the I/M program.

Our research is particularly conclusive with respect to I/M effectiveness in Arizona but the evidence for general I/M effectiveness is inconclusive to say the least. The conventional wisdom (especially among EPA and DHS personnel) is that the Tiao et al. evaluation of Oregon's I/M program "proves" the effectiveness of I/M programs generally. It is quite possible for a program such as I/M to "work" elsewhere but not in Arizona, of course, but Tiao et al. do not make the strong conclusions attributed to them.[10] In our opinion, no scientifically valid research has demonstrated the effectiveness of I/M. The empirical evidence, in fact, supports the opposite conclusion. Finally, even if an I/M program were marginally effective --- say a one-time CO reduction of five percent --- I/M would still not be a cost-effective means of improving CO air quality.

There are a number of programs which could, theoretically, deliver real improvements in air quality. Most evaluators would agree, for example, that federal new car standards have affected substantial reductions in ambient CO concentrations nationally. Arizona could conceivably use the funds that it now spends on I/M to improve the vehicle fleet. This could be accomplished through a tax credit program similar to the programs which support household energy conservation (improved insulation, solar hot water, etc.). Lacking cost figures, we could not recommend a specific program. But we do recommend that alternatives to I/M be studied.

Recommendation #2: We recommend that the Arizona DHS take immediate positive action to guarantee that future air pollution control programs be evaluated as a routine matter.

While it is a relatively simple matter to test null hypotheses of program effectiveness, it is quite another matter to address the more difficult "why" and "how" questions. Why was this program ineffective and how could it have been made effective? The evaluation described in this report deals only with impacts, not with the more important issues of process. Could the operational processes of the I/M program been adjusted somehow so as to be effective? Our research deals only with the impact of the status quo program on ambient CO, so we cannot answer this question. But if this question is at all important, the greatest failure of the I/M program is that it was not routinely monitored and evaluated at any time during its operation.

Whatever pollution control program is eventually implemented in Arizona, it is essential that the program be monitored and evaluated as an operational routine. Rossi et al. (1979) distinguish between "impact" and "process" evaluations. Impact evaluations (our research, for example) are typically ad hoc performance audits conducted by independent evaluators. Process evaluations, on the other hand, consist of routine monitoring activities aimed at adjusting the ongoing program. Since process evaluations address fundamentally different questions, the monitoring and evaluation of a pollution control program will require substantial planning. The most pressing problem, however, will be data collection. Specific recommendations in this respect include

1) CO Record Management: The specific CO indicators used in this evaluation were dictated by practicalities. Ledolter et al. (1979) in New Jersey and Tiao et al. (1981) in Oregon used mean monthly 8-hour and 1-hour CO concentrations for theoretical reasons. While different evaluative contexts and questions will require different statistics, a flexible, computerized record retrieval system could accomodate all evaluation and monitoring needs. The cost of this system (less than twenty thousand dollars) would be relatively small compared to the overall cost of the program. Special software would permit aggregation of CO readings into any statistic required for the process evaluation or for general management decision-making.

2) Site-specific Traffic: VMT statistics are not available for CO monitor sites. Secondary traffic data are available from the Department of Transportation only as highly aggregated monthly totals recorded on major arterials (I-10 and I-17, e.g.) far removed from the CO monitor sites. At a minimum, the Bureau of Emissions Inspection should be required to sample daily traffic passing the monitoring sites. These statistics are absolutely essential for any process evaluation.

3) Data Reconstruction: Comparing Arizona CO series with data from New Jersey and Oregon, it is clear that Arizona data are more seasonal and, otherwise, less structured. As a practical consequence, this means that relatively long time series will be required for any evaluation of an Arizona program. Future evaluations cannot be based solely on future data. Time series must

instead be extended into the past and these data must be reconstructed from existing records. The problem again is information retrieval. Although historical CO records exist, there is no simple method for retrieving these data as time series. Some attention must be paid to entering existing historical data into the recommended retrieval system.

4) Design Issues: A real problem with interpreting future evaluations is that there is no status quo control. In an evaluation of the Oregon I/M program, for example, Tiao et al. (1981) were able to compare CO levels in an I/M city with CO levels in a non-I/M city. Comparisons of this sort ensure the validity of evaluation findings and, under ideal circumstances, increase the statistical power of analytic models, allowing precise measures of marginal impacts. For this purpose, a CO monitoring site should be established in a desert city not serviced by the program.

5) Statistical Methodology: Finally, the Arizona program should be evaluated with a proven statistical methodology. At present, the most acceptable methodology is the one used in this research and in the successful evaluations of the New Jersey and Oregon I/M programs. An advantage of time series analysis in the process evaluation context is that it permits forecasting and targeting. The Bureau of Emissions Inspection presently has no forecasting capability.

The costs of routine evaluation will not be insubstantial but the benefits will be proportional. Responsive, effective programming requires ongoing monitoring and evaluation to optimize impacts and conserve scarce resources.

Footnotes

[1] The Ledolter et al. (1979) and Tiao et al. (1981) evaluation reports are not easily accessible to a lay audience. In a December 9, 1982 letter to the Auditor General, James E. Sarn of DHS argues that New Jersey and Oregon evaluations found significant I/M impacts on CO. Our reading of Tiao et al. (1981) does not support this interpretation. Tiao et al. report a 10.4 percent drop in CO at one Portland site (CAMS). The CAMS monitor probe was moved twice in the course of the quasi-experiment, however. When Tiao et al. correct this instrumentation threat, the reduction is estimated at 5.1 percent and this is not statistically different than zero; see their Table 4.5. No statistically significant reductions were found at the other Portland sites.

[2] The EPA MOBILE2 model appears to give overly optimistic estimates of I/M impacts. See Tiao et al. (1981, Table 4.4) for an assessment of MOBILE2 I/M impact estimates in Portland.

[3] Neuroth and Poynter argue that this assumption is warranted because CO levels in Phoenix are influenced more by emissions over a 100 square-mile area than by local traffic. As evidence, they note that the highest hourly CO concentrations occur at 10 P.M., well after peak commuter hours. CO dispersion is further complicated by a west-to-east wind shift at peak CO hours which blows the polluted air mass back over the area. We have no comments on the plausibility of these assumptions.

[4] In a January 17, 1983 memorandum to the Auditor General, Mr. R. Fred Iacobelli of DHS responds that "Neither of the referenced reports are biased..." We do not refer to the reports, of course, but rather, to the estimates of I/M impact. In simple terms, w^* is an unbiased estimator of the parameter w if $E(w^*) = w$. See Kmenta (1971: Chapter 6) for a discussion of estimation theory generally and bias specifically; see Hibbs (1974; 1977) for a discussion of bias in a longitudinal context.

[5] We have no information to support this hypothesis. Analyses of Maricopa and Pima County retest series suggest nonetheless that the I/M program shifted (in terms of retests) from 1980 to 1981.

[6] Natural logarithm results are presented throughout this report for the sake of exposition only; logged parameters can be conveniently interpreted as percentages. All analyses were replicated with the raw series as well as with the "best" Box-Cox (1964) transformation. Our findings are consistent across transformations. The two lowest variance explained statistics in Table 4a are due to poor model fits for the logged series. Both model fits exceed .9 when the raw data are considered.

[7] This exercise was suggested by Dr. John Trijonis, a DHS consultant, as a means of demonstrating statistical power. We are naturally sensitive to the idiosyncracy of this interpretation but, to the extent that one accepts this interpretation, it appears that the I/M program could not possibly have a salutary impact of more than a fraction of one percent. The reader is invited to replicate this exercise with the impact estimates in Tables 4c-5c. Note that the change in "power" from 84 to 96 observations is substantial.

8 There are several ways to incorporate retests into a model. The probability of a retest, for example, might be used as a measure of program standards. Analyses in Table 4c were replicated with retest probabilities but the results were consistently the same; null hypotheses were not rejected. Nominal retests are preferred because, presumably, each retest reduces CO emissions. Tiao et al. (1981) used retests expanded as a power series to reflect a 9-month maintenance life. The effects shown in Table 4c are derived from an empirically estimated maintenance life which varies slightly across series but which, generally, amounts to 9-12 months. All analyses using retests as the independent variable lead us to the same consistent results.

9 We are indebted to the Arizona Department of Transportation and to Mr. Gary Neuroth of DHS for these data. Tiao et al. used several models but most were variations of the general form

$$CO_t = ke^{bt}(1 + wR^*_t)VMT_tT_t$$

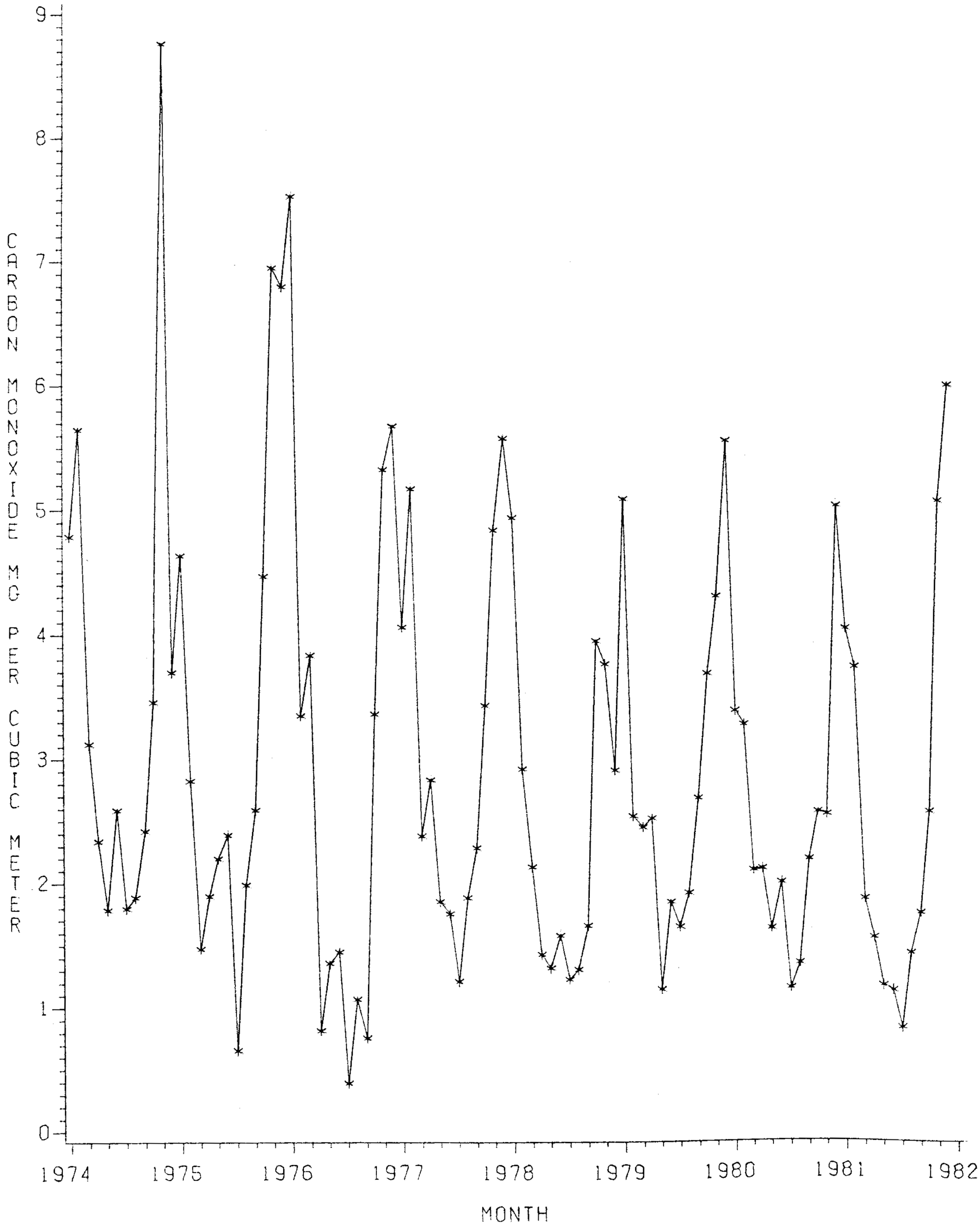
or

$$CO_t(VMT_tT_t)^{-1} = ke^{bt}(1 + wR^*_t)$$

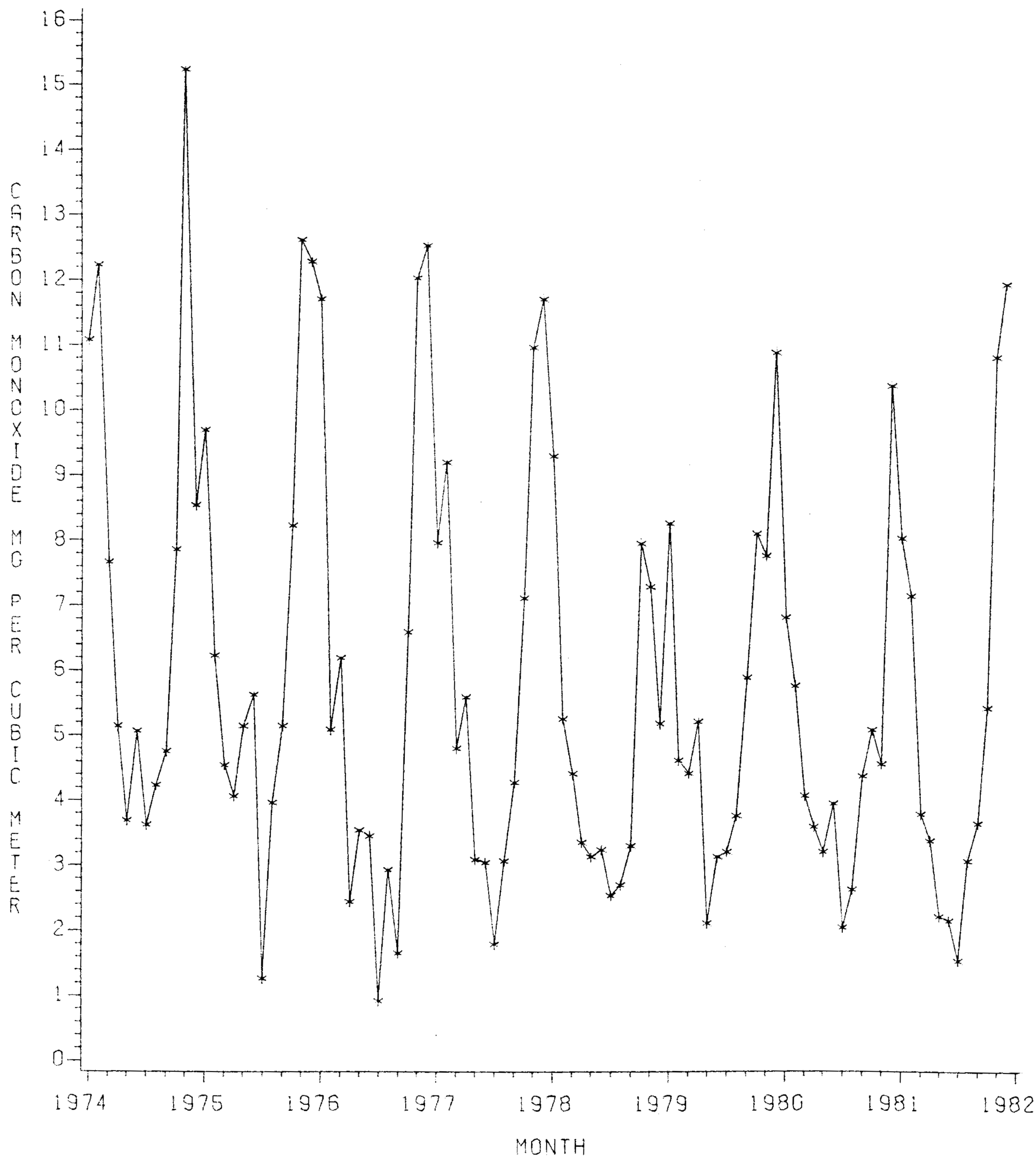
where R^*_t is the expanded retest series. The results shown in Table 5b then are derived from a model that is nearly identical to the general Tiao et al. model.

10 Cf. footnote #1. To set the matter straight, Tiao et al. reported an I/M impact at one monitor site (CAMS) only. When this impact was adjusted for two instrumentation threats, the impact was not statistically different than zero. The Tiao et al. work is undoubtedly the most thorough I/M evaluation in the literature.

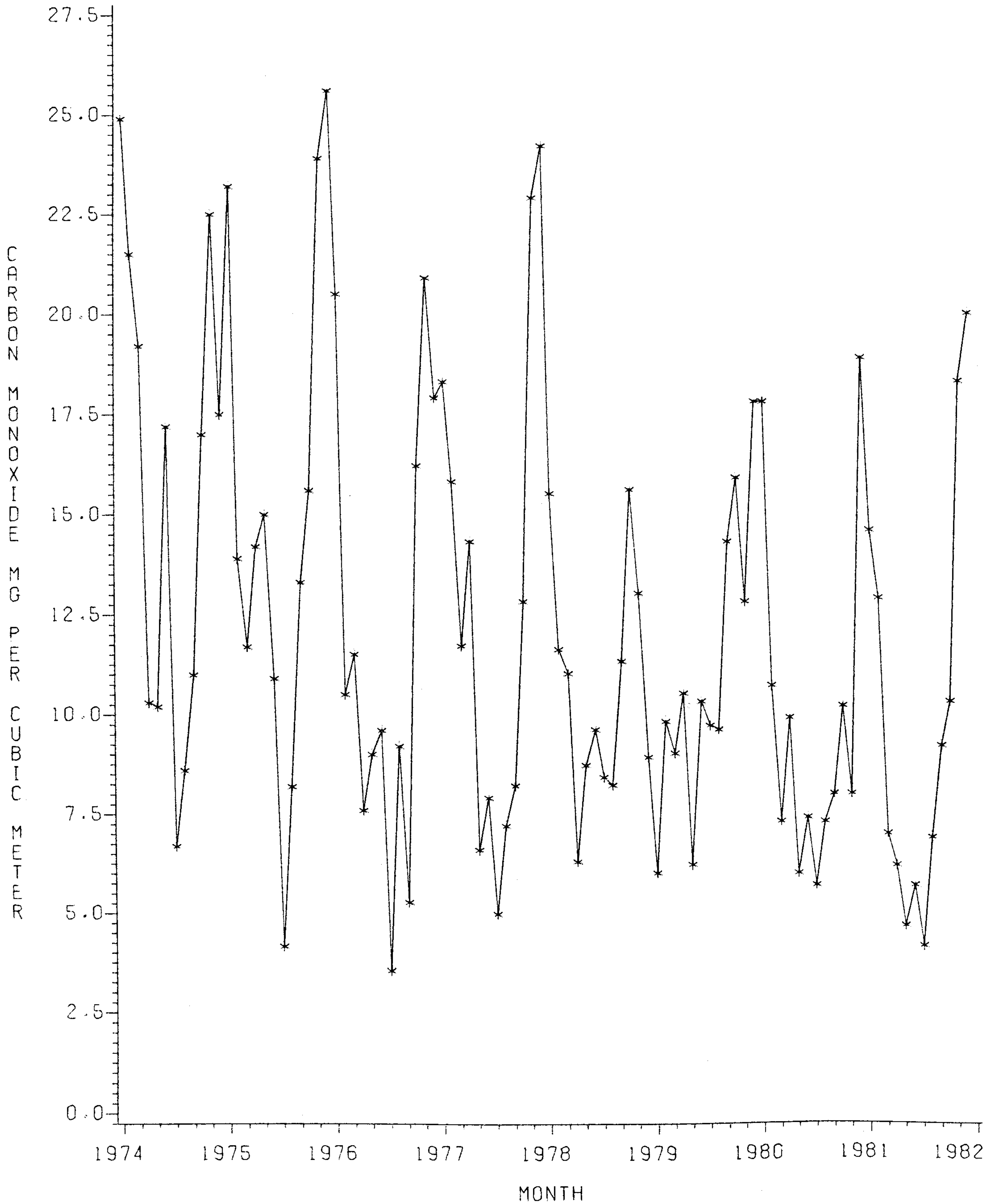
PHOENIX CENTRAL MO. AVERAGE CO LEVELS 1974/1981



PHOENIX CENTRAL CO 8 HOUR DAILY HIGH 1974/1981

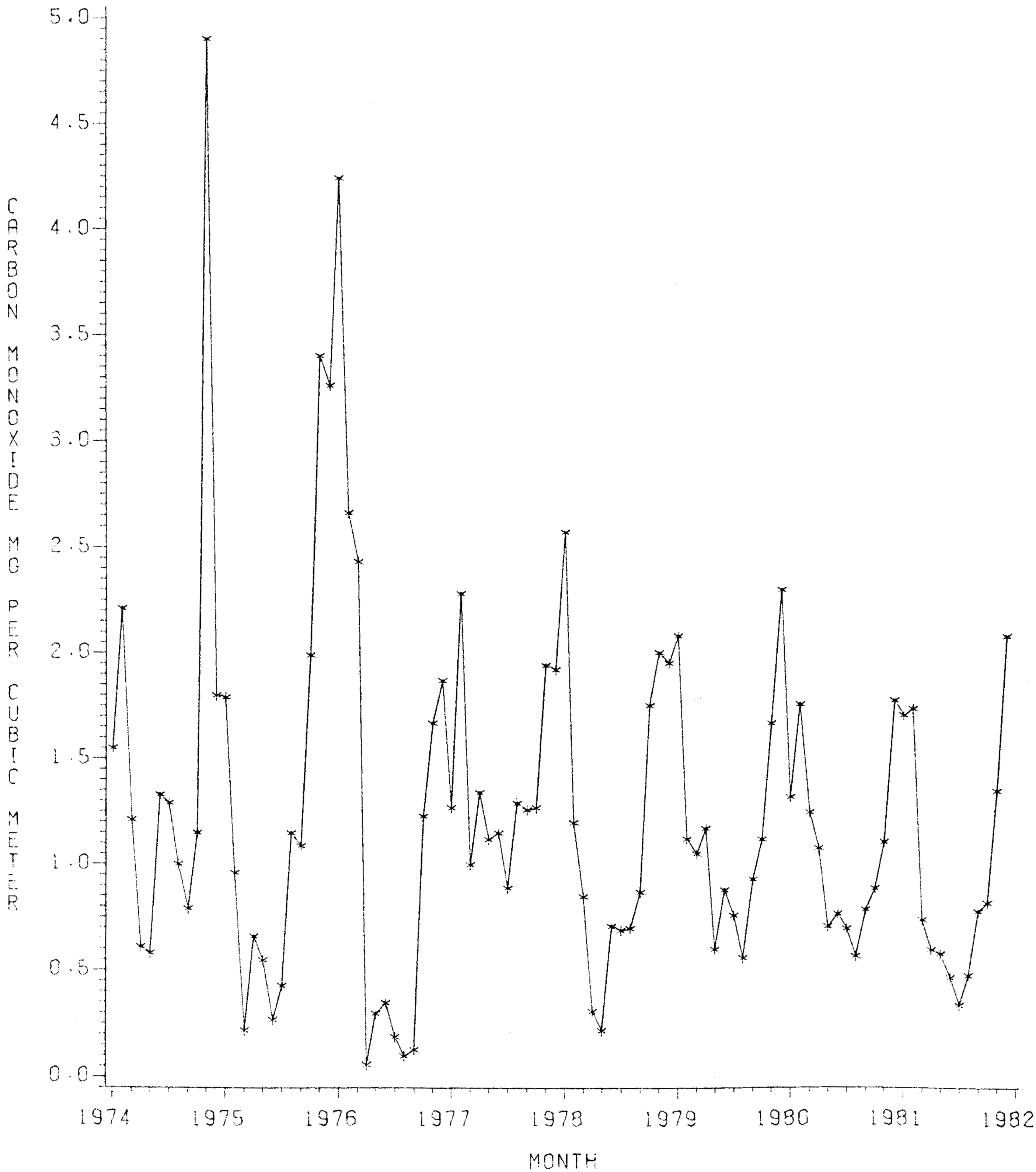


PHOENIX*8 HOUR HIGH MO. CO LEVELS 1974/1981

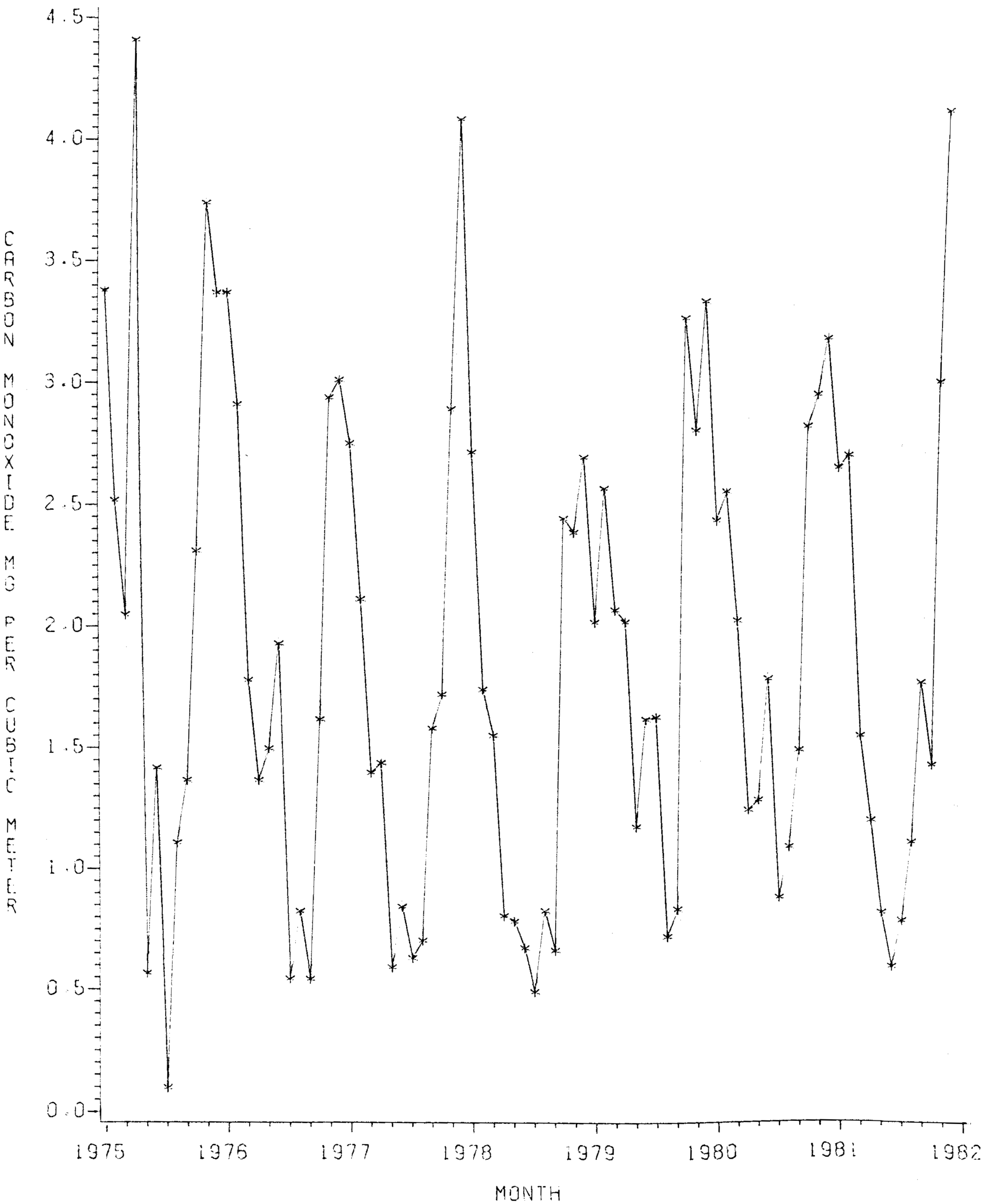


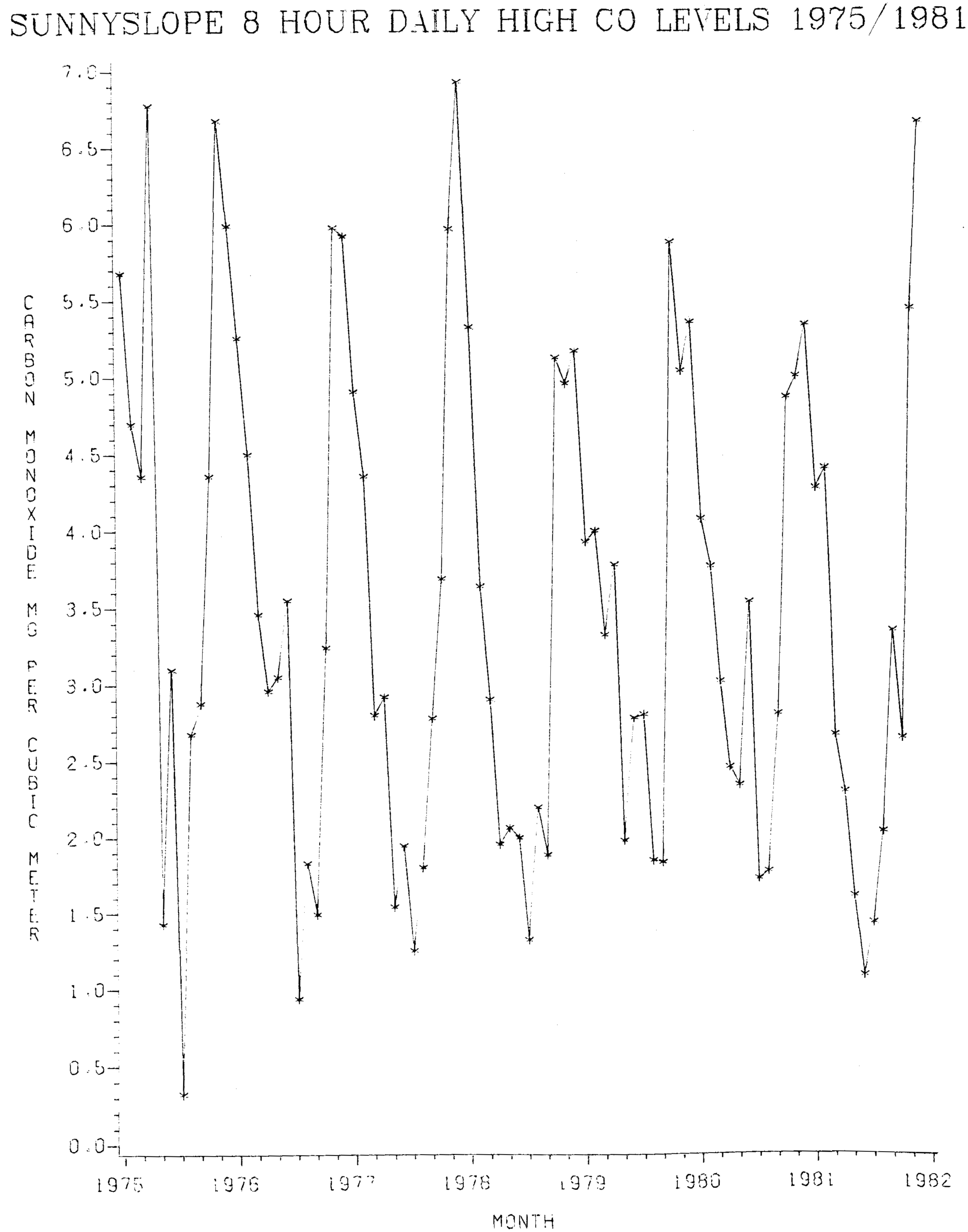
* Central Monitoring Station

PHOENIX CENTRAL 5 PM CO LEVELS 1974/1981

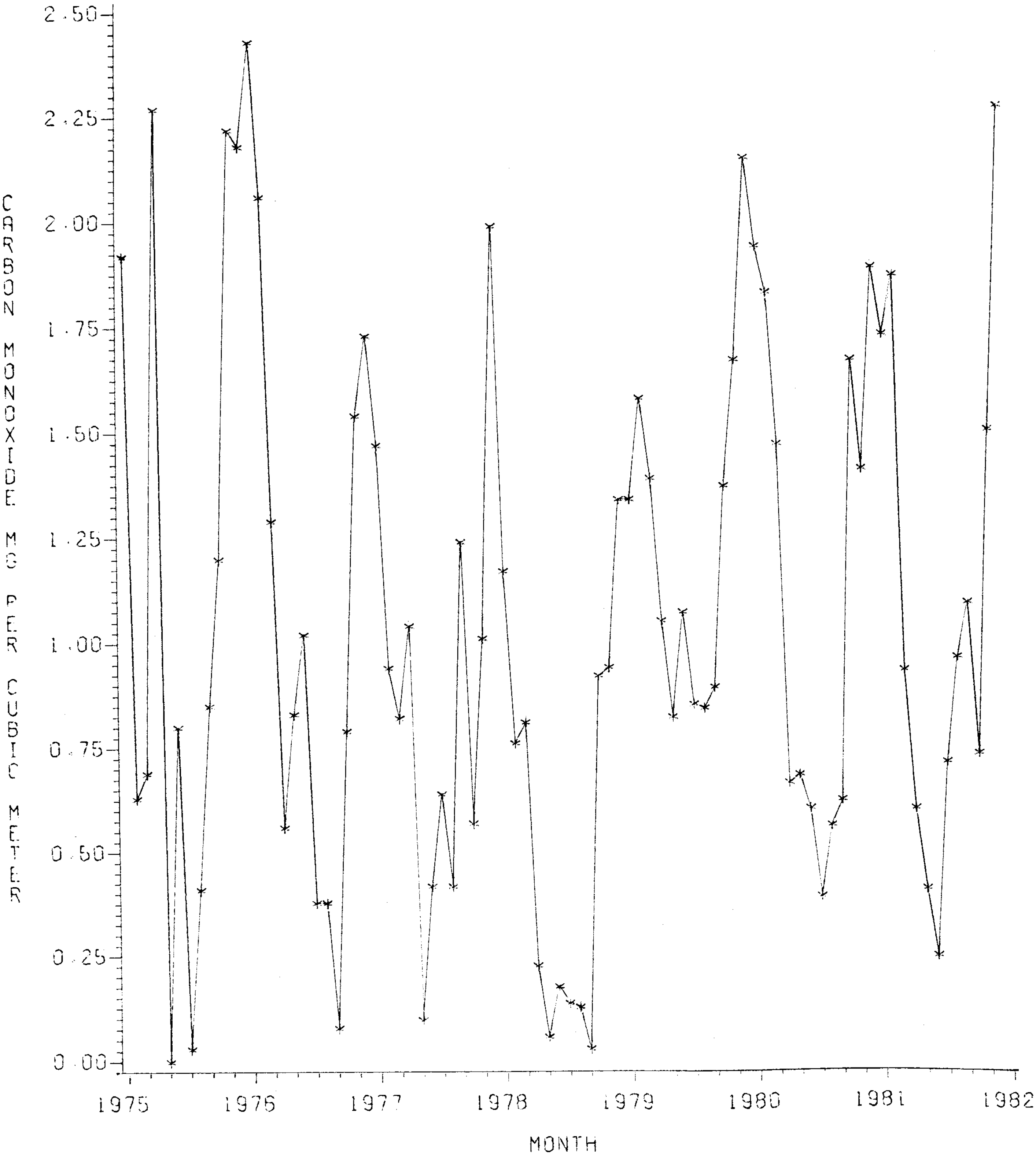


SUNNYSLOPE MO. AVERAGE CO LEVELS 1975/1981

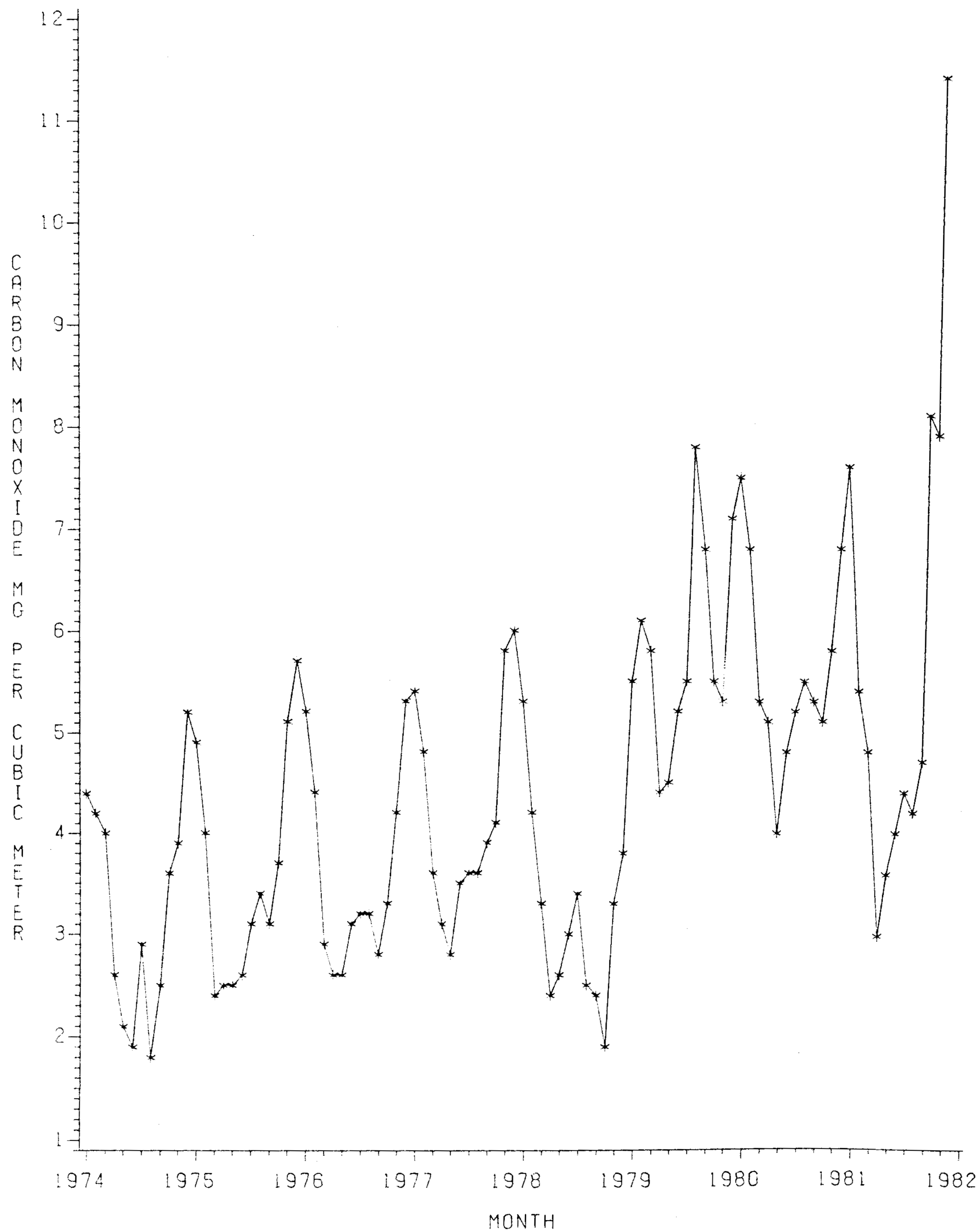



SUNNYSLOPE 8 HOUR DAILY HIGH CO LEVELS 1975/1981
CARBON MONOXIDE MG PER CUBIC METER
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1975
1976
1977
1978
1979
1980
1981
1982
MONTH

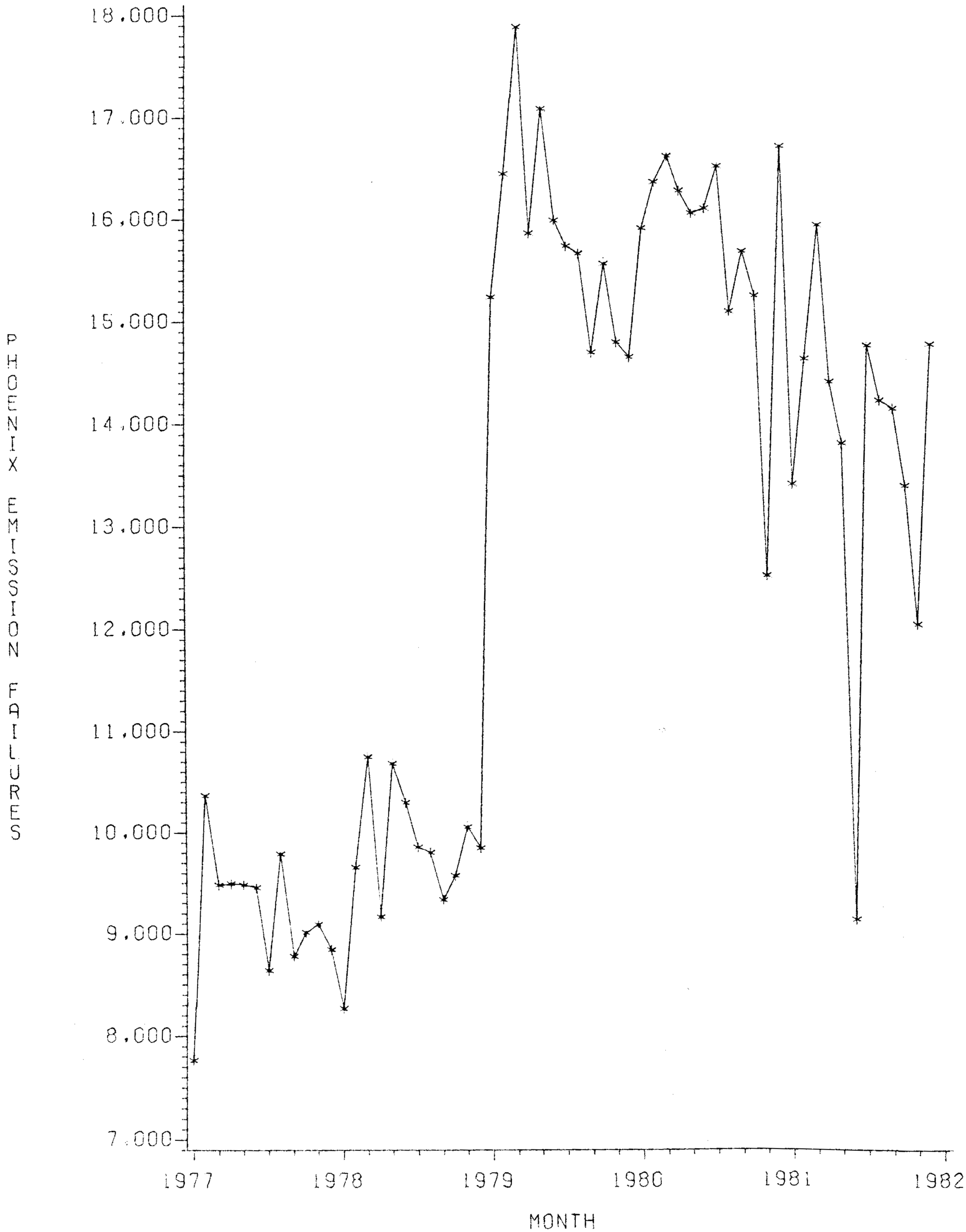
SUNNYSLOPE 5 PM CO LEVELS 1975/1981



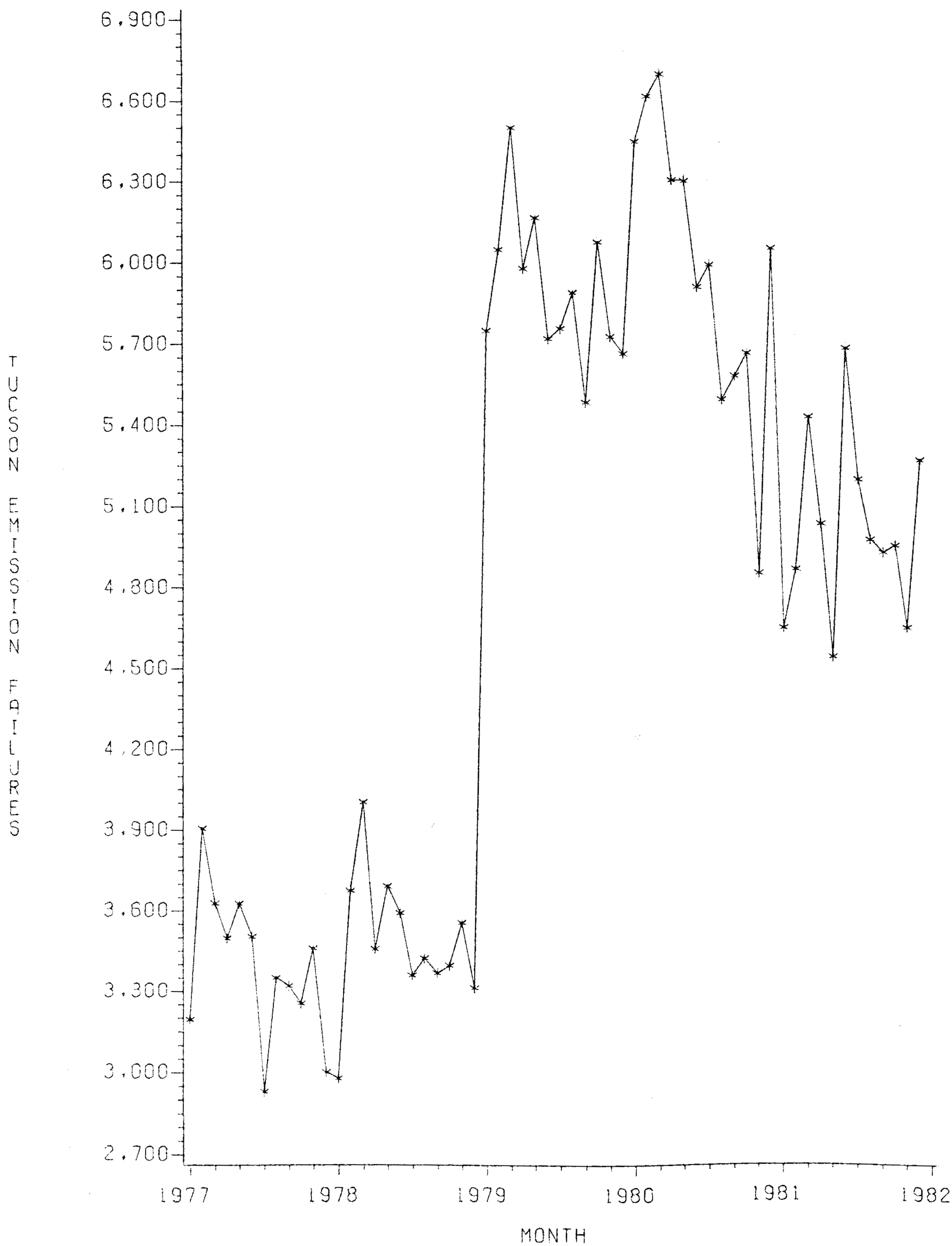
TUCSON 8 HOUR HIGH MO. CO LEVELS 1974/1981



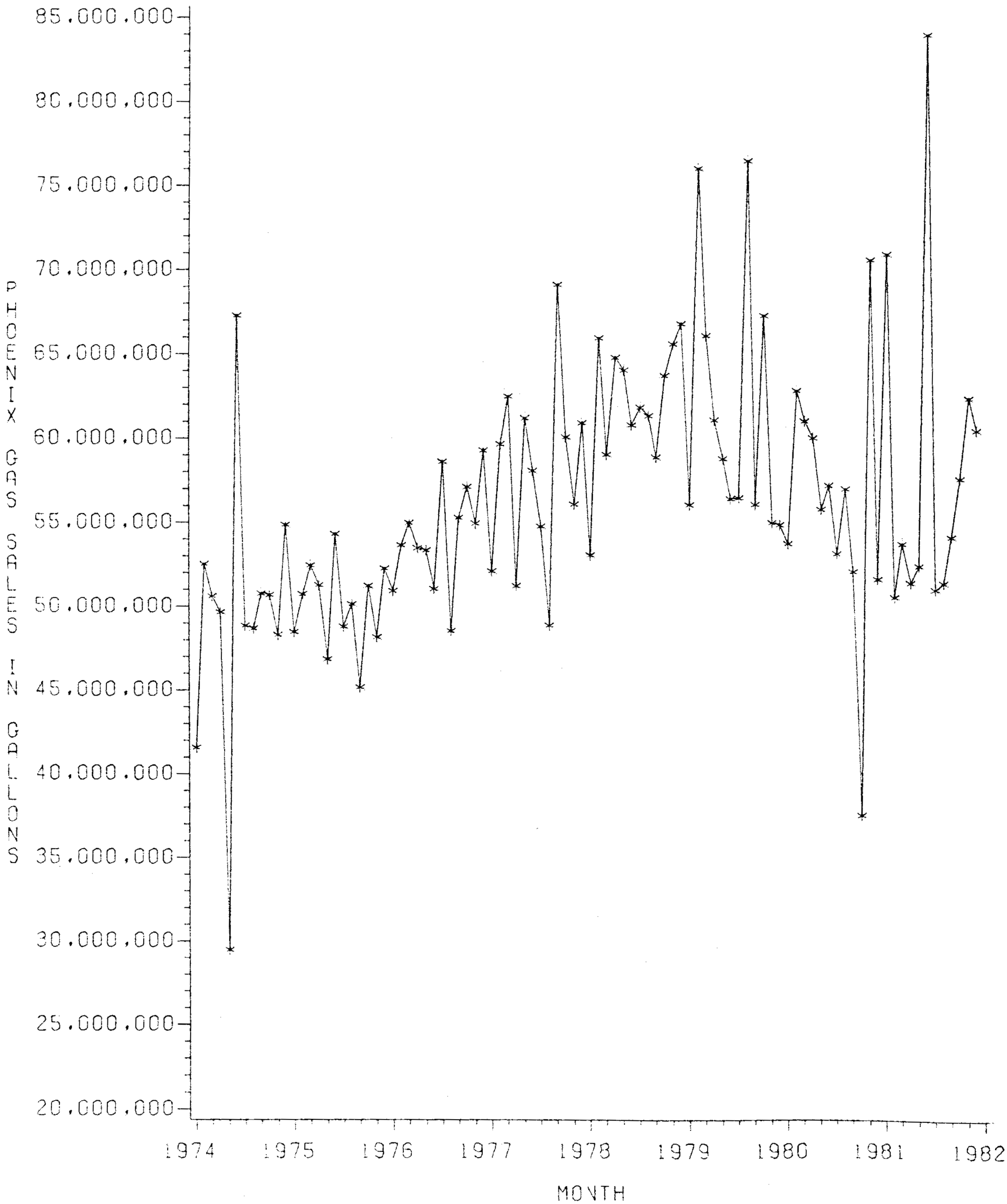
PHOENIX AUTO EMISSION FAILURES 1977/1981



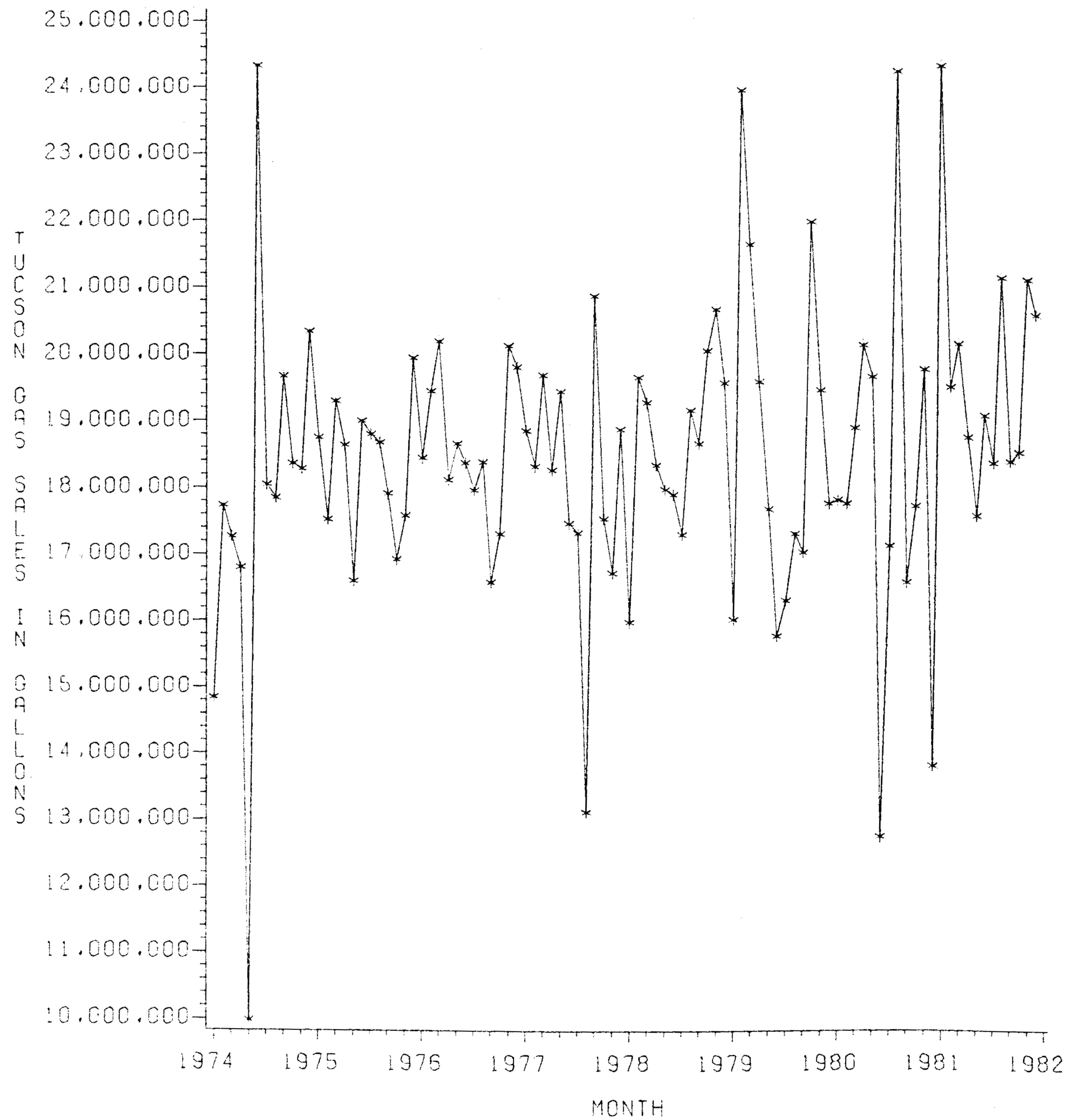
TUCSON AUTO EMISSION FAILURES 1977/1981



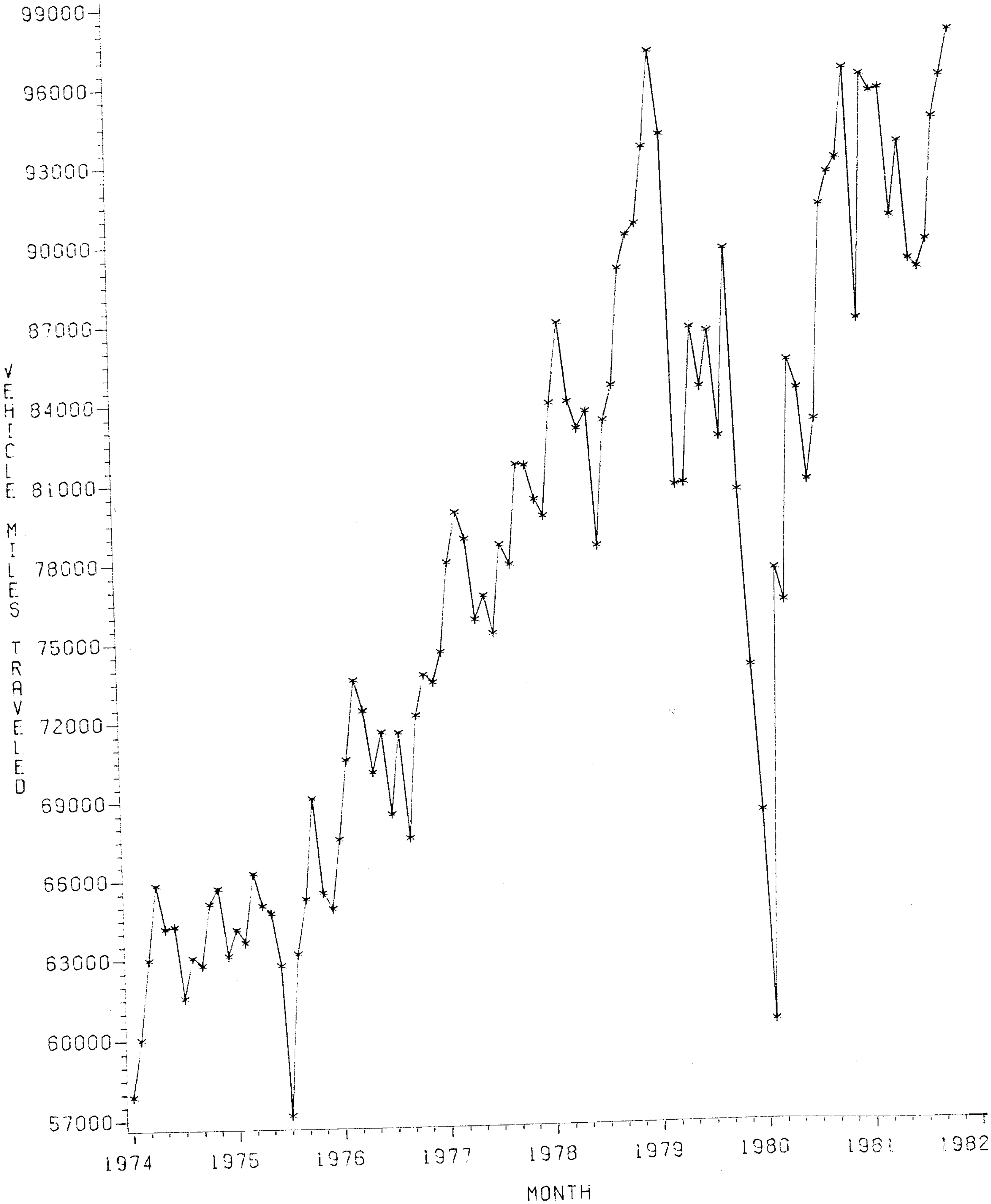
PHOENIX GAS SALES IN GALLONS 1974/1981



TUCSON GAS SALES IN GALLONS 1974/1981



PHOENIX VEHICLE MILES TRAVELED 1974/1981



Bibliography

Arizona Department of Health Services (1979). An Analysis of the Impact of the Vehicular Inspection/Maintenance Program on Ambient Carbon Monoxide.

Arthur, Cathy D. (1979). A Comparison of Theoretical Emission Inventories and Actual Carbon Monoxide Trends in Phoenix. Phoenix: Systems modeling and Reseach, Arizona Department of Transportation.

Box, G.E.P. and D.R. Cox (1964) "An analysis of transformations." Journal of the Royal Statistical Society, B, 26, 211-243.

Box, G.E.P. and G.M. Jenkins (1976). Time Series Analysis: Forecasting and Control, 2nd Revised Edition. San Francisco: Holden-Day.

Box, G.E.P. and G.C. Tiao (1965). "A change in level of a nonstationary time series." Biometrika, 52, 181-192.

Box, G.E.P. and G.C. Tiao (1975). "Intervention analysis with applications to economic and environmental problems." Journal of the American Statistical Association, 70, 70-92.

Campbell, D.T. and J.C. Stanley (1966). Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNally.

Cook, T.D. and D.T. Campbell (1979). Quasi-Experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand-McNally.

Hibbs, D.A., Jr. "Problems of statistical estimation and causal inference in time-series regression models." Pp. 252-308 in H.L. Costner (ed.), Sociological Methodology, 1973-74 (1974). San Francisco: Jossey-Bass.

Hibbs, D.A., Jr. "On analyzing the effects of policy interventions: Box-Jenkins and Box-Tiao vs. structural equation models." Pp. 137-179 in D.R. Heise (ed.), Sociological Methodology, 1977 (1977). San Francisco: Jossey-Bass.

Kmenta, J. Elements of Econometrics (1971). New York: MacMillan.

McCleary, R. and R.A. Hay, Jr. (1980). Applied Time Series Analysis for the Social Sciences. Beverly Hills and London: Sage.

Neuroth, Gary R. and W. Douglas Poynter (1982). An Analysis and Projection of Ambient Carbon Monoxide Air Quality in the Phoenix Area. Phoenix: Arizona Department of Health Services.

Office of the Legislative Analyst (1980). Air Quality Impacts of Advanced Inspection and Maintenance Programs in California. Sacramento, February 28.

Office of the Legislative Analyst (1980). Inspection/Maintenance Options for California. Sacramento, March 21.

Rossi, P.H., H.E. Freeman, and S.R. Wright (1979). Evaluation: A Systematic Approach. Beverly Hills: Sage.

Tiao, G.C., J. Ledolter and G. Hudak (1981). Statistical Analysis of the Effect of Inspection and Maintenance on Carbon Monoxide Air Quality in Portland, Oregon. Ann Arbor: U.S. Environmental Protection Agency.

APPENDIX II

STATE OF ARIZONA
OFFICE OF THE AUDITOR GENERAL

Metropolitan Arizona Vehicle Emissions Inspection Program Survey

October 7, 1982

prepared for

State of Arizona
Office of the Auditor General
111 W. Monroe, Suite 600
Phoenix, Arizona 85003

prepared by

Behavior Research Center, Inc.
1117 North Third Street
Phoenix, Arizona 85004

TABLE OF CONTENTS

INTRODUCTION

This study was commissioned by the Office of the Auditor General as one component of the Auditor General's performance audit of the Arizona Vehicle Emissions Inspection Program. The general purpose of this study was to provide input on the public's attitude toward the Vehicle Emissions Program. The specific purposes of the study were as follows:

A. To determine attitudes toward air pollution in metropolitan Arizona.

B. To determine public support of the emissions inspection program.

C. To measure the public's experiences in complying with the emissions inspection program.

D. To determine the extent of circumvention of the program.

The information generated from this study is presented in two sections in this report. The first section provides a detailed review of the findings and study methodology. The second section (appendix) consists of annotated computer tables presenting the responses to each study question cross-tablulated by a wide variety of variables. Throughout the first section of the report, table indices are listed down the right hand column of the pages. These table indices refer the reader to the detailed tables from which the data under discussion was drawn.

The Behavior Research Center has presented all of the data it believes germane to the basis research objectives of this study. However, if additional data retrieval or analysis is required by the Auditor General, we stand ready to provide such input.

BEHAVIOR RESEARCH CENTER, INC.

SUMMARY OF THE FINDINGS

1. By a ratio of 2.3 to one (or 66% to 29%), respondents favor continuation of the auto emissions testing program in Tucson and Phoenix. Even among those who initially failed the test, a majority favor program continuation. (Base=1483) 19.0-21.0

2. In addition, by ratios of better than three to one, respondents favor expanding the testing program to cover all areas of the state (rather than just the urban areas), and favor requiring older as well as newer vehicles to be tested. Opinion is sharply divided, however, as to the advisability of making the emissions test harder to pass. (Base=1483) 22.0-24.0

3. Support for continuation of the program rests on public perceptions that: (Base=1483) 19.0-21.0

   (a) Air pollution is a problem in the urban areas (83%);
   (b) Current air pollution laws are not too strong (only 6% hold this view);
   (c) Air pollution problems are more serious than five years ago (believed by 53%), and as serious my an additional 21 percent; and
   (d) If the program were to be discontinued, air quality would deteriorate (58%).

4. In addition, 57 percent believe automobiles are a major source of air pollution and an additional 36 percent classify automobiles as a minor source. (Base=1483) 7.0-9.1

5. In evaluating the emissions testing program, the following findings were recorded:

   (a) 96 percent had no difficulty in obtaining needed information about the program (hours, locations, etc.). 25.0-27.0

   (b) Respondents gave strongly favorable ratings to the process with respect to: (1) the speed of service; (2) the skills and demeanor of station personnel; (3) station hours; and (4) station locations. In addition, the $5.44 charge for the test was viewed as "reasonable" by a majority of respondents. 28.0-48.0

   (c) Eight out of ten respondents evaluated the overall service they received at the station as "excellent" or "good". Thirteen percent rated it as "fair" and only four percent rated it as "poor" or "very poor". Even among those who initially failed the test, the "excellent" or "good" ratings registered at 65 percent. 52.0-54.0

(d) Persons with prior experience in taking a vehicle through the emission test (88%) rated this most recent experience as follows: 35 percent "better", 58 percent "no change" and six percent "not as good". 58.0-63.0

6. Before going in for the emissions test, half of vehicle owners took steps to prepare the vehicle for the test. (e.g. tuned the engine, checked the air pollution control devices, changed the air filter, etc.) Taking such steps, however, does not relate to whether one has a higher or lower probability of passing the test. (Base=1483) 64.0-66.0

7. Eighty two percent passed the test on their initial visit and of those who did not, 81 percent passed on the second visit. Eventually, 98 percent passed the test. (Base=1483) 70.0-72.0

8. Among those who failed the initial test, seven out of ten spoke with someone at the station regarding what they should do next, and of those people, over 80 percent reported that their contact was willing to answer their question, and answered their question correctly. 88.0-90.0 The most common adjustments made prior to bringing the vehicle back for a retest involved carburetor and timing adjustments or having a tune up performed. (Base=269) 73.0-84.0

9. Consumer expenditures associated with the testing program (aside from the testing fee) are presented in the table below: (Base=1483) 130.0-132.0 Special Sort

	Average dollars spent: As % of All Respondents	Average dollars spent: Among those Actually Spending
Passed test first time	$13.86	$47.16
Passed test second time	48.46	55.15
Passed test third time	47.30	60.44
Never passed	115.17	155.17
Average: all respondents	21.69	54.83

---

10. Overall, 20% of vehicle owners have, (at some point in the past) after an emissions test, readjusted their engine to make it run more to their liking (Base = 1483). In the past year, the percent who have made such post-test adjustments appears to be about 11 percent. This practice is most common among people who fail the initial test and most adjustments involve adjusting or rebuilding the carburetor. Fifty-two percent of these adjustments are done by the respondent or a friend/relative while 46 percent are done professionally by a mechanic (Base=301). 100.0-108.0

ATTITUDES TOWARD AIR POLLUTION IN METROPOLITAN ARIZONA

A. Perceived Seriousness of Air Pollution

Only 15 percent of those surveyed said they did not think air pollution was a problem in their area. In contrast, 83 per cent perceive it as a problem and just over 40 percent rate it as 1.0-
a major problem. 3.0

"On the whole, would you say that air pollution is a major problem, a minor problem, or not a problem in your area?"

	Total	Phoenix	Tucson
Major	39%	41%	33%
Minor	44	43	49
Not a problem	15	14	16
Not sure	2	2	2

The strongest concerns about air pollution were recorded in the Phoenix area and among women, persons under 55 years of age, longer term residents and in upper income groups. In no population sub-group did the proportion classifying air pollution as a non-problem exceed 25 percent.

When asked whether they thought air pollution in their area today is more serious, as serious, or less serious than five years ago, a majority answered "more serious". In addition, of those expressing an opinion on this issue, the proportion answer- 4.0-
ing "more serious" rose to 62 percent. 6.0

"Would you say that air pollution problems in your area today are more serious, as serious, or less serious than they were five years ago?"

	All Respondents	Those With An Opinion
More serious	53%	62%
As serious	21	24
Less serious	12	14
Not sure	5	0
Not here 5 years ago	9	0

In effect, three out of four urban Arizonans believe the air pollution problem in their area is as serious or more serious than it was just five years ago and of those with an opinion, the percentage approaches 90 percent.

In addition, belief that the problem is worsening peaks in the Tucson area, among women, persons under 55 years of age, longer term residents. Interestingly, even among residents who are relatively new to the state most were willing to express an opinion on this question and very strongly believe air quality is worsening whether they derive their opinion from earlier visits to Arizona or from things they have read, heard or seen since arriving is not clear, but they most strongly believe that air pollution is a worsening problem in the community.

	As % of Those Expressing An Opinion		
	More Serious	As Serious	Less Serious
Total	62%	24%	14%
Area			
Phoenix	61	26	13
Tucson	62	20	18
Age			
Under 35	68	19	13
35 to 54	61	25	14
55 +	49	33	18
Residency:			
Under 3 years	74	16	10
4 to 6 years	63	22	15
7 +	60	26	14

B. Perceived Sources of Pollution

Automobiles, windblown dust and commercial vehicles are seen by most Arizonans as the major culprits contributing to air pollution in their area of residence. General business and industry, as well as mine smelters, are not widely perceived as major sources of air pollution in their communities. 7.0-9.1

"Would you say that each of the following is a major source, a minor source, or not a source of air pollution is your area?"

	Major	Minor	Not a Source	Not Sure
Automobiles	57%	36%	6%	1%
Windblown dust	55	38	5	2
Commercial vehicles	44	38	17	1
Business and industry	20	42	35	3
Mine smelters	22	26	44	8

Opinions on this subject vary somewhat within the two major communities. Thus, Tucsonians are more likely than their Phoenix counterparts to see windblown dust as the major air pollution source and are somewhat less likely to blame either automobiles or commercial vehicles. Tucsonians are also nearly twice as likely as Phoenicians to see mine smelters as a major source of pollution.

Percent Perceiving Each as a Major Source of Air Pollution

	Automobiles	Windblown Dust	Commercial Vehicles	Business & Industry	Mine Smelters
Total	57%	55%	44%	20%	22%
Phoenix	59	52	47	21	18
Tucson	50	64	38	18	30

C. Arizona's Air Pollution Control Laws

The bulk of urban residents interviewed believe air pollution control laws in Arizona are about as tough as they should be 10.0- and fewer than ten percent describe them as "too strong". Over a 12.0 third think the laws are too weak -- a belief that tends to grow in strength the longer one has lived in the state.

"Do you think air pollution control laws in Arizona are too strong, about right, or too weak?"

	Too strong	About right	Too weak	Not sure
Total	6%	47%	35%	12%
Those with an opinion (total)	7	53	40	0
Phoenix	7	54	39	0
Tucson	7	52	41	0

D. The Impact of the Vehicle Emissions Inspection Program on Air Pollution

Eight out of ten respondents believe the vehicle emissions testing program has helped keep the air clean in the Tucson and Phoenix areas. Thirty-two percent believe the program has helped "a lot" and 47 percent think it has helped "a little". 13.0- 15.0

"As part of its air pollution control program, Arizona requires that 1969 and newer vehicles in the Phoenix and Tucson areas pass an emissions test before they will be licensed in the state. Do you think the vehicle emissions testing program has done a lot, a little, or nothing at all to help keep the air clean in these areas?"

	Total	Phoenix	Tucson
A lot	32%	33%	30%
A little	47	46	50
Nothing	17	17	17
Not sure	4	4	3

Belief that the program has been effective or done "a lot" was fairly uniform in most sub-groups, but did tend to be stronger among persons over 55 years of age. On the other hand, longer term residents were among the least sanguine as to the efficiency of the program.

	Expressed as % of Those With an Opinion		
	A lot	A little	Nothing
Total	33%	49%	18%
Phoenix	34	48	18
Tucson	31	52	17
Age			
Under 35	31	50	19
35 to 54	28	49	23
55 +	40	41	19
Residency			
Less than 3 years	37	49	14
4 to 6 years	43	46	11
7 + years	29	47	24

Close to 60 percent of respondents believe the quality of air in the Tucson and Phoenix areas would worsen if the vehicle emissions testing program were to be discontinued. For the most part, the balance believe termination of the program would have no impact whatever on the quality of air in the urban areas. This division of opinion was quite uniform around the state and in no sub-group did the proportion who believe air quality would 16.0-
deteriorate drop below 50 percent. 18.0

"If the vehicle emissions testing program was to be discontinued, do you think the quality of air in the Phoenix and Tucson areas would improve, remain about the same, or get worse?"

	Total	Phoenix	Tucson
Improve	1%	1%	1%
No change	40	41	38
Get worse	58	57	60
Not sure	1	1	1

Respondents were also questioned regarding whether they would change their engine maintenance activities should the emissions testing requirement be eliminated. Ninty-three percent 124.0
answered they would not alter their maintenance practices, while 126.0
three percent felt they would cut back on tuning and adjusting their engines. Although variations were minor, people whose vehicle initially failed the emission test were more inclined to think they would cut back on engine tune-ups. The dominant pattern, however, is one in which most believe they would not alter their practices.

"If the emission test were not required, would you have your engine tuned or adjusted more often, about as often, or less often than you do now?"

	More	Same	Less	Not sure
Total	2%	93%	3%	2%
Phoenix	2	93	3	2
Tucson	4	92	3	1
Initial Test Result				
Pass	2	94	2	2
Fail	4	90	4	2
Second Test Result				
Pass	4	89	5	2
Fail	3	95	3	0
Third Test Result				
Pass	5	95	0	0
Fail	0	95	5	0

PUBLIC SUPPORT FOR THE ARIZONA VEHICLE EMISSIONS INSPECTION PROGRAM

By a ratio of 2.3 to 1, residents of the state's urban areas favor continuation of the auto emissions program in Phoenix and Tucson. In no population sub-group did support for its continuation drop below 60 percent. 19.0-21.0

"Do you strongly favor, favor, oppose or strongly oppose continuing the auto emission testing program in the Phoenix and Tucson areas?"

	Total	Phoenix	Tucson
Strongly favor	17%	16%	21%
Favor	49	51	46
Oppose	20	20	21
Strongly oppose	9	8	9
Neutral/Not sure	5	5	3

Examination was also made of responses with respect to the most recent experience respondents had with the testing program as well as in terms of their views on air pollution.

Looking first at attitudes toward continuing the program relative to people who passed or failed the test, we found, as expected, strongest support for its continuation among those who passed on their initial visit to the testing station. In addition, however, a majority of those who failed the first time also favored its continuation. In fact, a majority or a plurality favored its continuation regardless of whether they had to return a second or third time and regardless of whether they passed or failed.

	Position on Continuation of Emissions Testing Program	
	Favor	Opposed
Initial Visit to Testing Station		
Passed	70%	26%
Failed	52	42
Second Visit to Testing Station		
Passed	53	41
Failed	49	44
Third Visit to Testing Station		
Passed	47	41
Failed	52	48

Table Reads: Of those who passed the test on their initial visit to a testing station, 70 percent favor retaining the program while 26 percent are opposed.

Finally, attitudes toward the desirability of continuing the program correlate to other general attitudes about air pollution along lines that would be expected. Thus, those who think the state's air pollution laws are too strong tend to favor dropping the testing program. On the other hand, those who see the laws as either "about right" or "too weak" strongly favor retaining the program. Similarly, those who feel the program is having an impact on curbing urban air pollution favor retention of the program while 80 percent of those who see no impact would prefer it be discontinued. Unexpectedly, however, those who do not believe air pollution is a problem in their community nevertheless favor retention of the program.

Air Pollution Perceived as:	Position on Continuation of the Emissions Testing Program: Favor	Opposed
Major problem (39%)	80%	17%
Minor Problem (44%)	61	35
Not a problem (15%)	50	43
Air Pollution Laws Thought to Be:		
Too strong (6%)	19%	81%
About right (47%)	68	27
Too weak(35%)	74	22
Testing Program's Effect on Keeping Air Clear:		
Does a lot (32%)	93%	6%
Does a little (47%)	67	27
Does nothing (17%)	15	81

Table Reads: Of the 39 percent who believe air pollution is a major problem in their area, 80 percent favor retention of the testing program, while 17 percent favor discontinuing it.

Even if it helped keep Arizona's air cleaner in the future, 72 percent of those interviewed oppose a proposal to make it "a lot harder" for cars to pass the test. On the other hand, half favor making the test somewhat harder (although 44 percent oppose). Finally, three-quarters of urban residents support proposals to require older vehicles and automobiles registered in areas outside of Phoenix and Tucson to also be included in the manditory testing program. 22.0 24.0

"If it helped keep Arizona's air clean in the future, would you favor or oppose each of the following proposals?"

	Favor	Oppose
A. Expanding the emissions testing program so that all areas of the state are included in it, rather than only the Phoenix and Tucson areas.	76%	22%
B. Require older as well as newer vehicles to go the emissions testing program.	75	24
C. Make it somewhat harder for vehicles to pass the test.	49	44
D. Make it a lot harder for vehicles to pass the test.	23	72

Undoubtedly, the above results reflect the truth in the old adage that one's position on an issue is often but a reflection of whose ox is likely to be gored.

At the end of the questionnaire, and after respondents had been taken through all of the questions regarding their experiences with the testing program, they were again asked whether they thought the program should be continued in Arizona. The re- 127.0
sponse pattern is interesting because the proportion in favor of 129.0
continuation increased modestly and the proportion opposed decreased modestly. The net shift in favor of continuation of the program was recorded at approximately eight percent.

"All things considered, do you think Arizona should or should not continue the emissions testing program?"

	Total	Phoenix	Tucson
Yes, should continue	69%	68%	73%
Yes (if made tougher; applied statewide; covers all vehicles; includes safety check)	2	3	2
No, should discontinue	26	26	24
Not sure	3	3	1

EVALUATION OF THE VEHICLE EMISSIONS TESTING PROCESS

Survey respondents were next asked to rate the emissions testing process on several dimensions: the availability of information about the program, the actual testing process, and the quality of services received at the testing site. Generally speaking, all aspects of each of these variables were rated quite high by the majority of respondents.

The following sections address the ratings given the emissions testing process on these three dimensions.

A. Availability of Program Information

The emissions testing program was rated on three information-related subjects: station locations, testing hours, and other program requirements. As shown below, more than 95 percent 25.0-
of respondents reported having no difficulties getting informa- 27.0
tion on these subjects.

> "The last time you took your vehicle in for its emissions test, did you have any difficulties getting information on the following things?"

	Any Difficulties	
	Yes	No
Finding out where the inspection stations are located.	1%	99%
Finding out when the inspection stations are open.	3	97
Finding out about any other program requirements.	4	96

With one exception, there are no observed deviations from these findings among demographic or other population sub-groups. The exception is short-term residents of Arizona (one year or less in the state), who are more likely than others to report having difficulties finding where the stations are located (12%, vs 1% of the total sample). In addition, to some extent, this same group reported having somewhat more difficulty learning about other program requirements (8%, vs 4% of the total).

B. Rating of Testing Process on Selected Features

Respondents were also asked to provide a scalar rating of the emissions testing program on five factors: station location, length of time it takes to be tested, station personnel, cost, and hours of operation. As shown below, each of these five factors received an overall rating of 6.4 or better on a scale where ten is the best. A score of five or less would indicate, on balance, a negative evaluation.

28.0-
48.0

> "Now I'd like to have you rate the vehicle emissions inspection program on a number of different factors. As I read each one, please rate it on a scale of 1 to 10, where "one" means your evaluation is very negative, and "ten" means your evaluation is very positive. If your evaluation is somewhere in between very positive and very negative, you would use one of the numbers in between."

Factor	Mean Rating
How would you rate the length of time it took you to go through the testing process, I mean from the time you first arrived at the testing station until the time you left?	8.0
How would you rate the testing station personnel with regard to:	
...their professionalism	7.9
...their courtesy	7.9
...their friendliness	7.8
The stations open at 8:00 am each weekday and close at 3:30 pm on Monday, Wednesday and Friday. On Tuesday and Thursday they stay open until 7:00 pm. How would you rate these hours of operation?	7.6
How would you rate the distance you had to travel to the last testing station you visited?	7.3
It costs $5.44 for the emissions test. How would you rate the reasonableness of this charge?	6.4

Overall, the length of time it takes to be tested is given the highest rating, while the $5.44 cost is rated lowest. Looking closer at the data, the following variations are noted for population sub-groups:

-- Individuals 55 years of age and older tend to give station personnel higher ratings on courtesy, friendliness and professionalism than do younger persons, especially those under 35;

-- This same group (55 and over), tends to give a lower rating to "station location" than do other groups;

-- A respondent's total household income appears to have no relation to the rating given the cost for emissions testing (6.6 average from those annually earning less than $10,000, compared with a 6.5 average for those earning $30,000 and more).

-- As might be expected, respondents who believe: (a) that air pollution is not a problem, (b) that Arizona's air pollution laws are too strong, (c) that the emissions program is ineffective; or (d) those who oppose the program in general, tend to give lower readings on each dimension rated. Nonetheless, even these respondents gave ratings on the positive side of the scale.

In a related question, individuals who rated station hours of operation as poor (ratings of 1 to 4), were asked what times would be more convenient. As the summary table below illustrates, the majority (74%) who give a poor rating to stations hours would prefer the stations stayed open later on weekdays.

"What times would be more convenient for you?" (Asked of those who rated hours of operation 1 through 4).

Later on weekdays	74%
Open Saturday	15
Open weekends (day not specified)	13
Earlier on Weekdays	13
Open Sunday	*
Not Sure	10

* Indicates percent less than 1.

By way of further analysis, it is observed that most (91%) of those who failed the emissions test on the first try (and also gave a 1 to 4 rating to the hours of operation) would prefer that the stations stay open later on weekdays. Additionally, it is noted that a greater percentage of those answering this question who live in Tucson would prefer the stations open earlier on weekdays (24%, vs 9% of Phoenix residents).

C. Evaluation of Services Received at Testing Station

Eighty-two percent of respondents say they would rate the overall service they received at the testing stations as good or excellent. Females and older individuals tend to give more positive ratings to the overall service than do males and younger 52.0-
persons. 54.0

"Whether your vehicle passed or failed the test, would you describe the service you received at the testing station as excellent, good, fair, poor or very poor?"

		Gender		Age		
	Total	Male	Female	Under 35	35-54	55+
Excellent	35%	29%	38%	24%	37%	50%
Good	47	48	46	52	46	39
Fair	13	18	10	19	11	5
Poor	3	3	4	3	3	4
Very Poor	1	1	1	1	2	1
Not sure	1	1	1	1	1	1
	100%	100%	100%	100%	100%	100%

Not surprisingly, a larger percentage of those who initially passed the emissions test rated the service as excellent or good (85%, vs 65% of those who initially failed).

To gain a better understanding of the reasons for an individual's general satisfaction or dissatisfaction with the emissions testing service, they were asked to elaborate on their overall rating. As the next table reveals, 87 percent of the respondents rendered a positive response, while only 15 percent 55.0-
rendered a negative one. 57.3

The primary reasons given for their ratings by those respondents who rated the service in a positive nature revolved around the speed of service received and the quality of the station personnel. Station personnel were also given as the primary reason for giving the service a negative rating, however in total, positive personnel comments outweighed negative personnel comments by a seven to one margin. It may also be seen in the table that 46 percent of those respondents who rated the service as only "fair", gave it a positive comment.

	Service Evaluation			
	Total	Excellent/ Good	Fair	Poor/ Very Poor
(Base)	(100%)	(82%)	(13%)	(4%)
Net Positive Responses	87%	98%	46%	6%
Fast Service - not waiting in line, in and out	56%	66%	16%	2%
Personnel - Professional, know their jobs, efficient, organized, knowledgeable	26	30	12	0
Personnel - Courteous	26	30	7	4
Personnel - Friendly	18	21	7	4
Personnel - Informative: told me what was wrong with the vehicle	7	8	5	0
Personnel - Informative: explained test to me and what was going on	6	7	1	0
Personnel - Helpful (not specific	5	6	1	0
Personnel - Informative: answered my questions	2	3	1	0
Convenience - Station location	3	4	0	0
Convenience - Hours of operation	1	1	1	0
General Positive - no hassle, easy, good, painless, I passed	26	29	16	0
Miscellaneous Positive	5	6	2	2

Net Negative Responses	15%	3%	60%	96%
Personnel - Not courteous	3	*	12	21
Personnel - Not friendly	3	1	10	11
Personnel - Not professional, don't know their jobs, not knowledgeable	2	*	5	32
Personnel - Not informative: did not tell me what was wrong with vehicle	2	0	9	12
Personnel - Not informative: did not explain test to me	1	0	3	15
Personnel - Not helpful (not specific)	1	0	4	12
Personnel - Not informative: did not answer my questions	1	0	3	8
Program is a political rip-off, doesn't improve air	3	*	10	16
Slow service, long wait	2	1	5	9
Inaccurate test equipment	1	0	0	17
Mistreated car	*	0	0	8
Test does not measure proper variables	*	0	3	0
Miscellaneous negative	*	0	*	0
General negative, service only mediocre	3	*	19	8
Don't know	2	1	7	0
	206%	215%	158%	181%

* Indicates percent less than 1.
Totals exceed 100% due to multiple responses.

The final question in this section asked individuals who had prior experience at one of the testing stations (88%) to compare the service they received this time with previous visits. As the table below illustrates, nearly 95 percent of the respondents say the service they received on their most recent visit to the testing station was either as good (58%) or better (35%) than the service they received on earlier visits. The only notable exception to this service rating pattern was observed among those respondents who failed the test on the first or follow-up visit. 58.0-63.0

		Last Test Experience			
		Initial		Follow-up	
Quality of Service	Total	Pass	Fail	Pass	Fail
Better	35%	37%	27%	28%	24%
About the same	58	58	59	61	45
Not as good	6	5	13	11	27
Don't know	1	0	1	0	4

THE PUBLIC'S MOST RECENT EXPERIENCE WITH THE VEHICLE EMISSIONS TESTING PROGRAM

Survey respondents were asked a series of questions to determine: a) what steps they took to prepare for their most recent emissions test; b) the outcome (pass/fail) of the test; c) the steps they took if they failed the test, and; d) the amount of money they spent in order to comply with the emissions program. This section of the report details these findings.

A. Steps Taken to Prepare for Emissions Test

As the next table indicates, over one-half of the survey respondents indicated they did "nothing" to their vehicle prior to taking the emissions test. Females, lower and upper income, older, and newer residents to Arizona indicated the highest propensity to do "nothing" prior to the test. The most common step respondents took prior to the test was to have it tuned up (42%), 64.0-
with slightly more respondents performing their own tune up (24%) 66.0
as opposed to paying someone else to do it (18%).

Steps Taken Prior to Test
(% Indicating Each Was Done)

	Tune-up Paid For	Tune-up Self/ Family	Check/Make Sure Air Pollution Devices Hooked Up/Working	Make any Other Engine Adjustments to Improve Its Chances of Passing	% of Respondents Doing Some-thing	% of Respondents Doing Noth-ing
Total	18%	24%	34%	17%	49%	51%
Sex						
Male	15	32	38	21	54	46
Female	20	19	31	14	44	56
Income						
Under $15K	20	24	34	21	49	51
$15-24.9K	22	26	39	15	56	44
$25K+	16	25	33	17	46	54
Age						
Under 25	17	40	33	25	60	40
25-34	19	31	32	21	53	47
35-54	15	25	41	18	49	51
55+	22	6	28	6	35	65
AZ Residence						
Under 7 yrs.	14	22	28	16	41	59
7+ yrs.	20	26	36	18	52	48

NOTE: Combined step totals exceed 100% due to multiple steps taken.

Taking the above analysis one step further, it may be seen in the next table that the majority of respondents who check their pollution devices or make other adjustment prior to taking the test generally do so in concert with a tune up (78%) and not as an isolated step (22%).

Steps Taken Prior to Test	All Survey Respondents	As Percent of Respondents Doing Something	As Percent of Respondents Taking "Other" steps
Nothing	51%	N/A	N/A
Tune up only	11	22%	N/A
Tune up & pollution device check	15	30	39%
Tune up & other adjustment	2	5	7
Tune up & pollution device & other adjustment	12	25	32
Pollution device check only	6	12	16
Other adjustment only	2	4	4
Pollution device check & other adjustment	1	2	2
	100%	100%	100%

B. Results of Initial Emissions Test

Eighteen percent of the respondents surveyed indicated they failed the vehicle emissions test on their first try with Tucson residents (20%), females (20%), and low income households(23%) revealing the highest failure levels.

Of particular interest in the following table is the failure rate difference between those respondents who did "nothing" prior to the test (13%) and those who did "something" (26%). This variation could lead one to speculate that the best thing to do prior to taking the test would be "nothing", however, this would be misleading since this study was not designed to evaluate vehicle condition prior to the test -- a factor which could have a significant impact on the test results. Nonetheless, it is interesting that respondents who only checked the pollution device or made other adjustments without performing a full tune up had the highest failure rates. 70.0-72.0

	Initial Emissions Test	
	Passed	Failed
Total	82%	18%
Area		
Phoenix	83	17
Tucson	80	20
Sex		
Male	85	15
Female	80	20
Income		
Under $15K	77	23
$15 to 24.9K	82	18
$25K+	86	14
Steps Taken Prior to Test		
No	87	13
Yes	74	26
Tune-up only	79	21
Tune-up & Pollution device check	81	19
Tune-up & other adjustments	74	26
Tune-up & pollution device check & other adjustments	66	34
Pollution device check only	72	28
Other adjustments only	65	35
Pollution device check & other adjustments	64	36

C. Steps Taken By Respondents Failing the Initial Emissions Test

Respondents who failed the initial emissions test were asked a series of questions to determine what steps they took after failing the test. The first set of questions probed respondents on whether they spoke with any station personnel after they failed the test and whether they left the station with a clear understanding of what they were to do before they returned for a retest.

As the next set of tables reveals, 68 percent of the respondents who failed the test indicated they spoke with someone at the station regarding what they were supposed to do. Of this group, 82 percent said the person they spoke with was willing to answer their questions and that their questions were answered correctly.

It is also evident from the following tables that 72 percent of those persons who failed the initial test, left the test station with a clear understanding of what they were to do before returning for a retest. Those persons who left the test station without a clear understanding were more likely to be Phoenix residents than Tucson residents (28% vs. 20%), females than males (30% vs. 17%), and between 35 and 54 years of age than in any other age group (40% vs. 21%). 73.0-75.0

"Did you speak with anyone at the testing station regarding what you were supposed to do when your vehicle failed the test?"

(Base)	(269)
Yes	68%
No	30
Don't Recall	2
	100%

"Was the person you spoke with...
(Among persons who spoke with station personnel)

(Base)	(184)
Willing to answer your questions	
Yes	82%
No	15
Don't Recall	3
	100%

76.0-78.0

Answer your questions correctly	
Yes	82%
No	10
Don't Recall	8
	100%

"Did you get clear information regarding what to do about your vehicle before bringing it back in for retesting?" 79.0-81.0

(Base)	(269)
Yes	72%
No	26
Don't recall	2
	100%

Next, those persons who failed the test were asked what they did to their vehicle before they returned for a re-test. As may be seen, carburetor adjustments and full tune-ups lead the list. 82.0-84.0

"What did you do to your vehicle before you took it in to have it inspected again?"

(Base)	(269)
Adjusted carburetor	33%
Had tune-up performed	27
Adjusted timing/points	12
Cleaned/replaced spark plugs	10
Cleaned/replaced air filter	6
Adjusted emissions control devices	3
Changed oil	1
Got waiver	1
Other, non-specified second party adjustments made	21
Nothing	7
Don't Know	4
	125%*

* Percent exceeds 100 due to multiple responses

Finally in this section, owners of failing vehicles were probed on the outcome of their retest and what they did if their vehicle failed for a second time. As may be seen, 81 percent of owners indicated their vehicle passed the retest while 15 percent indicated their vehicle failed for a second time. In addition, it appears that Tucson residents had a somewhat higher failure rate on the retest than did Phoenix residents.

The most common step respondents took after failing the test a second time was to obtain a waiver. Tucson residents were more likely than Phoenix residents to obtain a waiver while Phoenix residents were more likely to either adjust their vehicles engine or simply go to another testing station. While these variations between Phoenix and Tucson appear to be quite distinct, they should not be overstated because of the small sample bases they are derived from.

Nearly half (46%) of all vehicle owners who did not pass the second test revealed that their vehicle never did pass the test. Once again, Tucson residents were more likely than Phoenix residents to have never passed the test.

"When you took your vehicle in to be retested, did it pass or fail the emissions test?" 88.0-90.0

	Total	Phoenix	Tucson
(Base)	(269)	(188)	(81)
Passed	81%	89%	63%
Failed	15	10	27
Don't Recall	4	1	10
	100%	100%	100%

"What did you do then?" 91.0-93.0
(Among persons who failed 2nd test)

	Total	Phoenix	Tucson
(Base)	(41)	(19)	(22)
Obtained waiver	41%	14%	64%
Had motor adjusted	28	43	17
Took vehicle to another station where it passed	13	29	0
Had DHS personnel adjust vehicle	12	14	9
Junked vehicle	3	0	5
Don't recall	3	0	5
	100%	100%	100%

"Did the vehicle ever pass the test?" 94.0-96.0

	Total	Phoenix	Tucson
(Base)	(41)	(19)	(22)
Yes	54%	86%	27%
No	46	14	73
	100%	100%	100%

Now, in order to review the complete testing sequence the following table is presented. As may be seen, statewide, 98 percent of all vehicles required to meet the state's minimum vehicle emissions standards do so. Phoenix residents appear to have a somewhat higher compliance rate than Tucson residents, however, this can not be definitely stated since the variation between the two areas is not large enough to indicate a significant difference, given the statistical parameters of this study.

Emissions Test Recap

	Total	Phoenix	Tucson
Passed test first time	81.9%	82.6%	79.7%
Passed test second time	14.7	15.4	12.8
Passed test third time	1.5	1.5	1.5
Summary			
Ever passed test	98.1%	99.5%	94.0%
Never passed test	1.3	.3	4.0
Unknown	.6	.2	2.0
	100.0%	100.0%	100.0%

D. Spending to Prepare for the Vehicle Emissions Test

An overall average of $21.69 was spent by survey respondents to prepare their vehicles for the emissions inspection test. Among those respondents who spent "something" to prepare for the test the overall average was $54.83. As might be expected, the amount of money spent to prepare for the test tended to increase in direct relationship to the number of times an individual had to take the test.

130.0-
132.0
Special
Sort

Average (Mean) Respondent Spending

	Among Total Sample	(Base)*	Among Respondents Spending "Something"	(Base)
Overall	$21.69	(1446)	$54.83	(572)
Passed test 1st time	13.86	(1194)	47.16	(351)
Passed test 2nd time	48.46	(206)	55.15	(181)
Passed test 3rd time	47.30	(23)	60.44	(18)
Never passed test	115.17	(18)	155.17	(18)

*Excludes respondents who could not recall amount spent.

CIRCUMVENTION AND MISUSE OF THE VEHICLE EMISSION TESTING PROGRAM

Over a fifth of respondents indicated that even though their engine could be adjusted so that it would pass the test, when so adjusted the vehicle doesn't run as well as it should. Nearly the same proportion said that they have had to readjust their engine after the test and roughly one in ten vehicle owners made such post-testing adjustments to their engines last year. The questions asked in this series were as follows:

"Some people have told us that although their vehicle engine can be adjusted so that it will pass the test, it really doesn't run as well as it should. Is this true in the case of your vehicle?" 100.0-102.0

(Base)	(1483)
Yes	22%
No	68
Not sure	10

"Have you ever had to readjust your engine after the test so that it runs the way you think it should?" 103.0-105.0

(Base)	(1483)
Yes	20%
No	78
Not sure	2

(Asked of those answering "yes" to the previous question:) "Did you have to do that the last time you took your vehicle in to be tested?" 106.0-108.0

(Base)	(301)
Yes	11%
No	89
Not sure	*

In general, the likelihood that one will make post-testing adjustments is strongest among younger vehicle owners as well as among those who view the state's air pollution laws as being too strong -- and importantly, among those whose vehicle failed the emissions test.

	Have Readjusted Their Engine After It Was Tested	
	At Any Time in the Past	After it Was Last Tested
Total	20%	11%
Age:		
Under 35	23	14
35 to 54	23	12
55+	13	5

Readjustment of the engine after the emissions test has occured among roughly seven percent of vehicle owners who passed the test on the first attempt. Thereafter, readjusting the engine after the second and third attempt to pass the test jumps to between 30 and 50 percent.

	Was Engine Readjusted After It Was Tested Last Time? (% Yes)	% Each Represents of Total Sample
Passed initial test	7%	5.4
Passed second test	31	4.5
Passed third test	32	.4
Failed third test; obtained waiver	53	.7

Any vehicle owner who made post-testing adjustments to their engine was next asked who had made the adjustments and what type of adjustments were made. As may be seen in the next table a majority of these owners either did the work themselves or had a friend or relative do it. Just under half paid a mechanic to do the work. 109.0-111.0

"Did you make the adjustments yourself, or did you have a mechanic do it?"

	Total	Phoenix	Tucson
Did it myself	42%	46%	32%
Friend/relative	10	10	10
Mechanic	46	42	56
Can't recall	2	2	2

In addition, it appears that paying a mechanic to complete the readjustment increases among those owners whose vehicles fail to pass the test either the first or second time.

Finally, most of those who make post-test engine adjustments focus on carburetor or timing adjustments. 112.0-114.0

"What type of adjustments were made?"

	Total	Phoenix	Tucson
Adjusted/rebuilt carburetor	72%	76%	63%
Adjusted timing/points	17	15	21
Had it tuned	5	5	5
Adjusted emission control devices	4	4	4
Cleaned/replaced plugs	3	2	5
Replaced air filter	2	1	5
Rebuilt the engine	1	1	0
Can't recall	10	7	16

Note: totals exceed 100% due to multiple responses.

Finally, we also asked respondents whether they had ever used the testing stations just to see how their car was running. Roughly one percent admitted they had so used the stations in the past. 115.0-117.0

"Some people say they occasionally use the testing stations just to see how their car is running -- that is, they take it to the emissions inspection station at times other than for the test required to get a license renewal. Have you ever taken your vehicle in for that purpose?"

	% Answering Yes
Total	.9
Area	
Phoenix	.7
Tucson	1.3

Persons who have used the station in this manner took their car in an average of 1.6 times each during the last 12 months. 118.0-120.0

Verbatim Comments Regarding Program Circumvention

"Mechanic adjusted carb so it would pass, then readjusted it later so it would run right. A lot of places are telling people up front that they'll do that now."

"My engine kept stalling all the time after my last tune-up...I assume the adjustments were made with the test in mind because I told them I was going when I took it in to be tuned. Mechanic opened up carb and set up idle after the test."

METHODOLOGY

The information contained in this report is based on a statistically valid telephone sample conducted among 805 vehicle-owning households in urban Arizona. Of the 805 total interviews, 406 were conducted among residents of the metropolitan Phoenix area and 399 among residents of the metropolitan Tucson area. Prior to generating the final detailed tables appended to this report, the sample was computer-weighted to reflect the actual proportional distribution of Phoenix (73%) and Tucson (27%) area households.

Household selection on this project was accomplished via a computer-generated random digit dial telephone sample which selected households on the basis of telephone prefix. This method was used because it ensures a randomly selected sample of area households proportionately allocated throughout the sample universe. This method also ensured that all unlisted and newly listed telephone households were included in the sample.

Up to three separate attempts -- on different days and at different times of day -- were made to contact each selected household. Only after three separate attempts was a selected household substituted in the sample.

Once a selected household was contacted, individuals were screened to ensure they were a valid survey respondent. Only individuals who met the following criteria were included in the survey:

- Their household had to own, or co-own, or lease a motor vehicle newer than 1968.

- They had to have personally taken, within the past 12 months, one of their household motor vehicles to an Arizona Motor Vehicle Inspection Station for an emissions test.

All of the interviewing on this project was conducted during September, 1982, at Behavior Research Center's central location telephone facility. Each interviewer working on this project was under direct BRC supervision 100 percent of the time. One hundred percent of the interviews were edited and any containing errors were pulled, the respondent re-called, and the error corrected. In addition, 15 percent of each interviewer's work was randomly selected for validation to ensure its authenticity and correctness.

The questionnaire used in this study was designed by BRC in consultation with the Auditor General's Office. After the Auditor General's Office approved the preliminary questionnaire, it was pre-tested with a randomly selected cross-section of 15 qualified households. The pre-test focused on the value and understandability of the questions, adequacy of response categories, questions for which probes were necessary, and the like. Several changes were made in the questionnaire after the pre-test. Following this, the amended questionnaire was again pre-tested and then re-submitted to the Auditor General's Office where it received final approval.

As the data collection segment of this study was undertaken, the completed interviews were turned over to the BRC coding department. The coding department edited, validated and coded the interviews. Each interview that received final coding department approval was then transferred to keypunching where all were 100 percent key verified.

Following completion of keypunching, a series of validity and logic checks were run on the data to ensure it was "clean". Following these checks, the computer tables presented in the appendix of this report were generated.

When analyzing the results of this survey it should be kept in mind that all surveys are subject to sampling error. Sampling error, stated simply, is the difference between the results obtained from a sample, and those which would be obtained by surveying the entire population under consideration. The size of sampling error varies, to some extent, with the number of interviews completed and with the division of opinion on a particular question. The sampling error for this study (at a 95 percent confidence level), is approximately ± 3.5 percent for the total sample, and ± 5.0 percent for the individual Phoenix and Tucson sub-samples.