LBNL-47754

Can More Frequent Testing Improve I/M Program Effectiveness? Analysis of the Effect of Annual Testing in the Phoenix IM240 Program

Final Report to Environmental Systems Products, Inc.

June 11, 2001

Tom Wenzel and Nancy J. Brown
Lawrence Berkeley National Laboratory
1 Cyclotron Road
Berkeley, CA 94720

This work was supported by Environmental Systems Products, Inc. Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Abstract

In this report we examine the potential for additional emission reductions from adopting annual rather than biennial testing in vehicle emission inspection and maintenance (I/M) programs. We tracked a fleet of cars reporting for testing in two biennial cycles in the Phoenix IM240 program. Using an assumption of linear deterioration following a passing I/M test, we found that annual testing in the Phoenix program in 1996 would have resulted in an additional 45%, 48%, and 27% reduction in HC, CO, and NOx emissions, respectively, over what occurred in the biennial program. More stringent cut points under a biennial program would have resulted in greater emission reductions, but with higher fail rates. Tighter cut points can also be applied to the annual program. In fact, an annual program might ease the progression to tighter cut points by resulting in relatively lower failure rates per test than biennial programs.

A small number of cars were identified that were given three I/M cycles, each roughly one year apart, in Phoenix. This group of vehicles had higher overall emissions and may not be representative of the whole fleet. However, these cars appear to have the same HC and CO emissions, but slightly lower NOx emissions, after two years as cars tested biennially. This result suggests that our assumption that emissions after annual testing would deteriorate at the same rate as observed in biennial testing may be optimistic. The findings from our analyses suggest that test-to-test emissions variability is a limitation of existing I/M programs, and is preventing them from properly identifying some vehicles with broken or malfunctioning emissions controls and ensuring that they are repaired. In theory, more frequent testing of suspected high emitters could help address this problem. One strategy that may work is to use remote sensing to identify suspected high emitters, and require that they come in for off-cycle testing, as frequently as necessary.

Arizona remote sensing data provide additional evidence that a significant benefit results from annual testing of older cars. Remote sensing measurements made in Phoenix of vehicles tested biennially under the Phoenix enhanced IM240 program and of vehicles tested annually under the Tucson two-speed idle program show interesting results. While newer vehicles from Phoenix have lower emissions than those from Tucson, as one might expect from the enhanced program, the older vehicles from Phoenix have higher emissions than those from Tucson. Phoenix vehicles over seven years old had emissions 37% to 47% higher than Tucson vehicles of the same age.

We conclude that more attention should be paid to promptly identifying and properly repairing the high emitters found in the older portion of the vehicle fleet.

Introduction

An evaluation of the California Enhanced Smog Check program by the California Air Resources Board (CARB) found that the program is achieving only 60% of the emission reductions the CARB claimed for the program in its State Implementation Plan, or SIP (CARB, 2000). CARB attributes this shortfall to changes made to the program by the state legislature: i.e., less stringent cut points, exemptions for 1966 to 1973 vehicles, fewer vehicles directed to Test Only stations, smaller geographical areas included in the enhanced area, etc.[1] To make up for this shortfall, CARB is proposing several changes to the program, including tightening NOx cut points to midway between their current level and the levels included in the SIP benefit estimate.[2]

However, Lawrence Berkeley National Laboratory's recent evaluation of the Enhanced Smog Check program for the Inspection and Maintenance Review Committee (IMRC) found that current cut points already capture the worst high emitters (Wenzel et al., 2000). In addition, over 85% of the total program benefit comes from 1989 and older vehicles. And significant numbers of vehicles failed their next Smog Check only one month after passing a previous Smog Check (20% of vehicles that failed then passed, and 6% of vehicles that initially passed, in the previous cycle). Previous analysis of the Phoenix IM240 program indicates that 40% of vehicles that failed their initial test, and passed a retest, in 1995 failed their next initial test two years later. These results suggest that more frequent testing of vehicles, particularly older vehicles, may be more effective in improving I/M programs than simply ratcheting down cut points on all vehicles. In theory, more frequent testing of vehicles would identify and repair high emitters sooner, resulting in substantial emission reductions from these vehicles before their next scheduled biennial test.

In this report we use data from the Phoenix IM240 program to simulate the potential emission reductions from requiring annual I/M testing, as well as from using tighter cut points. We use data from the Phoenix program because it is considered the benchmark Enhanced I/M program by EPA. We compare both of the simulations to the actual emission reductions observed in the 1995 and 1997 cycles of the Phoenix IM240 program. We also examine other data from the Phoenix IM240 and Tucson annual two-speed idle programs to determine how effective annual testing might be in reducing emissions over the medium term (one year after repair).

Observed Emission Reductions in 1995 and 1997

We first calculate the emissions and emission reductions from two I/M cycles, using the fleet of passenger cars that reported for testing in both 1995 and 1997. Because program cut points vary by vehicle type, we limit our analysis to passenger cars, which represent about 60% of the vehicles tested. 40% of the cars tested in 1995 did not report for testing in 1997; the portion of the fleet that did not return in 1997 was older, had a higher failure rate, and had higher initial emissions than the fleet tested in both years. It appears that 20% of these cars were tested "off-cycle" in 1995, either by mistake or because of a voluntary change-of-ownership test, and were

1. Another explanation for the shortfall not considered by CARB is that the EMFAC model over-predicted the emissions benefit of the full program originally proposed by CARB and BAR.
2. The original NOx cut points included in the SIP benefit estimate are unrealistic, in that the cut points for 1980 to 1985 vehicles are 20% lower than those for 1986 and newer vehicles.

not required to be tested in 1997. However, the remainder, or one-third of all cars tested in 1995, were not tested in 1997 because they either were removed from the Enhanced I/M area or managed to get registered without reporting for I/M testing (Wenzel, in press).

In order to calculate emission reductions from the actual program, and to compare them with the two simulations, we need to adjust emissions of cars that pass or fail I/M testing in less than 240 seconds to their estimated emissions if they were given a full IM240 test. We use a method developed by Resources for the Future (RFF) to make this adjustment (Ando et al., 1998). When this adjustment is made, a small of percentage of vehicles that passed inspection are projected to have full-test emissions higher than the cut point.[3]

We divide the fleet of cars reporting for I/M testing into four subfleets, based on the overall result of their I/M test in 1995, including visual and functional tests as well as tailpipe emissions:

- initial-pass cars, which pass their initial I/M inspection in 1995 (IP);
- fail-pass vehicles, which fail their initial 1995 test but pass a subsequent retest (FP); this group includes cars that pass a retest without any repairs being made;
- no-final-pass cars, which fail their initial 1995 and subsequent retest, and do not have a passing test in the timeframe studied (FF);
- no-second-test cars, which fail their initial 1995 test and do not receive a subsequent retest in the timeframe studied (F_).

Table 1 shows the number of cars in each subfleet. We frequently refer to the no-final-pass and no-second-test subfleets collectively as no-final-pass (NFP) cars.

In 1995, 10% of these cars failed their initial I/M test. Of these failures, 27% never passed a subsequent retest. This "no-final-pass rate" depends somewhat on the amount of time a car has in which to pass a retest; for instance, because we use inspection results from 1995 only, cars initially tested in December 1995 have only up to one month in which to pass a retest. Allowing cars more time in which to pass lowers the no-final-pass rate. In an earlier analysis we found that including data through March 1996 reduced the number of 1995 no-final-pass vehicles by 7%; this would reduce the no-final-pass rate in Table 1 from 27% to 25%. In addition, about 4% of all vehicles that fail their initial test receive a waiver; therefore the actual no-final-pass rate for cars tested in both 1995 and 1997 is likely 21%.

---

3 This adjustment results in 4% of the cars that passed the IM240 tailpipe test having projected full test emissions for at least one pollutant in excess of the start-up cut points that were in place at the time. These cars represent 6 to 9% of all excess emissions (depending on pollutant), as measured in initial 1995 I/M tests. Although these cars represent almost 5% of all cars passing their initial I/M test, they account for only 0.4% to 1.3% of all emissions of initial pass cars, depending on pollutant. There are two possible causes for a car that passed its I/M test having projected full test emissions in excess of the I/M cut points: 1) the algorithm Gordon-Darby uses to fast pass vehicles allows some vehicles that would have failed a full IM240 to pass in less than 240 seconds; and 2) the method RFF developed to adjust short test emissions to full test emissions results in higher emissions than would have occurred under full IM240 testing. In an earlier analysis we found that 2% of MY83-90 passenger cars that pass after only 30 seconds of testing using Gordon-Darby fast-pass criteria fail start-up IM240 NOx cut points (0.6% fail for HC and 1.2% fail for CO). In practice, if fast-fail standards were not in use in Phoenix, many of these vehicles may have passed the phase 2 cut points of the IM240 (or may have passed a second chance test as allowed in other states such as Colorado and Wisconsin).

Table 1. Number of cars tested in both 1995 and 1997, by I/M Result

I/M Result	Number	Percent of Total	Percent of Initial Fails
1) Initial pass	221,047	90.2%	
2) Fail-pass	17,605	7.2%	73.3%
3) No-final-pass	4,619	1.9%	19.2%
4) No-second-test	1,807	0.7%	7.5%
Subtotal 3 and 4	*6,426*	*2.6%*	*26.7%*
Total	245,078	100.0%	

Figure 1 shows the average per car emissions, in grams per mile, by 1995 I/M result (CO emissions are shown divided by 10). In 1995, fail-pass cars show a large reduction in emissions between their initial and final test, presumably because of repairs made to them. No-final-pass cars have higher initial emissions, in part because they tend to be older; their emissions are reduced somewhat due to repairs that were not sufficient to pass their retest. Between 1995 and 1997, emissions of initial pass and fail-pass cars increase, due to natural degradation of emissions controls and larger numbers of vehicles becoming high emitters over time. In 1997, 40% of the 1995 fail-pass fleet failed their initial I/M test (Wenzel, in press).

Figure 1 implies that the deterioration in emissions is linear over the two-year period, but results from "off-cycle" I/M tests in California (where either the first or second I/M cycle is triggered by a change of vehicle owner) indicate that this may not be the case. Analysis of California vehicles that initially failed and then passed a retest revealed that 20% of these "repaired" vehicles (and 6% of initial pass vehicles) failed a subsequent I/M test within the next three months. . The failure rate remained fairly constant for vehicles given an off-cycle test between three months and nine months after completing the initial I/M cycle (Wenzel et al., 2000).

Emissions of no-final-pass and no-second-test cars in Figure 1 decrease between their 1995 and 1997 tests, and the emissions of these cars are further reduced as some of them fail their 1997 test and presumably are repaired.

Figure 1. Average Emissions over Two I/M Cycles
Passenger cars tested in both 1995 and 1997, Arizona I

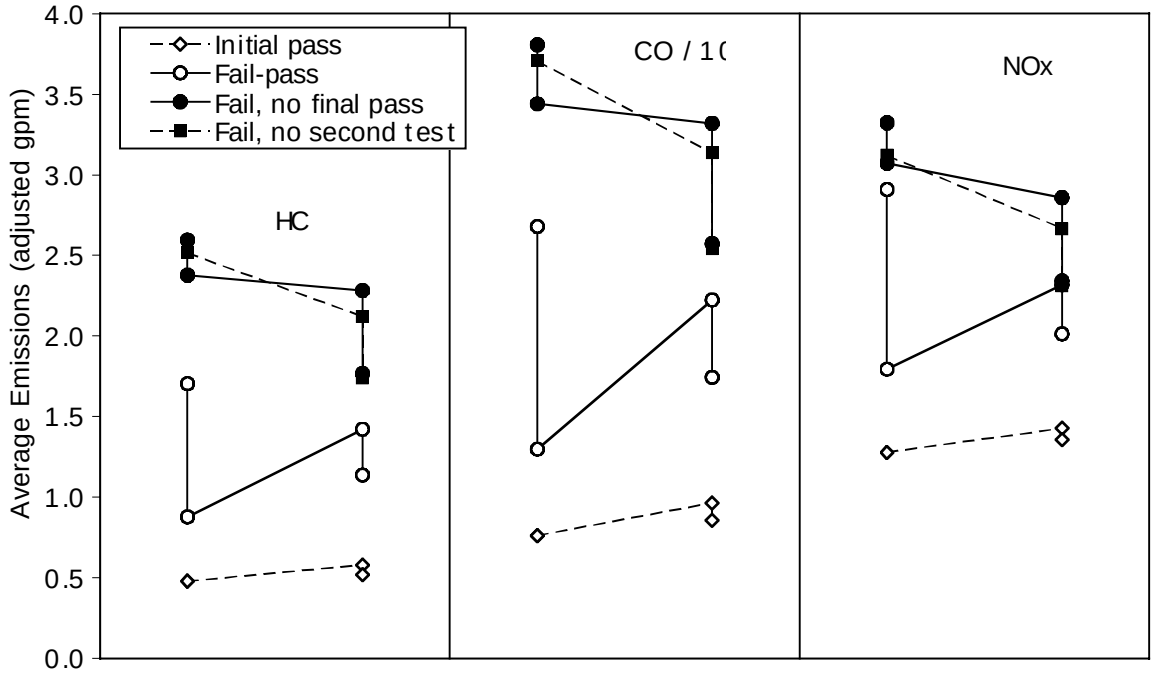


Table 2 and Figure 2 aggregate the emissions of the four subfleets of vehicles and show total emissions, in tons per day, of the fleet of passenger cars that reported for testing in both 1995 and 1997. Total emissions are determined by multiplying each car's gram per mile emissions by an assumed average number of miles driven each year to obtain annual grams of emissions; grams are then converted to tons, and then divided by 365 days to obtain ton per day estimates. The mile per year assumptions, which vary by vehicle model year and type, are taken from a recent analysis conducted for EPA's MOBILE6 (Acurex, 1997). Since only about half of the fleet eligible for I/M testing is tested in a given year, the ton per day estimates in Table 2 should be multiplied by two to estimate the total inventory of the entire I/M fleet.

The inventory tons estimated in Table 2 are a subset of the tons forecast by emission inventory models such as EPA's MOBILE and CARB's EMFAC. We have only used tailpipe emissions measured by the driving cycle in the IM240 test. The driving cycles assumed by the inventory models result in higher tonnages of emissions. The inventories forecast by models also include the effects of cold starts, air-conditioning loads, and evaporative emissions. Finally, the estimates in Table 2 are based on all cars that returned for testing in 1997, rather than all vehicles (including light duty trucks) tested in 1995.

Table 2. Fleet Emissions Weighted by Annual VMT (tons per day), by I/M Test and Year, All Cars Tested in Both 1995 and 1997

Test	HC (tpd)	CO (tpd)	NOx (tpd)
1995 initial I/M test	4.84	76.69	11.71
1995 final I/M test	4.39	69.05	11.10
1997 initial I/M test	5.31	87.15	12.43
1997 final I/M test	4.66	76.04	11.70

Figure 2. Weighted Fleet Emissions over Two I/M Cycles
Passenger cars tested in both 1995 and 1997, Arizona IM

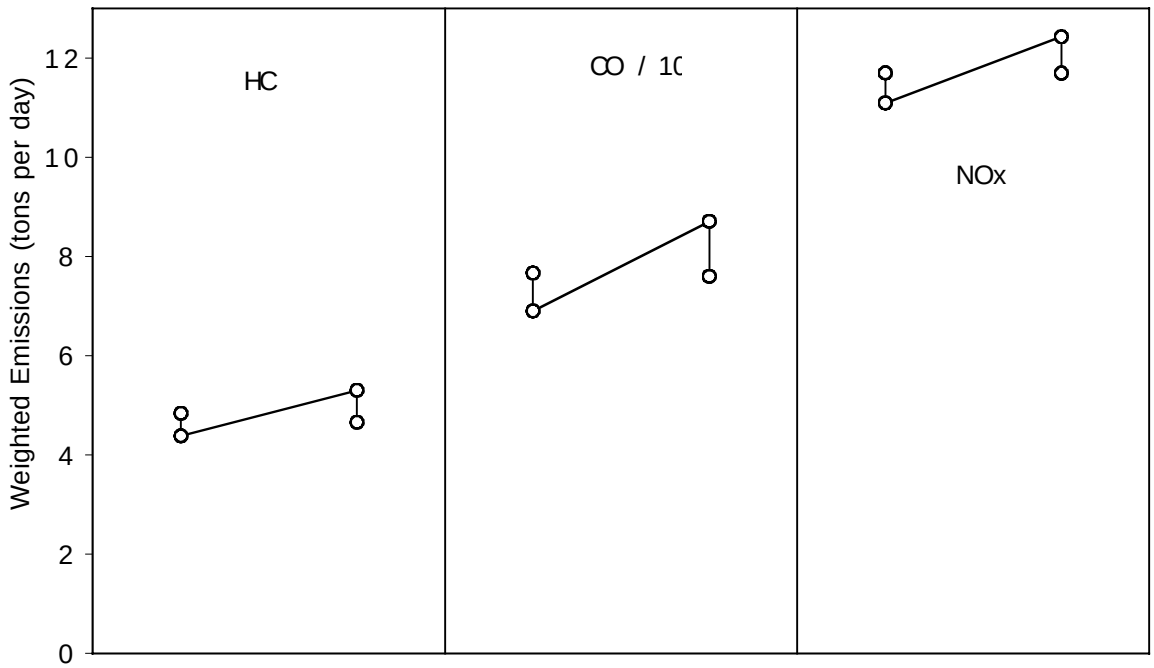


Table 3 indicates that the increase in emissions from vehicle deterioration between 1995 and 1997 is about two times the reduction from either of the 1995 and 1997 I/M cycles. After the 1997 I/M cycle, HC and CO emissions are 4% and 1%, respectively, lower than emissions prior to the 1995 cycle, while NOx emissions are the same. Because Tables 2 and 3, and Figure 2, are based only on cars reporting for testing in both years, they do not account for the effect of fleet turnover, the migration of cars into and out of the Phoenix area, or the inclusion of late model year cars in the I/M eligible fleet over time. In addition, they do not account for any emission reductions due to no-final-pass cars being removed from the Phoenix area, since all 1995 no-final-pass cars in this study returned for testing in 1997. Finally, they do not account for any emission reduction that occurred as a result of maintenance and repairs done immediately prior to initial testing in 1995.

Table 3. Percent Change in Weighted Fleet Emissions, All Cars Tested in Both 1995 and 1997

	HC	CO	NOx
Effect of 1995 I/M program (1995 final / 1995 initial)	-9%	-10%	-5%
Effect of 2 years of deterioration (1997 initial / 1995 final)	21%	26%	12%
Effect of 1997 I/M program (1997 final / 1997 initial)	-12%	-13%	-6%
Cumulative effect of two I/M cycles (1997 final / 1995 initial)	-4%	-1%	0%

Figures 3 through 5 show the distribution of total emissions, in tons per day, from passenger cars by model year, as measured on the IM240 in 1995. Starting from the top of each column, the first section shows the emissions from 1995 no-final-pass cars. The second section shows the emission reductions from repairing fail-pass cars to start-up cut points, as observed in the 1995 program. The third section shows the potential emission reductions from repairing cars to EPA's final cut points. These potential reductions are estimated assuming that all cars exceeding the final cut points are repaired to 95% of the final cut point.[4] The remaining emissions are shown in white. The figures indicate that most of the emission reduction potential, whether from repairing cars to final EPA cut points or removing 1995 no-final-pass cars from service, comes from older cars. Enforcement of EPA final cut points, coupled with removal of 1995 no-final-pass cars from service, could reduce HC 35%, CO 27%, and NOx 17% from their initial 1995 levels.

Figure 6 shows the cumulative distribution of 1995 observed emission reductions, and cars, by model year. 1984 and older cars account for 10% of the fleet and 40% of the emission reductions, while 1989 and older cars account for half of the fleet and 90% of the emission reductions. Although light duty trucks have the same age distribution as cars, more of the total emission reductions, especially NOx, from trucks comes from newer trucks. For example, 1985 and older trucks account for only 25% of the HC and CO emission reductions, and only 11% of the NOx reductions, while 1989 and older trucks account for only 80% to 85% of the HC and CO emission reductions and only 60% of the NOx emission reductions.

We expect that the profile of excess and repairable emissions by vehicle age will change in the future as vehicles equipped with on-board diagnostic (OBD) II systems (i.e. 1996 and newer) begin to dominate the on-road fleet, and are included in the fleet of vehicles subject to I/M testing.

4. As discussed below, applying EPA's final cut points to the Phoenix IM240 fleet would result in very high failure rates: 40% overall, and as high as 90% for the oldest cars.

Figure 3. Total HC Reduction Potential by Model Year
Passenger cars tested in both 1995 and 1997, Arizona IM240

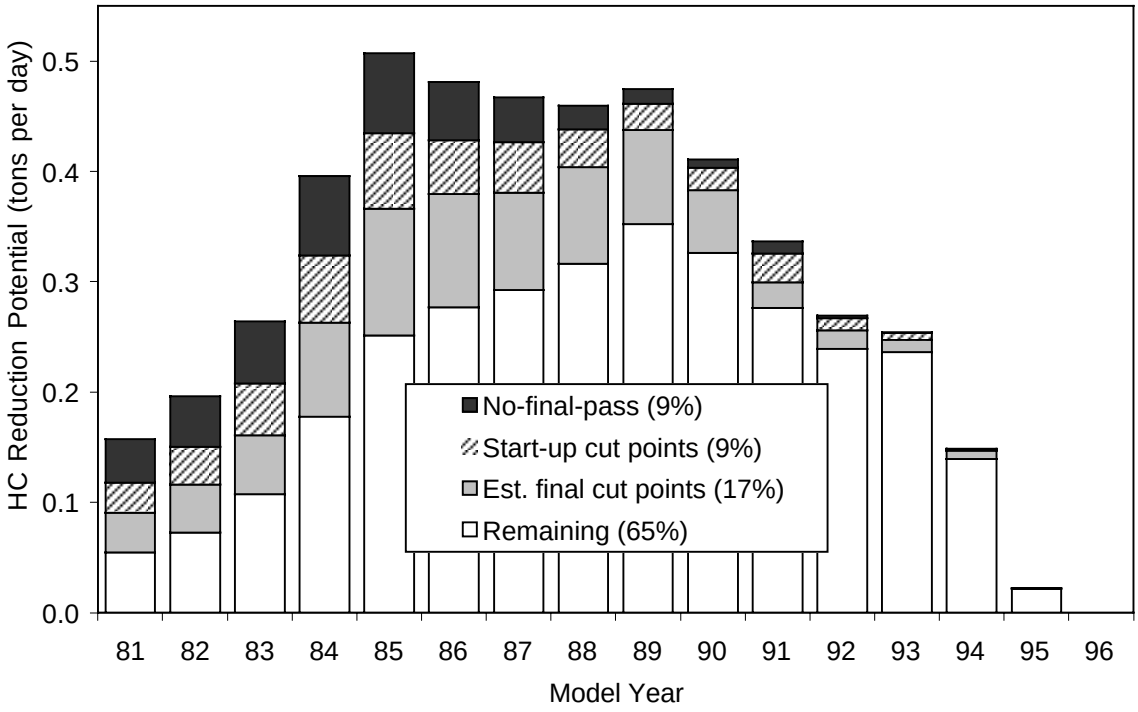


Figure 4. Total CO Reduction Potential by Model Year
Passenger cars tested in both 1995 and 1997, Arizona IM240

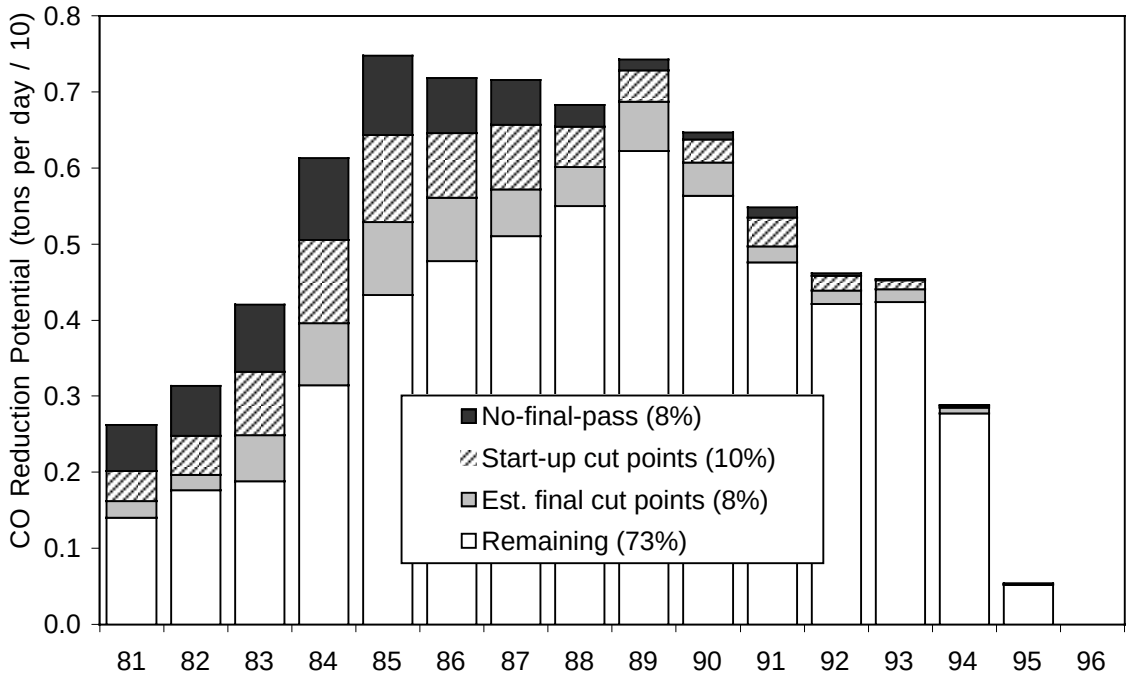


Figure 5. Total NOx Reduction Potential by Model Year
Passenger cars tested in both 1995 and 1997, Arizona IM240

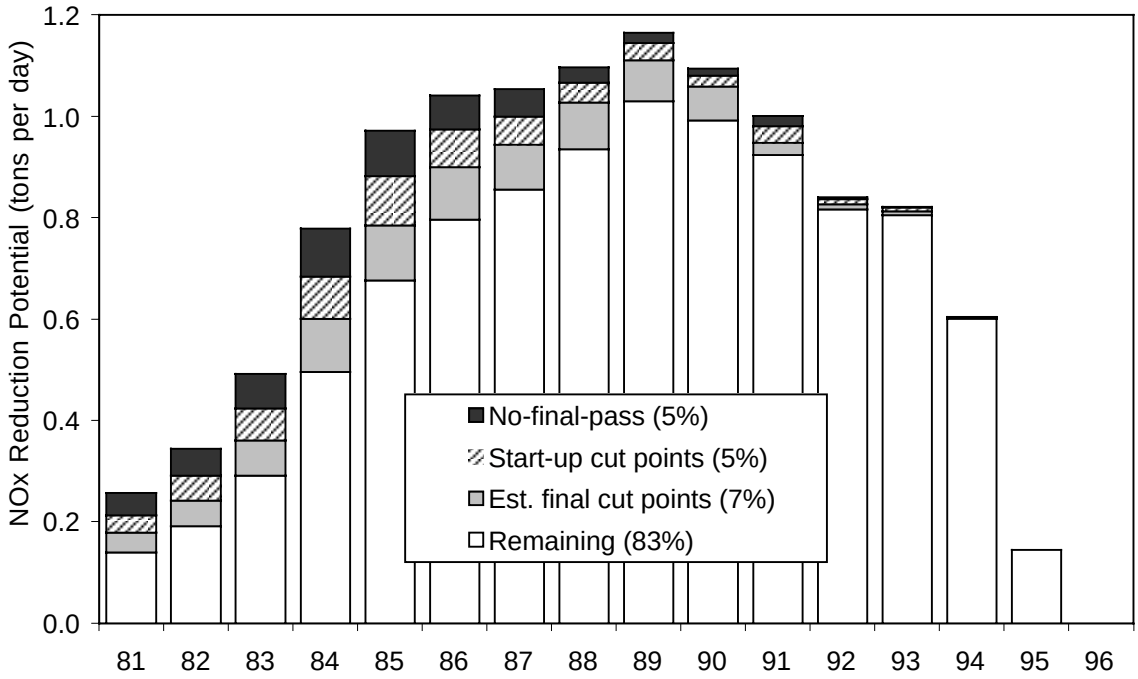
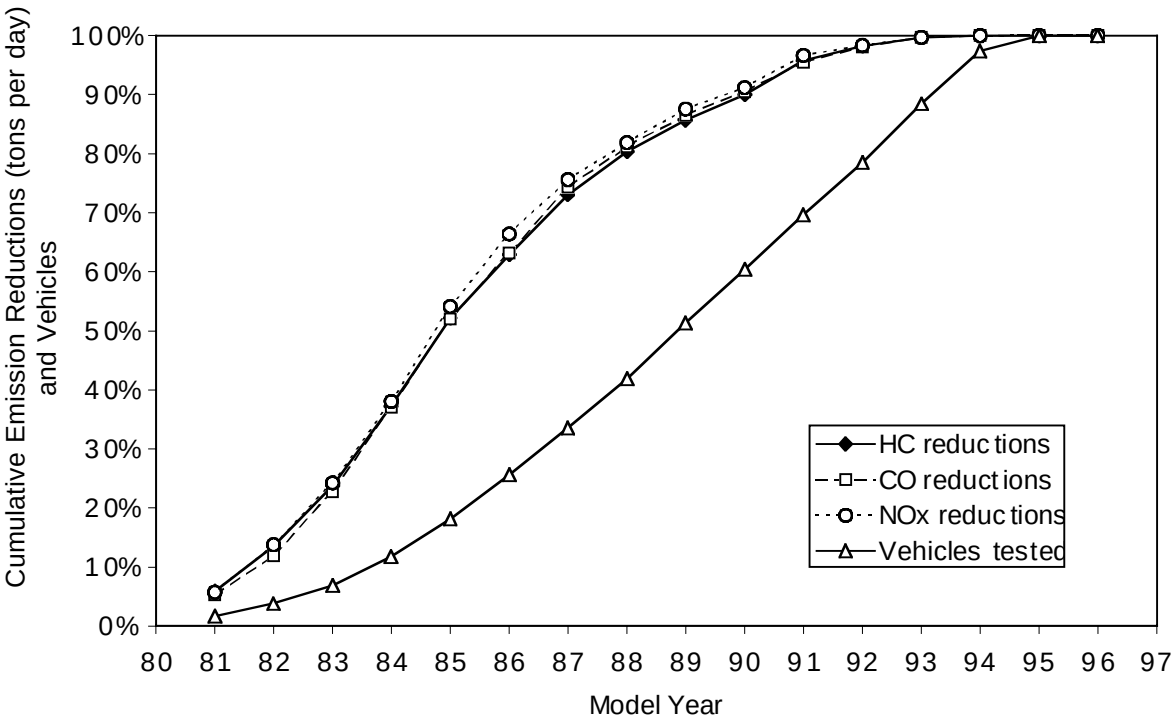


Figure 6. Cumulative Emission Reductions and Vehicles by Model Y
Passenger cars tested in both 1995 and 1997, Arizona IM24(



Figures 7 through 12 show the excess emissions and emission reductions by model year and I/M result, in both the 1995 and 1997 I/M cycles. 1995 and 1997 figures are shown on the same page for each pollutant, to allow easy comparison between I/M cycles (the 1997 figures group the cars by the result of their 1995 I/M cycle). The thick columns in each figure show the excess emissions by I/M result, while the thin columns show the actual emission reductions reported in the I/M test results from fail-pass cars. The emission reductions are greater than the excess emissions because the after-repair emissions levels are lower than the emissions cut point.

Several observations can be made about these figures. First, 1984 and 1985 cars account for the most excess emissions and emission reductions for each pollutant (CO excess emissions and reductions for 1981 and 1982 cars are smaller than those for 1983 cars in Figures 9 and 10 because of the tighter cut points for 1983 cars). Second, large portions of the excess emissions in 1995 come from no-final-pass cars, particularly for older model years. In fact, 1985 and older no-final-pass cars account for about half of the excess HC emissions from those model years (Figure 7). The large number of no-final-pass cars, and their contribution to remaining emissions, indicates that the Phoenix I/M program could be improved if more vigorously enforced. Third, for a given model year, emission reductions from fail-pass cars are about equal to the total excess emissions from all cars. This is because the emissions of fail-pass cars are being reduced to levels below the start-up cut points. However, this indicates that there is potential for more emission reductions from repairing the no-final-pass cars. Fourth, by comparing emissions from 1995 and 1997, one can see that the total excess emissions are greater for each pollutant in 1997 than in 1995. Most of the additional excess emissions in 1997 come from cars that passed their initial 1995 test. For the most part, emission reductions are greater in 1997 than 1995 (the exception is 1981 to 1985 cars, which exhibit a lower reduction in NOx emissions in 1997 than in 1995, Figures 11 and 12). Note that in some cases there are excess emissions from cars that pass their initial 1995 test, i.e. HC emissions from 1989 and 1994 cars in Figure 7. These excess emissions from passing cars are the result of projecting full IM240 equivalent results from fast-pass/fail tests.

Figure 7. Distribution of 1995 HC Excess Emissions and Reductions by Passenger cars tested in both 1995 and 1997, Arizona IM240

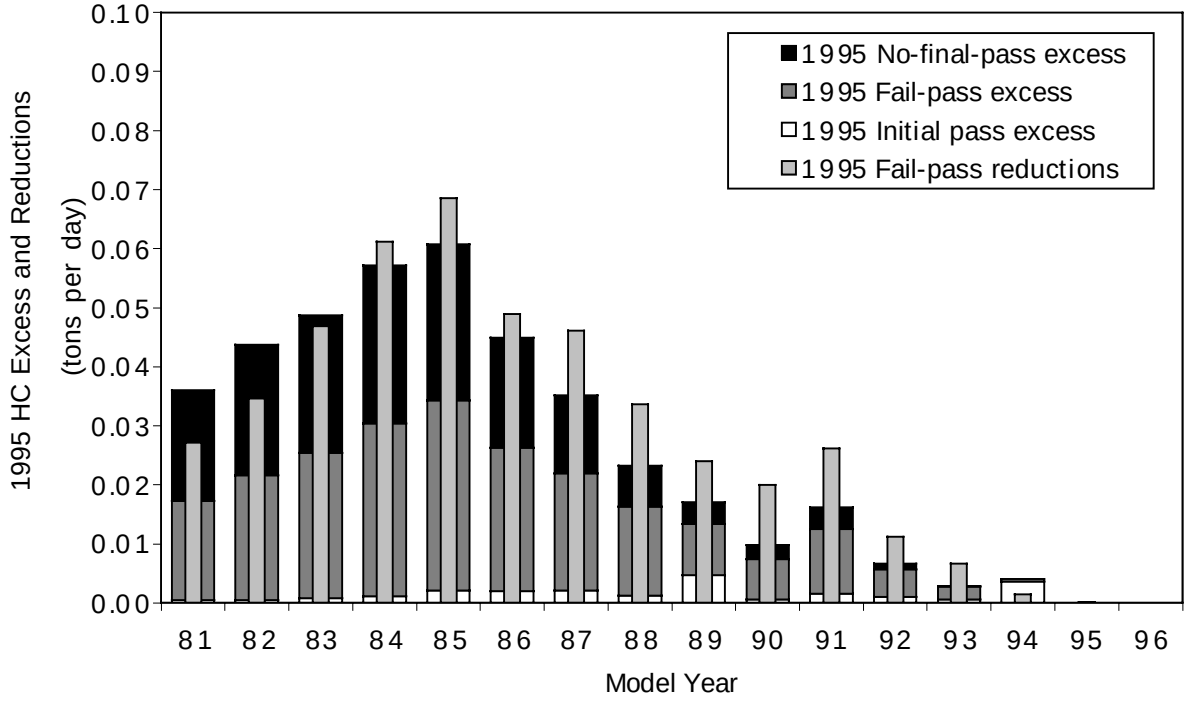


Figure 8. Distribution of 1997 HC Excess Emissions and Reductions by Passenger cars tested in both 1995 and 1997, Arizona IM240

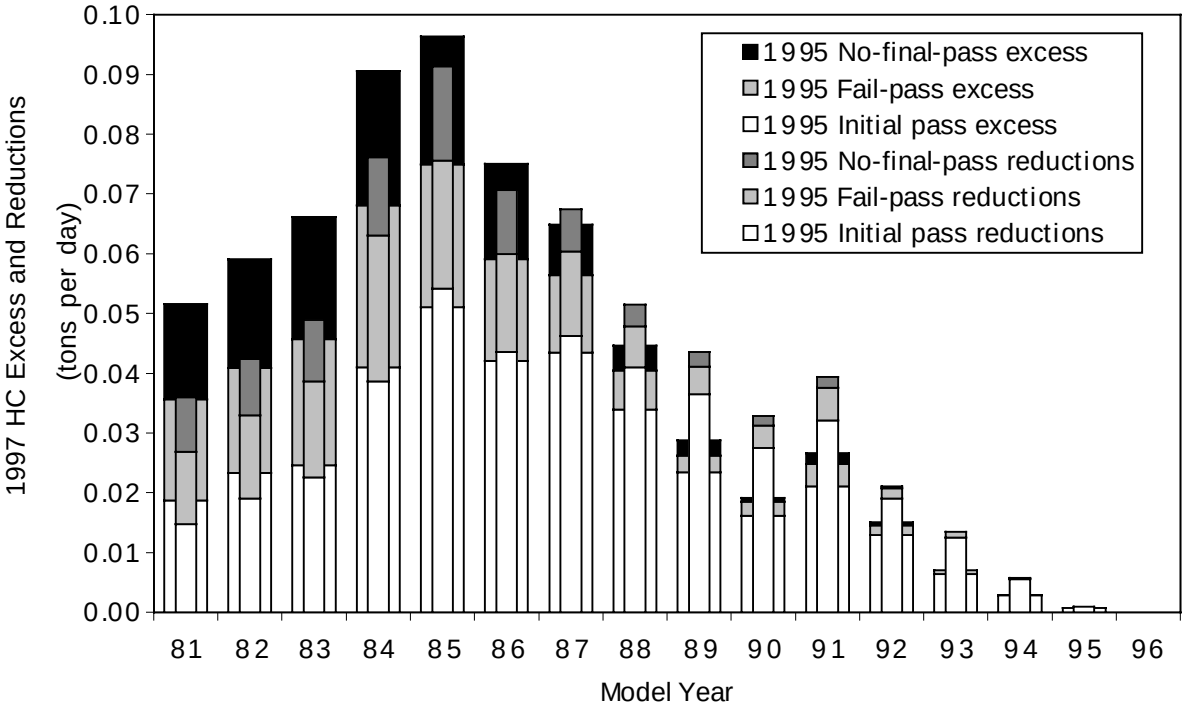


Figure 9. Distribution of 1995 CO Excess Emissions and Reductions by
Passenger cars tested in both 1995 and 1997, Arizona IM240

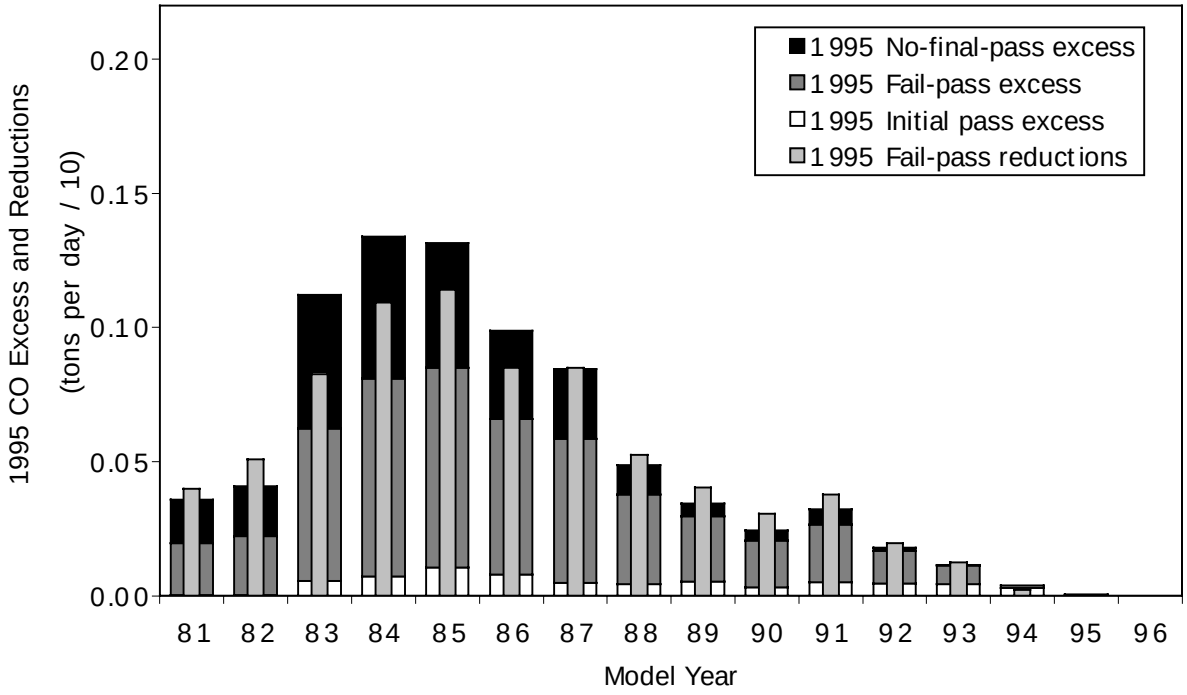


Figure 10. Distribution of 1997 CO Excess Emissions and Reductions by
Passenger cars tested in both 1995 and 1997, Arizona IM240

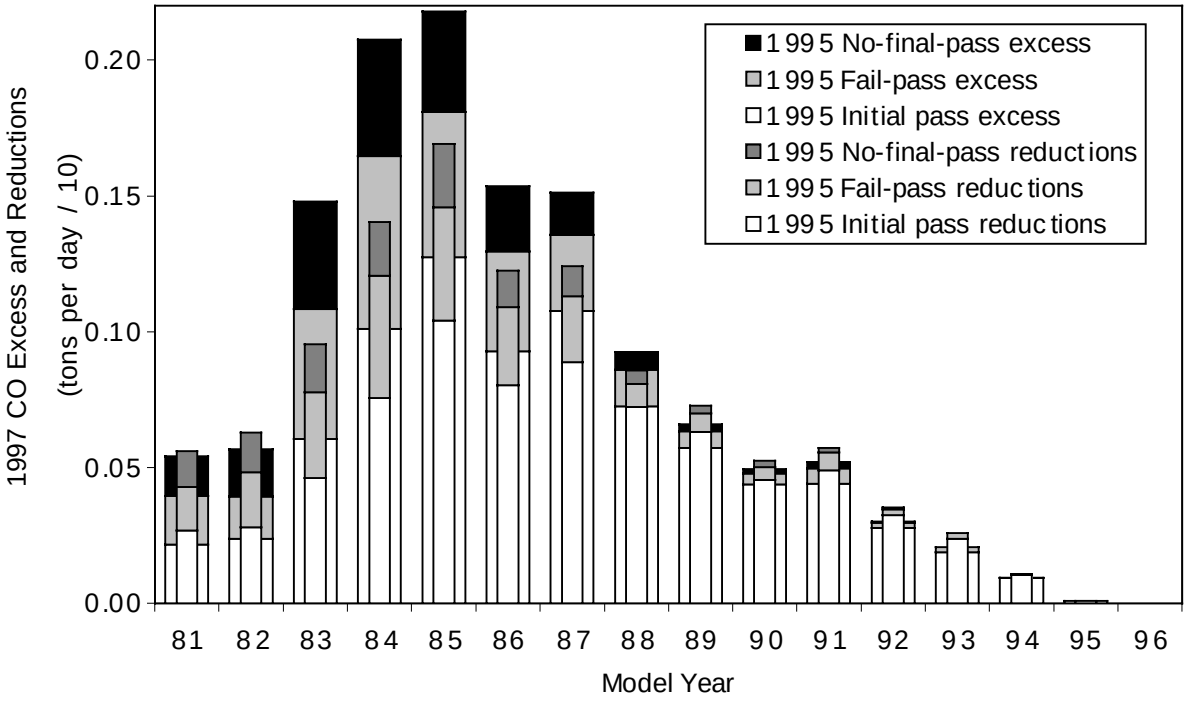


Figure 11. Distribution of 1995 NOx Excess Emissions and Reductions by
Passenger cars tested in both 1995 and 1997, Arizona IM240

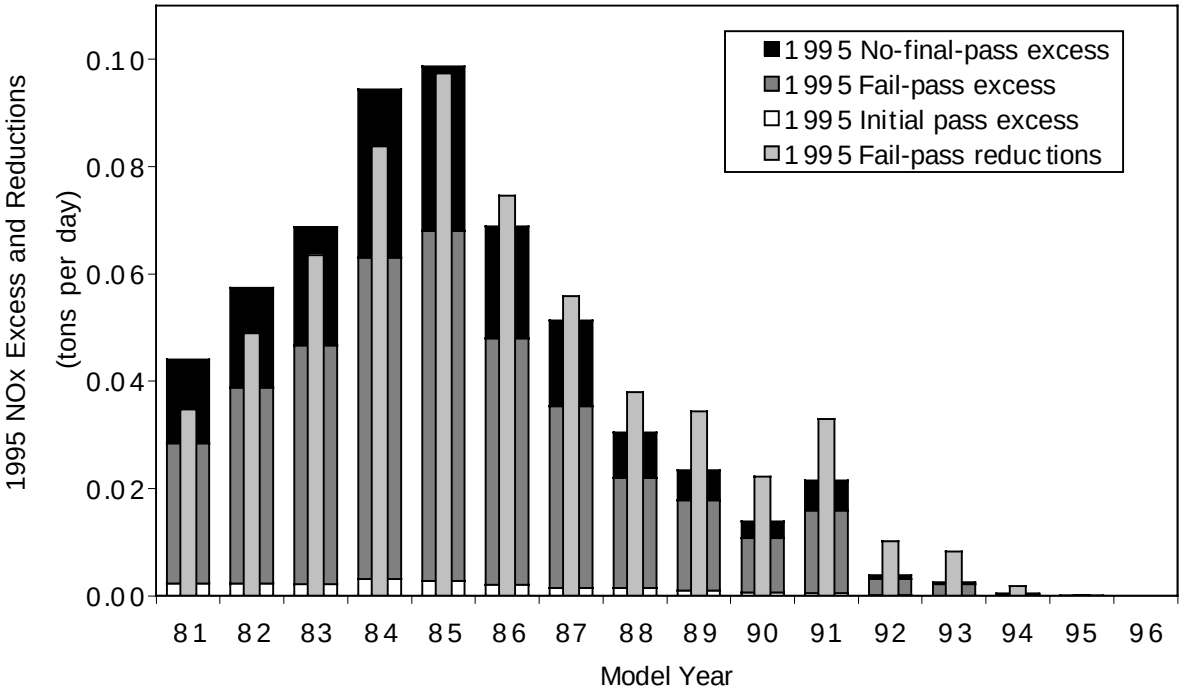


Figure 12. Distribution of 1997 NOx Excess Emissions and Reductions by
Passenger cars tested in both 1995 and 1997, Arizona IM240

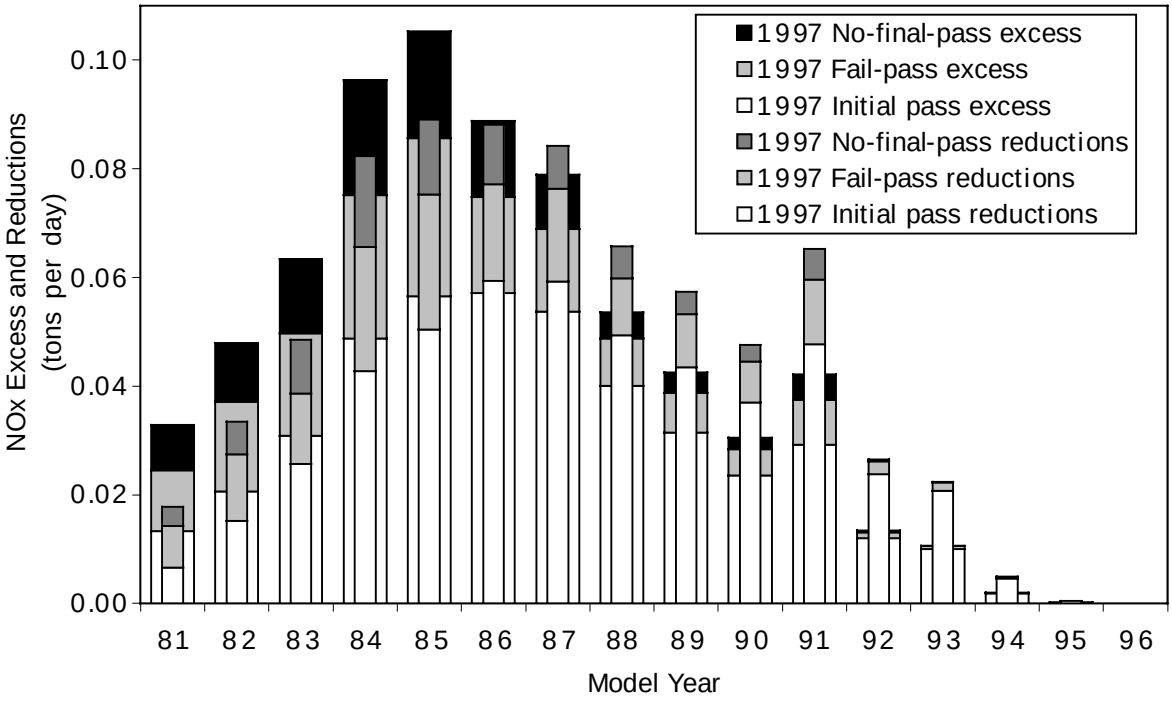


Figure 13 shows the overall failure rate by model year for the same cars tested in 1995 and 1997. The initial failure rate increases from 9.8% in 1995 to 15.5% in 1997. Two factors are driving the increase in failure rate and excess emissions from the 1995 to the 1997 cycle: many of the 1995 fail-pass cars fail their initial 1997 test, and a large number of 1995 initial pass cars fail their initial 1997 test.

Figures 14 and 15 show the HC emissions deterioration between I/M cycles of 1995 fail-pass (Figure 14) and initial pass (Figure 15) cars, by model year. The figures have the same vertical scale. Figure 14 indicates that the between-cycle deterioration in HC emissions is quite dramatic for fail-pass cars, particularly for 1987 and older cars. Figure 15 shows a similar trend for initial pass cars. Although the deterioration rates are lower for initial pass cars than for fail-pass cars, the sheer number of initial pass cars (90% of the fleet) greatly influences the overall fleet emissions deterioration. Note that 1993 through 1995 model year initial pass cars have virtually no emissions deterioration (figures for CO and NOx emissions deterioration by model year look similar to Figures 14 and 15 for HC).

Figure 13. Overall Failure Rates in Each Cycle, by Model Ye
Passenger cars tested in both 1995 and 1997, Arizona IM

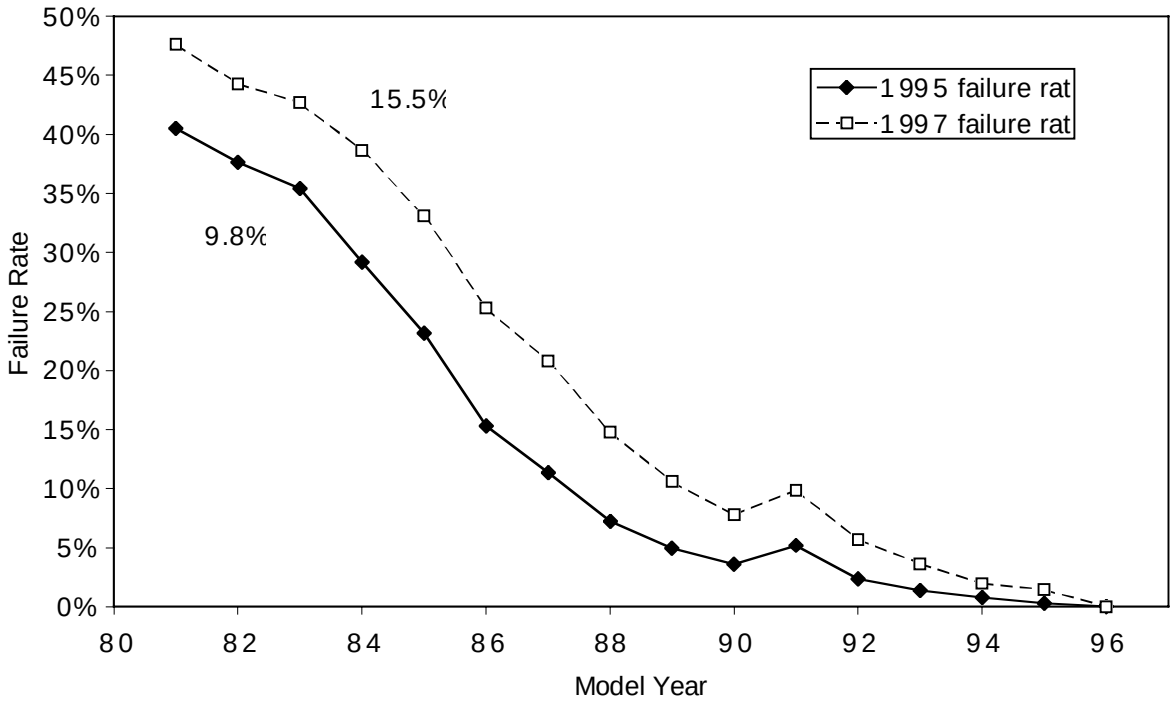


Figure 14. Average 2-year Deterioration in HC Emissions by Vehicle Age
Fail-Pass cars tested in 1995 and 1997, Arizona IM240

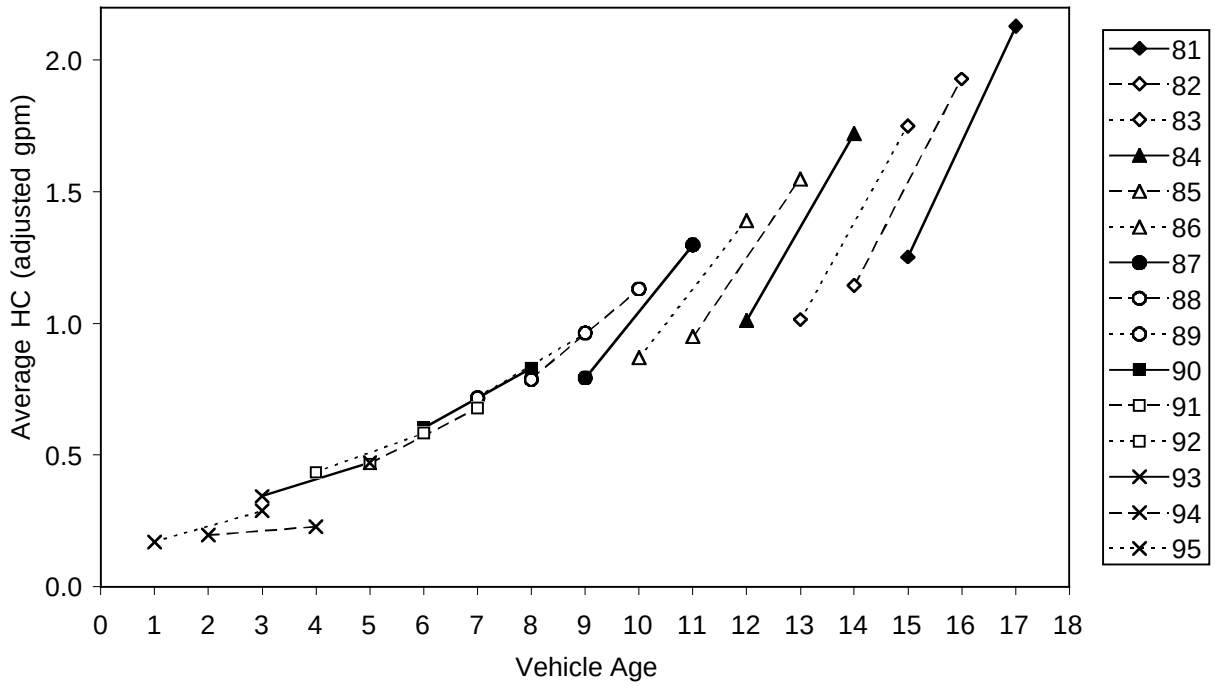
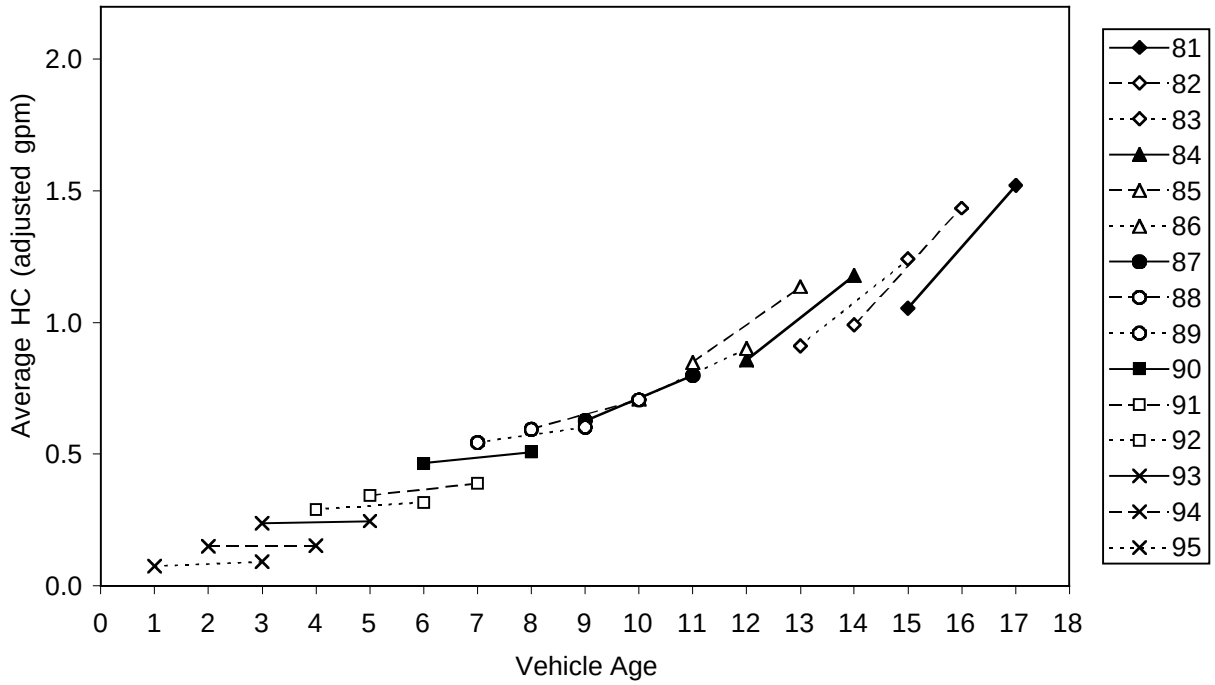


Figure 15. Average 2-year Deterioration in HC Emissions by Vehicle Age
Initial Pass cars tested in 1995 and 1997, Arizona IM240



Simulations

Next we simulate the emission reductions from two major changes to the 1995 I/M program: 1) annual rather than biennial testing of all vehicles (an additional test in 1996); and 2) application of final cut points recommended by EPA, rather than the start-up cut points actually used, in the 1995 test. We note that this analysis only considers the tailpipe emissions benefits from the two simulations; either simulation may result in changes in the number of vehicles avoiding program requirements, the number of vehicles applying for repair cost waivers, and the cost of making repairs. For example, we assume that neither of the changes affects the number of vehicles that never complete I/M requirements after failing their initial test; in other words, we assume that all additional vehicles that fail under each simulation are repaired and pass a subsequent retest. In addition, we do not attempt to simulate what would happen to visual and functional failure rates, and their effect on emissions, under each simulation.

We make the following assumptions for our simulation of annual testing:

- Observed 1995 to 1997 emissions deterioration is linear (simulated 1996 test emissions are the average of final 1995 and initial 1997 emissions);
- Emissions of cars with simulated 1996 emissions in excess of start-up cut points are reduced to final 1995 levels; pollutants with simulated 1996 emissions below cut points are unaffected by any repairs;
- Cars failing simulated 1996 test have simulated 1997 initial emissions equal to simulated initial 1996 emissions (1995 no-final-pass cars failing 1996 test have simulated 1997 initial emissions equal to their observed 1997 initial emissions);
- Cars passing simulated 1996 test have same 1997 emissions as observed.
- Emissions of cars failing the simulated 1996 test deteriorate between 1996 and 1997 at the same rate as observed between 1995 and 1997.

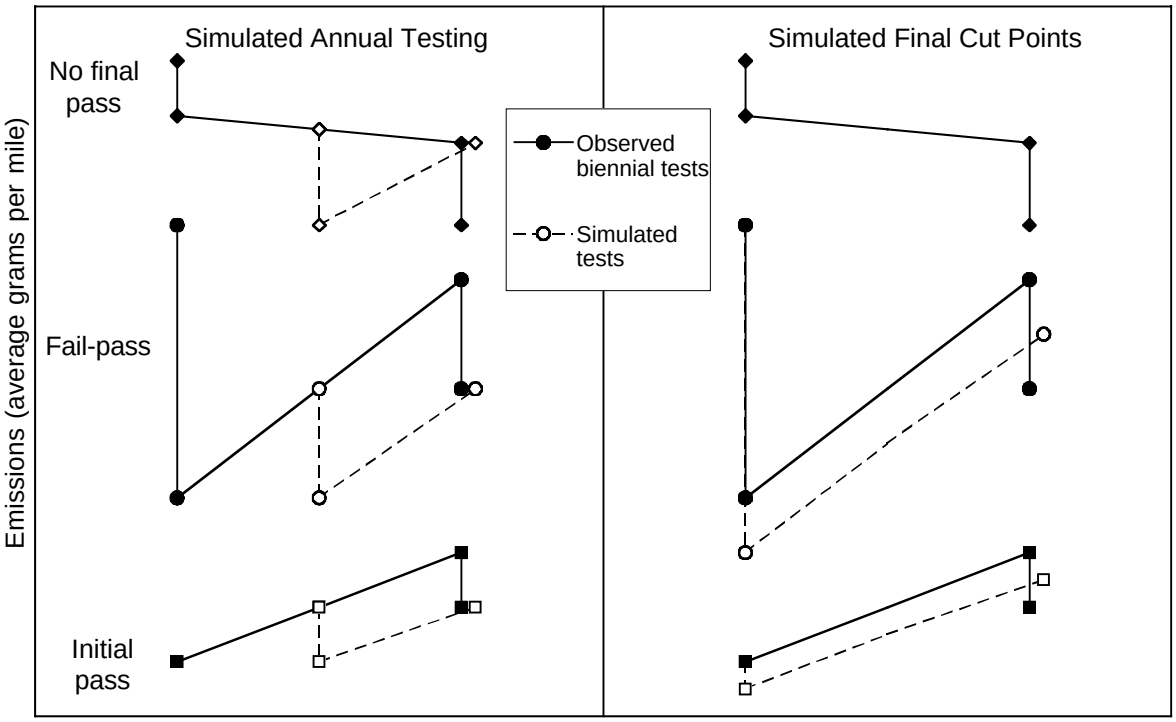
We make the following assumptions for our simulation of final cut points:

- 1995 initial pass and fail-pass cars with 1995 emissions in excess of final cut points fail testing;
- Emissions of cars failing final cut points are reduced to 95% of final cut point. This figure includes excess of emissions of cars that never are repaired down to final cut points. Pollutants with 1995 emissions below final cut points are unaffected by any repairs;
- Emissions of 1995 no-final-pass cars are unaffected.
- Emissions of cars failing final cut points in 1995 deteriorate to 1997 at same rate as observed, and is assumed to be linear;
- Cars passing final cut points in 1995 have same 1997 emissions as observed.
- Emissions of cars failing final cut points deteriorate between 1995 and 1997 at the same rate as observed under actual cut points (parallel deterioration).
- Alternate scenario has same deterioration assumption, but uses alternate cut points that are higher (looser) than EPA final cut points (described below).

Figure 16 presents a sketch of the assumptions we make under each scenario. The figure shows the hypothetical average gram per mile emissions of three subfleets of vehicles, no-final-pass (NFP), final-pass (FP), and initial pass (IP) vehicles, under the annual testing and final cut point

simulations. (The relative position of average emissions is not necessarily drawn to scale.) In each simulation, the closed symbols represent average emissions as observed in the actual 1995 and 1997 I/M tests of cars tested in both years. The open symbols represent simulated average emissions, with the long dashed lines representing the worst case scenario (least emission reduction, most emissions deterioration) and the short dashed line the best case scenario (most emission reductions, least emissions deterioration). Note that in all cases we assume that emissions deterioration is linear over time.

Figure 16. Idealized Sketch of I/M Simulation Assumptions (not to scale)



As described above, for the simulation of annual testing, we take the average emissions of the final 1995 test and the initial 1997 test for each vehicle as the emissions if the vehicle were tested in 1996. This assumes that emissions deteriorate linearly over the two years after a vehicle passes a test in 1995. We then determine how many of the IP and FP vehicles would fail an annual I/M test midway between their final 1995 and initial 1997 test. For cars that would fail a 1996 test, we assume that their 1996 post-repair emissions would be the same as their final 1995 emissions. We assume that the initial 1997 emissions of these vehicles would be the same as their initial 1996 emissions, i.e. the same linear deterioration rate.

For the simulation of more stringent cut points, we first determine how many of the initial IP and FP vehicles would fail EPA final cut points. We then conservatively assume that the post-repair emissions of these vehicles are only 5% lower than the final cut points. (Most vehicles will already have low emissions for pollutants that do not fail initial testing; therefore we assume that repairs do not affect emissions of pollutants that are already below the final cut points. For instance, take a vehicle that fails for NOx, but passes for HC and CO. We assume that the post-

repair HC and CO emissions of the vehicle are the same as measured under the start-up cut points in place in 1995, while the post-repair NOx emissions are 5% below the final NOx cut point.) This assumption is conservative, in that, under a real program, it is possible that cars could be repaired to much lower emissions levels, approaching those of cars that pass their initial test under final cut points. Finally, we assume that emissions of a vehicle repaired under the simulation deteriorates at the same rate as observed under the start-up cut points.

Figure 17 shows the 1995 failure rates by model year of passenger cars in four cases: 1) observed in all cars; 2) observed in the subset of cars given a random full IM240; 3) simulated using final cut points; and 4) observed after applying final cut points to the random sample. The actual 1995 failure rate is 9.8% of all cars; the actual failure rate of the random sample of cars is slightly higher, 10.8%. Older cars in the random sample have a higher failure rate than the overall fleet, perhaps because of relatively small numbers of old cars in the random sample. The figure also shows that 40% of all cars would fail EPA final cut points; this is nearly four times the actual 1995 failure rate. The simulated failure rate under final cut points ranges from 90% for 1981 and 1982 vehicles to 5% for 1994 vehicles. However, only 30% of the random sample would fail EPA final cut points. The figure suggests that our simulation overstates the failure rate, and emission reductions, under EPA final cut points. As discussed above, our simulation results in more failures than one would expect in a real program with final cut points, for several reasons: a real program would likely falsely fast-pass a small number of vehicles that would fail if given a full IM240; the RFF method we use to adjust short tests to full IM240s may overstate full test emissions; and, if the vehicles had actually failed their initial test, many would pass a second chance test (allowed in other I/M programs).

Raising EPA final cut points by 20% would lower the overall failure rate for each model year to roughly match that when the final cut points are applied to the random sample. However, these cut points would still result in extremely high failure rates for older cars, and are therefore unlikely to be adopted. We developed a more realistic set of cut points, shown in Table 4, which would result in the failure rates shown in Figure 18. We chose cut points that would result in an overall failure rate of 25%, and a failure rate less than 55% for any given model year. We simulate emission reductions using both EPA final cut points and these alternate cut points.

Table 4. Start-up, final, and alternate cut points for passenger cars

	Start-up			Final			Alternate		
Model Year	HC	CO	NOx	HC	CO	NOx	HC	CO	NOx
1981-82	2.0	60	3.0	0.8	30	2.0	2.0	60	3.0
1983	2.0	30	3.0	0.8	15	2.0	1.8	28	2.8
1984	2.0	30	3.0	0.8	15	2.0	1.6	26	2.6
1985	2.0	30	3.0	0.8	15	2.0	1.4	20	2.6
1986-90	2.0	30	3.0	0.8	15	2.0	1.0	18	2.4
1991-95	1.2	20	2.5	0.8	15	2.0	1.0	18	2.4

Figure 17. Overall 1995 Failure Rates of Passenger Cars, by Model Y

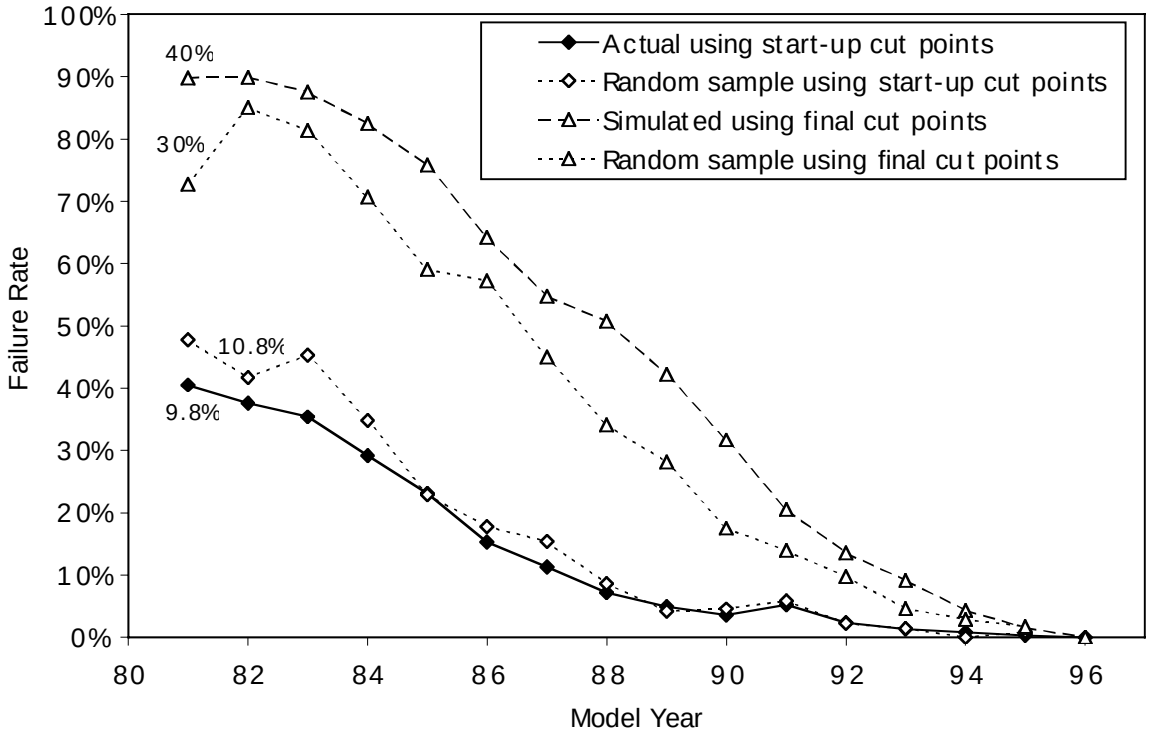
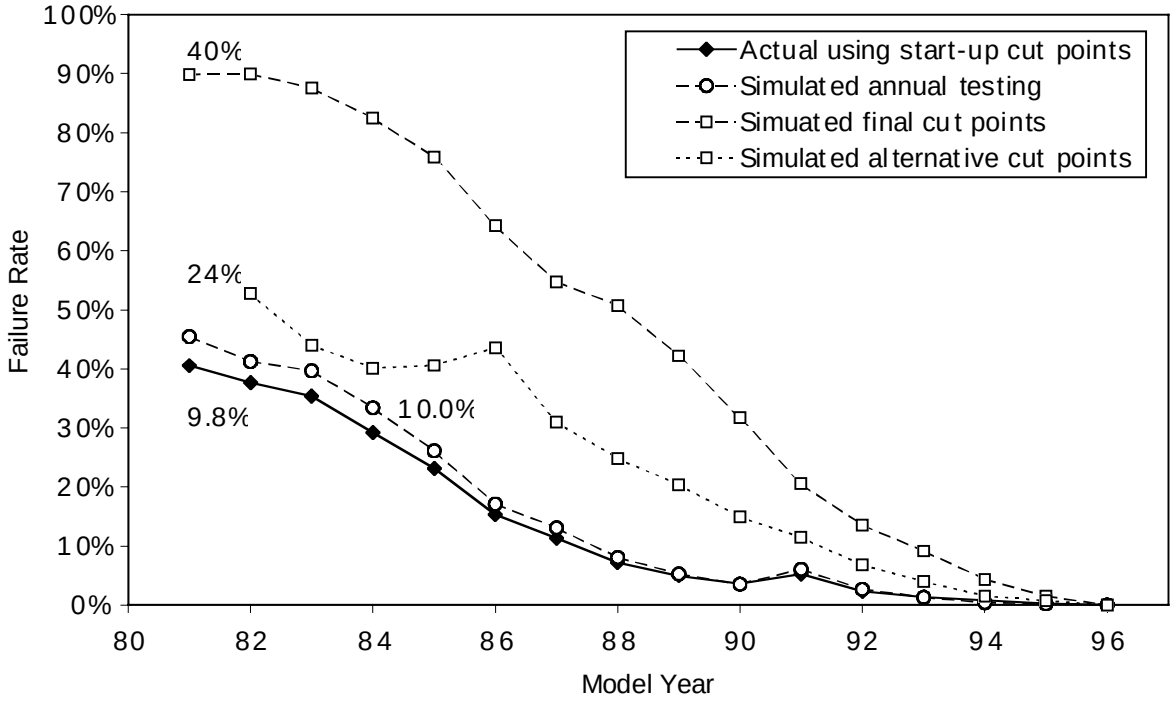


Figure 18. Overall Failure Rates under Simulations, by Model Y
Passenger cars tested in both 1995 and 1997, Arizona IM2



Before we estimate emission reductions from our simulations we need to make an assumption regarding what emissions would be in 1997 if the current I/M program were discontinued. We examine three possibilities:

1) overall fleet emissions would deteriorate to 1997 at the same rate as observed after the 1995 I/M cycle (parallel deterioration);
2) fleet emissions of initial pass and no-final-pass cars would deteriorate as observed, while emissions of fail-pass cars would deteriorate at the same rate as observed; and
3) same as 2), but emissions of fail-pass cars deteriorate to the initial 1995 emissions level of no-final-pass cars by 1997.

Table 5 and Figure 19 show the actual emissions in 1995 and the estimated emissions deterioration by 1997 under these three assumptions. The table and figure indicate that the assumption of fleetwide deterioration parallel to that observed is reasonable. If the I/M program had been ended prior to 1995, and emissions of the overall fleet deteriorated at the same rate as observed after the 1995 I/M cycle, fleet emissions would have been 10% higher for HC, 11% higher for CO, and 6% higher for NOx in 1997 than observed under the program. If emissions deterioration were estimated separately by I/M result, fleet emissions would be slightly higher than the estimate based on parallel fleet deterioration. If fail-pass car emissions deteriorated to the initial 1995 level of no-final-pass cars, the overall fleet emissions would be slightly less than under parallel fleet deterioration, and substantially less than if fail-pass car emissions deteriorated at the same rate as observed. This is because the observed deterioration rate of fail-pass cars applied to their initial 1995 emissions results in higher 1997 emissions than the initial 1995 emissions of no-final-pass cars. Because there is relatively little difference under the three estimates, we use the simple assumption of parallel fleet emissions deterioration for our estimate of what emissions would be if the I/M program were ended before 1995 in our estimates of total emission reductions below.[5] (Although our assumption affects the total amount of emission reductions under each simulation, it does not affect the emission reductions of the simulations relative to each other or the observed reductions from the actual program.)

Table 5. 1997 No-I/M emissions based on different assumptions regarding emissions deterioration

1995 to 1997 No-I/M Emissions	Initial 1997 Tons per Day			Percent above Actual 1997		
Deterioration Assumption	HC	CO	NOx	HC	CO	NOx
Actual	5.31	8.72	12.43	--	--	--
1) Parallel fleet deterioration	5.86	9.68	13.11	10.4%	11.1%	5.5%
2) By I/M result, FP parallel	6.03	9.99	13.24	13.7%	14.6%	6.5%
3) By I/M result, FP to NFP	5.87	9.43	12.95	10.6%	8.2%	4.2%

5. There are two other possibilities regarding emissions deterioration in the absence of an I/M program. One is that emissions from fail-pass cars would not increase at all. This seems unlikely, as the emissions of the fail-pass cars are observed to increase above their initial 1995 level, even after undergoing a cycle of I/M. The other is that the emissions from fail-pass cars would increase at a rate faster than observed under the program, or to levels higher than the initial 1995 emissions of no-final-pass cars.

Figure 19. Estimated Fleet Emissions Deterioration without I/M Pro
Passenger cars tested in both 1995 and 1997, Arizona IM24 C

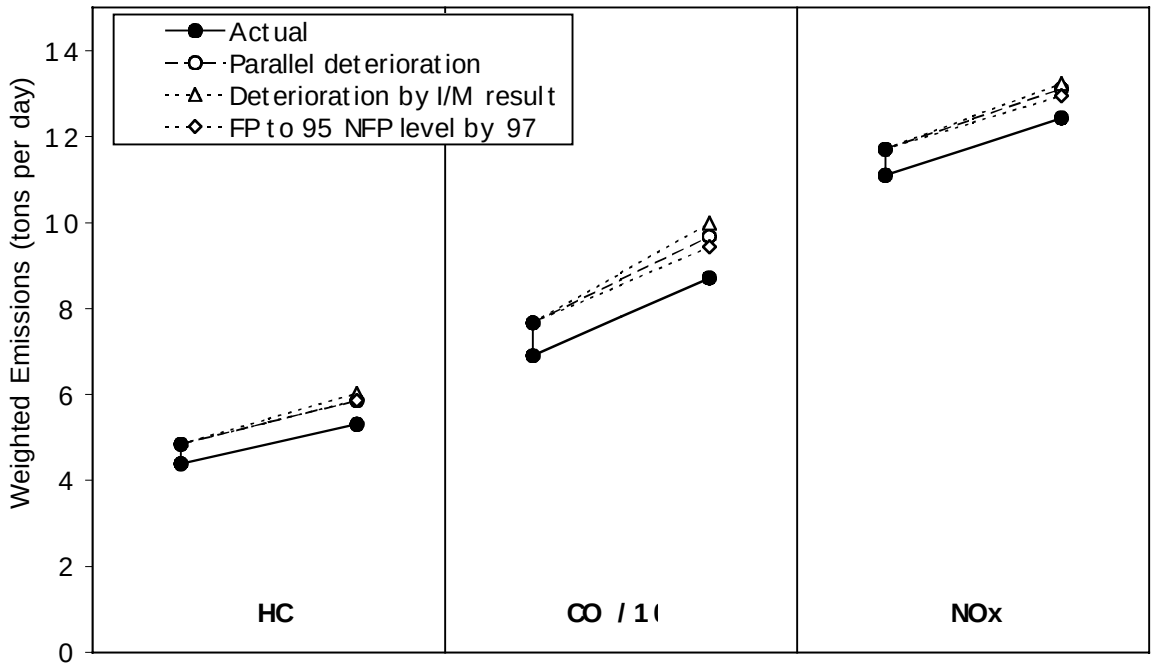


Table 6 and Figure 20 compare the total and ton per day emissions measured under the actual 1995 biennial program with those simulated under annual testing and under final cut points. The actual emission reductions shown in the table assume that, in the absence of the I/M program, emissions would have continued to deteriorate at the same rate as observed for the overall fleet under the I/M program. Table 7 indicates that the simulation of annual testing results in emission reductions 27% to 48% greater than those realized by the actual program in place in 1995. The simulated application of EPA final cut points results in reductions 115% to 227% greater than the actual program, while application of the alternate cut points results in reductions that are 65% to 117% greater. (The percentage reductions are the same whether they are based on total tons or tons per day.) Tighter cut points can also be applied to the annual program. In fact, an annual program might ease the progression to tighter cut points by resulting in relatively lower failure rates per test than biennial programs.

Table 6. Emission reductions observed in actual program and estimated under two simulations

	Total Tons			Tons per Day		
Simulation and scenario	HC	CO	NOx	HC	CO	NOx
Actual	333	557	443	0.46	0.76	0.61
Annual Testing	483	826	561	0.66	1.13	0.77
EPA Final Cut Points	1,090	1,196	1,230	1.49	1.64	1.69
Alternate Cut Points	723	930	732	0.99	1.27	1.00

Table 7. Percent additional reduction from simulations over reduction observed in actual program

Simulation and scenario	Percent Additional Reduction		
	HC	CO	NOx
Percent improvement over actual program:			
Annual Testing	45%	48%	27%
EPA Final Cut Points	227%	115%	178%
Alternate Cut Points	117%	67%	65%

Figure 20 shows the emission reductions observed under the actual program and simulated under annual testing and under alternate cut points. As in Table 6, the figure indicates that either biennial testing using the alternate, more stringent cut points, or annual testing with existing cut points would result in larger emission reductions than those in the current program.

Figure 20. Simulated Overall Fleet Ton Per Day Emissions over 2 I/M Cycles
Passenger cars tested in both 1995 and 1997, Arizona IM240

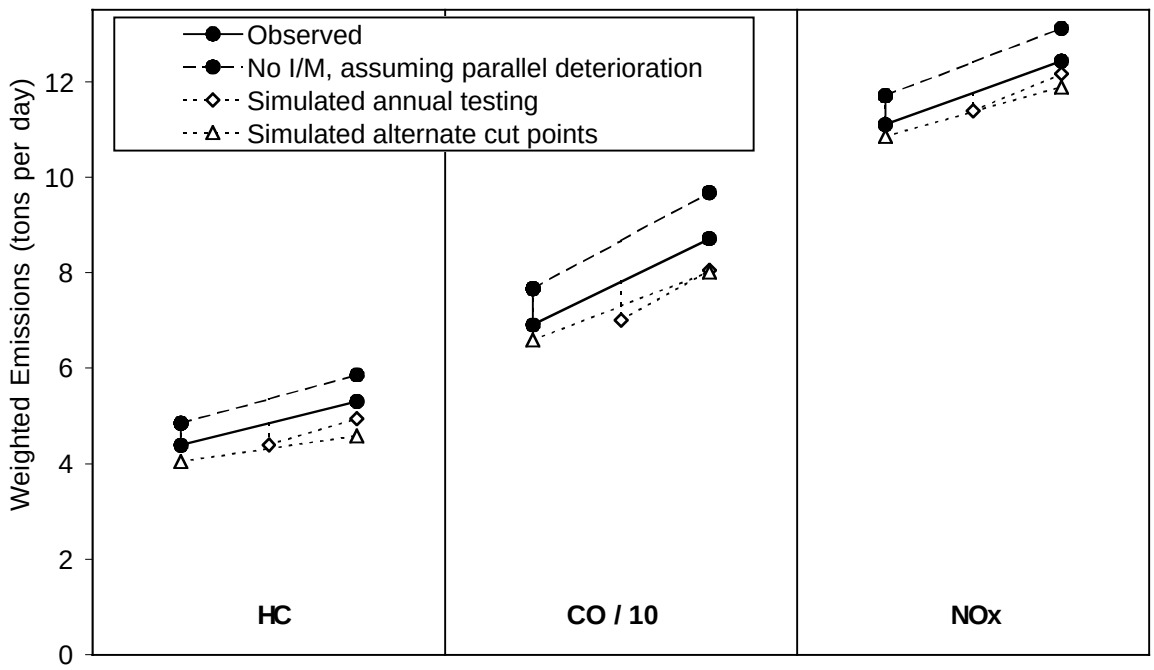


Table 8 shows the estimated emission reductions from 1981 to 1989 cars only, which represent half of the cars tested. The right-hand side of the table compares the tons reduced from the oldest cars with the tons reduced from all cars in Table 6. The simulated annual test results in the same fraction of total emission reductions from the oldest half of the fleet, about 87%, whereas the tighter cut point simulations result in slightly less of the reduction coming from the oldest vehicles (76% to 81%).

Table 8. Total tons reduced from 1981 to 1989 cars, and fraction of reduction from all cars

Simulation and scenario	Total Tons from MY81-89			MY81-89 Fraction of Total		
	HC	CO	NOx	HC	CO	NOx
Actual	286	482	388	86%	87%	88%
Annual Testing	417	718	489	86%	87%	87%
EPA Final Cut Points	617	497	635	82%	78%	81%
Alternate Cut Points	313	282	234	80%	76%	81%

Although the application of tighter cut points appears to offer greater reductions than moving to annual testing alone, it is likely that the benefits of tighter cut points are somewhat overstated. First, as discussed above, our application of final cut points may overstate the failure rate, and therefore the amount of excess emissions available to be reduced. However, it is not clear how to overcome this potential limitation of our analysis, short of applying somewhat less stringent final cut points in order to match the failure rate observed after applying final cut points to the random sample of cars given the full IM240 (as we do in our alternate cut point simulation).

Second, our assumption that emissions deterioration under final cut points is parallel to that observed under start-up cut points is optimistic; it is likely more difficult to make proper, lasting repairs to meet the tighter final cut points than the less stringent start-up cut points. If so, emissions should deteriorate more rapidly under final cut points than observed under start-up cut points. In the worst case, the application of tighter cut points would have no additional long-term benefit beyond the start-up cut points; that is, emissions of all cars would deteriorate to the level observed in their initial 1997 test.

In the simulation of annual testing, we assume that emissions of cars repaired in 1996 deteriorate at the same rate as observed between 1995 and 1997 (assuming linear deterioration over each period). It is possible that annual testing may result in more complete and durable repairs, and eliminate emissions deterioration in repaired vehicles. On the other hand, owners of vehicles that repeatedly fail I/M testing every year may be more likely to avoid completing I/M requirements, without making repairs that reduce emissions. Analysis of vehicles undergoing two I/M cycles within a two-year period may reveal whether annual testing results in lower emissions deterioration.

Finally, actual program performance depends on the number of vehicles that are forced to be scrapped or otherwise removed from the I/M area. To the extent that either alternative would result in greater numbers of high emitting vehicles being removed from the I/M area, the relative effectiveness would change. On the other hand, the extent to which each alternative results in vehicles avoiding I/M program requirements while continuing to be driven in the I/M area also would affect the relative effectiveness of each alternative.

It is possible that changing all of these assumptions might make the annual testing simulation as effective, or more effective, as tightening cut points.

Analysis of Arizona Multi-cycle Fleet

As discussed above, Figure 1 implies that the deterioration in emissions over the two-year period between I/M cycles is linear. However, analysis of vehicles receiving off-cycle I/M tests in California's decentralized program indicates that this may not be the case. (About 8% of vehicles in California's Enhanced program receive an off-cycle test because an I/M test is required on change of ownership.) In California 20% of fail-pass vehicles (and 6% of initial pass vehicles) failed a second I/M cycle within 3 months of passing their initial I/M cycle. There are four possible causes for a vehicle to fail an I/M test soon after completing an I/M cycle: 1) the passing result in the initial cycle is a fraudulent test result; 2) the vehicle owner or mechanic knows how to adjust the vehicle to pass a test, and then readjusts the vehicle after it passes; 3) the vehicle has an intermittent emissions control problem which causes it to fail a test and pass a retest without any repairs being made; or 4) the vehicle has a problem that can be temporarily improved by a repair but the effects are short lived (e.g. replacing spark plugs in an engine with one or more leaky piston rings that result in excessive carbon build-up on the spark plugs).

Arizona requires that a vehicle must pass an I/M test prior to being resold into the state I/M areas. However, dealers with a fleet vehicle inspection permit can self-inspect such vehicles. The self-inspection consists of an idle test and a visual inspection. As this inspection is not as comprehensive as an IM240, the state recommends that consumers purchasing a used vehicle perform a voluntary IM240 test. Analysis of vehicles given two I/M tests in less than a two-year period allows us to study short-term emissions deterioration (we call this group of vehicles the "multi-cycle" fleet). We can then compare repeat failure rates of the Phoenix multi-cycle with those observed in California. The Phoenix IM240 program is run by a centralized test-only contractor, whereas the majority of vehicles tested in California are tested at decentralized stations that can also perform repairs. Observers suspect that there is a large degree of test fraud occurring at decentralized California I/M stations. If that is the case, then there should be a lower repeat failure rate immediately after the initial I/M cycle in the centralized Phoenix program than in the decentralized California program.

The analysis of the Phoenix multi-cycle fleet is complicated in that, when a vehicle that does not receive a passing test within 5 months of their initial test returns later for an I/M test, that test is coded as an initial test rather than a retest, even though the vehicle never completed I/M requirements from its initial cycle. To account for this we disregard whether a test is coded as an initial test or retest, and we determine the end of an I/M cycle after a vehicle has fully passed an I/M test.

We examine only those vehicles with an initial I/M cycle in 1995; 87,000 vehicles (12% of all initially tested in 1995) had a second I/M cycle before 1997. Figure 21 shows the distribution of this multi-cycle fleet, by time since the initial test in their first I/M cycle. The figure indicates that 40% of the multi-cycle fleet had their second I/M cycle 12 to 14 months after the initial test in their initial cycle. We suspect that the owners of many of these vehicles brought them in for testing because they were confused about the switch from annual to biennial I/M testing in 1995, and not that they intended to sell their vehicle.[6] We call this subset of the multi-cycle fleet the

6. Arizona DEQ confirms this suspicion, and can provide no other reason for why so many vehicles were tested 12 months after their initial IM240 cycle.

"annual" fleet, as they received a second I/M cycle roughly one year after their initial cycle. There is a slight increase in vehicles tested 24 months after their previous I/M cycle. Many of these tests are regular biennial tests of vehicles that come in early for their registration renewal (vehicles whose renewal is due in January 1997 that come in for testing in December 1996).

Figure 22 compares the model year distribution of the overall I/M fleet,[7] the entire multi-cycle fleet, and the annual fleet; the figure indicates that both the overall multi-cycle fleet and the annual fleet are older than the overall I/M fleet, and that the annual fleet is slightly younger than the overall multi-cycle fleet. The overall I/M fleet has a lower failure rate by model year than the multi-cycle fleet or the annual fleet, as shown in Figure 23. Figures 24 through 26 show the initial and final emissions in the first I/M cycle by model year, for the overall I/M and multi-cycle fleets. Both the overall multi-cycle fleet and the annual fleet tend to have slightly higher initial HC (8%) and CO (5%) emissions than the overall I/M fleet; however, the final emissions (presumably after repair) of the two fleets are comparable.[8] The annual fleet has slightly lower initial emissions than the entire multi-cycle fleet. One explanation for the multi-cycle fleet being older, having a higher failure rate, and slightly higher initial emissions than the I/M fleet is that an owner may be more likely to try to sell a vehicle that has an emissions problem. Another is that a potential buyer may be more likely to request a voluntary I/M inspection prior to purchasing a vehicle that he or she suspects may later fail an I/M test. Because most of the annual fleet is not being tested because of a pending sale, we expect that this fleet to be more like the overall fleet than the multi-cycle fleet (i.e., to be somewhat younger, and to have a lower fail rate and lower initial emissions, than the multi-cycle fleet). However, the annual fleet still has slightly higher initial HC (4%) and CO (2%) emissions than the overall I/M fleet.

7. What we call the "overall I/M" fleet excludes vehicles with two initial I/M tests in 1995, but includes vehicles with an initial test in 1995 and a second initial test in 1996. Thus the overall I/M fleet includes all of the annual fleet, and some of the multi-cycle fleet.

8. Figures 24 through 26 are based on all vehicles tested in the Phoenix I/M program in 1995. If only those vehicles tested in both 1995 and 1997 are used, the average emissions of the overall fleet are lower, and the difference between the emissions of the multi-cycle and overall fleets are greater.

Figure 21. Distribution of Vehicles by Months since Previous C
1995 Arizona IM240

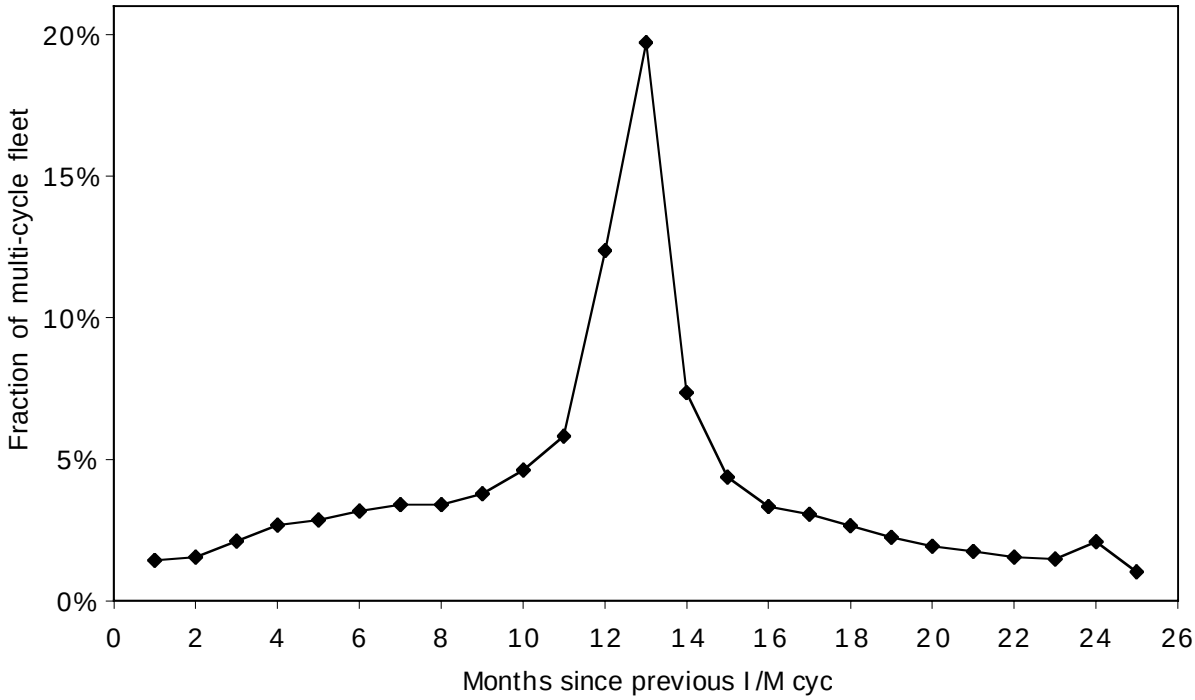


Figure 22. Distribution of Cars by Model Year and Test F
1995-97 Arizona IM240

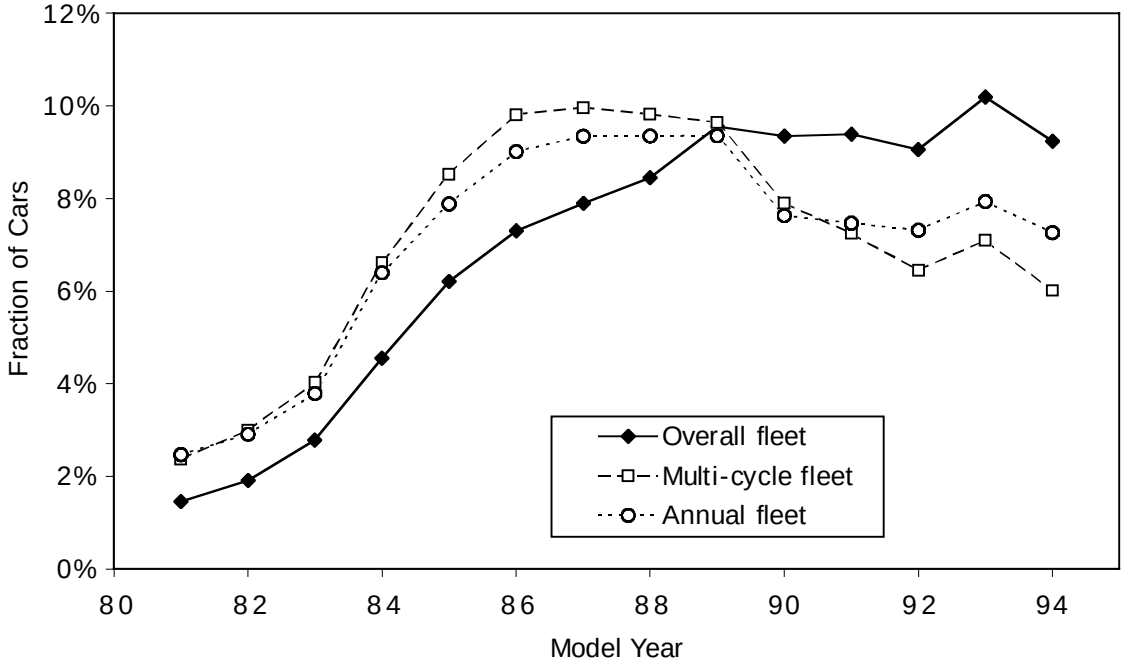


Figure 23. Fail Rate of Cars by Model Year and Fle

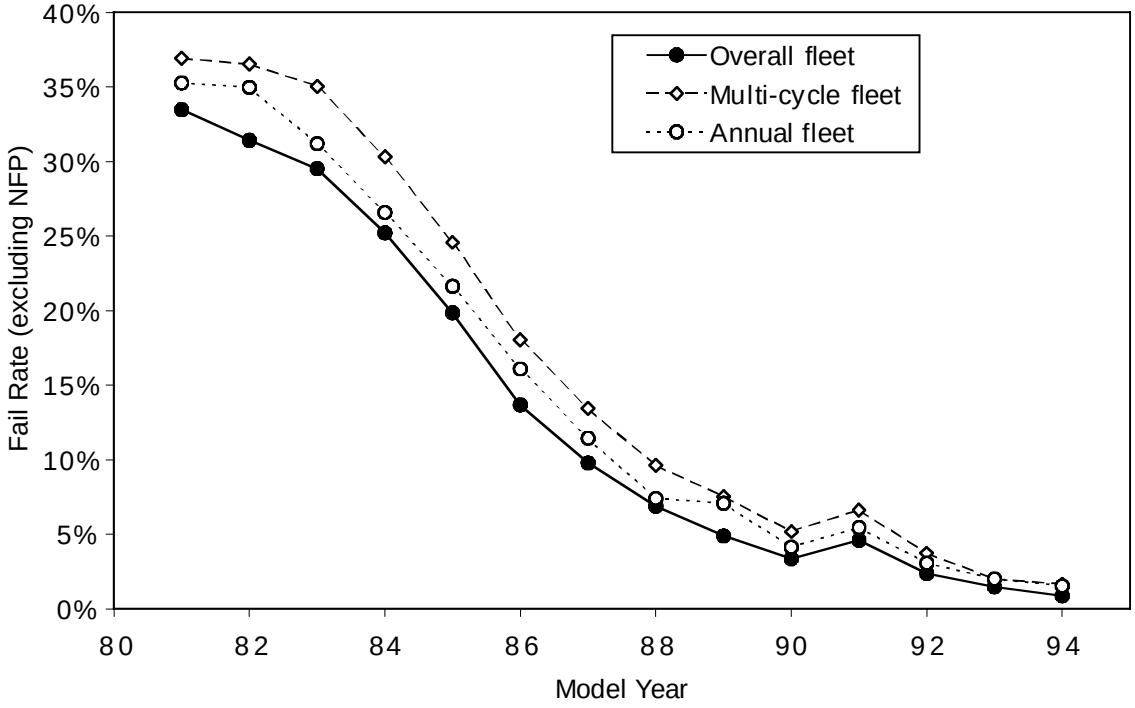


Figure 24. Average HC Emissions of Fail-Pass Vehicles by Fleet and T

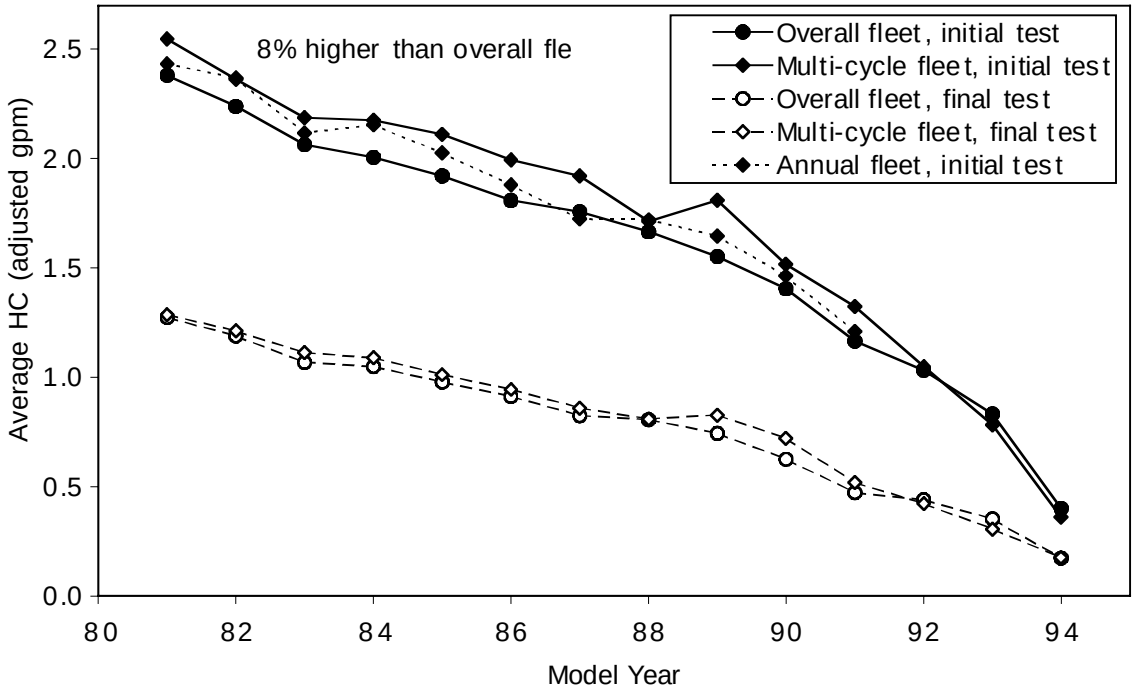


Figure 25. Average CO Emissions of Fail-Pass Vehicles by Fleet and T

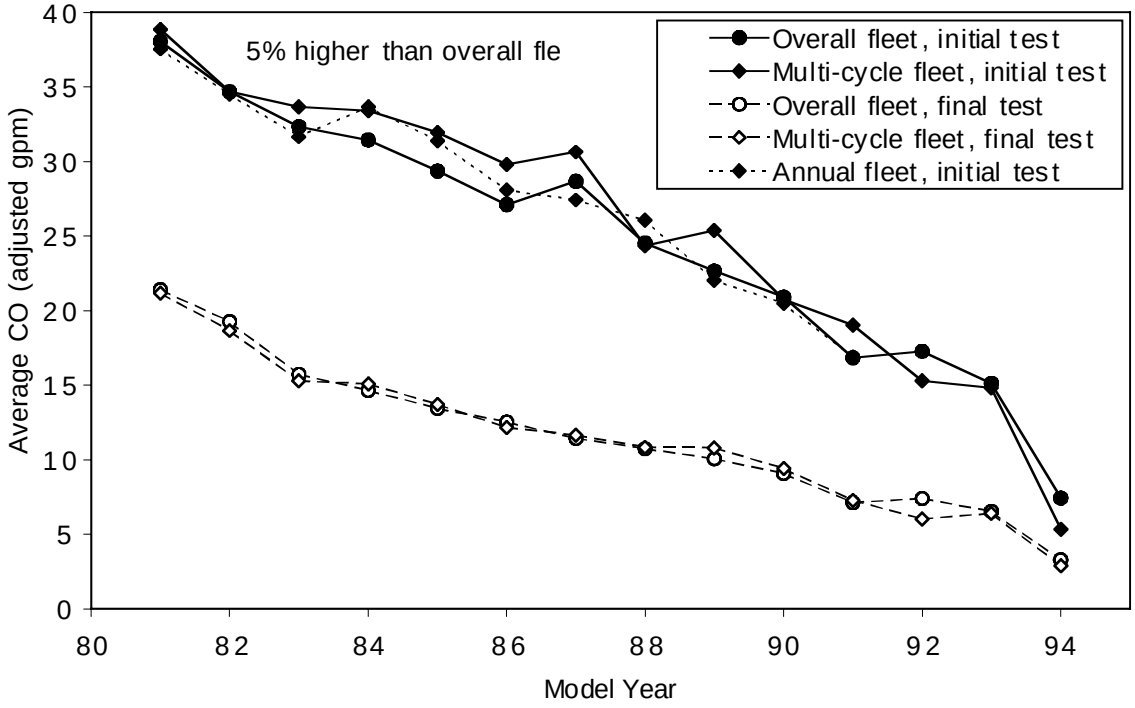


Figure 26. Average NOx Emissions of Fail-Pass Vehicles by Fleet and T

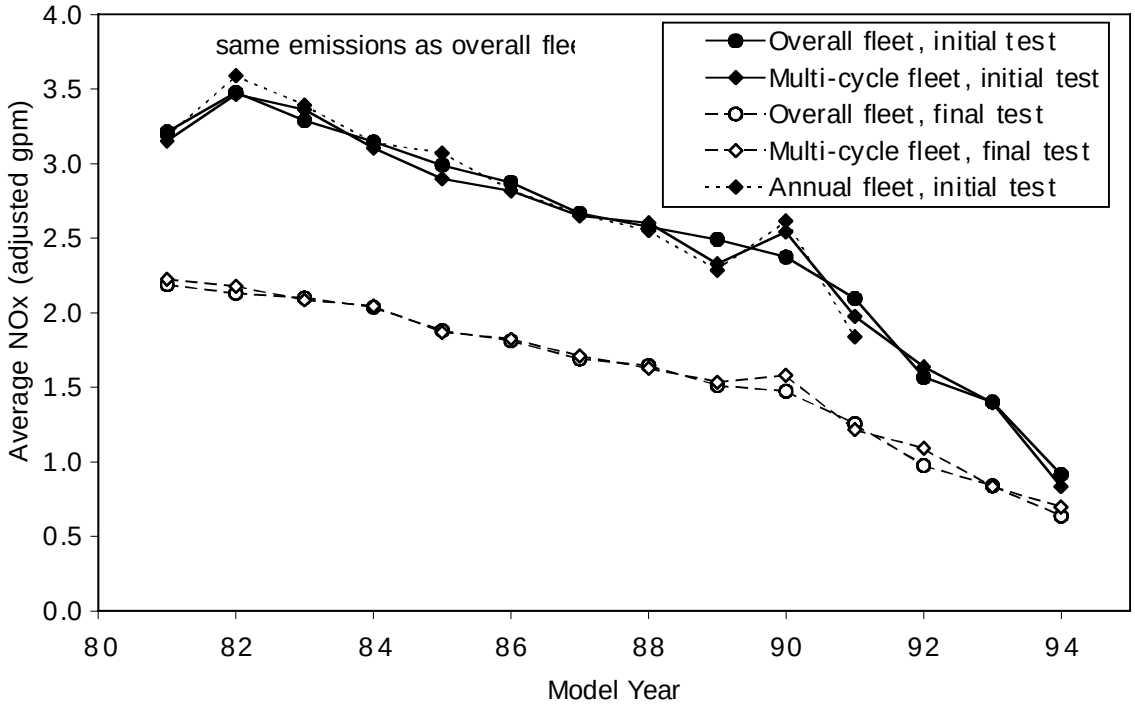
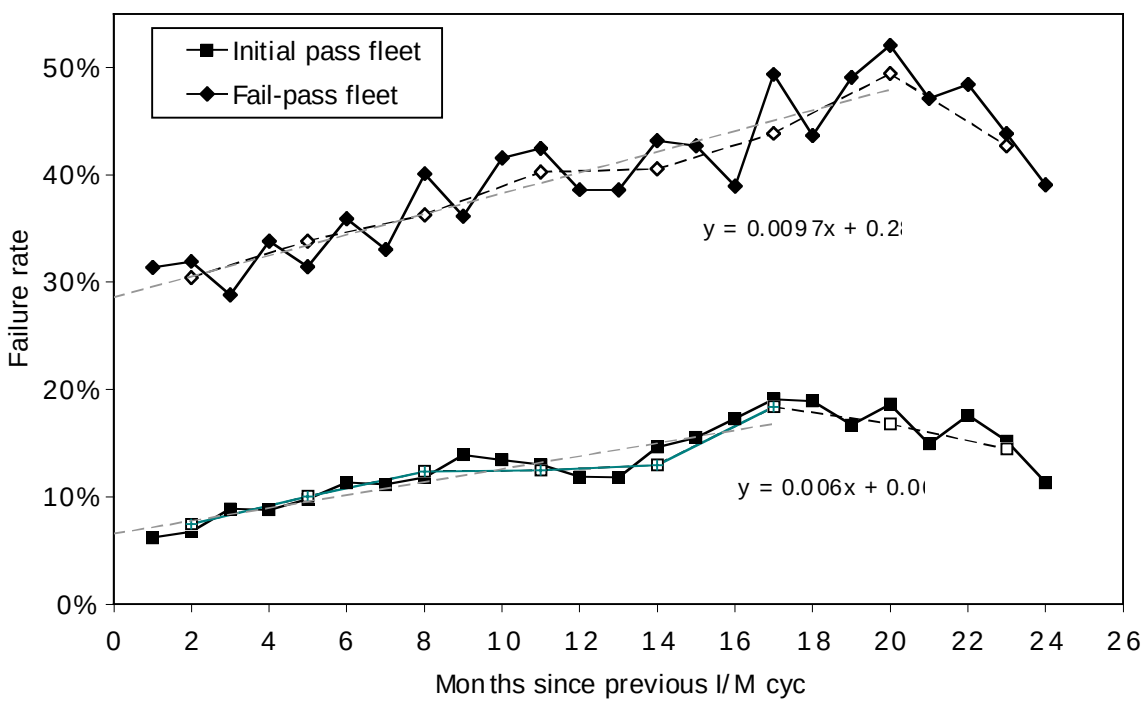


Figure 27 shows the fraction of vehicles that fail their initial test in their second I/M cycle, by the number of months from their initial test in their first I/M cycle (the filled symbols connected by solid lines show the trend by month since the first cycle, while the open symbols connected by dashed lines aggregate the data into three-month time periods). The failure rates for vehicles that pass their initial test in their first I/M cycle, and those that fail their initial test but pass a retest, are shown separately. It is important to note that each time period represents a different group of vehicles, tested at different times since their previous I/M cycle. The figure indicates that 30% of fail-pass vehicles fail their next I/M test up to 3 months after passing their previous test. This repeat failure rate increases to over 45% 21 months after their previous test, but then decreases. Similarly, 6% of initial pass vehicles fail their next I/M test within 3 months of passing, and nearly 20% fail within 18 months of passing. The trend lines in the figure indicate that, if all of the multi-cycle fleet were tested immediately after passing their initial I/M cycle, 29% of the fail-pass vehicles and 7% of the initial pass vehicles would fail. These rates imply that, if all vehicles were tested immediately after completing their initial I/M cycle, the same percentage, 10%, would fail as failed the initial cycle (10% fail-pass in cycle 1 * 29% fail cycle 2 = 3%; 90% initial pass in cycle 1 * 7% fail cycle 2 = 6%; 3% + 6% = 9%). However, two-thirds of these failing vehicles would be new failures that passed in the initial I/M cycle.

Figure 27. Failure Rate in Next Cycle by Months since Previous C
1995 Arizona IM240



There are at least four possible causes for a vehicle to fail so soon after passing an I/M test:

1. The vehicle was not properly warmed up prior to the failing tests in the first and/or second cycles. This is a relatively serious problem in the Phoenix program, as vehicles are allowed to fast fail after only 94 seconds of testing, and are not allowed a second IM240 if their

emissions are only marginally higher than the cut point (as in the Colorado and Wisconsin IM240 programs).

2. A legitimate repair made to the vehicle did not last or did not address the underlying cause of the initial failure.
3. The second cycle failure was due to an emissions problem unrelated to the problem identified and repaired in the initial cycle.
4. The vehicle has an intermittent problem which causes it to fail one test and pass a subsequent test, without any repairs being made.

The high failure rates so soon after passing an I/M cycle indicates that the Phoenix program is not identifying and repairing large numbers of vehicles with high emissions. It is important to note that the Phoenix multi-cycle fleet may not be representative of the overall I/M fleet. As discussed above, the multi-cycle fleet is older, has a higher initial failure rate, and slightly higher initial emissions than the overall fleet. It is quite likely, therefore, that the multi-cycle fleet has a higher repeat failure rate than the overall fleet. Again, this may be because many of the multi-cycle fleet vehicles have been recently sold (either after the initial I/M cycle or the second I/M cycle), perhaps because of an emissions problem; and, because an I/M test is not required when a vehicle is sold, a potential buyer may be more likely to request a voluntary I/M inspection for a vehicle that he or she suspects may later fail an I/M test.

The failure rates for both the fail-pass and initial pass vehicles begin to decrease after 20 months after their initial I/M cycle. One possible explanation for this decrease is that the population of vehicles varies in each time period. Figure 28 shows that the average model year changes substantially by time period; the initial pass fleet tested up to 3 months, 11 to 14 months, and over 21 months after the previous I/M cycle is younger than the initial pass fleet tested at different time periods. We would expect younger vehicles to have lower emissions and lower failure rates than older vehicles, and to explain some of the trend of lower repeat failure rates more than 21 months after the previous I/M cycle in Figure 27. In addition, we have to be careful that we are not observing a shift among vehicle types in each time period. Figure 29 shows that the fraction of cars (versus light duty trucks, which are subject to less stringent cut points) remains fairly constant over each time period.

Another explanation for the lower fail rate more than 20 months after the previous I/M cycle (in Figure 27) is that many of these tests, particularly those 24 months after the previous cycle, are early biennial renewals, rather than I/M tests associated with a change of vehicle ownership.

Figure 28. Distribution of Vehicles by Months since Previous C
1995 Arizona IM240

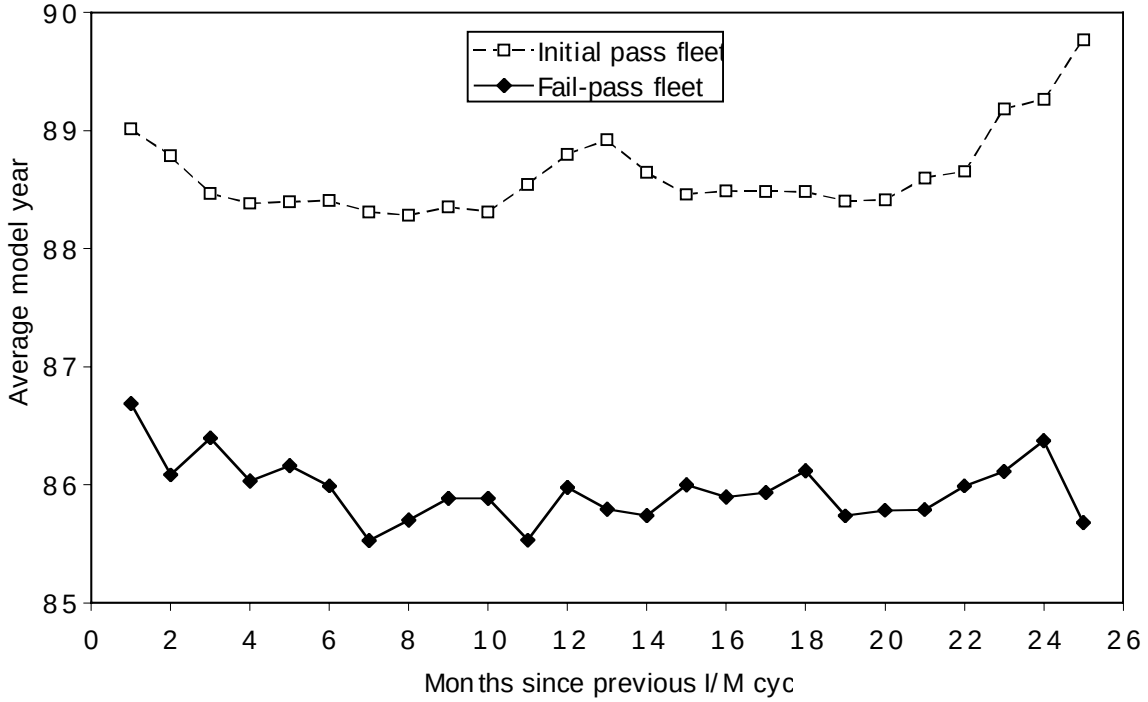
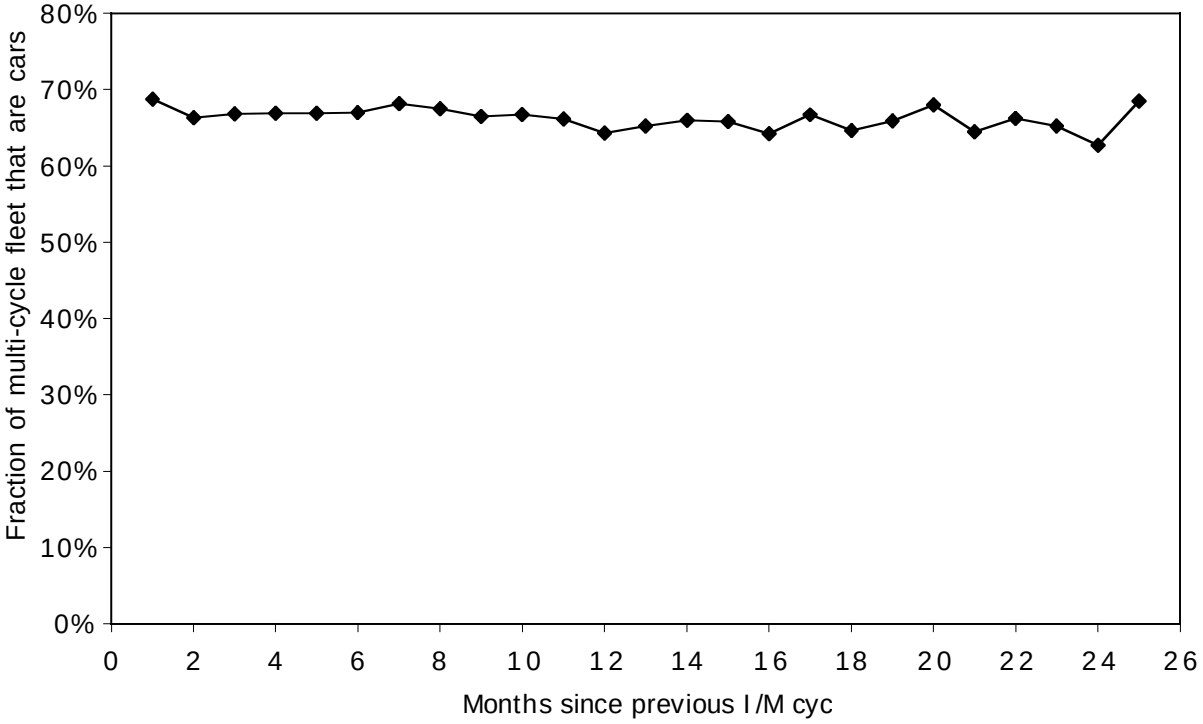


Figure 29. Fraction Cars by Months since Previous Cy



Analysis of Arizona "Annual" Fleet

There is a small number of cars in the multi-cycle fleet (3,700, or 6%) that had a second I/M cycle 12 to 14 months after, and a third I/M cycle 24 to 26 months after, their initial cycle. Of these, 368 (10%) were fail-pass cars in their initial I/M cycle in 1995. The emissions of these cars can be tracked over three I/M cycles, and compared with the emissions of the overall I/M fleet, to gauge the effect annual testing had on emissions deterioration. Figure 30 presents this comparison in terms of average gram per mile emissions (emissions of the annual fleet are offset slightly to make the figure easier to read). The model year distribution of cars in the annual and overall fleets are virtually identical, so none of the differences between these two fleets' emissions in Figure 30 can be explained by vehicle age. If annual testing had a long-term effect on emissions, one would expect the emissions deterioration rate between I/M cycles to decrease over time. However, Figure 30 indicates that the emissions deterioration of the 1995 fail-pass fleet between 1996 and 1997 is greater than the deterioration between 1995 and 1996, at least for HC and CO. On the other hand, the NOx emissions deterioration rate is about the same in between the three cycles.

Figure 30. Average Emissions over Two Years by Fl
1995 Fail-Pass passenger cars tested again in 199

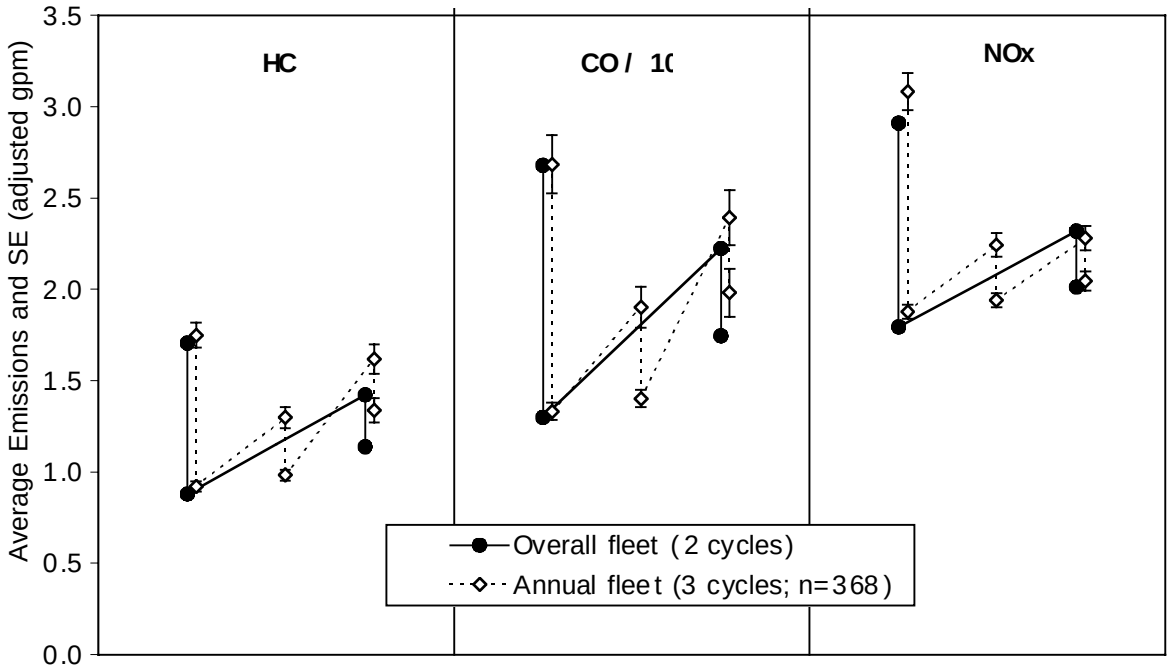


Figure 31. Average Emissions in Two Years by Fle
1995 passenger cars tested again in 1997

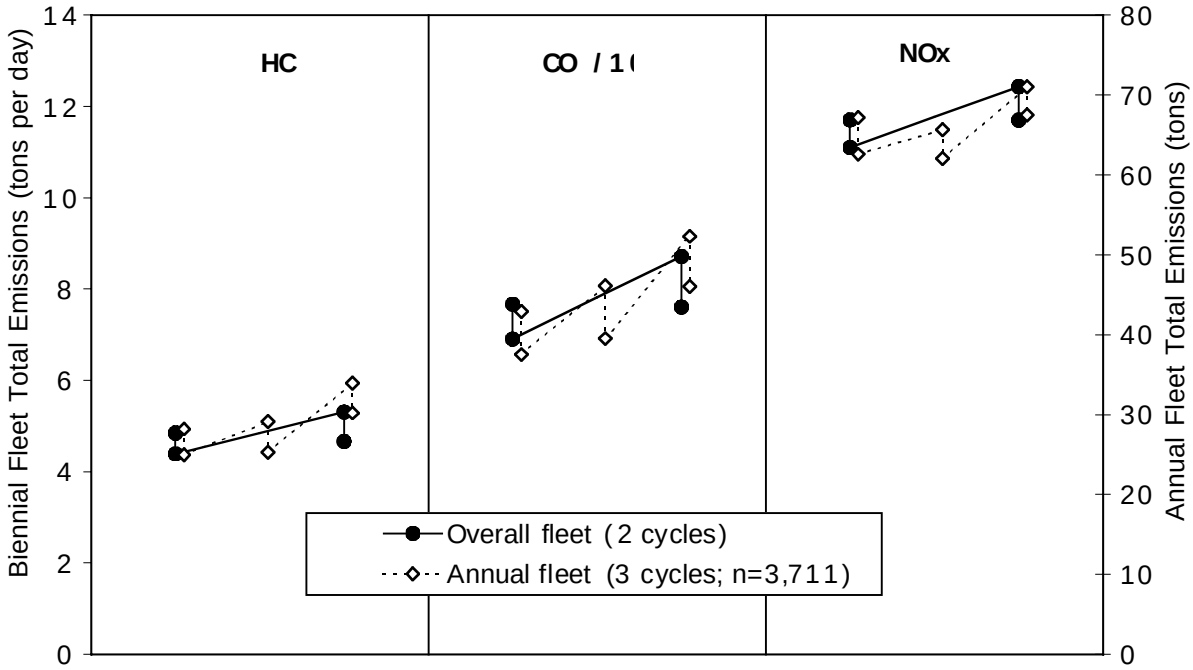


Figure 31 combines the initial pass cars with the fail-pass cars of Figure 30, and compares the total emissions of the annual and overall fleets. Note that because the annual fleet is much smaller than the overall fleet, emissions of the biennial fleet are shown on the left-hand axis in tons per day, while the emissions of the annual fleet are shown on the right-hand axis in total tons. The figure indicates that emissions after an annual test in 1996 deteriorate at a faster rate than after the initial 1995 test. This suggests that our assumption of parallel deterioration in the simulation of emission reductions from annual testing (shown schematically in Figure 16) is overly optimistic, and that actual emission reductions from annual testing, at least for HC and CO, are likely to be lower than either scenario shown in Table 6.

Comparison of Phoenix and Tucson I/M Programs

We have one other source of data to investigate the effect of annual I/M testing: comparison of on road measurements of vehicles participating in the Phoenix and Tucson I/M programs. Vehicles in the Tucson program receive two-speed idle testing every year. A small number of vehicles (9,500) participating in the Tucson annual two-speed idle I/M program were measured in the Phoenix area between January 1996 and August 1997. Figures 32 and 33 compare the on-road emissions measured by remote sensing of the Tucson and Phoenix I/M fleets, by model year.[9] The newer vehicles in the Tucson fleet have higher on-road emissions than the newer vehicles in the Phoenix fleet; however, the older Tucson vehicles have much lower emissions

9. All light duty vehicle types are shown, as the Tucson program does not define cars and trucks the same way as the Phoenix program. The distribution of light duty trucks with gvw greater than 6,000 lbs is similar between the two programs.

than the older Phoenix vehicles. The average on-road emissions of the overall Phoenix fleet are 13% higher for HC and 9% higher for CO than those of the comparable Tucson fleet. For 8-year old and older cars (model years 1981 through 1988), on-road emissions of Phoenix vehicles are 47% and 37% higher, for HC and CO respectively, than on-road emissions of Tucson vehicles

One possible explanation for this result is that the Tucson vehicles measured in Phoenix are better maintained than the average vehicle in the Tucson I/M fleet. However, comparison of the I/M emissions by model year of each fleet seen by remote sensing with those of the overall fleets in each I/M area indicate that the I/M emissions of each fleet measured by remote sensing are representative of the overall I/M fleet in each area. Another possible explanation is that repairs made to the Phoenix fleet to meet NOx cut points result in slightly higher HC and CO emissions than the Tucson fleet, which is not subject to NOx cut points. A third possibility is that the Tucson program idle cut points for older vehicles are relatively more stringent than the Phoenix IM240 cut points for older vehicles. However, the most likely explanation is that annual testing in Tucson is resulting in lower emissions for older vehicles than the biennial program in Phoenix, simply because the vehicles are tested (and repaired) more frequently.

Figure 32. Average Remote Sensing CO by Program Type
1996-97 Arizona I/M

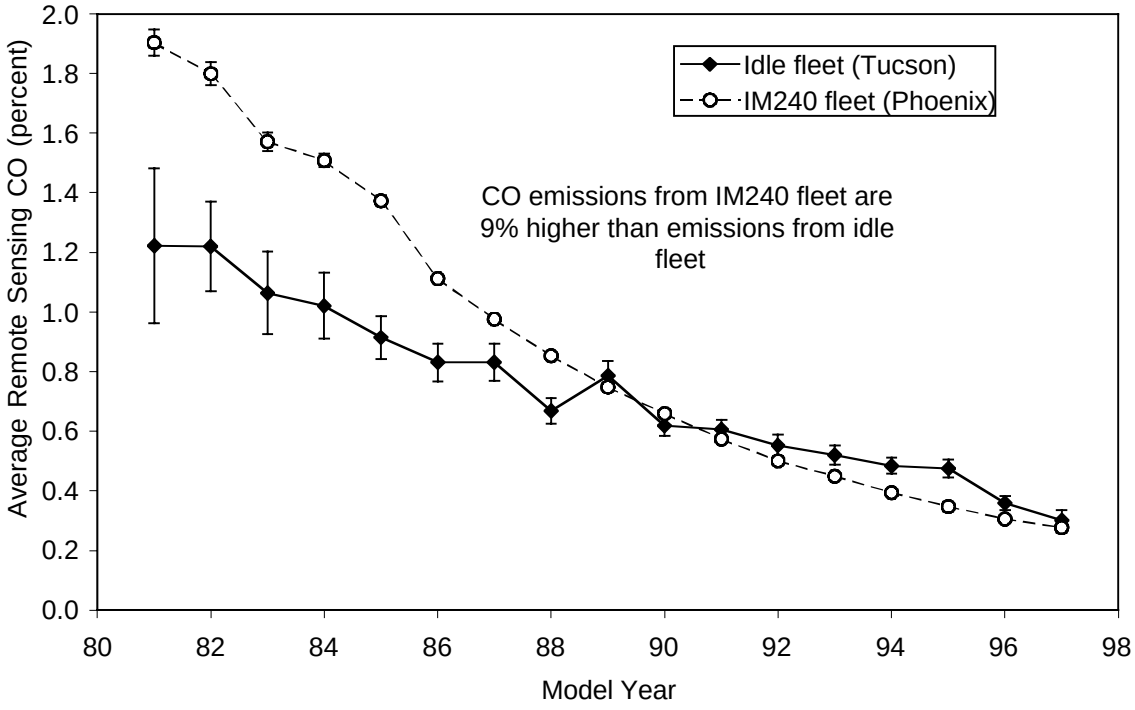
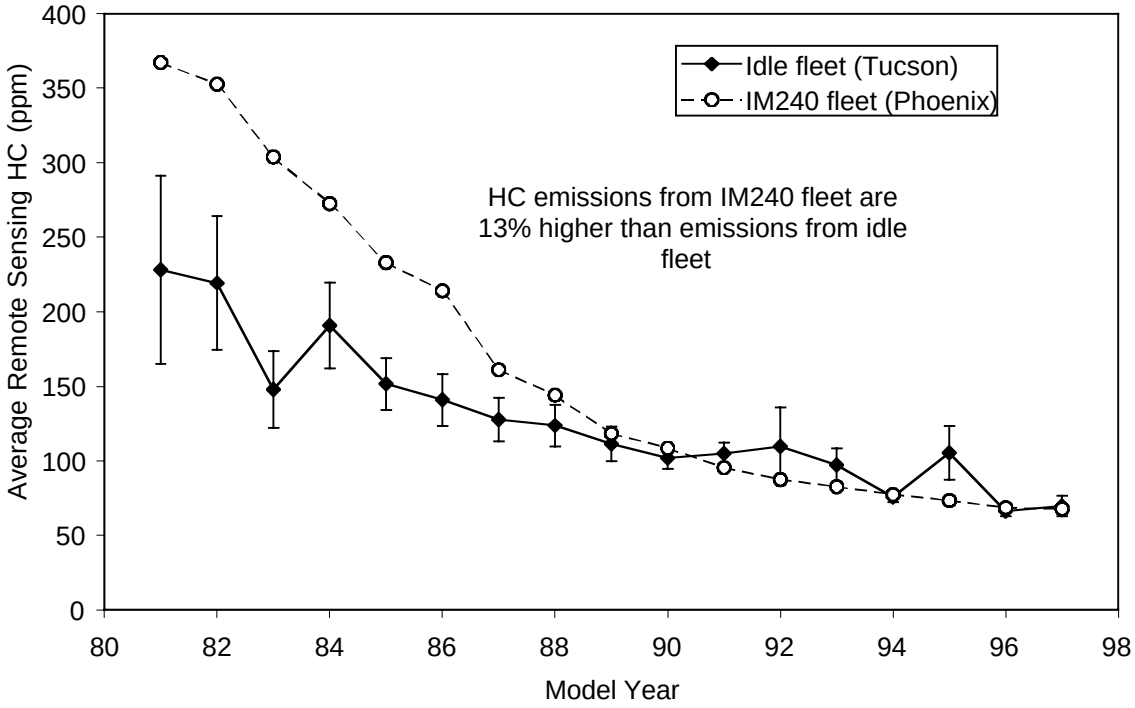


Figure 33. Average Remote Sensing HC by Program Type
1996-97 Arizona I/M



Conclusions

In this report we have examined the potential for additional emission reductions from adopting annual rather than biennial testing in I/M programs. We tracked a fleet of cars reporting for testing in two biennial cycles in the Phoenix IM240 program. We found that the amount of emissions in excess of program cut points increases from the first to the second I/M cycle. In addition, approximately one year of emissions deterioration between I/M cycles equals the emission reductions resulting from each I/M cycle. There are two causes for this increase in emissions over time: 40% of cars that fail their initial test, but pass a subsequent retest, in the first cycle fail their initial test two years later; and emissions of cars that pass their initial 1995 test increase over time as more of them become high emitters. The emissions deterioration between cycles is much higher for older cars than for younger cars.

Our simulation of an annual test in the Phoenix program in 1996 results in an additional 45%, 48%, and 27% reduction in HC, CO, and NOx emissions, respectively, over what occurred in the biennial program. These results assume that emissions after the second cycle would deteriorate at the same rate as after the initial I/M cycle. Applying more stringent cut points under a biennial program would have increased the failure rate to 24% and resulted in even larger additional reductions in HC (120%) and NOx (65%) emissions than under our annual testing simulation. Nearly 90% of the actual emission reductions observed in the program, and the simulated reductions under annual testing, come from 1981 to 1989 cars, which represent about half of the cars tested. Slightly less (80%) of the total reductions from applying tighter cut points come from these older cars.

Analysis of cars given an off-cycle test in Phoenix reveals that 7% of vehicles that initially pass their first I/M cycle fail a second I/M cycle up to 3 months later. Vehicles that fail and then pass in the first I/M cycle have a much higher repeat failure rate up to 3 months later: 30%. These high failure rates so soon after a passing I/M test should be of great concern to managers of I/M programs. It is possible that these failure rates are due to inadequate preconditioning, inadequate repair of broken vehicles, or new problems on repaired vehicles. However, these high failure rates may indicate that inherent variability in a vehicle's emissions, particularly high-emitter vehicles, accounts for some of the apparent emission reduction attributed to I/M programs. In addition, it suggests that large fractions of the I/M fleet are not being identified as high emitters and properly repaired.

Cars that were given three I/M cycles, each roughly one year apart, in Phoenix provide an opportunity to assess the effect annual testing would have on fleet emissions. The data indicate that the deterioration rate between the second and third I/M cycles is greater than the deterioration rate between the first and second I/M cycles. This result suggests that our simulation of annual testing overstates the potential additional emission reduction from annual testing; actual additional benefits are likely to be less. In contrast, comparison of remote sensing measurements in Phoenix of the Phoenix and Tucson I/M fleets indicates that newer vehicles have lower emissions in the Phoenix biennial IM240 program than the Tucson annual idle program, but that older vehicles in Phoenix have much higher emissions than their Tucson counterparts. More frequent I/M testing, and repair, in the Tucson I/M program is a likely explanation for why older vehicles have lower emissions in Tucson than in Phoenix.

Since the vast majority of increased emission reductions in the simulation of annual testing are coming from vehicles over eight years old, the combination of more frequent testing of the oldest vehicles combined with tighter cut points could greatly increase the overall program reductions.

The findings from our analyses suggest that test-to-test emissions variability is a major problem of I/M programs, and is preventing them from properly identifying vehicles with broken or malfunctioning emissions controls and ensuring that they are repaired. The only improvement to I/M programs that can potentially address this problem is to require more frequent testing of suspected high emitters. However, our analysis of a small number of cars tested every year suggests that more research is needed to better estimate the potential benefit from annual testing.

Another strategy is to use remote sensing to identify suspected high emitters, and require that they come in for off-cycle testing, as frequently as necessary. This strategy may succeed not by requiring one or two additional off-cycle tests, but by providing enough of a nuisance to owners of problem vehicles so that they remove them from the I/M area or ensure that they are properly repaired. However, if enforcement of an I/M program is weak (as appears to be the case in Phoenix), this strategy runs the risk of merely inducing these vehicle owners to find ways to avoid program requirements, unless enforcement is improved. Some vehicle owners may learn how to avoid detection by remote sensors, either by obstructing license plates or by avoiding locations where remote sensors are sited. On the other hand, a significant remote sensing

presence also provides the means for making program avoidance more difficult if sufficient penalties can be levied to make enforcement worthwhile for the agencies involved.

We believe more research is necessary to better understand the Phoenix multi-cycle fleet, and whether it is representative of the overall Phoenix fleet. In addition, more research is needed on the test-to-test variability in vehicle emissions. Similar analyses of multi-cycle fleets in other I/M programs would help determine if the high failure rates shortly after passing an I/M cycle are unique to the Phoenix program, or characteristic of I/M programs in general. And tracking of individual vehicles over three biennial cycles, as well as matching more remote sensing data with multiple biennial cycle test results, would reveal more information about the long-term benefit of the Phoenix I/M programs. Finally, these analyses could be applied in other states to determine whether the problems uncovered in Phoenix are found in other programs.

References

Acurex Environmental Corporation. (1997) *Update of Fleet Characterization Data for Use in MOBILE6*. Prepared for US EPA under contract 68-C6-0068. Acurex Environmental report 97-1050. May 16.

Ando, Amy W., Winston Harrington, and Virginia McConnell. 1999. "Estimating Full IM240 Emissions from Partial Test Results: Evidence from Arizona." *J. Air & Waste Manage. Assoc.* **49**:1153-1167. October.

California Air Resources Board. 2000. *Evaluation of California's Enhanced Vehicle Inspection and Maintenance Program (Smog Check II)*.

Wenzel, Tom, Brett Singer, Etan Gumerman, Robert Slott, and Robert Sawyer. 2000. *Section III: Evaluation of the Enhanced Smog Check Program: A Report to the California Inspection and Maintenance Review Committee*. Final report submitted to the California Inspection and Maintenance Review Committee, June 19. LBNL-46117

Wenzel, Tom. "Evaluating the Long-Term Effectiveness of the Phoenix IM240 Program." *Environmental Science and Policy*, in press.