EPA-AA-TSS-I/M-87-1

Technical Report

A Discussion of Possible Causes of Low Failure Rates in Decentralized I/M Programs

By

Eugene J. Tierney

January 1987

NOTICE

Technical Reports do not necessarily represent final EPA decisions or positions. They are intended to present technical analysis of issues using data which are currently available. The purpose in the release of such reports is to facilitate the exchange of technical information and to inform the public of technical developments which may form the basis for a final EPA decision, position or regulatory action.

Technical Support Staff
Emission Control Technology Division
Office of Mobile Sources
Office of Air and Radiation
U. S. Environmental Protection Agency

ABSTRACT

This technical report reviews six possible explanations for low reported failure rates in manual, decentralized I/M programs. The report analyzes and discusses random roadside idle survey data, reported I/M program data and data collected during audits of I/M programs. The data indicate that five of the explanations: quality control, fleet maintenance, differences in fleet mix or emission standards, anticipatory maintenance, and pre-inspection repair, do not sufficiently explain low reported failure rates. The report concludes that the major problem contributing to low reported failure rates in decentralized, manual I/M programs is improper inspections by test station personnel.

TABLE OF CONTENTS

SECTION	PAGE
Abstract	i
Table of Contents	ii
List of Tables	iii
List of Figures	iv
Introduction	1
Quality control issues	4
Better maintained fleet	6
Standards and coverage differences	6
Anticipatory maintenance	7
Pre-inspection repair	9
Improper Inspection	9
Conclusions	25
References	26

LIST OF TABLES

TABLE	TITLE	PAGE
1	Failure Rates in I/M Programs	2
2	Potential Causes For Lower Than Expected Failure Rates	3
3	Potential Impact of Quality Control Deficiencies on I/M Failure Rates	5
4	Non-I/M Vehicle Failure Rates Using Consistent Emission Standards	6
5	Reported Failure Rates vs. Survey Failure Rates	10

LIST OF FIGURES

FIGURE	TITLE	PAGE
1	Comparison of Idle Surveys	11
2	Frequency Distribution of Emission Scores	13
3	Cumulative Distribution of Emission Scores	14
4	Initial Test Scores Cumulative Distribution Model Year 1977	15
5	Retest Scores for Connecticut versus Initial Test Scores, Model Year 1977	16
6	Initial Test Scores Cumulative Distribution Model Year 1980	17
7	Retest Scores for Connecticut versus Initial Test Scores, Model Year 1980	18
8	Initial Test Scores Cumulative Distribution Model Year 1982	19
9	Retest Scores for Connecticut versus Initial Test Scores, Model Year 1982	20
10	Repeat Index - Colorado	22
11	Repeat Index - Virginia	23

INTRODUCTION

Inspection and maintenance (I/M) programs are currently operating in thirty-one states and affect approximately one third of all light duty cars and light duty trucks in the country. The annual inspection cost of these programs is in the neighborhood of $500 million. Given the significant impact, it is important to carefully assess the outcomes of these programs on an individual basis with an eye toward effectiveness and cost-efficiency. The object of an I/M program is to identify vehicles that are "gross emitters" of hydrocarbons (HC) and/or carbon monoxide (CO) and require emission-related repairs of those vehicles such that emissions are reduced. Thus, one important indicator of program effectiveness is the percentage of vehicles that are identified as gross emitters (i.e., fail the emissions test).

Two basic approaches to inspecting vehicles have been implemented among I/M programs: centralized and decentralized. In centralized programs, motorists bring their vehicles to high volume test facilities operated by the state or local government or by a contractor hired by the state or local government. The repair function is independent of the test function and the centralized facilities are generally highly automated and systematic. Decentralized I/M programs generally have few or no high volume stations. The state or local government licenses service stations, automobile dealerships and the like to do inspections. Motorists have the option of obtaining repairs at the licensed facility or going elsewhere. Two distinct types of decentralized inspection programs exist: ones that use manual emission analyzers and ones that use computerized analyzers. In the latter case, a computer is built into the analyzer that controls the test procedure, the selection of emission standards, the pass/fail decision, data recording, and quality control. In the case of manual analyzers, no computer is available; so, the inspector is responsible for quality control, chooses emission standards, reads meters or digital displays for emission levels, decides pass/fail status, and records the test data.

In its role as an oversight agency, EPA has been conducting audits under the National Air Audit System guidelines and has been gathering data collected by individual I/M programs. EPA has also been conducting random roadside tampering and idle surveys in cities throughout the country for many years. Analysis of these data has revealed some significant findings: first, emission test failure rates in I/M programs vary widely, from a low of about 2% to a high of 28% (see Table One); second, failure rates vary by program type. Decentralized programs with manual analyzers tend to have very low failure rates while centralized programs and decentralized programs with computerized analyzers tend to have much higher failure rates.

Table One

EMISSION TEST FAILURE RATES IN I/M PROGRAMS*

	REPORTED %	EXPECTED %	RATIO
CENTRALIZED (CC)			
Arizona	20.2	36.8	.55
Connecticut	17.2	33.0	.52
Delaware	13.7	7.7	1.00
Kentucky	15.7	16.2	.97
Maryland	14.6	14.0	1.00
Memphis, TN	8.1	3.7	1.00
Nashville, TN	24.5	25.4	.97
New Jersey	26.1	27.8	.94
Oregon	24.0	38.3	.63
Washington, D.C.	18.4	13.4	1.00
Washington	19.0	28.1	.68
Wisconsin	15.3	19.3	.79
DECENTRALIZED			
Computerized Analyzers (DC)			
Alaska			
Fairbanks	19.4	22.7	.85
Anchorage	15.7	24.7	.63
California	27.7	28.7	.96
Michigan	15.8	12.9	1.00
New York**	5.1	33.4	.15
Pennsylvania	17.6	19.5	.90
Manual Analyzers (DM)			
Georgia	6.6	25.0	.26
Idaho	9.8	16.9	.58
Missouri	6.7	20.5	.33
North Carolina	5.6	21.1	.27
Nevada			
Clark County	9.5	29.4	.32
Washoe County	11.0	29.4	.37
Utah			
Davis County	8.7	21.3	.41
Salt Lake County	10.0	21.3	.47
Virginia	2.3	15.6	.15

* For all model years, including light duty trucks.

** New York's analyzers are only partially computerized.

Table One lists reported failure rates from the most recent year available for each program, expected failure rates and the ratio of reported to expected failure rates. The reported failure rates are provided to EPA in several ways: as lump sum failure rates for all model years, as failure rates by model year, or as failure rates by model year group. In the latter two instances, national default registration distributions were used to weight separate model year or group failure rates together into one overall failure rate. As a result, the reported overall failure rate here may differ slightly from the actual overall failure rate experienced in each program. EPA does not have available to it the distribution of vehicles tested to make more precise calculations. The expected failure rates are based on the emission standards used in the program applied to the 1984 Louisville I/M data base. This data base was chosen because it is the best available data base for this purpose. It represents the non-I/M fleet and covers light-duty cars and light-duty trucks. The failure rate ratio is the reported failure rate divided by the expected failure rate (in the few cases where this yields a number greater than one, the result was rounded down to one).

This report discusses six potential causes for the differences in reported failure rates among I/M programs (see Table Two) and cites available data to support or cast doubt upon them. The data come from a variety of sources: audit reports conducted under the National Air Audit System guidelines; reported I/M program failure rates, emissions scores and other data; random roadside idle emission surveys conducted by EPA; and, various contractor studies conducted for EPA. Analysis of I/M program data is limited by the program design: how data is handled and reported to EPA limits the kinds of analyses that can be conducted. Since resource constraints naturally exist, contractor studies have not generally involved analyses of all I/M programs but rather selected programs that are representative of the different types of programs. As a result, the findings must be extrapolated to other programs of a similar type.

Table Two

POTENTIAL CAUSES FOR LOWER THAN EXPECTED FAILURE RATES

1) Quality control issues
2) Better maintained fleet
3) Standards and coverage differences
4) Anticipatory maintenance
5) Pre-inspection repair
6) Improper inspection

QUALITY CONTROL ISSUES

Non-dispersive infra-red analyzers are used in all I/M programs to determine emission levels of hydrocarbons and carbon monoxide from motor vehicles subject to the program requirements. Proper calibration of emission analyzers is essential for obtaining accurate test results. Quality control requirements vary somewhat from program to program but several common elements exist:

1) Weekly calibration of analyzers
2) Low range calibration gas
3) Weekly leak check
4) Periodic audit of analyzers by program officials
5) Zero and span within one hour of each test

In programs with computerized analyzers, some quality control functions are done automatically and software protection exists that prevents use of the analyzer unless the leak check and gas calibration check have been conducted and passed within the past seven days. The analyzer software guides the inspector through the steps necessary to complete quality control checks thereby insuring consistency and accuracy. EPA audits of computerized analyzer programs usually show few analyzers failing quality control checks.

In programs with manual analyzers, quality control is done manually and it is up to the inspector to insure that it gets done. Typically, programs specify that the weekly gas calibration and leak check take place each Monday morning. Nothing exists to prevent use of the analyzer if the quality control functions are not performed except periodic audits by program officials. EPA audits have shown that analyzers in manual programs are frequently out of calibration, possess leaks, or have other problems that can severely compromise test quality (e.g. clogged filters).

Quality control for calibration gas is accomplished two ways in I/M programs: through periodic station audits and calibration gas specifications. Most programs conduct monthly audits and the auditors carry gas cylinders to check analyzer accuracy. This also accomplishes a check on calibration gas accuracy because once erroneous calibration is eliminated continued failure of an analyzer usually indicates a gas problem. Calibration gas specifications are fairly consistent among I/M programs. Most specify an accuracy of $\pm$ 2% and require specific concentrations (i.e., a zero blend tolerance) of 1.6% CO and 600 ppm HC in nitrogen.

Quality control lapses will diminish the accuracy of emission scores and, to some degree, will alter the pass/fail outcome. The question is whether quality control lapses could explain the low reported failure rates experienced in some I/M programs. There are three sources of data that will help answer this question: audit data, idle survey data, and operating data.

If analyzers are typically reading low or have significant leaks in the sample system, vehicles with emission scores close to standards may pass the test when they should fail. A review of the gas audit results obtained during EPA audits in Georgia, North Carolina, Idaho and Missouri(1) shows that, on average, analyzers were 2% out of calibration on the high side. This means that vehicles that have emission scores within that range will tend to incorrectly fail rather than incorrectly pass.

An analysis of the Virginia I/M data was also conducted to determine the impact of increasing all emission scores by 5%. Five percent was chosen because the audit data showed that over 80% of the emission analyzers checked were within the 5% tolerance; it is also the audit tolerance used in most I/M programs. The overall failure rate for the 1975 through 1984 vehicles sampled for this analysis increased from 3% to 4.2% when emission scores were increased 5%.

Finally, random, roadside idle survey data(2) were analyzed to determine the impact of lowering emission scores by 100 ppm HC and 1% CO to simulate the results which would have been obtained if the analyzers were severely out of calibration or had gross leaks in the sampling system. Table Three shows, for all model years, the idle survey failure rates and the idle survey failure rates with the cushion added. Note that no dramatic drop in failure rates occurs as a result of a cushion and that survey failure rates with the cushion are still much higher than reported rates in manual I/M programs. These three analyses show that low reported failure rates are not explained by typical quality control deficiencies in manual I/M programs.

Table Three

POTENTIAL IMPACT OF QUALITY CONTROL DEFICIENCIES ON I/M FAILURE RATES

STATE	PROGRAM TYPE	REPORTED FAIL RATE	SURVEY FAIL RATE	SURVEY RATE WITH CUSHION
Connecticut	CC	17.2%	16.0%	13.2%
Missouri	DM	6.7%	16.2%	11.2%
New Jersey	CC	26.1%	34.1%	27.1%
New York	DC	5.1%	22.2%	19.1%
North Carolina	DM	5.6%	17.6%	14.7%
Pennsylvania	DC	17.6%	18.4%	15.3%
Virginia	DM	2.3%	16.0%	13.5%

BETTER MAINTAINED FLEET

There is no evidence to indicate that mechanic effectiveness varies significantly from region to region or that vehicle owners are more conscientious about getting repairs in one state or another. Nevertheless, it is conceivable that better general maintenance in an area could result in a cleaner fleet, overall, and lower than expected I/M failure rates. If this were the case, then there should be a significant difference between failure rates of non-I/M survey vehicles (i.e. those registered outside the program boundaries but surveyed while operating within the I/M boundaries) among I/M areas. Table Four illustrates the survey failure rates at constant cutpoints (pre-1981: 3.0% CO/300 ppm HC; post-80: 1.2% CO/220 ppm HC) for non-I/M vehicles in two groups of areas where EPA has conducted random, roadside idle surveys. The data were grouped due to the small sample size of non-I/M vehicles in I/M areas. The members of the group were determined based on reported failure rates, VA, et.al. being low and PA, et.al. being high. Note that the failure rates are very similar between the two groups. This indicates that maintenance differences between areas does not seem to be influential on non-I/M vehicle failure rates. By extension, this is likely to be true of I/M vehicles in these areas as well.

Table Four

NON I/M VEHICLE FAILURE RATES
USING CONSISTENT EMISSION STANDARDS

MODEL YEARS	VA,NC,MO,NY	PA,CT,NJ,OR
Post 1980	7.1%	6.4%
Pre-1981	50.6%	54.7%
Overall	20.7%	24.2%

EMISSION STANDARD AND VEHICLE COVERAGE DIFFERENCES

Emission standards are established by each program and are used to determine whether a vehicle must be subjected to repairs to bring emission levels down to acceptable levels. No two I/M programs have identical emission standards for all model years, although most programs use the same standards for 1981 and later vehicles (1.2% CO and 220ppm HC). The difference in standards from program to program should result in different failure rates. Two other important factors that contribute to expected differences in overall failure rates among programs are model year coverage and vehicle type coverage. Thus, it is conceivable that low failure rates in some I/M programs may be due to one or a combination of these factors.

To evaluate this question, the emission standards and vehicle coverage from every I/M program were applied to a common data base consisting of emission distributions from the Louisville I/M program[3]. The Louisville program data is for calendar year 1984 and represents the non-I/M fleet since the Louisville program started January 1, 1984. Table One lists the expected failure rates for twenty-one I/M programs. The expected failure rates range from a low of 3.7% in Memphis, Tennessee to a high of 38.3% in Portland, Oregon.

Expected failure rates for most I/M programs fall into the 20% to 40% range. In particular, the expected failure rates for manual decentralized programs also fall into this range, with the exception of Virginia and Idaho which have expected failure rates of 15.6% and 16.9%, respectively. It is clear that, based on this analysis, the combination of emission standards, vehicle coverage and model year coverage yields a range of expected failure rates but does not explain the low reported failure rates of decentralized manual I/M programs.

To evaluate this question further, the ratios of reported failure rates to the expected failure rates were calculated. The third column in Table One shows the ratio of reported to expected failure rates, hereafter referred to as failure rate fractions. The failure rate fractions in manual decentralized programs are all under 0.4, except Idaho and Utah. Most other I/M programs have failure fractions over 0.7. Some differences in failure rate fractions are anticipated due to variations in historical emission standards, tampering program coverage, waiver rates, pre-conditioning, and length of operation of the program. For example, the emission standards used throughout the life of the program are very significant. Note in Table One that, among the centralized programs, the three with the lowest failure rate fractions also have the highest expected failure rates. These programs have a history of tight emission standards which has led to lower current failure rates than otherwise expected. The failure rate fractions in centralized programs reflect normal program variations. However, these variations are never large enough to cause the low failure rates in manual, decentralized programs. Thus, the normal range of failure rate fractions also shows that differences in emission standards and model year coverage do not explain low failure rates in manual decentralized I/M programs.

ANTICIPATORY MAINTENANCE

Anticipatory maintenance occurs when a motorist obtains repairs on a vehicle due to have an inspection in the near future with the intent of avoiding test failure but without the knowledge that the vehicle would, in fact, fail the test. There is no doubt that anticipatory maintenance occurs, but it is not clear to what extent it occurs and there is no evidence to show it occurs more frequently in one program or another. It is conceivable that anticipatory maintenance could contribute to low initial test failure rates.

Data collected by EPA indicate that anticipatory maintenance may not achieve its intended goal; in fact, some evidence indicates it will increase chances of test failure! In 1979 and 1980, EPA conducted several studies[4,5,6] to determine the potential emission benefits of a mandatory vehicle maintenance program. In these studies, mechanics were asked to adjust vehicles to manufacturer specifications. The mechanics were not aware that they were being tested; some of the vehicles were out of adjustment (i.e., Federal Test Procedure (FTP) failures) and some were not. In Houston, adjustment by mechanics resulted in a 2.3% increase in FTP mass HC emissions and a 2.7% increase in FTP mass CO emissions. In a St. Louis study, an 87% increase in idle HC and a 30% increase in idle CO emissions were observed in 83 shops. In general, vehicles that were FTP failures before "repairs" showed some reduction in emissions while clean cars typically suffered emission increases. The two situations studied here are analogous to that of anticipatory maintenance: mechanics are not being asked to fix the vehicle in response to an I/M failure.

Even when repairs occur in response to an I/M failure, successful emission reductions are not assured. Retest failure rates in centralized I/M programs are typically in the 30-40% range. Thus, it is reasonable to assume that anticipatory maintenance may not be very successful, especially when applied to "clean" cars.

Another reason anticipatory maintenance does not seem to be a satisfactory explanation for low failure rates in decentralized manual I/M programs is that its effect, if important, should be felt in other types of programs as well. In fact, it is arguable that the effect should be greater in centralized programs. In decentralized programs, the testing and repair functions are combined in the service station environment. The normal sequence of events is for a customer to visit the garage, get an emissions test and if a failure occurs, to obtain repairs at that facility. Given this scenario, there is little motive for anticipatory maintenance. In centralized programs, where repair functions are separate from testing functions, the motorist has to make at least three trips if an initial test failure occurs. This provides an incentive to avoid initial test failure especially when a failure was experienced in a previous year. Data from the Arizona and Seattle, Washington I/M programs show that, in those centralized programs, vehicles that failed in the previous year fail at higher than average rates. This implies that motorists are not attempting to avoid failures through anticipatory maintenance or that such maintenance is unsuccessful.

PRE-INSPECTION REPAIR

Pre-inspection repair may occur in two basic ways: first, a vehicle is brought in for an emission test and an unofficial initial test is conducted to determine if the vehicle will pass. If the vehicle fails, repairs are conducted such that it will pass and then the official initial test follows. The second scenario occurs when a vehicle is brought in for a tune-up plus an emission test and, again, repairs precede the official test. It is likely that this phenomenon occurs in both manual and computerized decentralized I/M programs. To the extent that this is the case, it would lower initial test failure rates but is a phenomenon confined to decentralized programs.

The structure and requirements in manual and computerized decentralized I/M programs are essentially the same with exception of the analyzer. Thus, the opportunities and incentives for pre-repair are about the same in both types of program. So, the impact of pre-repair in terms of failure rates should be the same as well. However, the data in Table One show that decentralized programs with fully computerized analyzers do not experience the very low failure rates of decentralized manual programs. This indicates that pre-inspection repair does not seem to be a big issue in computerized programs. The software prompts in the Michigan computerized analyzer include a question regarding repairs within a week preceding the initial test. The data[7] for the first and second quarters of 1986 show that for vehicles passing the initial test, about 6-7% were known to have received repairs within the past week. It is not known how many of these vehicles received anticipatory maintenance as opposed to pre-inspection repair. It is also unknown whether the repairs were actually needed to pass the initial test or effective at reducing emissions. In any case, these data indicate that, taken together, pre-inspection repair and anticipatory maintenance are not common phenomena and therefore are not satisfactory explanations for low reported failure rates.

IMPROPER INSPECTION

Improper inspection is believed to occur in several ways: inspectors skip the emissions test and invent passing emission scores; inspectors conduct the test but still invent passing emission scores without doing repairs; inspectors conduct the test, do repairs but, failing to bring the vehicle into compliance, they invent passing emission scores. We can imagine a host of variations on these three basic scenarios, but all have one important factor in common: emission reductions are not achieved. Improper inspection has been found to occur in I/M programs through covert audits. What is unclear is the magnitude of the problem.

One way to assess whether emission reductions are not actually occurring is to randomly test I/M vehicles. EPA has conducted random idle surveys throughout the country. Figure One, on the next page, illustrates the results for four eastern states: North Carolina, Connecticut, Pennsylvania and Virginia. The graph lines illustrate the failure rate of 1981 and later vehicles (using each program's own cutpoints) over time since last inspection. Note that the rates for North Carolina and Virginia are relatively flat - the failure rates start high and end high. On the other hand, the failure rates for Pennsylvania and Connecticut start low and end high. The symbols on the right Y axis are the reported failure rates for the four programs. The reported failure rates in Pennsylvania and Connecticut are essentially identical to the 12 month survey failure rates. The reported rates for Virginia and North Carolina are much lower than the survey rates. Two things are apparent from these results: vehicles sampled in the survey in Virginia and North Carolina do not seem to have their emissions lowered after inspection, and the reported failure rates do not accurately reflect the actual idle emissions of sampled vehicles.

The data presented in Table Five show a comparison of overall I/M failure rates and reported failure rates for cars in the same four states. The reported failure rate in Pennsylvania and Connecticut is very similar to the survey failure rate. In North Carolina and Virginia, the reported failure rates are much lower than the survey failure rates. Again, the data indicate that I/M cars are not achieving significant emission reductions.

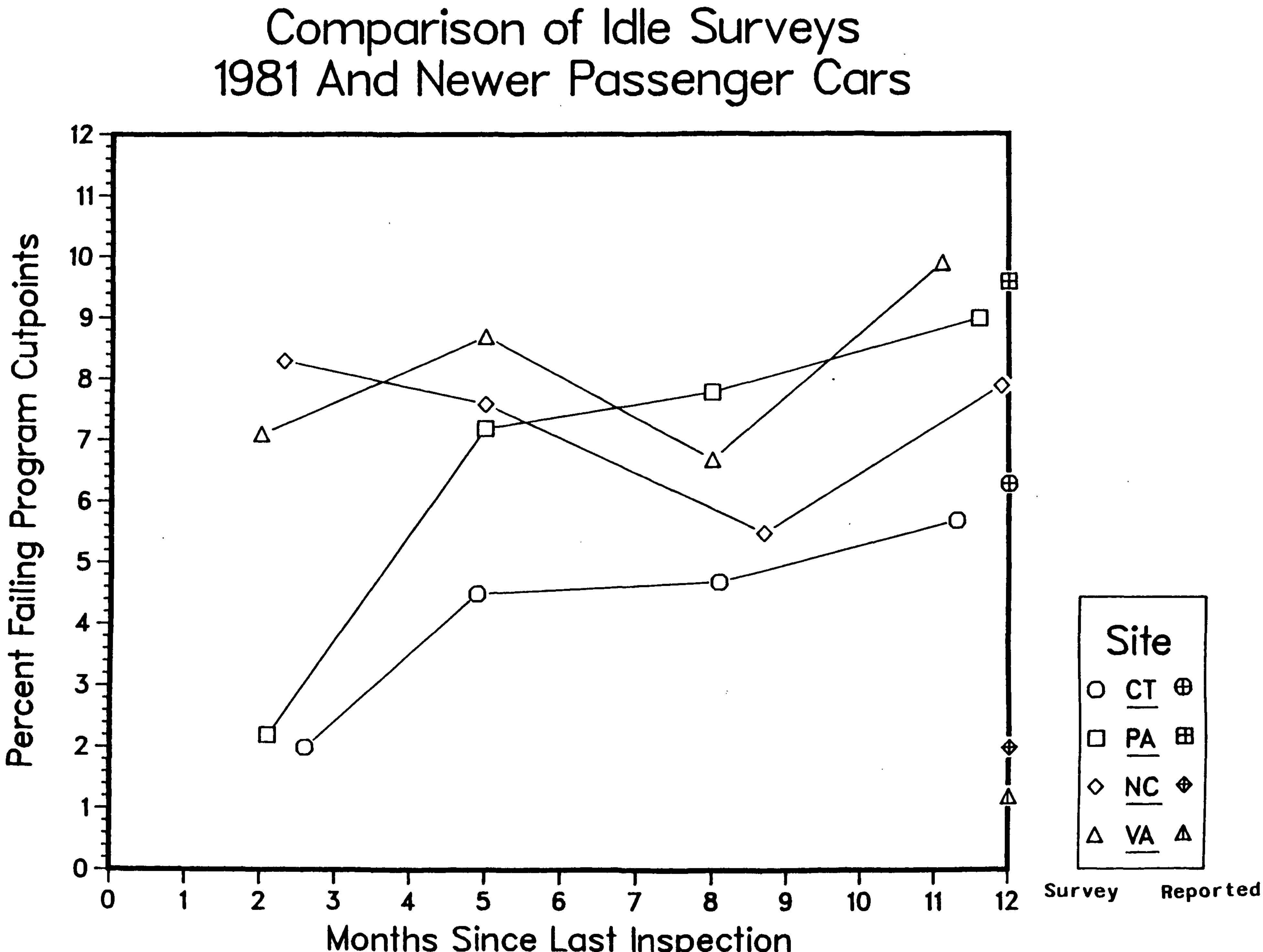
Table Five

REPORTED FAILURE RATES VS. SURVEY FAILURE RATES

STATE	PROGRAM TYPE	REPORTED	LOCAL SURVEY I/M CARS
North Carolina	DM	5.6%	17.8%
Virginia	DM	2.3%	14.9%
Pennsylvania	DC	17.8%	18.2%
Connecticut	CC	17.2%	16.4%

An EPA contractor was given a work assignment in 1986 to analyze I/M program data to determine if significant differences were present between different I/M programs. The draft report[8] from this work has been completed. The contractor analyzed and compared reported data from I/M programs in Washington, Virginia, New York, Colorado and Connecticut (an analysis of Massachusetts data is in progress and will be available in early 1987).

Figure One


Comparison of Idle Surveys
1981 And Newer Passenger Cars
Percent Failing Program Cutpoints
0
1
2
3
4
5
6
7
8
9
10
11
12
Months Since Last Inspection
Site
CT
PA
NC
VA
Survey
Reported

Three representative model years were analyzed to reduce the enormity of the task: 1977, 1980 and 1982. Manufacturers were broken up into three groups: G1 consists of Chrysler, Ford and AMC; G2 consists of only GM vehicles; and, G3 consists of all imports. Figure Two provides an example frequency distribution of carbon monoxide failure rates from four programs. The X axis is the CO emission score and the Y axis is the percent with that score in the sample. The Virginia and Colorado data are very "spiky" since the emission readings are manually recorded and mechanics appear to round off the readings. Note also that the decentralized programs show a step change in the distributions exactly at the program cutpoints. Additionally, the distributions show a very large number of vehicles (relative to Washington) just below the cutpoints. This distribution is typical of what was found in other model years, in other vehicle groups, and for HC.

The contractor also produced cumulative distributions to overcome the distortion created by the spikes in the data. In these curves, the gradient of the curve is proportional to the number of vehicles at the particular emission level. Figure 3 shows the distinctive kink in the curves at the program cutpoints except in Washington, which shows a smooth distribution. To further assess the question of whether the kink in the curves could be due to pre-repair, the contractor compared the first test results from the decentralized programs with two sets of results:

1) the first test results of centralized programs.
2) the first test results of centralized programs for passing vehicles and the retest results for failed vehicles combined into one distribution.

The latter comparison is intended to simulate the distribution resulting from pre-repair. Figures 4 through 9 show the results of these analyses. The contractor's analysis of these figures succinctly states the case:

> From Figure Five it is obvious that Connecticut's vehicles have much lower emissions after repair, and the New York, Colorado and Virginia curves lie between the Connecticut first test (Figure Four) and after repair distributions (Figure Five). This may indicate that a fraction of the cars are being pre-repaired and that the average lies between the two Connecticut distributions. However, examination of the 1980 and 1982 distributions in Figures Six to Nine shows that this hypothesis is unlikely to be correct. For both model years the Connecticut's initial test distributions shows lower CO values than the decentralized program distributions. Moreover, a substantial portion of the population in each decentralized program appears to have CO emissions just below the cutpoint (as indicated by the steepness of the line just below the CO cutpoint). It is [usually] not possible to repair 1980 and 1982 cars to "just meet" the

Figure Two

FREQUENCY DISTRIBUTION OF REPORTED EMISSION SCORES IN THREE DECENTRALIZED AND ONE CENTRALIZED I/M PROGRAM

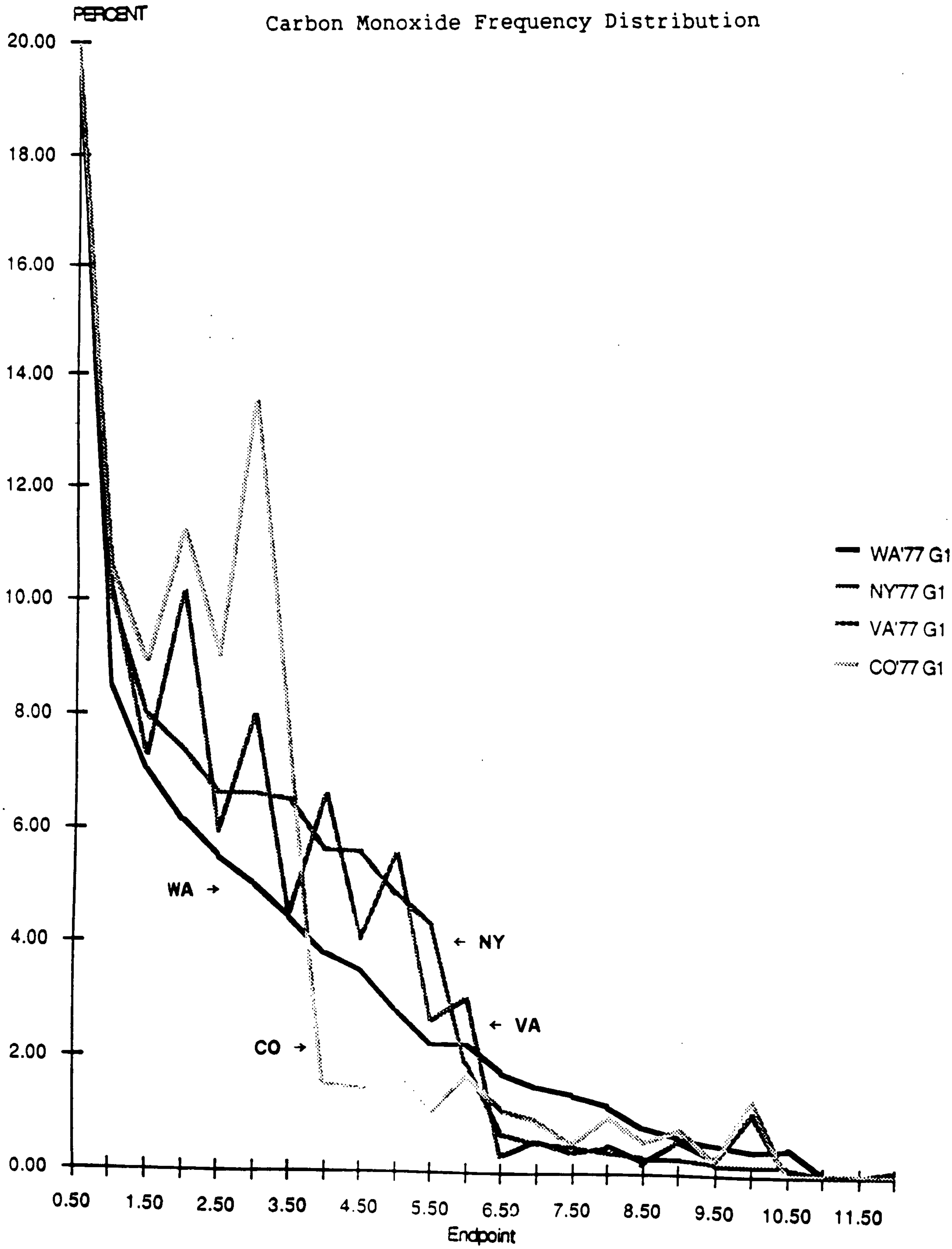


Figure Three

CUMULATIVE DISTRIBUTION OF REPORTED EMISSION SCORES

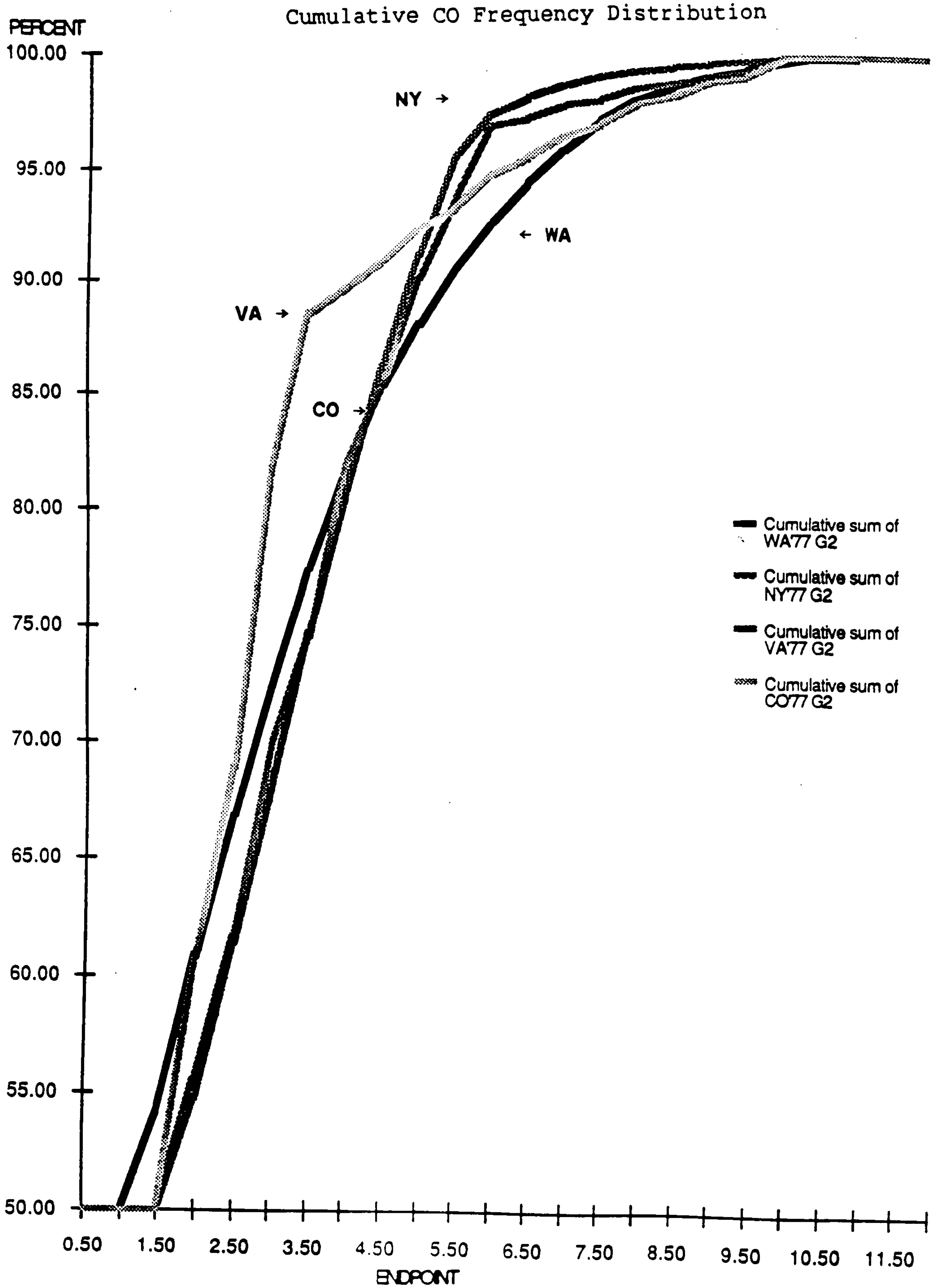


Figure Four

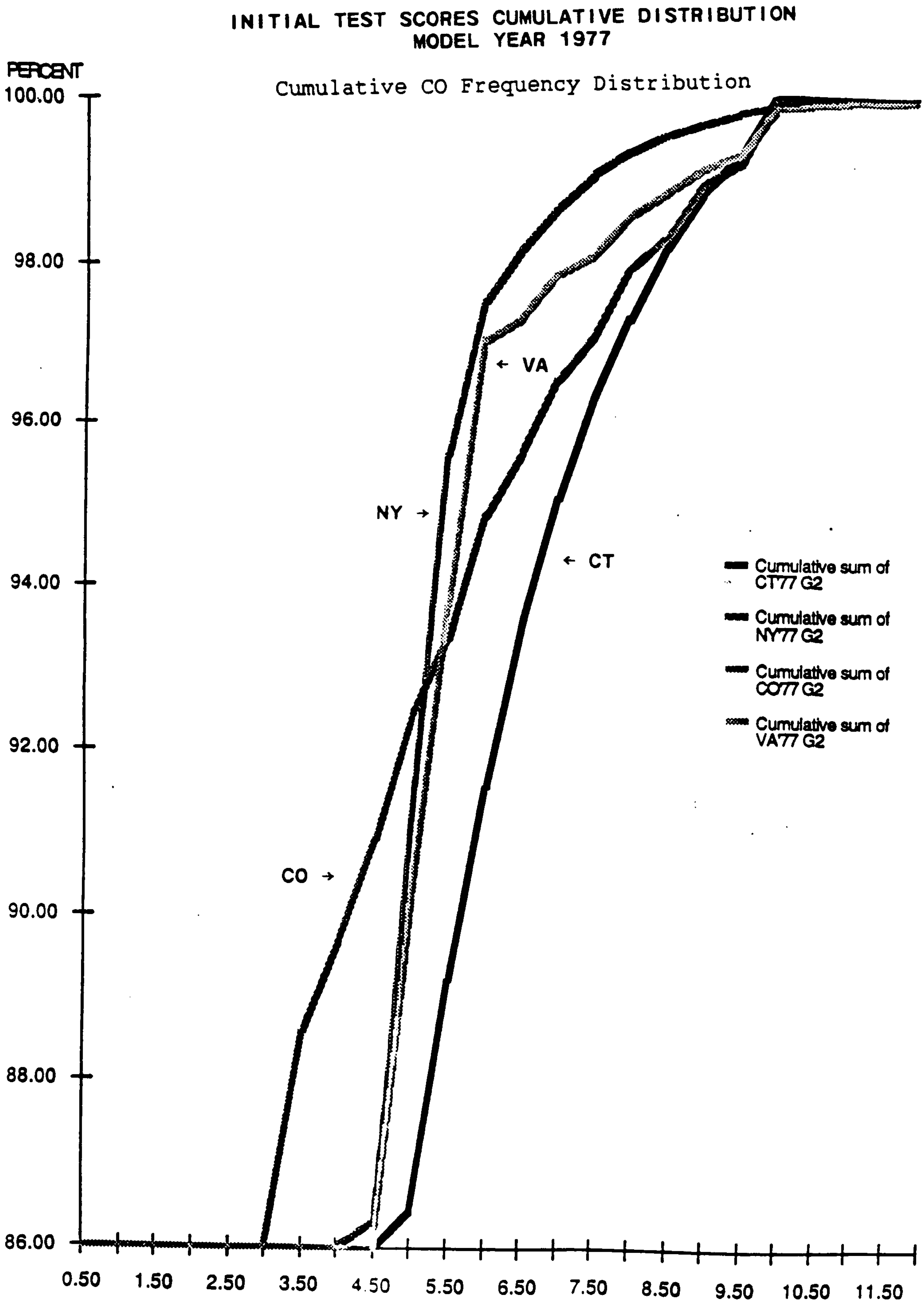

INITIAL TEST SCORES CUMULATIVE DISTRIBUTION
MODEL YEAR 1977
PERCENT
Cumulative CO Frequency Distribution
100.00
98.00
96.00
94.00
92.00
90.00
88.00
86.00
0.50
1.50
2.50
3.50
4.50
5.50
6.50
7.50
8.50
9.50
10.50
11.50
VA
NY
CT
CO
Cumulative sum of CT'77 G2
Cumulative sum of NY'77 G2
Cumulative sum of CO'77 G2
Cumulative sum of VA'77 G2

Figure Five

CONNECTICUT RETEST SCORES VERSUS
INITIAL TEST SCORES IN DECENTRALIZED PROGRAMS
MODEL YEAR 1977

Cumulative CO Frequency Distribution

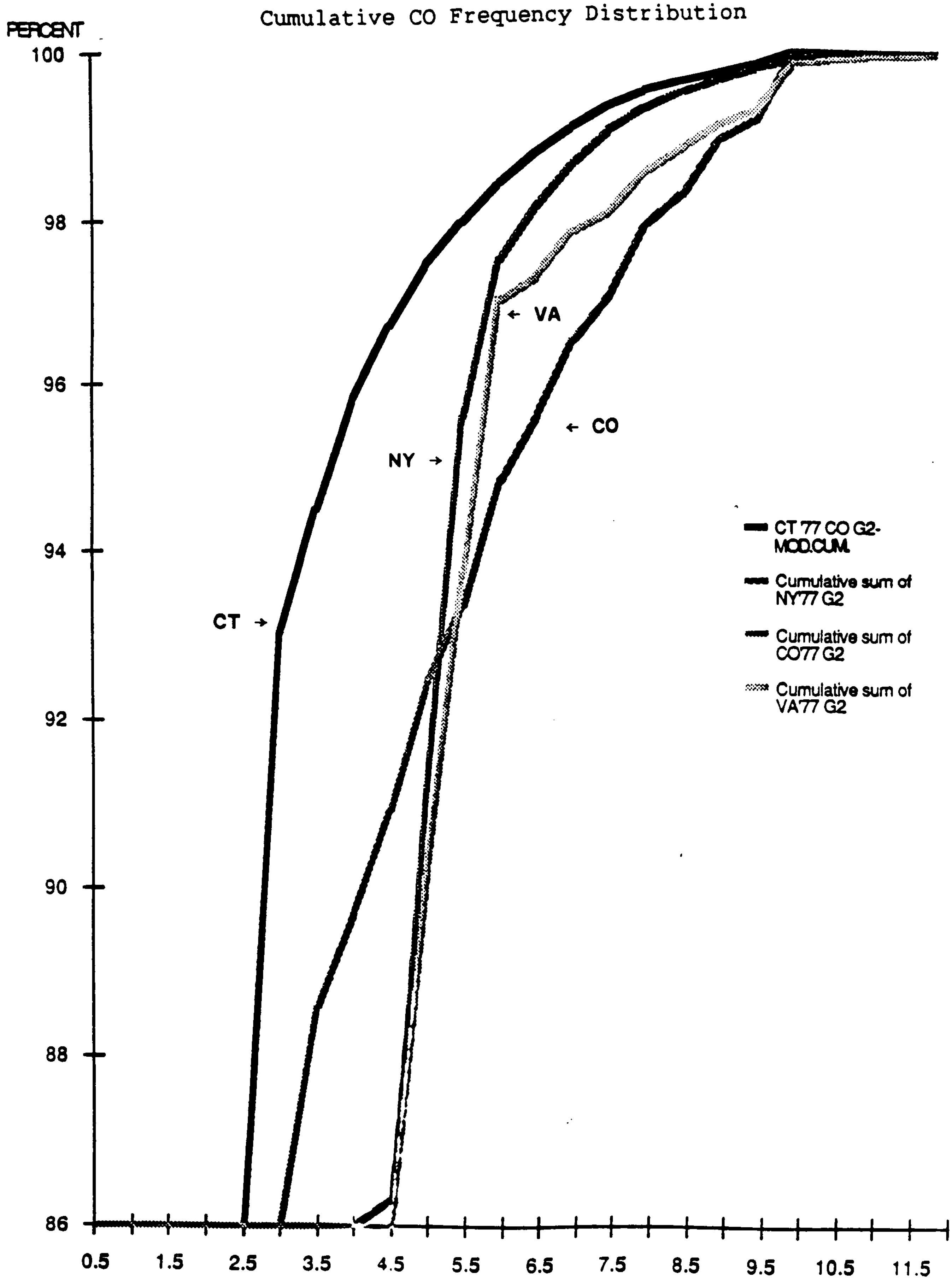


Figure Six

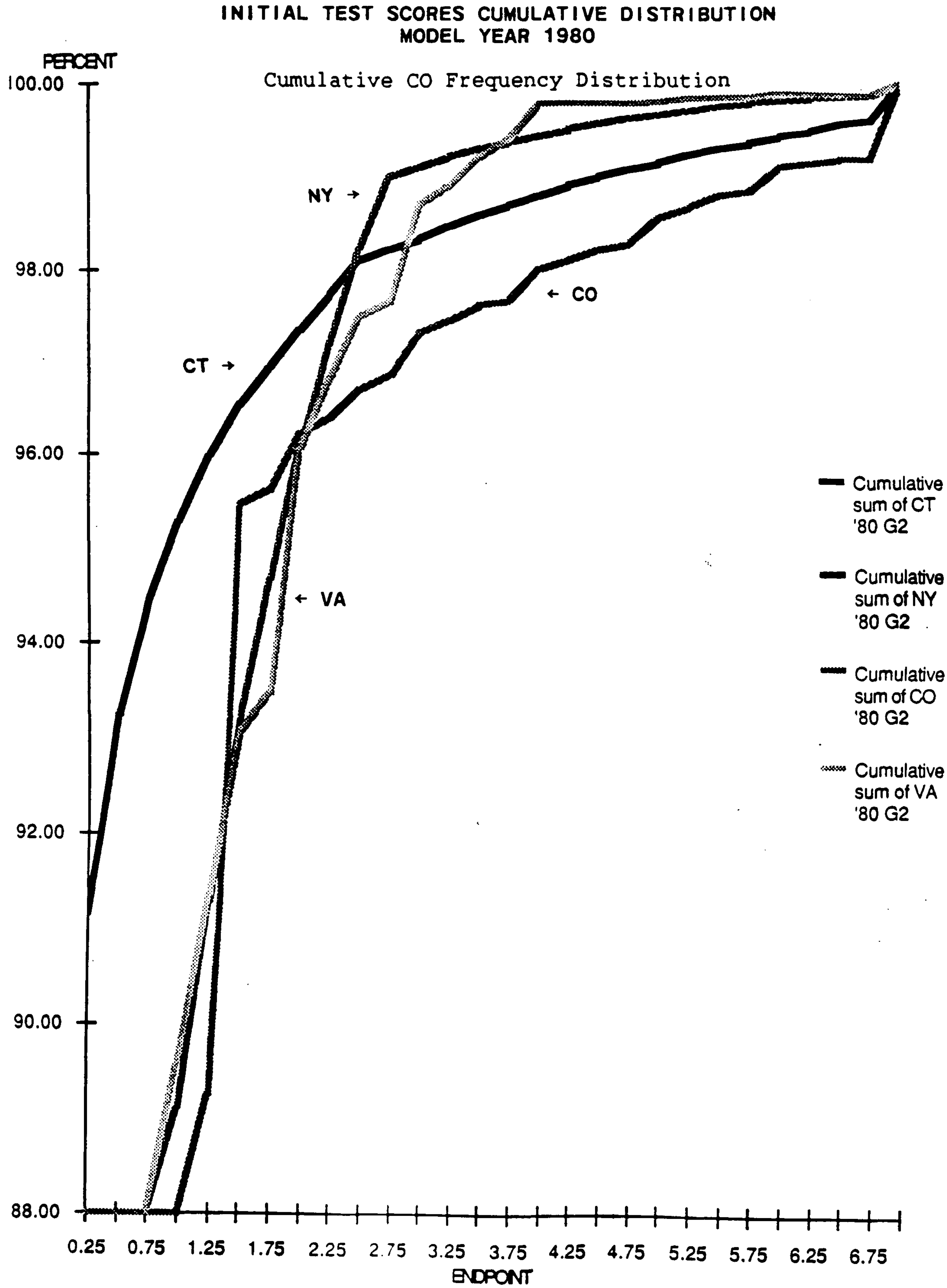

INITIAL TEST SCORES CUMULATIVE DISTRIBUTION
MODEL YEAR 1980
Cumulative CO Frequency Distribution
PERCENT
100.00
98.00
96.00
94.00
92.00
90.00
88.00
NY →
CT →
← CO
← VA
Cumulative sum of CT '80 G2
Cumulative sum of NY '80 G2
Cumulative sum of CO '80 G2
Cumulative sum of VA '80 G2
0.25
0.75
1.25
1.75
2.25
2.75
3.25
3.75
4.25
4.75
5.25
5.75
6.25
6.75
ENDPOINT

Figure Seven

CONNECTICUT RETEST SCORES VERSUS
INITIAL TEST SCORES IN DECENTRALIZED PROGRAMS
MODEL YEAR 1980

Cumulative CO Frequency Distribution

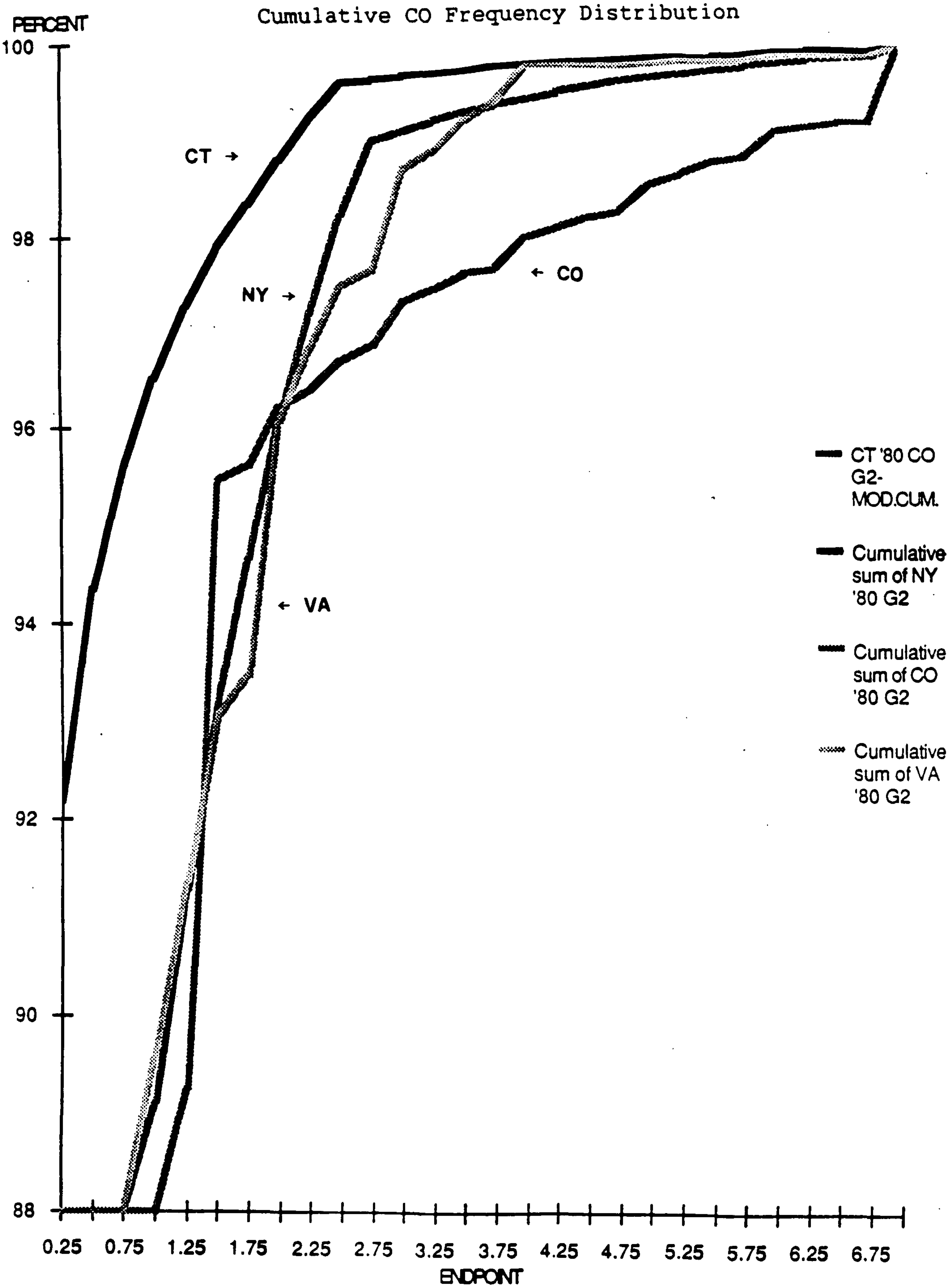


Figure Eight

INITIAL TEST SCORES CUMULATIVE DISTRIBUTION
MODEL YEAR 1982

Cumulative CO Frequency Distribution

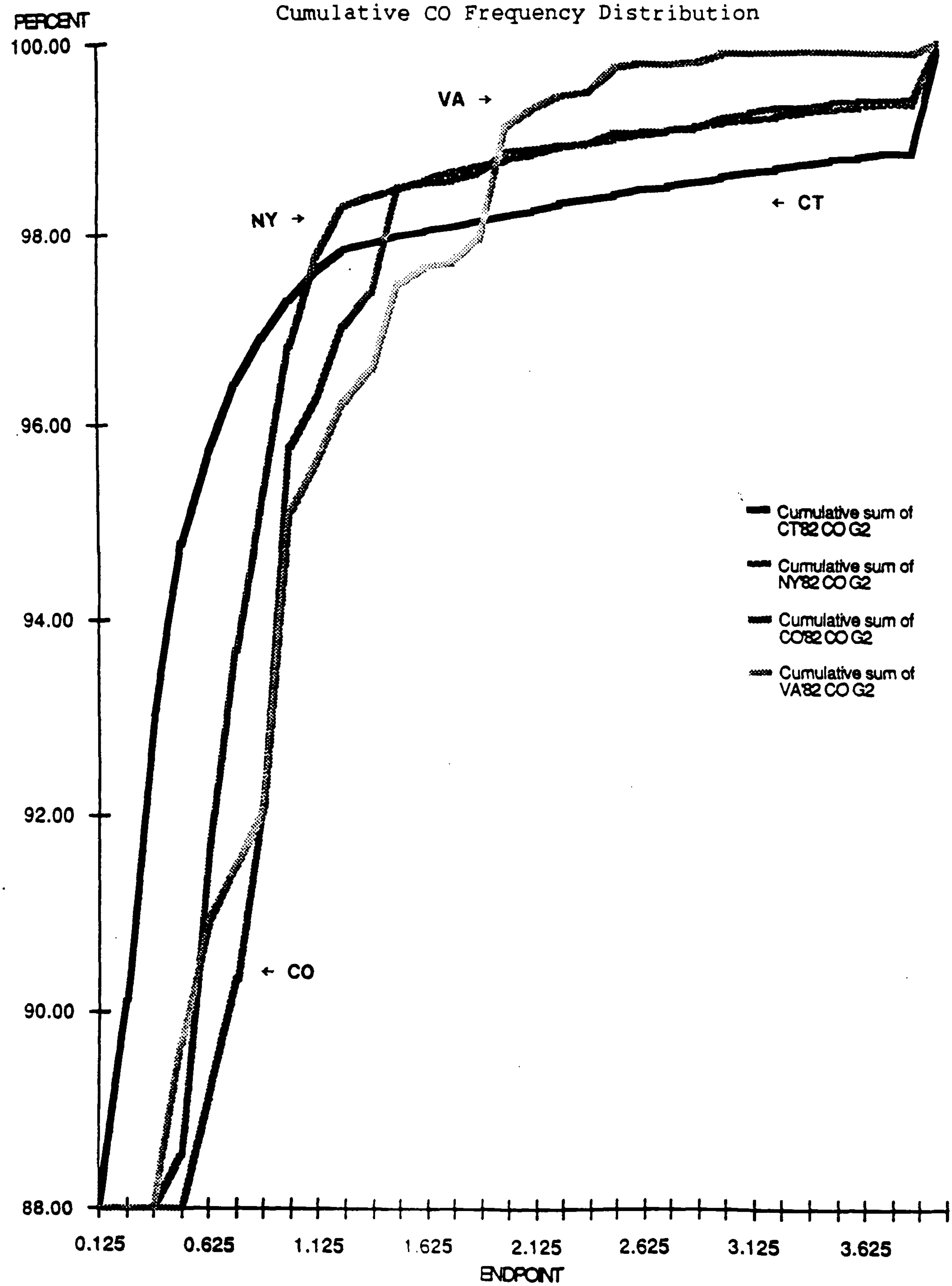
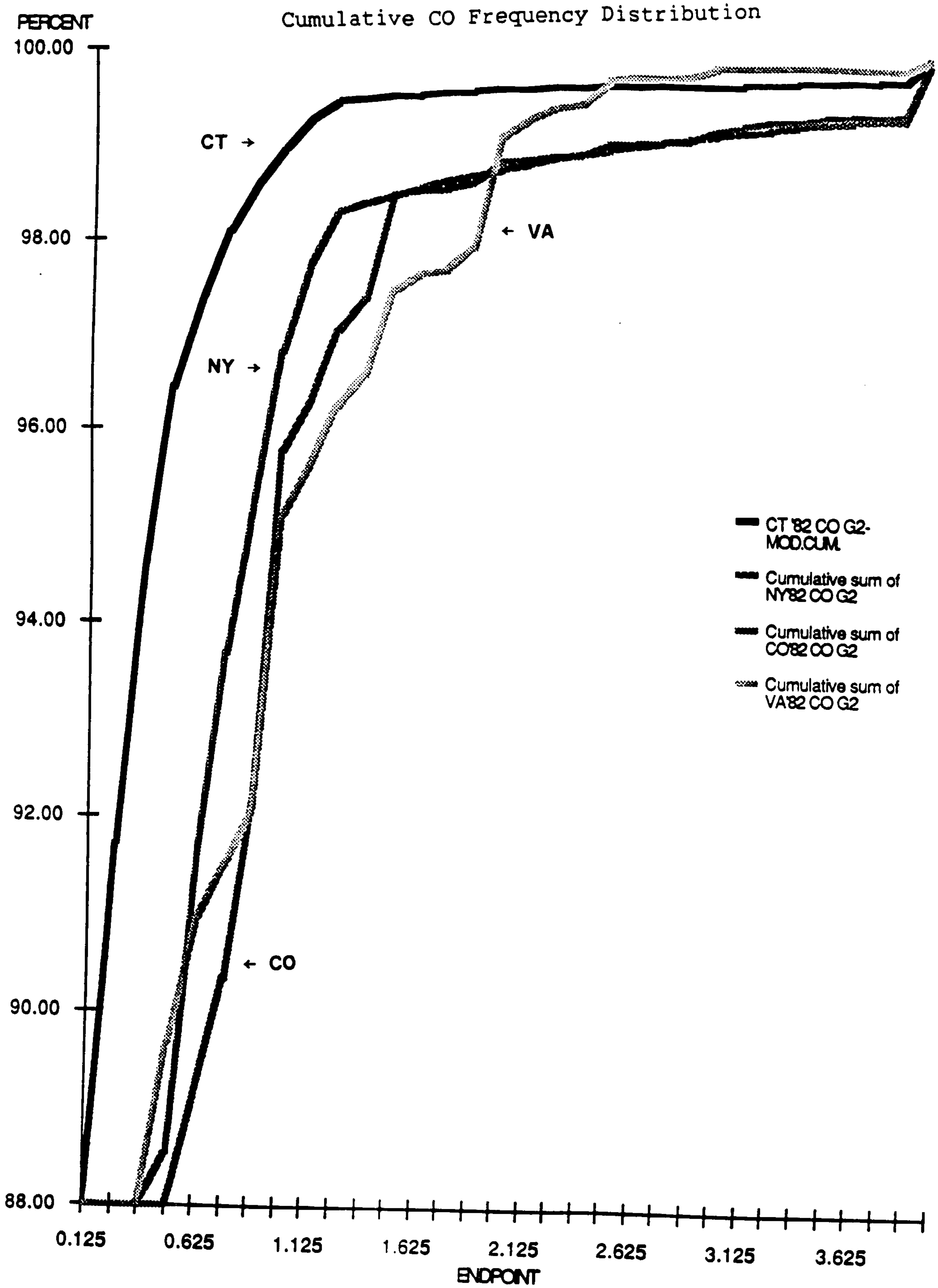


Figure Nine

CONNECTICUT RETEST SCORES VERSUS INITIAL TEST SCORES IN DECENTRALIZED PROGRAMS MODEL YEAR 1982

Cumulative CO Frequency Distribution



CO standard as they have sealed adjustments for the carburetor and, in many cases, sophisticated electronic controls that are either operational or malfunctioning, with no "in-between" states. Finally, it must be noted that Connecticut's post-repair distribution for any of three model years considered do not show a large group of cars below the applicable cutpoint. This indicated that even if some form of pre-repair occurs in decentralized programs, the repairs do not appear to be complete or related to the defects in the emission control system.

Two important conclusions drawn from this data analysis are that:

1) Emission distributions from decentralized programs do not resemble those from centralized programs. In decentralized programs the distributions exhibit a distinctive discontinuity at program cutpoints.

2) Pre-repair is not a satisfactory explanation for the shape of the emission distribution curves, particularly for newer model years, in decentralized programs.

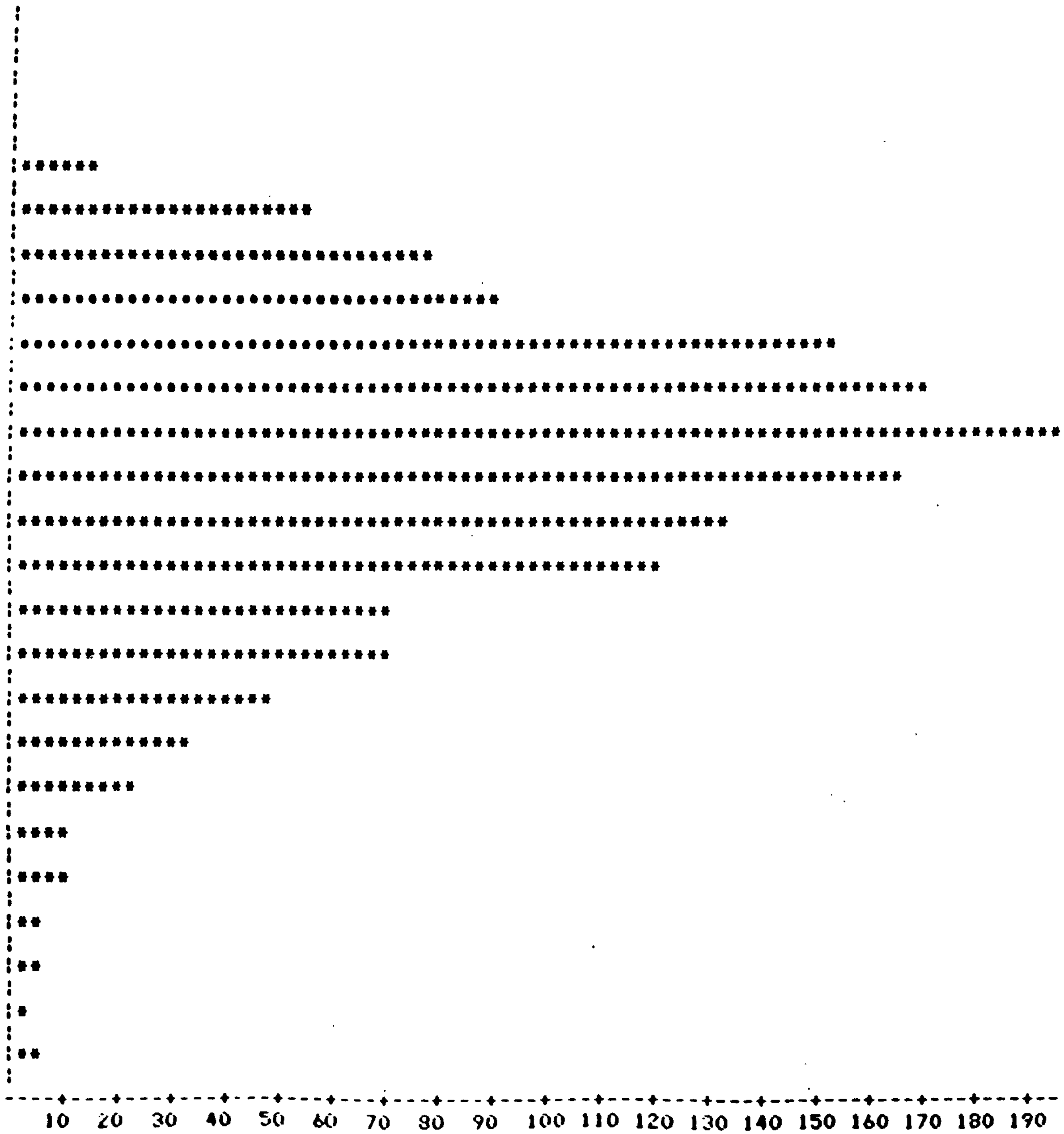
The contractor devised two indices to attempt to further distinguish pre-repair and improper inspection. The abnormal frequency (ABF) index is the percent of cars that have emission scores 0.7 to 1.0 times the cutpoint. This index is based on the finding that an abnormally large percentage of the fleet in decentralized programs is reported to have emission scores in a narrow emissions range just below the cutpoints. The index for Washington and Connecticut was approximately 0.15 for 1975 to 1980 vehicles using the 3.0% CO/300 ppm HC cutpoints, dropping to 0.09 using 6.0% CO/600 ppm HC cutpoints. Analysis of New York, Colorado and Virginia data showed that 40-60% of stations analyzed have high ABF indices.

The second index devised is the repeat index. In reviewing I/M station records, EPA auditors have observed reported emission scores repeated again and again at a given station. The contractor devised a repeat index by counting the number of times each HC and CO value is repeated and the three highest counts for HC and CO (six in all) are added. This number is then normalized by the sample size to derive the Repeat Index. By way of example, if all emission readings are reported as one of any three values for HC and CO (an extreme case), the calculated Repeat Index would equal 2. The contractor suggests a level of 0.5 to 0.6 as criteria for identifying potentially dishonest stations. Analysis of the Colorado and Virginia data showed indices ranging as high as 1.2. Figures Ten and Eleven illustrate the results for all stations examined.

Figure Ten

REPEAT INDEX - COLORADO

Frequency Bar Chart

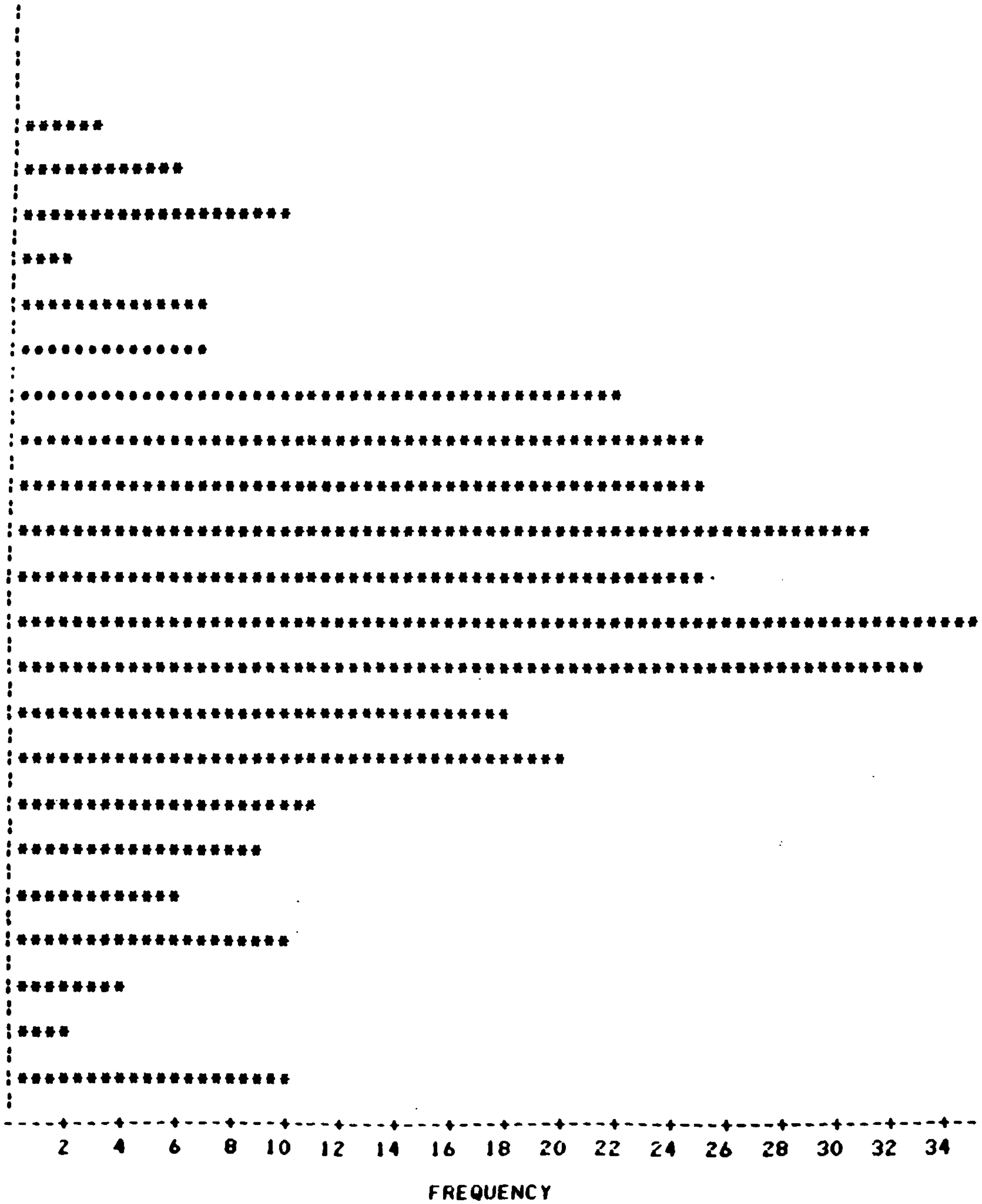


MIDPOINT RPT_IDX	FREQ	CUM. FREQ	PERCENT	CUM. PERCENT
0.025	0	0	0.00	0.00
0.075	0	0	0.00	0.00
0.125	1	1	0.07	0.07
0.175	16	17	1.10	1.17
0.225	56	73	3.85	5.02
0.275	78	151	5.37	10.39
0.325	91	242	6.26	16.66
0.375	152	394	10.46	27.12
0.425	170	564	11.70	38.82
0.475	195	759	13.42	52.24
0.525	164	923	11.29	63.52
0.575	132	1055	9.08	72.61
0.625	119	1174	8.19	80.80
0.675	70	1244	4.82	85.62
0.725	70	1314	4.82	90.43
0.775	47	1361	3.23	93.67
0.825	32	1393	2.20	95.87
0.875	23	1416	1.58	97.45
0.925	10	1426	0.69	98.14
0.975	9	1435	0.62	98.76
1.025	6	1441	0.41	99.17
1.075	5	1446	0.34	99.52
1.125	3	1449	0.21	99.72
1.175	4	1453	0.28	100.00

Figure Eleven

REPEAT INDEX - VIRGINIA

Frequency Bar Chart



MIDPOINT RPT_IDX	FREQ	CUM. FREQ	PERCENT	CUM. PERCENT
0.025	0	0	0.00	0.00
0.075	0	0	0.00	0.00
0.125	3	3	0.93	0.93
0.175	6	9	1.87	2.80
0.225	10	19	3.12	5.92
0.275	2	21	0.62	6.54
0.325	7	28	2.18	8.72
0.375	7	35	2.18	10.90
0.425	22	57	6.85	17.76
0.475	25	82	7.79	25.55
0.525	25	107	7.79	33.33
0.575	31	138	9.66	42.99
0.625	25	163	7.79	50.78
0.675	35	198	10.90	61.68
0.725	33	231	10.28	71.96
0.775	18	249	5.61	77.57
0.825	20	269	6.23	83.80
0.875	11	280	3.43	87.23
0.925	9	289	2.80	90.03
0.975	6	295	1.87	91.90
1.025	10	305	3.12	95.02
1.075	4	309	1.25	96.26
1.125	2	311	0.62	96.88
1.175	10	321	3.12	100.00

In addition to the extensive data analysis just reviewed, practical experience in the field leaves a strong impression that improper inspection is occurring. The procedures used by EPA to audit I/M programs include a review of records and usually a demonstration by station inspectors of test procedures. Auditors have often witnessed a lack of knowledge, on the part of inspectors, of how to correctly conduct the emission test or calibrate the instrument. Unfortunately, most decentralized programs have limited resources and place insufficient emphasis on identifying stations that improperly inspect vehicles. Nevertheless, there are examples in every manual I/M program where undercover work is conducted of inspectors who got caught issuing a certificate of compliance without conducting the emission test.

There are plenty of reasons why it is advantageous to improperly inspect rather than conduct inspections properly. Mechanic/ inspectors are in business to make money, not necessarily to reduce air pollution. Customers want to meet the program requirements with as little time and expenditure as possible. If a regular customer goes in for an emission test, there is strong incentive for the mechanic to report a passing result regardless of actual performance of the vehicle. If a tune-up had been done recently on the vehicle, reporting a failure would call into question the mechanic's ability and trustworthiness. There is also the situation where the station does not have any competent mechanics. These stations are in the business to collect the test fee and sell gasoline. They fail a few vehicles here and there to avoid attracting suspicion. Undercover work has verified that such stations exist; customers come to realize that they will always pass at this station, word spreads and business is up.

On the other hand, there is the competent mechanic who runs a good station and is faced with increasingly complex automotive systems for which training is difficult and time consuming to obtain. This mechanic may try very hard to fix the vehicle but in failing to do so may report a pass anyway. To say the least, customers would not be very happy with the mechanic charging them the test fee, repair costs and not getting the vehicle to pass! Good mechanics are in high demand these days, and when the workload is heavy, they may issue a certificate without doing the test, simply for lack of time. In that failure rates among new technology vehicles are very low, some mechanic/inspectors may come to believe that they never fail so they don't bother testing them after a while. These motives and incentives are very real problems. While there is little in the way of hard evidence to support these ideas, they are based on observations and discussions with mechanics and program officials during audits.

CONCLUSIONS

This paper has reviewed a variety of possible causes for the low reported failure rates found in manual, decentralized I/M programs. It was shown that quality control lapses could, at most, only make a small difference in the failure rate outcome in an I/M program. Another potential explanation, better local maintenance, does not seem to be the case when emission scores of non-I/M vehicles from different areas are analyzed using uniform cutpoints. Cars in one area seem to fail at rates similar to that of other areas. Differences in emission standards and model year coverage among I/M programs do not seem to explain the low reported failure rates. By estimating the expected failure rates using a common data base, it was shown that reported failure rates in manual, decentralized I/M programs differ radically from expected failure rates. This was not the case for other program types. Anticipatory maintenance and pre-inspection repair, while likely to exist to some limited extent, do not seem to be prevalent phenomena and therefore cannot contribute significantly to low failure rates.

Improper inspection seems to be the primary cause of low failure rates in decentralized, manual I/M programs. The random, roadside idle surveys conducted by EPA show that I/M vehicles are not "clean" after inspection. The surveys show that failure rates of these vehicles are much higher than the reported failure rates. Analysis of I/M data shows that the reported emission scores from manual decentralized I/M programs are very unusual. Instead of having a smooth distribution of emission scores, the distributions from manual programs show higher scores close to the cutpoints. A distinct kink in the distribution occurs right at the cutpoint, where the scores drop off dramatically. This indicates that inspectors are not entering real test scores; the scores they invent are more often right below the cutpoint. The abnormal frequency index attempts to quantify this phenomenon and shows that in fact, there are unusually large numbers of vehicles just below the cutpoints in these programs. Patterns of emission scores have been observed during audits of inspection stations in which scores are repeated again and again. The repeat index was devised to quantify this in relation to the cutpoints and it lends support to the observed pattern of emission scores. Finally, practical experience during the audit process and the results of undercover inspection work show that improper inspection is a problem.

In conclusion, low reported failure rates in manual, decentralized I/M programs seem to be primarily a function of improper inspection rather than simple variations in vehicle coverage, test procedures or other program characteristics. In order to correct this problem, programs will have to either eliminate the manual aspect of the program or put sufficient resources into oversight and undercover efforts.

REFERENCES

1 National Air Audit System Reports. 1984, 1985, 1986. U.S. EPA, Office of Air Quality Planning and Standards.

2 Motor Vehicle Tampering Surveys - 1983, 1984, 1985, 1986. U.S. EPA, Office of Air and Radiation.

3 Vehicle Exhaust Testing 1984 Annual Report, Air Pollution Control District of Jefferson County, January 25, 1985.

4 "Summary of Programs Simulating a Mandatory Maintenance Program." Memorandum from R. Bruce Michael, I/M Staff to Tom Cackette, Chief, I/M Staff. U.S. EPA, February 27, 1980.

5 Effectiveness of Idle Adjustment on Light Duty Trucks at Commercial Repair Facilities. John C. Shelton, Test and Evaluation Branch, U.S. EPA, June 1980, EPA-AA-TEB-81-17.

6 Effectiveness of Idle Adjustment on Passenger Cars at Commercial Repair Facilities. John C. Shelton, Test and Evaluation Branch, U.S. EPA, October 1980, EPA-AA-TEB-81-5.

7 Auto Exhaust Testing Inspection Statistics" First and Second Quarters 1986. Michigan Department of State.

8 Development of Data Analysis Systems for Decentralized I/M Programs, Energy and Environmental Analysis, Draft Report, September 1986.