On-Road Evaluation of an Automobile Emission Test Program

DONALD H. STEDMAN* AND GARY A. BISHOP
University of Denver, Denver, Colorado 80208-2436

PHILIP ALDRETE
Aurora Central High School, 11700 East 11th Avenue, Aurora, Colorado 80010

ROBERT S. SLOTT
Shell Oil Company, P.O. Box 2463, Houston, Texas 77252

In January 1995, the State of Colorado superseded an annual decentralized idle test program with an EPA recommended biennial centralized IM240 test program in the six-county metro Denver area. Odd model year 1982–1994 vehicles and any vehicles that entered the counties or changed ownership required testing in 1995. Even model year vehicles were to be tested in 1996. Five days of on-road remote sensing exhaust measurements in winter 1995–1996 provided 29,255 identified, unique, gasoline-fueled vehicle emission measurements to independently evaluate the marginal benefit of the 1-year-old inspection and maintenance program. Eligible, untested even model years form the control group to compare to tested odd model years. Benefits were significant for CO (between 4% ± 2% and 7% ± 2%). Benefits for HC and NO were not detectable. Error bars are obtained using each day as an independent analysis. An effect not observable by other methods shows odd model year cars outside the program jurisdiction with significantly higher average CO emissions than their even model year peers, possibly indicating out of registration for failing odd model year vehicles, thus avoiding the new emission test.

Introduction

The 1990 Clean Air Act Amendments were interpreted by the EPA for vehicle exhaust emission tests as requiring biennial centralized IM240 testing on 1982 and newer vehicles (7). This IM240 test (IM stands for inspection and maintenance, and the test was originally designed to be 240 s long) is an abbreviated version of the third phase of the urban dynamometer driving schedule contained within the Federal Test Procedure (the new car certification test) and has been engineered with the intent to provide comparable “warm” operation emission measurements to the longer, more expensive test. The exhaust gas sampling can occur for as long as 240 s, although many vehicles “fast pass” in 30 s or more. In 40 CFR Part 51, Vehicle Inspection Requirements for State Implementation Plans, the EPA states that IM240 “more closely reflects how vehicles perform under actual driving conditions than do current idle, 2500 rpm/idle, or loaded steady-state emission tests”. Some states chose not to implement the IM240 program, and in 1995, the U.S. Congress passed the National Highway Safety Designation Act allowing states to establish alternate I/M programs (2).

However, the states were required to conduct evaluations of I/M program effectiveness using any of the following “...the performance of any previous automobile emissions inspection and maintenance programs; the results of remote sensing or other roadside testing techniques; fleet and VMT profiles ...”.

In January 1995, the State of Colorado superseded an annual decentralized idle test program with the EPA recommended program in six counties (Adams, Arapahoe, Boulder, Denver, Douglas, and Jefferson). Colorado HB93-1340, May 1993, established the new emission testing program starting in January 1995. The state required IM240 testing in 1995 of odd model year 1982–1994 vehicles and any vehicle that entered the counties or changed ownership. Remaining vehicles were scheduled for testing in 1996. New vehicles purchased in 1994 before October 1 were required to be tested in 1995. New vehicles purchased after this date received a 4-year testing exemption. All 1981 and older vehicles receive annual idle tests in test-only facilities. Colorado’s IM240 program utilizes the EPA recommended high altitude phase-in emissions standards that can be interpreted to predict a failure rate of approximately 20% of tested vehicles and a program benefit for CO of approximately 23% (3). Vehicles are failed based on CO, HC, and visual inspections only. NOx tests are conducted, but the results are treated as “advisory”. Pressure/purge testing of the fuel system and on-board diagnostic interrogation are not conducted. The test costs $24.25, and the failure rate is reported to be approximately 6%, which translates into about $400 to identify each failed vehicle. In excess of 800,000 vehicles were tested in 1995 (4).

Carbon monoxide emission reductions have been estimated by the State of Colorado at 20–23%. Colorado Department of Public Health and Environment (CDPHE) Press Release November 1995 states “Carbon monoxide emissions are being reduced by 20.0%”; by December 14, 1995, the benefit “...confirmed a carbon monoxide reduction of 23 percent” (5). This claim was repeated in a report by Klausmeier prepared for CDPHE and in an annual report to the State Legislature (4, 6). The benefits were estimated by the state from the average emissions by model year (MY) of vehicles that first fail the test minus the average emissions by MY of retested failed vehicles (not always the same vehicles) that subsequently pass the test. The observed average emissions for failing light-duty gasoline-powered vehicles were 61, 3.2, and 1.1 g/mi for CO, HC, and NO, respectively (6). Upon passing a retest, these numbers became 16, 1.1, and 1.5 g/mi. When 1982–1994 gasoline-powered truck data are included, the overall percentage benefits are 23%, 18%, and 12% for CO, HC, and NO, respectively (6). These data compare vehicles eligible for the Colorado program before and after IM240 testing and provide the basis for comparison with our on-road emission measurements.

It has been pointed out that the state’s analysis is subject to several sources of overestimation (7, 8). (a) By not using paired data, the database contains many vehicles (47% of “first fail” vehicles) that initially fail the inspection and are not accounted for as ever passing the inspection, even though they may continue driving in the Denver area. (b) Analyses that focus on subsets of a distribution (all first fail vehicles with high emissions) are subject to regression toward the mean (9). High emitters that fail their first test but pass a retest without effective repair and high emitters that pass on their first test all cause overestimation of the benefit (10, 11). According to Radian Corp., Arizona officials carrying out IM240 tests "...have found that CO emissions can vary by almost an order of magnitude from test to test, even though the driver maintains the speed within the allowable 2 m.p.h.
tolerance" (12). (c) Scheduled maintenance, which effectively reduces emissions, will be counted as a program benefit though not initiated by the I/M program.

An independent method proposed to measure I/M program effectiveness is to use roadside remote sensing (13). Since much of the benefit of a new I/M program is expected to occur in the first year, the end of the first year of a biennial program offers a unique opportunity to assess the marginal decrease in on-road emissions resulting from the new control program (4).

I/M Evaluation by Means of On-Road Emissions Monitoring. On-road Remote sensing was developed in Denver. The methods and policy implications have been described elsewhere (14-18). On-road monitoring has been used to assess I/M program emission benefits within a program area (12, 15) and by making separate measurements inside and outside of I/M areas. Each have drawbacks. Separate measurement sites must be very carefully calibrated. Measurements at one site avoid any calibration problem. However, both methods are unable to distinguish causes of reduced emissions in I/M areas. Lower emissions could be due to successful I/M repairs following inspection failure, owner behavior, or some uncontrolled factor. Owner behavior can be constructive (repair prior to testing) or counterpro- ductive (registering gross polluters outside the I/M area or carrying out temporary repairs to "pass the test"). Other factors can cause vehicles of the same age to differ in their emissions. Vehicles of the same age, in the same I/M program, show lower average emissions when the owners live in higher income areas (19, 20).

In view of these problems, we have taken advantage of an unique opportunity 12 months into the Colorado biennial I/M program. Untested even MY vehicles could be used as the control group against which tested odd MY emissions are compared. Measurements of on-road emissions at a single site avoids any possible calibration effects. The untested even MY control group should have the same demographic factors as the tested odd MY group.

An on-road remote sensing study at a single site directly tests I/M program benefits observable at that site. If an I/M program is testing warmed-up emissions of on-road vehicles, then its on-road effectiveness should be independent of measurement site provided that warmed-up vehicles are measured. To the extent that the fleet average on-road data correlate with fleet average IM240 data, then an on-road study evaluates the overall program percentage benefits as effectively as an IM240-based study. The observed correlations are discussed in a later section. We strongly recommend this independent evaluation method to any state embarking on biennial programs because of the unique availability of a built-in statistically valid control group and the ability to observe the "real" on-road fleet of vehicles, not just the ones that choose to show up at test stations.

Results and Discussion

Measurements were made for 5 days on the exit ramp from northbound I-25 to westbound US6. This exit is a tightly curved, uphill ramp (8% grade) with an average measured speed and acceleration of 22 ± 4.5 mph and -0.2 ± 1.3 mph/s. Table 1 shows a summary of all the emissions data obtained. The instrument was calibrated using a certified cylinder containing CO, CO₂, propane, and NO in nitrogen (Scott Specialty Gases, Longmont, CO).

License plates for all vehicles identified as Colorado registered were matched by the Colorado DMV for make, MY, fuel type, emission test eligibility, emission test status, and date of test. The database is available from the authors. It contains information on 26 255 identified, unique, non-diesel-fueled vehicles. Of these, 19 011 have DMV records that indicate they are eligible for IM240 testing. Automobile emissions are γ-distributed, in this data 10% of the vehicles produce 63% of the measured on-road CO emissions (in g/gal of fuel used) (21).

Comparison with IM240 Data. I/M program evaluation has been proposed using measured and modeled IM240 emissions. The model used is MOBILE, the U.S. EPA Mobile Source Division computer model of mobile source emissions (current version MOBILE5a), which states are required to use in evaluating their I/M programs for their State Implementation Plans. The model, including a 50% discount applicable to I/M programs that combine testing and repairs in one facility, assumes that all vehicles of a given age travel the same annual mileage (22).

In order to show that remote sensing data can be as effective as IM240 data to monitor control strategy effectiveness, a correlation between remote sensing measurements and IM240 tests is required along with an evaluation of how well the remote sensing data can detect fleet emission differences.

The fleet sampled by remote sensing was operating under a range of speeds and loads contained in the current IM240 test, but with a much narrower distribution. For both tests, all vehicles are in a warmed-up operating condition. The State of Colorado collected a database (NONFAST.DAT) of 36 114 eligible odd and change of ownership even model year vehicles (average MY 1989.65) that took the whole 240-s test. The 10% highest emitters cause 49% of the measured IM240 CO emissions.

We compared the Colorado database to a remote sensing fleet of IM240 eligible odd and even model year vehicles (19 011; average MY 1989.56). For CO and HC, both data sets show low, slowly changing emissions from newer MY cars and trucks and larger, more steeply changing emissions with increasing age from vehicles 10-14 years old. For NO, both data sets show nearly linear increases with increasing age. Correlation of average fleet remote sensing data by MY against the NONFAST.DAT g/mi data is shown in Figure 1. The r² values are 0.96, 0.37, and 0.37 for CO, HC, and NO, respectively.

If the remote sensing data are converted to g/mi by multiplication by estimated fuel economy (by model year using CAFE standards), the correlations are essentially unchanged largely because fuel economy estimates vary only slightly (26-29 mpg) between 1982 and 1994 MY. The slopes for these correlations are not 1:1, and especially for HC, the remote sensing data reflect emissions from only the major hydrocarbon species (alkanes) found in gasoline (23, 24).

The high level of fleet average correlation by model year leads us to believe that the limited range of driving modes in the on-road emission study can be used to represent the relative change in the emissions benefits as measured by the IM240 test. We are most certain of the relationship for CO. For warm vehicles, driving mode makes little difference for CO as long as hard acceleration is avoided (20, 25). HC and NO emissions can be greatly affected by load and driving mode. We therefore would expect HC emissions to be minimized and NO emissions to be maximized at this loaded
mode site. In both cases, the observed linear correlations are excellent between on-road and IM240 averages by MY.

The ability of remote sensing data to distinguish emissions differences was explored by analyzing the differences in emissions for two DMV categories of vehicles. In Colorado, light-duty pickup trucks are given a separate registration designation (LTK) from the rest of the light-duty fleet (PAS). Figure 2 shows a PAS/LTK difference of 25 ± 9% based upon a total fleet of 15,813 PAS and 3,129 LTK. Error bars are standard error of the fleet mean and are estimated by treating each of the five measurement days as an independent sample. A t-test shows that this difference is statistically different from zero at better than the 1% level. When the impact of sample size is tested by randomly reducing the size of the truck fleet to one-fifth (625 records), we continue to find a statistically significant difference at the 5% level (two-tailed t-test). The random reduction in sample size was carried out by means of a random number generator and separately by taking each individual day’s data. The results were the same in both cases. The IM240 test protocol distinguishes three categories: light-duty gasoline vehicle and light-duty gasoline trucks 1 and 2, with different failure cut-points (6). The fact that these categories are not the same as the DMV categories does not in any way alter the above analysis. This analysis of the PAS/LTK emission differences shows that program benefits on the order of 20% should be observable even for small control fleets.

Program Benefit Analysis. Figure 3 is based upon a subset of the 19,011 (1982–1994 MY) on-road vehicle emissions for unique, gasoline-powered vehicles registered in counties requiring IM240 testing. To evaluate the program benefit, we show only odd MY vehicles with evidence that they had been tested in 1995 (before the on-road test (7906 records, filled circles) to compare with even MYs that definitely had not been tested in 1995 (6097 records, open circles). Qualitatively the program benefit should show up in Figure 3 as an up and down alternation between the untested (open circles) model years and the tested (filled circles). Only model years 1995 and 1987 show this effect. These two model years have IM240 failure rates in NONFAST.DAT of 20% and 8%, respectively. However, 1983 has a failure rate of 13% yet has emissions similar to untested 1982 and 1984 vehicles.

In order quantitatively to evaluate the program benefits, an age-adjusted control group was created by averaging the emissions of the two even model years that surround each odd MY, thus the average emissions of untested 1982 and 1984 MY vehicles was compared to the measured emissions of the 1983 MY fleet. Two calculations of the apparent benefit were carried out. The differences between the six average untested even MY pairs and the six tested odd MY were determined. And the averages were weighted by the total number of vehicles measured by model year. The weighting function did not significantly alter the result from the unweighted percentages. A 6.7% unweighted average % CO benefit became a 6.9% benefit when thus weighted.

In order to determine the precision of this estimate, the benefits were calculated, as above, for each of the five measurement days as an independent sample. Thus, the five mean benefits provide an overall sample mean and a standard error of the mean. The results of these analysis for CO, HC, and NO are shown in the first row of Table 2.

In the process of this analysis, we noticed that about 11% (969 vehicles) of the odd MY cars that were registered in the six I/M counties apparently had managed to be registered without passing the test in 1995. Twenty-eight had received waivers, and three were listed as failing. Some of the others had benefited from start-up problems in January and February of 1995 and been allowed to register without testing, some apparently had declared to the DMV that the vehicle was not in Colorado at the time of registration, and some had returned 1995 plates (as of April 1996). These vehicles had higher average CO emissions than their tested brethren for every MY except 1993. We reasoned that the untested even MY cars had no incentive to cheat on a test they had not yet been
This accidental observation is another reason why states with biennial programs should consider on-road evaluation at the end of the first year. Without on-road data, the apparent effect of the program on the ineligible fleet, observable from the odd—even alternation, would have been undetectable. This also leads to the prediction that at the end of 1996 there will be no detectable alternation in the on-road emissions of the ineligible fleet, because the owners of the even MY vehicles will have been subjected to the same motivations as the owners of the odd MY vehicles.

Because of change of ownership (and 1994 registrations), many even MY vehicles were actually tested in 1995. Comparing the on-road CO emissions of the tested to the untested vehicles leads to an apparent program benefit similar to the other comparisons (Table 2, fourth row).

Another independent analysis method compares the on-road emissions of vehicles tested during the four months before a measurement day with emissions from vehicles tested during the two months after (26). This analysis (Table 2, last row) leads to an apparent benefit of 8 ± 6% for CO. One would expect this result to be comparable to the first row of Table 2 because it compares on-road emissions of definitely IM240-tested vehicles to the control group of vehicles, which were IM240 tested in the two subsequent months.

All of the various analyses point to a modest benefit for CO emissions as a result of the new program. The measured hydrocarbon emissions benefit is smaller than for CO, and the program may be increasing NOx [as forecast by Coninx (27) and NRC (28, 29)]. The benefits observed in this analysis are consistent with earlier analyses by Slott (30) and Lawson et al. (31).

For the 1981 and older vehicles, eligible vehicles are tested annually at test-only facilities. A comparison of the average on-road CO emissions of 1581 tested 1981 and older MY vehicles (1.85 ± 0.07% CO) to 321 not tested vehicles of the same age (1.86 ± 0.16% CO) results in a statistically insignificant difference. HC and NO differences are also insignificant for these older vehicles.

Acknowledgments
We would like to thank various members of the Phillipson Group at the University of Denver who were helpful in this study, particularly Mr. Will McHale, Ms. Annette Bishop, and Ms. Beth Reeves.

Literature Cited


(7) Sherritt, J. Visible Smoke, an Analysis of an Analysis; Privately published report, Denver, CO, 1996.


(27) Coninx, P. Specific Scientific and Technical Issues Relating to the Effectiveness of I/M Programs; Prepared for the Automobile Protection Association, Toronto, Ontario, 1996.


Received for review September 16, 1996. Revised manuscript received December 13, 1996. Accepted December 16, 1996.*

ES960792Q