A Behavioral Analysis of EPA's MOBILE Emission Factor Model

Winston Harrington
Virginia D. McConnell
Matthew Cannon

Discussion Paper 98-47

September 1998

RESOURCES FOR THE FUTURE

1616 P Street, NW
Washington, DC 20036
Telephone 202-328-5000
Fax 202-939-3460

A Behavioral Analysis of EPA's MOBILE Emission Factor Model

Winston Harrington, Virginia D. McConnell, and Matthew Cannon

Abstract

This paper examines the behavioral and stochastic aspects of modeling emission reductions from vehicle Inspection and Maintenance (I/M) programs. Forecasts of the potential emission reductions from such programs have been modeled by the use of the Environmental Protection Agency's MOBILE Model, EPA's computer model for estimating emission factors for mobile sources. We examine the structure of this Model and review the way behavior of drivers, mechanics and state regulatory authorities is incorporated in the current generation of the Model. We focus particularly on assumptions about vehicle repair under I/M, compliance with I/M requirements, and the impact of test measurement error on predicted I/M effectiveness. We also include some preliminary comparisons of the Model's outcomes to results of the I/M program in place in Arizona. Finally, we perform some sensitivity analyses to determine the most influential underlying parameters of the Model.

We find that many of the assumptions of the I/M component of the Model are based on relatively small data sets on vehicle done in a laboratory setting, and that the output of the Model makes it difficult to compare the results against real world data from on-going state programs. In addition, the Model assumes that vehicles will either be repaired or receive a waiver. In the Arizona program there appears to be a third category of vehicles -- those which fail the test and do not receive passes. This share may be as high as a third of all failing vehicles. Vehicles which do not eventually pass the test would be treated in the Model as non-compliant. However, in current programs, states do not seem to be measuring and entering the compliance rate correctly. The paper also examines the evidence about whether emissions deteriorate over the life of vehicle in a grams per mile basis (as assumed by the Model) or a grams per gallon basis. It finds support for the argument that emissions deteriorate on a grams per gallon basis.

We find through sensitivity analysis that the repair effectiveness assumed by the Model to occur in an IM240 test are much greater than for the idle test, and that identification rates and repair effectiveness vary a great deal according to the cutpoint. These results are based on small numbers of vehicle tests in a laboratory setting and could be compared to real world evidence. Examining costs and cost-effectiveness of variations in I/M programs is important for determining improvements in I/M programs. States may not have incentives to develop cost-effective programs based on current Model that forecast emission reduction "credits" that are overly optimistic.

Key Words: emission reduction, vehicle inspection and maintenance program

Acknowledgments

This work was partially supported by a grant from the Federal Highway Administration. We would like to thank Michael Savonis, Kirk Fauver, John Byun, Jerry Everett and Dick Schoeneberg of FHWA for their support and encouragement, without which the research would not have been possible. The work also would not have been possible without the cooperation of EPA's Office of Mobile Sources in Ann Arbor, Michigan. We are especially grateful to Dave Brzezinski of OMS, who was enormously helpful in answering our questions and guiding us through the TECH and MOBILE models. We would also like to thank Bob Slott of Shell Oil Co., Dave Brzezinski, various folks from the FHWA, and our RFF colleagues Amy Ando and Alan Krupnick for their very helpful comments on a previous draft. The usual disclaimer applies.

Table of Contents

List of Figures and Tables

Executive Summary

This report provides a review and assessment of several important aspects of the MOBILE Model, EPA's computer model for estimating emission factors for mobile sources. Inventory models like MOBILE have many uses, but we focus primarily on the Model's role for estimation of emission reduction credits from I/M programs. Scrutiny of the I/M component is important for a number of reasons. First, there has been little discussion or oversight of this part of the Model because much of its structure and underlying assumptions are not evident to users. The I/M assumptions are imbedded in the program code in the TECH component of the model, and are opaque to the user. One purpose of this report, therefore, is to describe the structure and assumptions underlying the I/M Model. As part of this assessment, we focus on how the model incorporates behavioral responses to as well as technical aspects of I/M processes.

Second, there is growing concern over the way I/M credits are granted through the Model. States input a handful of parameters that characterize their fleet, region and program, and obtain I/M emission reduction credits into the future. This process makes it virtually unnecessary for users to examine the underlying assumptions in the Model or to try to benchmark actual reductions to those being forecast by the Model. As part of our analysis, then, we point out the importance of using evidence from on-going programs to inform and to validate the Model. We do some preliminary comparisons of the Model underlying assumptions and results to available evidence from other studies and I/M programs. Our work suggests that there may be a gap between emission reduction credits generated by the Model and the actual performance of I/M programs in the field. Although we have some preliminary suggestions for modifications to the Model, the larger issue of whether to revise the process of granting credits for I/M and other related emission reduction programs will need to take place among policy makers over a period of time.

This is an opportune time to examine the MOBILE Model and the determination of I/M credits. As a result of the National Highway System Designation Act, states are now considering ways to estimate their own emission credits for I/M programs. How this should be done, and the role of the Mobile Model in the process are still undecided. In addition, EPA is now reaching out to the "stakeholder" community for advice on the model. EPA's Office of Mobile Sources (OMS) is now at work on Version 6 of MOBILE, with the assistance of a technical subcommittee of the Clean Air Act Federal Advisory Committee. Also, this report ties into others that have been done recently. Within the past year, for example, review of various aspects of the MOBILE model have been prepared by the National Academy of Sciences (NAS, 1997) and the Government Accounting Office (GAO, 1997).

Our work differs from these other efforts in several ways. While other studies have focused on the inventory uses and the importance of emission correction factors in the Model, our main intent is to explain and evaluate the I/M structure and assumptions in the TECH component of the Model. In its origin and intent our work is perhaps most similar to a report prepared for the Federal Highway Administration not long ago by Sierra Research

Corporation (1994). That report provided much-needed documentation of the MOBILE Model, describing how it worked and documenting all the calculations used to generate the emission factors in the model. Unlike that study, which focused on the technical aspects, we pay particular attention to the ways behavior is included in the Model, and how it may be left out. In particular, we believe that the effectiveness of I/M programs in practice are likely to be strongly influenced by the behavior of motorists, mechanics and even state regulatory authorities. For example, we examine repair effectiveness assumptions, the definition and use of compliance rates, and the incorporation of tampering behavior in the Model. In addition we perform sensitivity analyses to determine how the Model responds to changes in different parameters, comparing the responses to changes in the behavioral parameters to changes in some of the more technical parameters.

As part of that analysis, we try to compare the predictions of the Model to evidence from programs in operation. However, while we are interested in empirical validation of the model, we do not, as others have done, compare aggregate emission estimates produced by the model with real-world emission estimates (derived from tunnel studies and other on-road studies). In this study we are more interested in the empirical validation of the assumptions and data used to generate Model outcomes. To do this we have concentrated on the I/M-related performance measures such as repair effectiveness, failure rates, numbers of retests, repair cost, and compliance measures. We compare empirical results from the Arizona I/M program, from EPA laboratory data, and from other studies of repair to the assumptions in the Model.

Behavioral and stochastic influences on vehicle emissions

To the extent that the Model does not account for behavior, it may only be able to forecast emissions reductions that could occur under some ideal setting, and not reflect the reductions that occur in practice. If the forecast reductions are too optimistic, other policies either within or outside the I/M program that have the potential to improve emission reductions may not be adopted.

Some of the behavioral aspects of emission factors, we find, affect the inputs to the MOBILE model and therefore can be accommodated in the existing MOBILE model structure. For example, with suitably modified inputs, MOBILE could be used to analyze the emission effects of policies that change the relative prices of vehicles of different types or different ages. But although some behavioral responses could be incorporated in model estimates, we find that very often they are not. States often find it easier and more advantageous to use default values of important inputs or parameters, like the age distribution of the fleet and the compliance rate, rather than using their own fleet age distribution or measuring actual compliance with their program. They may in fact be discouraged in some cases from using parameters based on their own analyses. For example, policies that would make the age distribution of the fleet newer over time are not allowed.

Other behavioral effects cannot be handled by changes in MOBILE inputs. Some of these involve behavioral assumptions that are made implicitly and hard-coded in MOBILE or in the TECH submodel, such as the repair effectiveness assumptions. In addition, there are behavioral aspects that cannot be dealt with at all given the current structure of the model. Most of these cases of hard-coded data and structural assumptions involve the estimation of I/M credits. To the extent that these issues cause the Model to be inaccurate, I/M credits generated from it will not reflect actual emission reductions from on-going programs.

We also examine the stochastic elements of vehicle emissions, primarily because the existence of uncertainty can have behavioral consequences. Probably the most important instance is the variation in emission test results--the fact that successive emission tests on the same vehicle can have different results, often dramatically so. MOBILE does not model such uncertainty explicitly, but it enters the model implicitly in the definition of "identification rates," which refer to the ability of emission tests to identify vehicles in need of emission repair. EPA's Office of Mobile Sources (OMS) asserts that all variation in emission test results is attributable to the test itself, and that the "true" emissions of the vehicle are constant. Researchers outside EPA have found that large variations in emissions are characteristic of some vehicles, which may have malfunctioning emission control systems but may nevertheless be fortunate enough to pass a single emission test. If such vehicles are not rare the implications for I/M programs are obvious and unsettling. We believe research into the sources of emission test variability is therefore warranted.

Summary of findings on I/M

We review the way the I/M component is handled in the model in some detail, identifying three areas where the behavior of motorists and mechanics are likely to be important--repair effectiveness, compliance and tampering. We examine how each of these are handled in either the TECH or MOBILE Models. In all of these areas, the underlying assumptions in the Model are not evident to the user, but are often embedded in the program code. We tried to identify the various assumptions, and when possible, describe the sources of the underlying data on which the assumptions are based. Then, because evidence from actual I/M programs reflect all aspects of the I/M process including the underlying behavior of motorists and mechanics, we summarize actual data from the Arizona I/M program and other repair studies and compare them to the EPA results.

Repair Effectiveness. The Model assumes that repair effectiveness under an I/M program depends solely on the test cutpoints and on the measurement method or type of test. For the IM240 test, the repair effectiveness factors are based on 266 vehicles which were repaired at EPA laboratories during the 1980s and early 1990s. Emissions were measured before and after repair using the IM240 test. In addition, emission tests using the Federal Test Procedure (FTP) were administered, in order to obtain the "true" emissions and emission reductions for each vehicle in the sample. Two procedures were used to get the emission reduction estimates for a given set of cutpoints. For those vehicles which passed the IM240 test after repair, the emission improvements were taken from the reductions in emissions determined on the FTP test. Otherwise, IM240 test results were extrapolated to the point where the test was passed and then the IM240 reading was converted to FTP results.

In other words, the actual repair data were used only when the vehicle was brought into compliance; when the vehicle was not successfully repaired (approximately 30% percent of the time in the EPA repair dataset), it was assumed that compliance would follow further repair. Although it was understood that some vehicles would not be able to be repaired, it was assumed that all such non-complying vehicles would be required to get an emission waiver, as specified in the Clean Air Act. Waivers were handled in an aggregate fashion elsewhere in the model. All of these assumptions are built into the TECH code, and not possible for the user to change. The vehicles are classified into "emitter categories" and average emission reductions are reported for each emitter category.

To examine the repair issue further, we obtained the EPA repair dataset for analysis and comparison with a number of other datasets collected in the field and providing evidence of repair cost and repair effectiveness that might be encountered by a motorist subject to an actual state I/M program. These other datasets include the following: Arizona's IM240 program results from 1995-96[1], California's I/M program results from 1992 (the "Pilot Project"), and a study of repairing gross-emitting vehicles conducted by the Sun Oil Company in 1994. Although the studies are based on different vehicle fleets, and different repair standards some of the comparisons are revealing.[2]

- The TECH Model repair effectiveness assumptions are based on evidence from small samples of vehicles repaired under EPA laboratory conditions. These assumptions are unlikely to capture accurately the repair effectiveness from on-going I/M programs. Available evidence suggests that repair may be considerably less effective than assumed by EPA, and that gross-emitting vehicles are particularly difficult to bring into compliance. The fact that cars do not get fully repaired in real world programs bears further scrutiny (see compliance section below).
- Repair costs are important to include in any assessment of repair effectiveness. A review of existing studies of the cost of repair shows that these costs range from an average of $330 in the Sun Oil Company study (which included a number of older vehicles) to about $175 in the Arizona program.[3] However, in the Arizona program, many vehicles were having some difficulty passing the test so these costs are not the cost of achieving compliance. Repair costs will influence not only the

[1] We use the 2% random sample of vehicles in the Arizona program that were given full IM240 tests in the first 5 months of 1995.

[2] The comparison to the EPA dataset must be done with caution because the EPA repair effects are supposed to capture the full impact of I/M compared to no-I/M, whereas the repair effects observed in some regions which have had on-going I/M may have smaller reductions in any one time period.

[3] EPA estimates of the average repair costs for a cost-effectiveness assessment of enhanced I/M was between $75 and $120 (EPA, 1991).

cost-effectiveness of I/M programs, but driver and mechanic behavior toward repair.

- As stated above, MOBILE assumes, as does the Clean Air Act, that all vehicles unable to pass the emission test will receive an emission waiver. In real I/M programs there appears to be a third category. In Arizona, it appears that not all vehicles that were unable to pass the test received the emission waiver. No one knows whether these vehicles were scrapped or sold elsewhere, or continue to be operated in the Phoenix area. It is difficult to tell exactly how many vehicles are in this category because waiver information is not available from the Arizona program. In Colorado, a remote sensing study has established that at least some of the "disappearing" vehicles are still being operated in the Denver area. This issue needs to be examined. This problem can dealt with in the Model as affecting repair effectiveness, or it can be addressed more as a compliance issue.

Compliance. This raises the important issue of how compliance rates are defined and used in the Model. In this paper, we examine two important issues about how compliance is handled in the Model. The first has to do with how emissions reductions are discounted as a result of non-complying vehicles. The second has to do with how the compliance rate is used in practice by the states. First, we review how the compliance discount is calculated. Emissions reductions resulting from I/M, as determined by the TECH Model (including repair effectiveness assumptions as discussed above), are brought into MOBILE and then reduced according to the share of non-complying vehicles in the fleet. The Model includes an emission reduction adjustment for non-compliance that is non-linear (the initial non-complying vehicles will be somewhat dirtier than the average vehicle), but the adjustment is not very large. It assumes the non-complying vehicles will be somewhat more likely to fail than the average vehicle in the fleet. The evidence from Arizona shows that many of the *failing* vehicles are the non-complying vehicles, so the emissions adjustment for non-compliance should be quite large. For example, if one quarter of the failing vehicles are not getting fully repaired, this may only be 4% of the fleet, but it represents up to 25% of the potential emissions reductions from repair under the I/M program.

The second and related problem is the way the compliance rate is actually used as an input to the Model by the states. States are not required to actually measure the compliance rate and compare the measured value to what they are using as an input to the Model. Many states use the default value currently set at 4% of the fleet, which results in a loss of emission reduction of about 8% (compared to emission reduction with 100% compliance). We discussed above that the 8% is probably too low an estimate of the emission reduction foregone. But use of the default value for compliance means that states do not have to attempt to measure what their compliance rate really is.

Compliance is difficult both to define and to measure. The Arizona data show that the extent of non-compliance may be quite large. We find that 25-35% (depending on the data sample) of the vehicles do not have passing retests after they fail. But some of this could be data collection or coding errors, and there are other possibilities as well. Vehicles could be receiving waivers, they could be scrapped, moved out of the region or be driven illegally in the region. The data also suggest that of the failing vehicles there are some that may be difficult to repair to the standards (and the standards are fairly lax in Arizona currently). We argue that the compliance component of the Model should be reconsidered, with a new measure for states to use that more closely reflects the emission reduction performance of the I/M program. Measured compliance could then be compared to the compliance assumptions the states use as input to the Model.

Non-compliance rates are also likely to depend on the type of I/M regime in place, and to vary over time as adjustments are made to the requirements of a program. An I/M program that is more difficult to pass, like the IM240, is likely to result in more non-compliance. And, over time, after the vehicles that have the most difficult time complying either get scrapped or move outside the region, compliance rates may improve. In MOBILE currently, there is a compliance rate that remains in effect for all forecast years, although it is possible that the user could input different compliance rates and run each year separately. In practice, they do not currently do that and would be unlikely to want to do it since they would lose credits for the I/M program.

Tampering. Tampering behavior is explicitly modeled in the Mobile Model in several ways. In the absence of any emission control program, there is assumed to be some base amount of tampering. Tampering can be then be reduced by I/M programs in two ways. The first is a reduction in tampering due to a deterrent effect which occurs just as a result of the presence of an I/M program. Second, I/M testing can include a separate anti-tampering program that specifically checks for the impact of certain types of tampering and reduces it. In the paper, we describe the tampering assumptions in the Model, and do a brief comparison to the evidence of tampering in the Arizona I/M program, the California Pilot Project, and the California I/M program. Tampering rates do seem to be lower in Arizona than in the other programs. However, Arizona records tampering from I/M records, while California

Sensitivity Analyses. We conducted two different sensitivity analyses in order to examine the relative importance of different parameters in MOBILE. The first looks at variations in two different parameters, identification rates and repair effectiveness, which both affect emissions reductions forecasts under different I/M test regimes. The second is a more general sensitivity analysis on a variety of both technical and behavioral parameters to see the relative impact of each. In order to perform sensitivity analyses on parameters contained in the TECH model, such as repair effectiveness, we modified the TECH model inputs and re-compiled the program. The resulting output was used as input for the MOBILE model. Analyses of other parameter changes, such as the I/M compliance rate, required changing user-specified inputs into the MOBILE model.

In the first analysis, we examine the effect of variations in repair effectiveness parameters and test identification rate assumptions across different test types. We find that the IM240 test obtains much higher emissions reductions than alternative test regimes like the 2500 idle test. However, the underlying sample of cars used to estimate the relevant parameters for both test regimes was quite small. **Give nos?** It would seem to be important to confirm this result by looking at evidence from the field as that evidence becomes available.

In addition, we find that the identification rates are shown to vary with the IM240 test cutpoint, particularly for the more recent model years. This could have important implications for tightening cutpoints in I/M programs over time. There is some conflicting evidence showing that vehicles close to the standard are difficult to repair, and that sometimes their repair results in higher rather than lower emissions. There has been very little empirical evidence focusing on the impact of tightening cutpoints in on-going I/M programs. Further investigation of this issue would be important for predicting the emissions impacts of tighter cutpoints.

The second sensitivity analysis examines the sensitivity of the emissions reductions for all three pollutants to variation in a range of different parameter and input changes. Many of these changes reflect the model components that have been discussed extensively in the report. Others, such as the absence of an I/M program, and variations in the speed of travel serve as points of comparison. We examine variation in I/M program test type, cutpoint variation, the assumption about the 50% reduction for decentralized programs, repair effectiveness and compliance rate assumptions, and finally, variation in the age distribution of the fleet. We find that variation of many of the parameters or assumptions can have large variation in the emissions predictions from the Model. The results are presented in Figure 10 of the paper. To note a few of the examples, we find that the variation in repair effectiveness and compliance rates over the approximate ranges we observe in Arizona produce substantially higher emissions predictions from the Model. A different age distribution of the fleet, which states are supposed to input to reflect their own regional fleets, can also have a large impact on forecast emission reductions.

Other General Findings

Our review of the MOBILE Model has led us to several more general conclusions.

Model calibration. Efforts at model calibration have been hampered, in our view, by the limits placed by EPA on the what constitutes acceptable data. This is especially true for the parts of the Model used to predict the effectiveness of I/M programs. For the measurement of repair effectiveness, for example, data collection efforts have been limited to studies of repair in EPA or contractor laboratories, without comparison to data available about repair from other analyses. Furthermore, EPA's repair studies may also give misleading indications of what can be expected from vehicle repair because the repairs were conducted in a highly artificial laboratory situation. In addition, for the purpose of setting the basic parameters of the model, the EPA will only consider emission data from FTP tests. As a result, the data sets that have been used by EPA to determine the basic input parameters of the Model are strikingly limited in both their size and scope--that is, the number of observations on which some important assumptions of the Model are based are often quite small, and the definition of what data are considered acceptable is limited. While FTP tests are perhaps the most precise and replicable of available emission tests (and certainly the most expensive), the FTP data sets may suffer from selection bias that could be a more serious problem than the measurement errors of less sophisticated emission tests.

Model validation. MOBILE was not developed and is not used with the idea of making it testable against evidence from the real world. There have been some broad attempts at validating the Model emissions outcomes through speciation studies from ambient air quality models, and by comparison of average emissions in the Model to emissions from tunnel studies. With few exceptions, however, the empirical studies have attempted to test the overall results of the Model; we are aware of only a few studies that have attempted to compare to real-world outcomes the predicted results of the I/M components. While aggregate emission measurement is the "gold standard," there are in fact a number of other ways the Model could be compared to real world results. This includes comparison of assumed failure rates to actual failure rates, and shares of vehicles in different emitter categories assumed in the model to the actual distribution, to name only a few.

Model structure. The I/M component of MOBILE is a *static* model. The user plugs in the user inputs and MOBILE generates I/M credits for each year over the planning horizon, regardless of other local mobile-source policies and regardless of how many years an I/M program has been in effect. More useful, we believe, is a *dynamic* approach, most likely a stock-flow model capable of simulating the deterioration and repair of vehicles of different ages and emission rates. A dynamic model would also be able to endogenize, to the extent that is considered appropriate or is supported by available data, the behavioral effects we discuss in this report. Such a model would have several advantages over the current model. Since it attempts to simulate the actual performance over time of I/M programs, it can generate hypotheses about the details of I/M programs that are assumed away in the current model: hypotheses concerning test failure rates, repair effectiveness, repair duration, identification rate and compliance. In addition, by running it under a variety of scenarios it can suggest which parameters most affect the results, thus providing a blueprint for data collection.

SIP Credits. Empirical validation of the model is particularly important in view of the current practice of basing approval of State Implementation Plans (SIPs) on outcomes generated by MOBILE. SIP approval depends on having credits equal the required emission reductions, which drives local air quality planners to compare alternative policies on the basis of the emission credits generated by the Model. Any policy that is not recognized in the emission model as reducing emissions does not generate credits and therefore is unlikely to receive much consideration. That is, the use of MOBILE in making judgments about attainment almost makes it inevitable that it will be used in making policy comparisons as well. It no longer matters what will happen in fact, but what MOBILE says will happen. Thus the EPA-approved model becomes the reality.

Even when it generates emission credits, a potentially useful program can be "crowded out" by inspection and maintenance programs, if its emission reductions overlap with the emission reductions that are credited to I/M by MOBILE. Crowding out results because

(a) Enhanced I/M is required in regions that are in serious nonattainment for mobile-source pollutants, and (b) MOBILE's estimates of the emission reductions achievable from enhanced I/M are very large. Once the emission reductions calculated by MOBILE are credited to Enhanced I/M, the maximum emission reductions available to any competing policy may be quite small. If I/M turns out to be roughly as effective as the MOBILE model predicts, then the problem is minor. On the other hand, if I/M is not as effective as EPA and MOBILE anticipate, then not only will resources be wasted on I/M programs, but potentially useful opportunities could be missed.

Fuel economy and emission deterioration. Finally, we wish to bring to readers' attention another research finding that concerns the emission deterioration rates in the model. At present the main issue with the deterioration rates is the existence of a "kink," that is, whether emission deterioration rates are higher for older vehicles. We suggest that an equally important issue may be the form of the rates. At present they are assumed to be constant in terms of grams of pollutant per mile for vehicles of a certain age. However, we have found that for HC and CO at least, emission rates deteriorate at a constant rate in terms of grams per gallon of fuel used. Thus, emissions of vehicles with better fuel economy deteriorate more slowly in gram-per-mile terms.

Table 1 summarizes our major findings and recommendations.

Executive Summary - Table I

Components of the Mobile Model	Issue	MOBILE5	Comment	Recommendation
Repair effectiveness	Repair data may not accurately reflect repair in the real world. There is some evidence that the EPA assumptions are optimistic about how much repair will be achieved in practice.	MOBILE5 is based on FTP data from 266 vehicles repaired at EPA labs by skilled technicians who were directed to fix everything wrong with the vehicles. Even then, many of the vehicles were not repaired to the standard. When they are not, emission reductions are extrapolated so that the standards are met.	Many other repair studies and evidence from on-going I/M programs show higher costs and lower emission reduction than assumed by the Model.	Reassess MOBILE repair assumptions for I/M. Allow/require states to compare the Model assumptions about repair to actual repair effectiveness in each program.
Limitation of data acceptable for MOBILE	Repair data that is the basis of assumptions about repair in the Model cannot be based on real world repair experience.	Only FTP data are allowed to be used in the model. FTP tests are expensive and are only available from special studies having limited number and type of vehicles; potential for selection bias.	IM240 test and repair data are becoming available from a number of states on large numbers of vehicles repaired under real world conditions.	Use IM240 based repair data from state I/M programs to compare to evidence from laboratory FTP data on repair effectiveness.
Cutpoints	Repair effectiveness changes with stricter cutpoints are assumed based on limited data.	MOBILE5 awards greater SIP credits for more strict cutpoints based on evidence from relatively small EPA repair dataset. Yet mechanics repairing these vehicles were not repairing to a particular cutpoint.	Cost-effectiveness of repair decreases with more severe cutpoints. Vehicles with emissions near cutpoints may even have higher emissions after repair. Stricter cutpoints could lead to more motorist resistance or less compliance with I/M programs. Cutpoint impacts and cost-effectiveness needs to be assessed.	Review evidence from real world I/M programs on emissions reductions under different cutpoints, and compare to assumptions in Model.

Components of the Mobile Model	Issue	MOBILE5	Comment	Recommendation
Compliance	The definition and use of the compliance rate in the Model.	The compliance rate in MOBILE5 adjusts the emissions reductions from I/M downward based on the proportion of the fleet which does not comply with the I/M program. That adjustment may be too small since it does not appear to take into account the extent to which non-complying vehicles are failing vehicles. Also, states do not have to measure compliance rates and compare their measured rates to that assumed in the Model.	The appropriate measure of compliance rate is complicated to define and may be difficult for states to measure. Nonetheless, it is an important variable in MOBILE and needs further consideration. The evidence from Arizona and other state programs is that a fairly large % of vehicles may not be getting fully repaired or may be disappearing from I/M records.	The definition of the compliance rate and the emission discount taken in the Model should both be reexamined. There should be some guidance for states on how to measure compliance, so states can compare their measured compliance rates to the rates assumed in the Model. Fate of disappearing vehicles should be determined.
Evaluating I/M programs	States have no incentive to improve I/M programs which may not be as effective as assumed by MOBILE. Other vehicle emission reduction programs or policies may be ignored as a result of overly optimistic I/M SIP credits.	States may use the MOBILE Model to generate credits using default values available in the Model. This assumes that all vehicles will have emissions reduced to the standard. And, the Model may not be accounting for compliance correctly. Also, it is difficult to modify the Model to include alternative policies that might cost-effectively reduce emissions.	The credits generated by the Model in terms of predicted emission reductions from I/M should accurately reflect the emissions changes actually occurring. Only then will states have the incentive to seek ways to improve the effectiveness and cost-effectiveness of emission reduction programs.	Ways to benchmark I/M program credits to emissions reductions that actually occur in practice should be evaluated. Emission reductions measurements based on both in-program test data and out-of-program data (such as random pullover or remote sensing) should be considered.

Components of the Mobile Model	Issue	MOBILE5	Comment	Recommendation
Variable emissions problem	To what extent are variable emissions from the same vehicle a problem for identifying and repairing high emitting vehicles in an I/M program?	MOBILE assumes that all variation in emissions for the same vehicle is due to test measurement error or improper pre-conditioning.	Emission variation on the same car may be due only to test measurement error, but the alternative explanation that there are some vehicles whose emissions are inherently variable needs to be examined. If some vehicles do have variable emissions, this is one explanation about why there is a difference between I/M performance as measured in I/M test lanes and on the road.	There should be further study of this issue to determine the extent and seriousness of the problem. To the extent that it is a problem, modeling and policy alternatives for dealing with it should be considered.
Fuel economy and emissions	There is some evidence that HC and CO emissions deteriorate on a gram-per-gallon basis, so that emissions in terms of grams per mile of vehicles with good fuel economy deteriorate more slowly.	MOBILE assumes all vehicles of the same age and vintage deteriorate at the same rate in terms of grams per mile.		There should be consideration and perhaps further study of whether emission deterioration rates should continue to be constant in gram-per-mile terms.

A Behavioral Analysis of EPA's MOBILE Emission Factor Model

Winston Harrington
Senior Fellow
Resources for the Future

Virginia D. McConnell
Senior Fellow
Resources for the Future

Matthew Cannon
Research Associate
Resources for the Future

Discussion Paper 98-47
September 1998

Resources for the Future
1616 P Street NW
Washington, DC 20036
(202) 328-5000
Fax (202) 939-3460
Internet: `http://www.rff.org`

CHAPTER 1

INTRODUCTION: SOME PRELIMINARIES ABOUT MODELS

In this report we review the MOBILE Model, EPA's computer model for estimating emission factors for mobile sources. An emission factor model takes emission data collected under laboratory conditions and uses it to construct emission factors, in terms of grams of pollutant per vehicle mile traveled (VMT), for a variety of vehicle types, operated under a variety of driving modes. Emission factor models are needed for this purpose because of the great difficulty of obtaining reliable, representative emission measurements under real-world conditions.

It is important to keep in mind that MOBILE is an "emission factor" model and not an "emission model"; unlike the latter, MOBILE cannot be used to generate aggregate emissions. Nonetheless, emission factor models are essential inputs to emission models, which combine emission factor estimates with estimates of the amount and type of vehicle use in order to estimate current and forecast future vehicle emissions inventories. Vehicle emission inventories, in turn, have several 7uses, including (i) assessing the relative contribution of stationary and mobile sources to air quality problems, (ii) providing source inputs to air quality models, (iii) assessing emission reduction strategies, (iv) determining whether air quality programs undertaken in nonattainment areas will reach attainment with ambient air quality standards, (v) assessing conformity in transportation planning, and (vi) calculating emission reduction "credits" under I/M programs for State Implementation Plans (SIPs).

The emission factors produced by MOBILE are better suited for some of these uses than others, and we discuss the attributes of the Model with regard to these uses below.[1] The primary (though not the exclusive) focus of this analysis is on the structure of the Model's I/M component and its underlying assumptions and data. We review the Model's assumptions regarding vehicle repair under I/M, compliance with I/M requirements, and the impact of test measurement error on predicted I/M effectiveness. As part of this analysis we also examine the extent to which behavioral and stochastic elements of I/M regulations are accounted for in the design of the I/M component of the Model. There is concern in the regulatory community that the Model design has focused on technical aspects of I/M performance and has not fully accounted for the behavioral responses that occur in real world implementation.

The influences of behavior on vehicle emission rates is the focus of our efforts. However, we need to be clear about what we mean by "behavior." To some close observers of MOBILE, it appears that this term "driver behavior" refers to the real-time behavior of drivers while behind the wheel, such as whether speed or acceleration are as assumed in MOBILE. Our definition of behavior includes these real-time considerations, but it includes

---

[1] For a thorough review of the various uses of MOBILE see "Big Picture Modeling Issues" Modeling Workgroup, FACA Mobile Sources Technical Review Subcommittee, EPA OMS, Ann Arbor, Michigan, 1997.

much more, namely all changes in activities of motorists, mechanics, or other relevant actors that are made in response to the incentives of various regulatory policies.

This report is one of several analyses that are appearing this year by observers outside the EPA about the structure and performance of EPA's MOBILE model. The other reports include reports by the National Academy of Sciences (NAS, 1997) and the Government Accounting Office (GAO, 1997). Our report differs from these others in its explicit focus on the behavioral aspects of the model and in the intensive examination of the I/M portions of the model. The NAS and GAO reports are in some respects complementary to our efforts. Whereas much of the analysis in those two reports is concerned with the definition and use of the "emission correction" factors in the model,[2] we have largely limited our analysis to issues surrounding MOBILE's determination of the so-called "base emission rates." The base rates with and without I/M are supposed to be average emissions in grams per mile that would be experienced if the vehicle fleet were operating under a given base set of conditions and assumptions. Those assumptions concern average vehicle speed, ambient temperature and a number of other variables that affect emissions. The emission correction factors are applied to generate estimates of emission rates when these variables take on other values.

Now is an opportune time for raising these questions for two reasons. First, EPA is in the midst of a major overhaul of the model. The current version, MOBILE5b, will soon be replaced by MOBILE6, and staff members of EPA's Office of Mobile Sources (OMS) are hard at work collecting and analyzing the latest data about vehicle emission performance. OMS is also reaching out, to an unprecedented degree, to the public for assistance and to get the reaction of interested parties to the Agency's preliminary findings before the new model is proposed. A workgroup reporting to the Mobile Sources Technical Advisory Subcommittee of the EPA's Clean Air Act Advisory Committee has been set up to consider issues involving the upcoming revisions to MOBILE model.

The second reason arises from the National Highway System Designation Act, which contained a provision that changed the rules affecting state I/M programs in three important ways. First, it prohibited EPA from requiring states to use the IM240 test. Second, it prohibited EPA from imposing an automatic 50 percent discount on the emission credits granted to test-and-repair (as opposed to centralized test-only) programs. Third, EPA was further prohibited from granting credits to state I/M programs simply on the basis of whether they have certain characteristics. Instead, the states were allowed to make their own estimates of emission credits, and EPA could review their methodology. EPA recently proposed rules concerning the evaluation methodology for checking state plans, but these issues are still very much undecided.

In the remainder of this chapter we make some general comments about models, with particular reference to emission factor models. We examine the close relationship between model characteristics and model uses and discuss briefly the problem of validating large-scale models like MOBILE. We expand on the notion that to be useful for policy analysis a model

---

[2] Additional work along these lines is being done by the national labs in the definition of the TRANSIM model.

must take explicit account of behavioral considerations. Since we find also that the behavioral responses that most strongly affect emission rates are related to I/M programs, we conclude Chapter 1 with a brief look at a particular use to which the MOBILE model has been put, namely the calculation of emission reduction credits available for implementing vehicle Inspection and Maintenance (I/M) programs.

Chapter 2 contains a brief description of the MOBILE model, with particular attention paid to the I/M component. In Chapter 3 we describe the principal mechanisms by which behavioral responses can affect emission rates. This chapter also shows how a structural approach to modeling I/M, in which emission measurement errors and uncertainty about repair effectiveness are explicitly modeled, can suggest new hypotheses about the performance of I/M, involving variables that are not now collected.

The main empirical results are found in Chapter 4. We carefully examine the data and methods used by EPA to set the repair effectiveness parameters in MOBILE. We then compare EPA's results to results from other repair data sets, notably the emission data collected as part of the Enhanced I/M program in Arizona. We also examine the parts of the I/M component of the model dealing with compliance and tampering and compare them to results from Arizona. Chapter 5 contains a sensitivity analysis, comparing the importance to model results of various parameters in the model, and our conclusions are presented in Chapter 6.

We first review the uses of emission factor models like MOBILE and explain what is meant by behavioral responses to regulations. We then describe the MOBILE Model focusing on the I/M component, examining both the possible technical and behavioral aspects. We then review, where possible, the empirical basis underlying the assumptions of the I/M component of MOBILE, and compare that to evidence from other sources. We then present some sensitivity analyses by varying some of the more important technical and behavioral parameters of the Model. Finally, we examine some possible modifications and alternatives to the Model and present some conclusions.

1.1. Emission Models and Emission Factor Models

A mobile-source emission model produces estimates of emissions under a wide variety of traffic conditions and roadway configurations. These emission estimates are made, essentially, by multiplying together two types of inputs: emission rates, expressed in grams per mile, and travel, expressed in miles. In every state except California, the MOBILE model generates the emission rates used in the development of emission inventories and in planning response rates.[3] MOBILE has been developed over the last 30 years by EPA's Office of Mobile Sources. The other component providing inputs is a local or regional travel demand model, which produces estimates of vehicle miles traveled (VMT).[4] Applying the emission

[3] California has developed its own emission factor model, the latest version of which is EMFAC7G.

[4] Travel demand models can also produce other outputs that may be relevant for emissions. For example, some can produce estimates of trips as well as VMT, which is important for emission estimation because of cold starts.

factors generated by MOBILE to these travel demands generates estimates of local mobile-source emissions.

For some applications (such as the input into local air quality models) the emission estimates generated by emission models must be quite specific to time and place. This means that both the travel demand models and the emission factor models must likewise be capable of generating time- and location-specific estimates. In the MOBILE model this requirement is handled by splitting the emission rate calculation into the calculation of "base" emission rates and "correction factors" applied to those base rates. The base rates are the emission rates applicable to a set of reference conditions--the conditions, in fact, that define the Federal Test Procedure (FTP). Once a "base" emission rate is calculated, MOBILE attempts to modify it to take into account variables found to exert a strong effect on emission rates: the temperature, average speed, fuel quality, and driving mode. This last variable is a vector that gives the frequency of the three different driving modes (hot start, cold start and normal operation).

The operations required to produce an estimate of emissions are roughly as follows. Base emission rates are generated by the TECH model. The TECH outputs are inputs to MOBILE proper, which takes the specific emission factors and computes a weighted average vehicle emission rate for each of six classes of vehicles--cars, small light-duty trucks, large light-duty trucks, and three kinds of heavy commercial vehicles. MOBILE also calculates the correction factors based on locally specific data and applies them to each of the classes of vehicles, to obtain emission factors applicable to the specific situation of interest. The last step is to take these specific emission rates and multiply them by the total distance traveled by all vehicles in each class.

This approach to emission estimation has the quite substantial virtues of flexibility and simplicity. The emission factors can be produced by MOBILE for almost any level of aggregation for which data (or estimated values) on vehicle use are available, from the entire metropolitan area for a year, all the way down to a single stretch of highway on a summer day. Furthermore, by separating the emission rate calculation from the travel demand estimate, emission estimates can be made by combining the results of a local travel demand model and national emission rate model. The advantages of this separation for local and state governments is that they are spared the expense of developing and maintaining their own emission factor models. From EPA's perspective, the existence of a single emission factor model applicable to all the states greatly simplifies the administrative burden associated with assuring nonattainment areas are living up to their statutory obligations. If each state developed and calibrated its own model, it would require substantial effort for EPA to review each one. On the other hand, the existence of a single emission factor model to be used by all the states imposed some inflexibilities of its own, including a lack of adaptability to local conditions and to certain types of policies and a discouragement of innovation. The development and use of MOBILE thus raises the familiar conflict between the virtues of standardization and its disadvantages.

1.2. Model Uses and Model Characteristics

The difficulty of modeling emission rates should not be underestimated. Emissions depend on a vast array of variables, including

- technical variables such as the emission control equipment in use on the vehicles in the region, as well as the performance and durability of that equipment;
- situational variables such as ambient temperature, vehicle speed, vehicle "mode" (cold start, hot start or running), and frequency and rate of acceleration. the characteristics of the road network, and the level of congestion; and
- behavioral variables, such as the number and types of vehicles in use, how much they are used, how they are driven, how well they are maintained, and whether or how well they are repaired when they malfunction.

The MOBILE model tends to concentrate on the first two categories, the technical and situational variables. The core of the model takes into account the emission control technology and other hardware-related factors that affect emissions, while the situational variables are taken into account through the use of the correction factors that modify the base emission rates. MOBILE is indeed intended to be a "technical" model, one that takes the technical characteristics of the vehicles in the fleet and computes an average emission rate.

Because behavioral factors affect emission rates in numerous ways, MOBILE can hardly avoid making implicit behavioral assumptions. The behavioral assumptions remain largely unidentified but are often reflected in the averages that are used to determine many supposedly "technical" parameters. For example, MOBILE incorporates a gradual increase in emission rates as vehicles age, at a rate that is specific to the type of vehicle. These "deterioration rates" are determined empirically from the "Emission Factors Database," which consists of emission tests conducted on a sample of vehicles each year. The observed rates are the result of an interaction between equipment durability and the average motorist's proclivity for vehicle maintenance. Motorists' habits regarding maintenance presumably reflect some balancing between the cost of maintenance now and the consequences of no maintenance later. This balancing is surely affected by their incomes and the relative prices of fuel, vehicle parts, and vehicle repair. It could also be affected by the nature of the local I/M program and perhaps other policy variables. In MOBILE there is no recognition of the contingent nature of these variables; they are considered to be constants fixed by the technology, just as the emission profiles associated with particular types of engines are considered.

Furthermore, if parameters are the way they are because they represent the aggregate response of actors in the system to prices and policies, then they could change in response to changes in those variables. But dynamic adjustments to changing parameters by consumers, mechanics, vehicle and parts manufacturers and other private actors are not built into the

Model. Certainly, constructing such models would be a big task, requiring great amounts of data that may not be readily available, and so it would be asking for too much to advocate that such adjustments be built into emission factor models. We wish to ask a simpler set of questions: To what extent does the failure of MOBILE to incorporate behavioral responses affect the results of the model and the policy implications that flow from it? How sensitive is the model to its implicit behavioral assumptions? Are there simple modifications or additions that can be made that will provide for a wider or at least different set of policy options to be considered? What would be the implications for data collection if a serious effort were made to incorporate behavior?

Even if it is true that behavioral adjustments affect emissions, it is not immediately obvious that they have to be included in an emission factor model. No model has to take into account all the relevant variables to be useful, and in fact it cannot. The whole point of a model is to zero in on the most relevant variables, and that depends on which questions are being asked and what the model is being used for. In thinking about whether emission factor models ought to incorporate behavioral elements, it is worth considering how the model is to be used. Consider how well suited MOBILE is for two important ways to use emission models.

Descriptive uses. An important use of emission factor models is to support development of emission inventories. The pattern of emissions can then be linked to ambient air quality, so that policy-makers can determine how much emission reduction is necessary to achieve ambient air quality objectives. When used for this purpose, the mobile source inventory must be combined with a stationary source inventory. Not only does this exercise generate an estimate of required emission reductions, but it also helps to generate a comparison of the relative importance of the various source categories.

Given an existing pattern of vehicle use, behavioral considerations are not important, because the behavioral variables of the greatest interest for emission rate determination can be considered fixed in the short run.

Observers generally agree that MOBILE now does a reasonably good job of estimating past and current aggregate average emission rates, although this wasn't always the case. Tunnel studies, remote sensing studies and roadside pullover studies conducted in the late 1980s suggested that MOBILE4 was drastically underestimating emissions, and as a result emission factor estimates in MOBILE5 were approximately doubled for HC and CO (Calvert et al. 1993). Several examinations of the accuracy of MOBILE were presented at the 1994 On-Road Vehicle Emission Workshop. These papers found generally that MOBILE5 predictions were an improvement over those of MOBILE4, although now emissions were being overpredicted (Gertler et al. 1994). The quality of MOBILE estimates of future emission rates is less certain and depends on the quality of the input data--that is, the accuracy of future projections of emission characteristics of new vehicles and the performance over time of the existing vehicle fleet. MOBILE is no exception to the rule that it is difficult to predict the future.

Policy comparison and evaluation. To estimate current emissions or current emission rates, one only needs a *static* model, one that estimates current emission rates and translates them into total emissions. Static models like MOBILE have no memory; they are not influenced by what has happened in previous years. In order to predict future emissions and more particularly to answer what-if questions, a *dynamic* model is required. The problem is not simply that input data changes over time (something that can be dealt with in a static model), but that the changes are endogenous to the emission reduction policies being examined. For example, a change in the emission policy this year could affect vehicle purchases and retirements and hence the structure of the fleet in coming years.

When emission models are used for policy comparison and instrument choice, behavioral responses are almost always important. For one thing, nearly every policy can have inadvertent effects on behavior, with subsequent effects on emissions. Consider a regulatory policy such as I/M. Even though this is not an "incentives" policy, it does offer the prospect of changing the behavior of motorists. If I/M is made more stringent, older vehicles will be reduced in value, thus accelerating their retirement. Greater stringency may also cause motorists to try harder to evade I/M. Both actions will affect emission factors, though one will raise and the other reduce them.

In addition, some policies work by giving consumers incentives that change behavior. Most economic incentive programs work in this way. For example, a policy that changes the tax structure for ownership of new and older vehicles would affect vehicle holdings of different ages and therefore the fleet emissions. Or, a high minimum expenditure for waivers combined with exemptions for new vehicles in an I/M program may induce motorists to more quickly replace old vehicles with newer ones. The Model currently has no way of analyzing such programs. (Adjustments to the vehicle stock could be made outside the model, and the results fed to the model. Several models have attempted to do this kind of analysis, and we discuss some of them in more detail below.)

1.3. Model Validation

As noted above, there have been some attempts to validate MOBILE by comparing model results to observed outcomes, primarily average emission rates. However, simply limiting model validation to emission rate comparisons is not sufficient, especially when evaluating the importance of behavioral responses. It is difficult to collect emission data sets that are useful for this purpose and that are acceptable to all the interested parties. Besides the tunnel studies alluded to above, it is also possible to estimate average emission rates from remote sensing studies, but these are not considered very reliable by some observers.

Emission rates are difficult to observe in any event, because it is difficult and expensive to find vehicle emission data from a representative sample of vehicles that fairly represents the emissions of the vehicle as it is actually used. The most accurate emission tests (FTP tests) are quite expensive and time-consuming, so any data set of FTP testing must be voluntary and thus raises questions of sampling bias. In addition, vehicle emission rates show large variation from one vehicle to another and for the same vehicle under different operating conditions. There are also several potentially important stochastic components. A vehicle's "true" emission rates could depend on variables that are difficult or impossible to observe, so that emission rates could differ at different times, even when ostensibly under the same operating conditions. In addition, the vehicle's measured emissions could be subject to measurement error. Any emissions test is likely to be imperfect and subject to some error in measurement. A further problem is that emission measurements are often made under a restricted set of vehicle operating conditions in a laboratory setting, and it is unclear how these measurements should or can be used to predict average vehicle emissions in use. There are many aspects of driver or mechanic behavior that may be different in real world settings compared to the lab results.

Fortunately, emission and emission factor models can be validated by examining other outcomes that do not involve comparisons to measured emission rates. These additional outcomes frequently involve I/M policy. Examples include the failure rate of the I/M program, the effectiveness of emission repair, the distribution of emissions across vehicles, and the relationship of emission repair to the emission test cutpoints. Now that I/M programs have been underway in several states, there is also information available from these sources for comparing to MOBILE's input assumptions and intermediate outcomes. We report on some of these comparisons in Chapter 4.

1.4. MOBILE and Inspection and Maintenance Policies

MOBILE is especially important in the evaluation of locally implemented policies that have a direct bearing on the average emission rate of mobile sources. Some of these programs are required in certain nonattainment areas, including I/M, reformulated gasoline and oxygenated fuel, although the local air quality control region may have some discretion over the details of these programs. The difference between the estimated emission reductions required for attainment and the emission reductions expected to be achieved by these mandatory programs is the number of emission "credits" that must be achieved by other programs. The local air quality authorities then makes choices from a large menu of alternative policies, each of which has emission reduction credits calculated by the emission model.

SIP approval depends on having credits equal to the required emission reductions, which drives local air quality planners to compare alternative policies on the basis of the emission credits generated by the model. Any policy that is not recognized in the emission model as reducing emissions does not generate credits and therefore is unlikely to receive much consideration. That is, the use of MOBILE in making judgments about attainment almost makes it inevitable that it will be used in making policy comparisons as well. It no longer matters what will happen in fact, but what MOBILE says will happen.

Even when it generates emission credits, a potentially useful program can be "crowded out" by inspection and maintenance programs, if its emission reductions overlap with the emission reductions that are credited to I/M by MOBILE. Crowding out results because (a) Enhanced I/M is required in regions that are in serious nonattainment for mobile-source pollutants and (b) MOBILE's estimates of the emission reductions achievable from enhanced

I/M are very large. Once the emission reductions calculated by MOBILE are credited to Enhanced I/M, the maximum emission reductions available to the competing policy may be quite small. If I/M turns out to be as effective as the MOBILE model predicts, then the problem is minor. On the other hand, if I/M is not as effective as anticipated, then not only will resources be wasted on I/M programs, but potentially useful opportunities could be missed. Thus, a lot is riding on the performance of the I/M components in the MOBILE model.

Given its importance, we examine the I/M component carefully. As we point out in greater detail below, the way MOBILE calculates emission reductions due to I/M does not attempt to take into account the stochastic and dynamic elements of I/M. All the calculations in the MOBILE model are done on the basis of average values. As we discuss below, the existence of uncertainty could provide opportunities for behavioral responses that are not present in a deterministic model. We show how explicit modeling of uncertainty, therefore, could lead to new hypotheses about what is important in I/M programs and what is not.

Likewise, we show how explicit consideration of the dynamic aspects of I/M enforcement and compliance can affect one's understanding of how I/M works. For several reasons an I/M program may require several years to reach a "steady state," and one may not be able to predict what that steady state will be without thinking about the dynamic elements, and perhaps doing empirical research. For example, states may wish to implement I/M slowly, by beginning with relatively lax emission standards and gradually bringing them down. Even if stringent standards are imposed right away, it may take several years to find and repair vehicles properly, and in that case it is likely that the number of repairs will be very high initially and fall rapidly as vehicles are repaired. In addition, learning by mechanics and motorists will provide dynamic effects. On the positive side, repair effectiveness may improve even as cost goes down, as mechanics become more adept at vehicle emission repair. On the other hand, motorists and mechanics may also become more proficient at finding ways of avoiding compliance.

Such dynamic considerations are generally beyond the scope of the MOBILE model. The model provides emission estimates for each year independently of what has happened before. With one relatively minor exception,[5] that is, MOBILE calculates the effects of I/M each year as if (i) that year were the first year of I/M and (ii) all the benefits of I/M are achieved in that first year. Furthermore, MOBILE reports only average emission rates. Not reported are other, more easily observable aspects of I/M, such as the fraction of vehicles failing the test each year or the number receiving emission repairs. Information of this sort would be very useful in checking the MOBILE results against the real world outcomes. Without model predictions of failure rates or repair effectiveness rates, there are only two outcomes available for validating the model, namely the overall emission factor as determined

---

[5] The exception is that in a biennial program the emission benefits are phased in over a two-year period, since only half the vehicles are tested each year, and in subsequent years the average emissions are somewhat higher than they would be in an annual program because of the additional time between tests allows greater emission deterioration.

from tunnel studies and the effect on ambient air quality. Furthermore, EPA appears to have adjusted past discrepancies between the model and calculated emission rates or ambient concentrations by adding overall "fudge factors," such as the tampering deterrence factor, the decentralized program penalty, and the noncompliance adjustment. Thus any agreement between the model and the real world is imposed; it does not emerge out of the structural properties of the model.

We have been told that the very earliest EPA models (c. 1979) of I/M did in fact embody a dynamic approach, but limitations of data caused the Agency to adopt the current static and non-stochastic approach. Unfortunately, without a model that dealt explicitly with uncertainty and behavioral responses, the lack of data that would allow the testing of the importance of these factors became self-perpetuating. In part, this was another legacy of the use of the MOBILE model for enforcement purposes: It essentially prevented any further work on competing models. One of the uses of models not mentioned above is that it imposes on the developer a structured way of thinking about a problem. Different models provide different structures, and there is no way of knowing, at the outset, which are the most useful ways to think about a problem. But if one model has already been selected as the representation of "reality" then it becomes awkward to entertain the possibility of alternatives.

These questions are especially pertinent, we believe, to those parts of MOBILE that analyze state Inspection and Maintenance (I/M) programs and are used to generate emission credits in the State Implementation Plan (SIP) process. MOBILE is more than just a tool for decision making; it has become the arbiter of whether a region in nonattainment is meeting its schedule of emission reductions. In our review we concentrate our efforts on the way in which the I/M component is designed, and how it accounts for behavioral factors.

CHAPTER 2

A SHORT DESCRIPTION OF THE MOBILE MODEL

In this section we provide a brief functional description of the essentials of the MOBILE model. The essentials, that is, for examining the behavioral aspects of emission modeling. We make no attempt to be comprehensive, and those seeking more information on how to run the model should look elsewhere. The best published description of the model is a report prepared for the Department of Transportation *Evaluation of MOBILE Vehicle Emission Model* by Sierra Research, Inc. (1994), which describes in great detail how the technical parameters in the model are derived and used. In order to make the discussion as precise as possible, we make occasional use of mathematical expressions. When expressions are not identical to the formulas actually used in the MOBILE model, it is because the actual expressions must address details that are not relevant to the point we are discussing.

MOBILE is an emission factor model that produces an estimate of the average emission rates in grams/mile for a set of vehicles under a particular set of circumstances.

The calculation is broken into three steps, as follows.

(i) Calculation of base emission rates. Base emission rates are estimated for a wide variety of vehicle types and are supposed to represent the emissions of an average vehicle of that type when used in average urban driving. Vehicles are classified by class (cars, small pickup trucks, large pickups and three types of heavy commercial vehicles), model year, age and engine technology. The driving pattern that is used to determine the base rates is the same as the Federal Test Procedure (FTP) used to certify compliance with emission standards for new vehicles.

(ii) Aggregation of base emission rates. The model-year- and age-specific base emission rates are aggregated into a base emission rate for the six vehicle classes. The base emission rates for each class are weighted averages of the vehicle types in each class, with the weights determined by the number of vehicles of each type and the estimated average use. These age weights can be location-specific, at least for cars.

That is, if $e(i,a)$ is the average emission rate for a vehicle of age *a* produced in model year *i*, then the base emission rate in 1992, say, is the weighted average emissions of new 1992 vehicles, one-year-old 1991 vehicles, etc.

$$B_{92} = w_0 e(92,0) + w_1 e(91,1) + \cdots + w_n e(92-n,n)\ .$$

The weights, essentially, are the estimated number of miles driven by all vehicles of each age.

(iii) Application of correction factors. Correction factors are applied to the average base emission rate produced by the TECH submodel to produce an average emission rate applicable to a specific set of real-time conditions. The variables taken into account by the correction factors are those that can vary considerably over short periods of time and have been found to exert a strong effect on emission rates: the ambient temperature, average speed,

fuel quality, and driving mode. This last variable refers to the frequency with which vehicles are in three different driving modes (hot start, cold start and normal operation).

If c_T, c_s, c_F and c_M are the correction factors for temperature, speed, fuel quality and driving mode, respectively, then the 1992 base emission rate corrected for the particular factors is

$$B_{92}^{c}(c_T, c_s, c_F, c_M) = c_T c_s c_F c_M B_{92}$$

The application of correction factors to base emissions is taken in part as a convenience to users but it requires some simplifying assumptions, namely that the correction factors are independent of fleet mix. That is, if the effects of speed differ across vehicle types, errors could be introduced if the stock of vehicles is markedly different from the one in place when the speed correction factor was derived. It also allows (though it does not require) users to assume that evaporative emissions are a multiple of mileage, even though such emissions occur in the summer regardless of whether the vehicle is in operation. If something were to cause daily vehicle mileage to grow or decline, for example, it is unlikely that the real change in evaporative emissions would be proportional although in MOBILE they would be.[6] The use of average rather than vehicle-type-specific correction factors is presumably not worth the additional model complexity that would be required, but it is an illustration of how endogenous changes in fleet composition could affect emissions without being picked up by the model.

Because most of the behavioral effects we are concerned with make their presence felt in the base emission rates, we discuss them in more detail in the following two sections.

2.1. Base Emission Rates without I/M

The base emission rates are produced in an auxiliary model, TECH, which does much of the work of preparing emission factors and computing the effects of I/M. The output of TECH is a set of emission factors that are embedded into the MOBILE model proper to produce the emission rates required for the particular scenario for which emissions are to be estimated.

TECH begins with the following inputs:

Zero-mile emission rates. The emission rates of new vehicles.

Deterioration rates. The rates of increase in emissions. The deterioration rates are expressed as changes in grams/mile per 10,000 miles of wear and are converted to time-based deterioration by means of average age-specific cumulative mileage estimates. The deterioration

[6] Although there is no user input to replace the default values for miles per day, the model has an option to express non-driving emissions in other units. However, it is unclear whether more locally specific mileage and trip assumptions are used by the states.

rates themselves are also mileage-specific, with a faster increase in emissions for vehicle models with more than 50,000 miles of use.[7]

The zero-mile emission rates and the deterioration rates are disaggregated by *vehicle class* (3 classes: cars small pickups and large pickups), *pollutant* (HC, CO and NOx for gasoline vehicles), *technology* (open-loop, carburetted, and two types of fuel injection for gasoline vehicles), and *emitter group*. Vehicles are designated as being in one of four emitter groups for HC and CO (normal, high, very high and super) and two groups for NOx. The zero-mile emission rates and the deterioration rates are derived from FTP testing of vehicles in use since 1972. Certification tests are performed on each new vehicle model. For recent model year vehicles EPA has also been conducting, since the late seventies, FTP tests on a sample of vehicles one to seven years old. Emission data on older vehicles comes from special programs, including IM240 testing at the Agency's test lanes in Indiana and Arizona. Collectively, these emission-test data sets comprise the "Emission Factors Database."

These data are combined to yield, as outputs, the average vehicle emission rates (the $e(i,a)$ above) by model year and vehicle age, for each vehicle class and pollutant. These outputs become inputs to the MOBILE model proper.

Inside MOBILE the fleet age distribution is as it was in 1990, regardless of the year, so that for every age and year, the number of vehicles of age a remains constant for each age a. Obviously, this is not quite right, since for all but new vehicles, the number of vehicles of age a in year i is the same as the number of vehicles of age $a-1$ in year $i-1$, less scrappage during the year.[8] MOBILE does have an add-on module that allows the user to generate and substitute a more realistic fleet that begins with the current fleet and a projection of new vehicle sales and then generates future fleet distributions based on the basic accounting identity above. At present this option is available for cars only, but future plans are to extend it to trucks.[9]

The emission rates are increasing in age for fixed model year, and in most cases decreasing in model year for fixed age. That is, as a set of vehicles produced in the same year get older their average emissions increase as parts wear out, etc. But the emission rates for vehicles the same age decline in newer vehicles, reflecting improvements in durability of emission control technology. Empirical support for such an improvement is clear. One of the major issues in the calculation of deterioration rates is the existence of a "kink," or an increased rate of deterioration after 50,000 miles, which was mentioned above. The kink is particularly prominent in early vintages. The existence of a kink may be entirely technology-driven or it could arise in part because of the structure of vehicle warranties. A 50,000-mile warranty used to be required for the emission control system. Data from very recent model-year vehicles fail

---

[7] It is expected that the deterioration rate change at 50,000 miles will be dropped in the MOBILE6 version of the Model.

[8] This formulation excludes imports and exports of used vehicles, which could be important in some contexts, such as along the Mexican border.

[9] It is unclear how many states actually use this option to age the fleet, or how many use the default of a constant distribution.

to show a kink; perhaps due to the extension, in the 1990 Clean Air Act, of the warranty period for the emission control system to 80,000 miles.[10]

Although deterioration rates are allowed to vary by technology type, it is assumed nonetheless that vehicle emissions deteriorate on a constant gram-per-mile basis. However, a recent empirical analysis at RFF showed that for HC and CO, at least, a better fit to the observed vehicle-specific emission data could be obtained if the deterioration rate were expressed at a constant rate in terms of grams per unit of fuel burned. This work is described briefly in Appendix B. One implication is that fuel economy is a determinant of emissions: Although all vehicles of a given vintage and class start out by regulation with the same emission rates, vehicle use drives up the emission rates of gas guzzlers faster than those of fuel-efficient vehicles. As discussed further in Appendix B, the cause of the difference is probably that when emission systems fail, the emissions approximate those of uncontrolled vehicles, which are higher on a gram-per-mile basis for gas guzzlers. The same improvements in emission system durability that are reducing the importance of the kink are also slowing the growth in emission differences attributable to fuel economy as vehicles age. However, the effect is still substantial in 1986 model year vehicles after 10 years.

2.2. The Calculation of Base Emission Rates with I/M.

The calculation of the base emission rates for each year, $e(i,a)$, represent a world without inspection and maintenance programs. The TECH model also has a component that computes I/M program effectiveness. Given the characteristics of an I/M program, the TECH model reports out a set of emission "credits," or percentage reductions in emissions for each pollutant and for each combination of model year and age. Based on fleet characteristics, MOBILE then calculates the emission reduction credits available from I/M.

Perhaps the most important determinant of the effectiveness of an I/M program, at least as far as the I/M calculation in the TECH model is concerned, is the type of emission test and the criteria for failing it (the cutpoints). Each emission test, in turn, has two important characteristics: a set of identification rates, one for each emitter group, and the repair effectiveness function.

Identification rate I. To define the identification rate, EPA first develops the concept of "excess emissions," which is the total of the "true" emissions from a set of vehicles in excess of the applicable emission standard(s). The true emissions are defined to be the results of an FTP test. For an emission test and for a particular set of emission "cutpoints," the identification rate is defined to be the fraction of total emissions from vehicles in the set that fail the emission test. The cutpoints are the pass-fail criteria, and clearly the higher these cutpoints, the less stringent the test and the lower the identification rate.

---

[10] The EMFAC model used by California in place of MOBILE also has a kinked specification of deterioration rates, but for the oldest vehicles it is reversed. That is, old vehicles deteriorate more slowly, perhaps because the surviving oldest vehicles are well-maintained and they may already have emissions at the uncontrolled level anyway.

Repair effectiveness R. The repair effectiveness is the percentage improvement in emissions of a repaired vehicle, compared to its emissions before repair. For a given emission test type, the repair effectiveness is a piecewise linear function of before-repair emission rates and test cutpoints, expressed as the percentage improvement of the initial emission rate. Lower (i.e., more stringent) cutpoints require more repairs and in the Model result in lower post-test emissions.[11] Test types examined in MOBILE include the idle test, 2500 idle test, loaded idle test, and IM240 test.

The empirical derivation of the repair effectiveness function is described below in Section 4.1 in Chapter 4 and the function is illustrated in Figure 3. Here we note several important aspects of that derivation. It is based on repairs made by EPA contractors in a laboratory setting, not on the performance of working mechanics in field situations. It is not clear what the marching orders of the contractors were, with respect to what the post-repair targets were or how much to spend. Also, some cars had emissions that still did not meet the emission standards after repairs were completed. For the purposes of estimating repair effectiveness, those vehicles were assigned emissions that met the standards, or actually emissions that were somewhat better than the cutpoints. In other words, the repair estimates are based to some degree on repair targets, rather than actual repair performance. This is a potential problem, since other repair studies have suggested that on average repair of super-emitters is quite costly and may not achieve emission reductions necessary to allow the vehicle to pass the test. This is discussed further below in the section 4.1.

Recall that vehicles at each age and vintage are classified into eight emitter groups according to their emissions, and that TECH calculates emissions by emitter group for each model year and each vehicle age. To get after repair emissions, and hence repair effectiveness, the before-repair emissions for each age/model-year/emitter group classification are read off the appropriate repair function in (Figure 3 shows an example of a repair function).

The effectiveness of I/M depends on both the identification rate of the I/M test for a given cutpoint and the effectiveness of repairs. Suppose e_a^0 are the base emissions for a group of vehicles with characteristics a (including emitter group, technology group and age), and let I_a and R_a denote the identification rate and repair effectiveness of the I/M test in question for these vehicles. The base emissions for these vehicles when subject to an I/M program are given by (Sierra Research, 1994):

$$e_a^{IM} = [(1 - I_a) + I_a(1 - R_a)]e_a^0 \tag{1}$$

The after-repair emissions in (1) is a weighted average of the original base emissions for the "unidentified" excess emissions and the after-repair emissions for the identified total emissions. Like other calculations in MOBILE, this calculation is done for many different

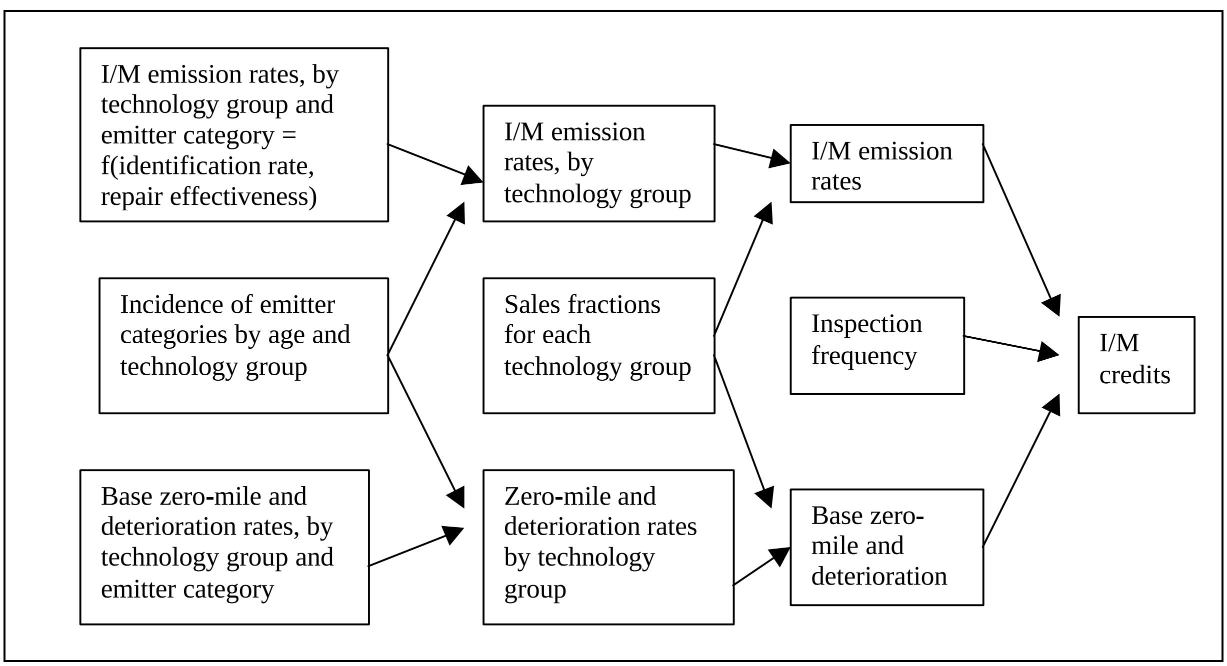
11 At least in some older I/M programs more stringent cutpoints may have led to increased emissions, as mechanics tried to repair marginal vehicles. See California I/M Review Committee, 1994.

subgroups of vehicles with particular characteristics (the a) and then a weighted average is taken.

Repair effectiveness gives the reduction in emissions, normalized by the before-repair emissions. The repair effectiveness is defined separately for each emission test and cutpoint, and is specific to technology and vehicle model year. It is based on empirical data on vehicles that are identified as violators by that test, and repaired. (Most but not all of the vehicles in the repair database achieved compliance after these repairs. It is not clear what distinguished those vehicles that did not. For the vehicles for which repair was not successful, the MOBILE model assumes that further repair would have reduced emissions below the cutpoints and assigns emission rates to those vehicles accordingly.) The main reason why test characteristics affect repair effectiveness are that there are repairs that enable the vehicle to pass certain tests without having more than a temporary effect on emissions (Lawson 1993). The IM240 test was developed largely to eliminate the ability of mechanics to "fix" the car to pass the test. The reason why the cutpoint is important is that tighter cutpoints can require more extensive repair (although the available evidence suggests that more extensive repair on marginally failing vehicles is as likely to increase as to reduce emissions (California I/M Review Committee 1996). The empirical development of the repair effectiveness function is discussed further below in Chapter 4.

Figure 1 shows an abbreviated flowchart of the TECH model. A more extensive flowchart is found in Appendix A.

Figure 1. TECH Model Flowchart



2.3. Input of TECH I/M Results into MOBILE

TECH produces two inputs into MOBILE that are relevant to the present discussion: (i) the base emission rates for each age and model year and (ii) the basic I/M credits available each calendar year, which are the percentage reductions in emissions for each pollutant attributable to the I/M program adopted by the state.

These credits are then combined in MOBILE with three further adjustments--the emission waiver rate, the compliance rate, and the decentralized-program discount--and used to calculate the Base Emission Rates ($BER_{I/M}$) under an I/M program, compared to the base rate B_0. Those adjustments are:

> *w* – the user-supplied waiver rate, the fraction of all vehicles receiving a waiver. A motorist can receive an emission waiver if he spends a certain amount on vehicle emission repair, even if the vehicle cannot pass the test. In the Clean Air Act of 1990, the waiver limit is set at $450, but new Enhanced I/M programs in some states appear to be setting waiver limits at much lower levels, at least initially.
>
> $f(c)$ – the adjustment factor for user-supplied compliance rate, the fraction of vehicles that do not comply with I/M. The compliance rate is the fraction of vehicles nominally subject to I/M that actually go through the testing process. The Model includes a disproportionate adjustment to emissions for non-compliant vehicles, based on the idea that the vehicles that avoid I/M are likely to have higher emissions than vehicles that do not. The compliance rate is a user input, presumably based on the state's I/M experience, but the adjustment function is hard-wired in MOBILE.
>
> *d* – the discount for decentralized test-and-repair programs. State I/M programs are of two basic types: decentralized programs and centralized (or test-only) programs. In a decentralized program the emission testing is done by privately owned service stations and repair shops certified by the state. Decentralized programs are often called test-and-repair programs since shops that test vehicles also repair them. (It would be possible to separate test and repair in a decentralized program, but no state does so.) In a centralized program the tests are done in a relatively small number of testing stations, usually operated by a contractor retained by the state. For decentralized programs EPA, until recently, applied a 50 percent discount to I/M credits, reflecting the Agency's skepticism about the efficacy of such programs. Recently, they adopted a more flexible policy that allows states to obtain credits for decentralized programs based on the actual emissions reductions obtained. (As noted above, the 50 percent discount has been eliminated by an Act of Congress.)

The formula used for calculating the basic emission rate with I/M is then

$$\mathrm{BER}_{\mathrm{I/M}} = B_0[1 - \mathrm{CRED}(1-w)f(c)d] \quad (2)$$

where CRED is the I/M credit determined by TECH.

CHAPTER 3

BEHAVIOR AND EMISSION RATES

In this chapter we discuss the potential influences of behavioral variables on emissions, with particular attention to the performance of I/M programs. Section 3.1 contains a short description of what seem to us to be the most important behavioral variables. Section 3.2 contains a closer look at behavioral effects in three important areas of I/M: emission identification, repair, and compliance. In this section we illustrate the importance of the stochastic and dynamic elements of I/M and show how thinking explicitly about these elements can suggest different hypotheses about I/M and, in some cases, data that would be interesting and useful to collect.

Behavioral responses can be incorporated in the MOBILE model in three ways:

(1) The Model inputs can be changed to reflect behavior. Examples include the vehicle age structure and the average speed.

(2) They are made implicitly and hard-coded in MOBILE or in the TECH submodel. For example, the emission test identification rates and the vehicle deterioration rates. Changes in these assumptions require the program code to be modified and the program re-compiled.

(3) There are some behaviors that cannot be easily captured by the current structure of the Model. The dynamic characteristics of I/M would be included here, (i.e., the dependence of I/M emission reductions in year t on the distribution of clean and dirty vehicles in year t-1). This would require a different kind of Model.

3.1. A Gallery of Behavioral Adjustments

Besides the example of deterioration rates given above, behavioral adjustments can affect other variables that in turn can affect emission rates.[12] To set the stage for what follows it is useful to lay out the principal mechanisms through which behavior can affect emission rates of vehicles in use. Some of these mechanisms affect the emissions of individual vehicles; others do not affect individual emission rates and change average emission rates by modifying the mix of vehicles in the fleet.

The type of vehicles owned. Different types of vehicles are subject to different emission standards even after adjusting for age. This means, for example, that the recent tendency of households to substitute sport utility vehicles (SUVs) and small trucks for

[12] Consumer behavioral adjustments can also affect emissions without affecting emission rates--by affecting vehicle use. Vehicle use changes might follow changes in fuel prices or changes in vehicle fuel economy, for example.

ordinary passenger cars may have raised average emissions above what they would have been if earlier buying patterns had persisted.

The age structure of the fleet. The gradual increase in the average age of the fleet has reduced the impact of lower emissions of new vehicles. Increased vehicle longevity is essentially a response of vehicle owners to a variety of factors, including higher prices of new cars, better durability, and lower prices of vehicle parts (see Hamilton and Macauley, 1997).

The distribution of vehicle use among vehicles. Vehicle use is highly variable and difficult to explain with easily-observable factors. The MOBILE model uses an age-specific use distribution to allocate miles to vehicles, but survey data from the 1990 Nationwide Personal Transportation Survey (NPTS) on vehicle use suggests that age explains about 4 percent of the variation in vehicle miles traveled (VMT) per vehicle. More elaborate models with larger numbers of variables explain somewhat more, but never more than about 70 percent, and those models contain data that are unavailable except in special surveys like the NPTS. In multiple-vehicle households (currently about 58 percent of U.S. households) mileage can easily be reallocated from one vehicle to another in response to changing prices or a changing regulatory framework. When we consider the opportunities for within household reallocation of travel, for example, it is conceivable that a gasoline tax, which has been put forward by numerous observers as an environmentally beneficial policy, would increase rather than decrease emissions, at least in the short run. The reason is that in the last few years vehicle fuel economy has slowly worsened in response to lower real fuel prices and the growth of the SUV segment of the new car market. A rise in the price of gasoline could cause a shift in use toward these older, more-polluting vehicles.

Repair effectiveness. How successful repairs are in real world settings would seem to depend critically on such factors as the incentives and ability of mechanics to learn, and on the cost of repair. If repair costs are high or higher than anticipated, there are a number of possible responses: mechanics may be inclined to do more partial, less effective repairs, state regulatory bodies may change requirements to lessen the burden, or motorists may scrap or sell their vehicles or register them outside the I/M area. There are also information asymmetries between mechanics and motorists that may result in less effective repair, as well as differences in the motives of mechanics and motorists that could have the same result. We discuss these issues in more detail below.

I/M compliance rate. The compliance rate is the fraction of vehicles that are subject to the emission test that are actually brought into compliance. Under this definition, compliance means that cars that fail must be repaired to the standards, and that cars which are outside the registration are brought into it. Compliance rates are likely to depend on the program stringency, the cost of repair and enforcement. For example, a more effective I/M program may reduce the number of vehicles that are able to avoid repair under the program, but it may increase the number of vehicles that attempt to avoid the program by not registering or registering outside the region.

Characteristics of I/M program. I/M programs are classified either as centralized test-only programs, in which a state or its contractor operates all the emission test stations and motorists have to go elsewhere for repairs; or decentralized test-and-repair programs, in which existing privately-owned repair shops are licensed to perform emission tests and may or may not repair the vehicles that fail the tests. As we discuss further below, the incentives facing mechanics in the two different types of programs are quite different I/M programs also differ in terms of test type. The IM240 test is forecast to be much more effective than its predecessors at identifying high-emitting vehicles and, on retest, at determining whether those vehicles have been successfully repaired.

Vehicle speed. Vehicle speed certainly makes a difference in emission rates. The California Air Resources Board estimates that a vehicle traveling 10 miles in 30 minutes will emit 2.5 times the running exhaust VOC emissions as one traveling the same distance in 11 minutes (see Burmich, 1989). In addition, vehicle speed is quite dependent on policy changes. Increasing the supply of roads or decreasing the demand for travel, for example through a policy of roadway pricing or parking fees, would very likely reduce the number of vehicles on the highway and, by improving traffic flow, reduce the emissions of those vehicles that remain. This adjustment, fortunately, is easy to handle in the current structure of the MOBILE model, because one of the important factors that users can adjust in the Model is the average vehicle speed.

The changes we consider here are incorporated into the emission factor model. Average speeds are inputs to MOBILE. In contrast, the behavioral effects that we describe affect model parameters that the user has little control over and that are often specified in the program code itself rather than as input data.

3.2. Behavioral Effects of Stochastic and Dynamic Elements of I/M

The three adjustments discussed at the end of the Chapter 2--for compliance rate, waiver rate and decentralized program discount--are largely behavioral, so MOBILE does recognize the importance of behavioral variables in I/M programs and attempts to take them into account. The problem is that these behavioral adjustments do not arise from structural elements that are built into the model, but are included in an *ad hoc* manner.

For example, a decentralized or test-and-repair I/M program (as opposed to a centralized test-only program) is subject to a 50 percent discount of its projected emission reductions, simply for being a decentralized program. The idea is that the incentives facing mechanics in test-and-repair programs are potentially perverse. It is easy to imagine how this could be so; on the one hand, shops would profit from requiring emission repairs, which the consumer would have no way of knowing were needed or not. On the other hand, a mechanic may have a long-term relationship with his customers and overlook excess emissions as a way to gain favor with them. In addition there are the incentives imposed by the "pass or don't pay" deals that are offered by many stations in California's Smog Check program.

But while it is reasonable to expect some problems with decentralized programs, MOBILE makes no attempt to model those problems explicitly or to consider whether there are measures that can be taken in a decentralized program to improve the incentive structure--for example, use of undercover repair vehicles, or comparison of failure rates across licensed stations. Instead it imposes a quite arbitrary discount. From a modeling standpoint the problem is that this is done as an across the board adjustment at the end of the program.

Because of the particular importance of vehicle identification, repair effectiveness, and the compliance rate we consider each of them in more detail.

3.2.1. Vehicle identification

As stated above, MOBILE and TECH include neither dynamic nor stochastic elements. Both models are in their essentials spreadsheets for calculating weighted averages of emissions under various conditions. Vehicle emissions are inherently stochastic, however, and it is important to understand the source and magnitude of the variation. In particular, we want to separate the variation in emissions due to the true variation in vehicle emissions and the variation due to errors in the emission test.

To facilitate the discussion we imagine a huge emission testing program involving multiple emission tests on a single vehicle, with tests being conducted on a number of different days during the year and with several tests being conducted on each of those days. Let X_{jk} denote the results of the kth test on the jth day (j=1, . . .T; k=1,…,n_j). The daily mean will be denoted $\overline{X}_{j.}$, i.e. $\overline{X}_{j.} = \sum X_{jk} \big/ n_j$, and the overall mean by $\overline{X}_{..}$. Each observation can be written in terms of the deviation from the daily mean:

$$X_{jk} = \overline{X}_{j.} + \varepsilon_{jk} .$$

Also each daily mean can be written as in terms of the deviation of the daily mean from the overall mean:

$$\overline{X}_{j.} = \overline{X}_{..} + \delta_j$$

Thus each observation can be written as the sum of the overall mean, the deviation of the daily mean from the overall mean, and the deviation of the observation from the daily mean.

$$\overline{X}_{jk} = \overline{X}_{..} + \delta_j + \varepsilon_{jk} \tag{3}$$

If there is high short-term correlation in true vehicle emissions, we can take true vehicle emissions to be constant on any given day, in which case ε_{jk} becomes a measure of the precision of the emission test. On the other hand δ_j is a measure of the differences in emissions from one day to another day, and most likely represent true variation in emissions from the vehicle. Each vehicle can thus be considered as subject to two kinds of errors: stochastic variation in vehicle emissions and test-specific measurement error, and (3) indicates what it might take to measure them. (There may be other specifications of the error that are not captured by (3); for example, the errors could be multiplicative rather than additive. That is a matter for empirical analysis.)

It would be useful to know the relative size of the variances of d_j and e_{jk} in (3). If d_j is relatively large, i.e. vehicle variation is large relative to test variation, then regulators should not place such a high degree of importance on test "identification rates." A perfect emission test could not prevent emission test results from being non-replicable. Furthermore, if one is to determine the mean emission rate of a vehicle, it may be more useful to have more numerous tests, rather than more accurate tests. Suppose the variance of d_j, the deviations around the true average emissions, is s_v^2. Suppose further that we can choose between two (unbiased) emission measurement methods: Method *i* (*i*=1,2) has a cost per test of c_i and measurement error variance of v_i. Finally suppose we have a fixed amount to spend on multiple tests using either method 1 or method 2. We want to compare the variances V_i of the mean test results, which can easily be shown to be the following:

$$\frac{V_1}{V_2} = \frac{c_1(v_1 + s_v^2)}{c_2(v_2 + s_v^2)} \tag{4}$$

If s_v^2 is large relative to the test variances, then the use of the less accurate but less expensive test is clearly more cost effective.

We are not aware whether EPA has ever conducted research to compare emission test variances with inherent emission variability. The implicit assumption is that vehicle variation is always small. That is certainly true for many if not most vehicles, since their emissions tend to be low under all conditions. And yet, the limited evidence provided by repeated tests on the same vehicle at approximately the same time shows that emission variation--on some cars, at any rate--cannot be explained by test variation alone. While vehicles with properly functioning emission control systems probably have consistently low emissions, the converse may not be true. Bishop, Stedman and Ashbaugh (1996) used emission test results from several sources, including FTP tests done as part of the Auto-Oil Program[13] to show that successive FTP tests on the same vehicles can have drastically different results. While some malfunctioning vehicles showed consistently high emission test results, in general, the greater the mean emission rate, the greater the variation as well. That is, vehicles with the greatest emission variability are the ones it is most important to identify in an I/M program.

The stochastic elements in emission testing suggest a number of important research topics and also suggest that there may be opportunities for behavioral responses by manufacturers, motorists and mechanics to emissions testing.

(1) Another reason for research into the relative size of vehicle and test variation is the oft-repeated observation that "x percent of the vehicles account for y percent of the

[13] This was the popular name of the Air Quality Improvement Research Program, a research effort undertaken in 1990 by a consortium of automobile and oil companies to examine the emission implications of fuel modifications specified in the 1990 Clean Air Act Amendments.

emissions." Actually there are three variances in play here: emission test variation, inter-vehicle variation and intra-vehicle variation. To understand how misleading statements of this sort can be, suppose that all the variation is intra-vehicle variation--i.e., emission measurements are perfect and all cars have the same emission mean and variance. Given a single emission measurement from each vehicle, we can sort the emission measurements in descending order and calculate the cumulative percentage of *measured* emissions accounted for by the first x percent of vehicles. It would be possible to conclude that vehicle emissions are highly skewed even though all vehicles are identical. Certainly, the variance in emission tests from one vehicle to another is far too great to be accounted for by either within-vehicle variation or measurement error. But while one can be confident that average emissions vary greatly across vehicles, the actual skewness of the fleet emission distribution cannot be determined with much precision until the relative magnitudes of these variances is better understood.

(2) The fact that EPA uses a predetermined driving cycle for the FTP and, to a lesser extent the IM240 test, gives manufacturers the opportunity and incentive to optimize their engines and emission control systems with respect to that particular driving cycle. A known and predictable driving cycle does provide a standard test that in principle at least allows replicability of test results. It is at least conceivable, however, that equipment optimized for this particular test cycle will not be the best design for a wider range of driving conditions, including conditions that motorists are more likely to encounter in everyday driving. Certainly part of the reason that enrichment events are now such a major cause of high emissions in new vehicles is that manufactures knew that they could design vehicles to a particular test cycle, and that high-acceleration events were not part of that cycle.

(3) If the test variation is large relative to the mean test result--i.e., a high signal-to-noise ratio--then motorists have a simple strategy for avoiding repair of high-emitting vehicles: Repeat the test until you pass. Given current practice in many states of not charging for a retest, motorists may repeat the test indefinitely; there is no way of determining at each visit to the testing station whether any serious repair attempts have been made. Obviously this strategy will not work for all vehicles, but in fact it is not known how often it will work. Examination of IM240 data for Arizona suggests that it is being employed on occasion, since there are vehicles that have appeared for testing more than five times. Moreover, of the vehicles that passed the first retest, ten percent retook the test within two hours, and most of those reported no costs of repair. What is not known is the number of ordinarily high-emitting vehicles that got lucky and passed a subsequent emission test without repair. Again, more precise emission tests may reduce the instance of this phenomenon, but it cannot eliminate it as long as vehicle emissions are themselves inherently variable.

(4) If the errors in emission tests are not correlated with vehicle characteristics, then repeated tests will uncover gross-emitting vehicles. This gives I/M programs a dynamic element: testing may not catch a vehicle during this testing cycle, but

another test in a year (or two years in a biennial program) will provide another chance to identify the vehicle. If these errors are truly independent, then simulations of I/M programs over time show that after several years, large differences in test identification rate do not translate into very large differences in program effectiveness (Harrington and McConnell 1994). In effect, the EPA method for estimating gains from I/M assumes the worst: that emission test errors and vehicle characteristics are perfectly correlated. If a vehicle is not identified as a gross emitter on one test, then it will never be so identified. Thus, the importance of the identification rate of the test may vary a great deal depending on the nature of the variation in emissions.

For these reasons it is important to understand the variability and error structure of emission testing and vehicle performance.

3.2.2. <u>Vehicle repair</u>

Repair of vehicles that have high emissions is a critical part of any program to reduce in-use emissions. However, there is little evidence about repair effectiveness even in most on-going I/M programs--data are difficult to collect, of uncertain quality, and not much analysis has been done on existing data sets. OMS has based all of the repair assumptions in MOBILE on repair done under controlled laboratory conditions, except when even laboratory repair did not bring the vehicle into compliance. In that case, EPA "adjusted" the repair results downward under the assumption that further repair would bring the vehicle into compliance. In fact, as we discuss in Chapter 4, studies of vehicle repair on gross-emitting vehicles find those vehicles to be very difficult to bring into compliance. Using laboratory data may represent some "ideal" in terms of repair potential, but may bear little resemblance to what happens in the real world.

EPA's reliance on laboratory data may also arise from a belief that existing data sets are irrelevant, because up to now mechanics have not been required to perform sophisticated repairs of emission systems, nor have they had access to the diagnostic information available from modern emission tests. While it may be true that mechanics in the field may not be trained as well as nor have the same experience *at present* as lab mechanics, it is not unreasonable to assume that after the onset of the I/M program, mechanic learning would soon bring the quality of repairs up to the "ideal" level. That is what might be expected to happen in a competitive economic environment: mechanics are driven by competitive pressure to learn how to repair vehicles, and those who do not learn will not fare as well in the repair market

However, there are a number of reasons to doubt this optimistic picture. First, it is not clear that "ordinary" auto repair--repairs that produce private benefits for the motorist--fits the competitive model very well. In several ways the economics of car repair resemble the economics of medical care, which is another market badly afflicted with market failures. In the competitive model, consumers decide on the amount and quality of the product to buy, but in medicine or car repair that decision, though made by the consumer, is strongly influenced by the expertise of the supplier of the service. That is, the information asymmetry generates an interaction between supply and demand that violates the competitive model. It is also difficult for consumers to observe differences in mechanic quality, since many repairs are *sui generis,* in which case the consumer cannot compare the outcome of using mechanic A with what the outcome would have been had mechanic B been used. In addition, there are several sources of uncertainty that will make it difficult to separate signal from noise. For example, sometimes it is not possible to arrive at a conclusive diagnosis, so that even good mechanics sometimes fix the wrong thing. Small wonder that auto repair is always ranked near the top of the list of industries generating consumer complaints.

Repair of emission systems are subject to all these difficulties plus a few additional ones. The main difference is that the motorist receives no private benefit from emission repair, or at least not enough to justify the repair cost. (Some repairs that reduce emissions also yield improved driveability or fuel economy.) In addition, the improvement in emissions can only be observed at the repair facility with an error that is likely to be greater than the error at the emission test station, inasmuch as repair shops are likely to have less sophisticated emission test equipment. This difficulty of determining whether the vehicle has been repaired satisfactorily can lead to repeated repair attempts and emission tests even if the motorist is trying honestly to get the vehicle repaired.

Finally, the relative unfamiliarity of emission repair, combined with the difficulty of distinguishing between good and bad repair facilities, mean that mechanics are likely to be poorly trained at the outset and improve more slowly than would be the case with other sorts of repair.

These characteristics of repair, and of emission repair in particular, affect the incentives and therefore the behavior of motorists and mechanics. But although real-world outcomes depend crucially on motorist and mechanic behavior, the interactions between motorist and mechanic and between each and the I/M test are complex and poorly understood. The motorist may have altruistic motives and would be eager to comply with emission regulations. However, this group is not why I/M is needed; I/M is directed against those who are motivated by private gain and who therefore need enforcement to comply with regulations.

In contrast to a laboratory mechanic, who may simply have emission reduction as the only objective, never mind the cost, mechanics in the field may operate from a mix of motives. Like motorists, they may be influenced by environmental motives or they may seek economic gain. Even if they pursue their private interest, it is not clear what the most profitable behavior is. For example, they may try to sell motorists on the most expensive repair under the waiver limit. Or, they may identify their interest with that of the motorist, especially for motorists with whom they have a long-term relationship.

These complexities mean that the data on repair costs under older I/M programs or the special repair programs conducted by several oil companies (discussed in the next section) are far from irrelevant. Now that Enhanced I/M has begun to be implemented in some states, it is even more important to collect data from the field on repair. For one thing, the resulting information can be disseminated to mechanics and result in more effective and lower cost repairs. However, repair facilities are unlikely to collect accurate repair data unless there is some incentive to do so.

The MOBILE model has tended to ignore these behavioral aspects of repair, and has instead based repair effectiveness assumptions on EPA laboratory test and repair results. The EPA has focused on technical solutions to problems of repair, such as mechanic training and the provision of information to mechanics. While information may be important, other incentives may also play a role in overall repair effectiveness.

3.2.3. Compliance

The extent to which motorists comply with an I/M program is an important determinant of the effectiveness of a program. The decision about whether or how to comply with I/M is clearly behavioral, and that decision is likely to depend on a number of different factors, some of which are a function of the I/M program itself. It would be likely to depend on how difficult or expensive it is to comply with the requirements of the program, and what the penalties are for non-compliance. Most states now use registration enforcement, requiring vehicles to be in compliance with the emission test in order to obtain their registrations. This is probably an effective enforcement mechanism since the penalties for driving without a registration are fairly high in most states. The expense or difficulty of complying are likely to depend on the cutpoints of the I/M test in place, the waiver rate, and the characteristics of the test system (whether is it centralized, the number of stations, ease of getting repairs, etc.).

A vehicle is noncompliant if it fails to pass the emission test, fails to receive an emission waiver, and yet continues to be driven in the area. In order to define compliance it is necessary to consider all the possible ways that motorists can respond to registration and I/M requirements. Figure 2 shows the possibilities: a noncompliant vehicle is either unregistered, never took the I/M test, or took it, failed and never passed on subsequent tests. Vehicle owners may simply be taking a chance that their failure to register or to obtain a valid I/M sticker will not be discovered. Some owners have the option of reducing the probability of discovery by registering the vehicle outside the I/M area, even though they continue to drive it there. Or they may sell the vehicle to someone who does the same thing. Non compliance does not include all the vehicles on the road that fail to meet emission standards. Some may be "false passes"; others may have failed and received a temporary repair that enables them to pass the retest even though the reduction in emissions is illusory or temporary. Presumably, more accurate emission tests like the IM240 test reduce the probability of such occurrences.[14]

---

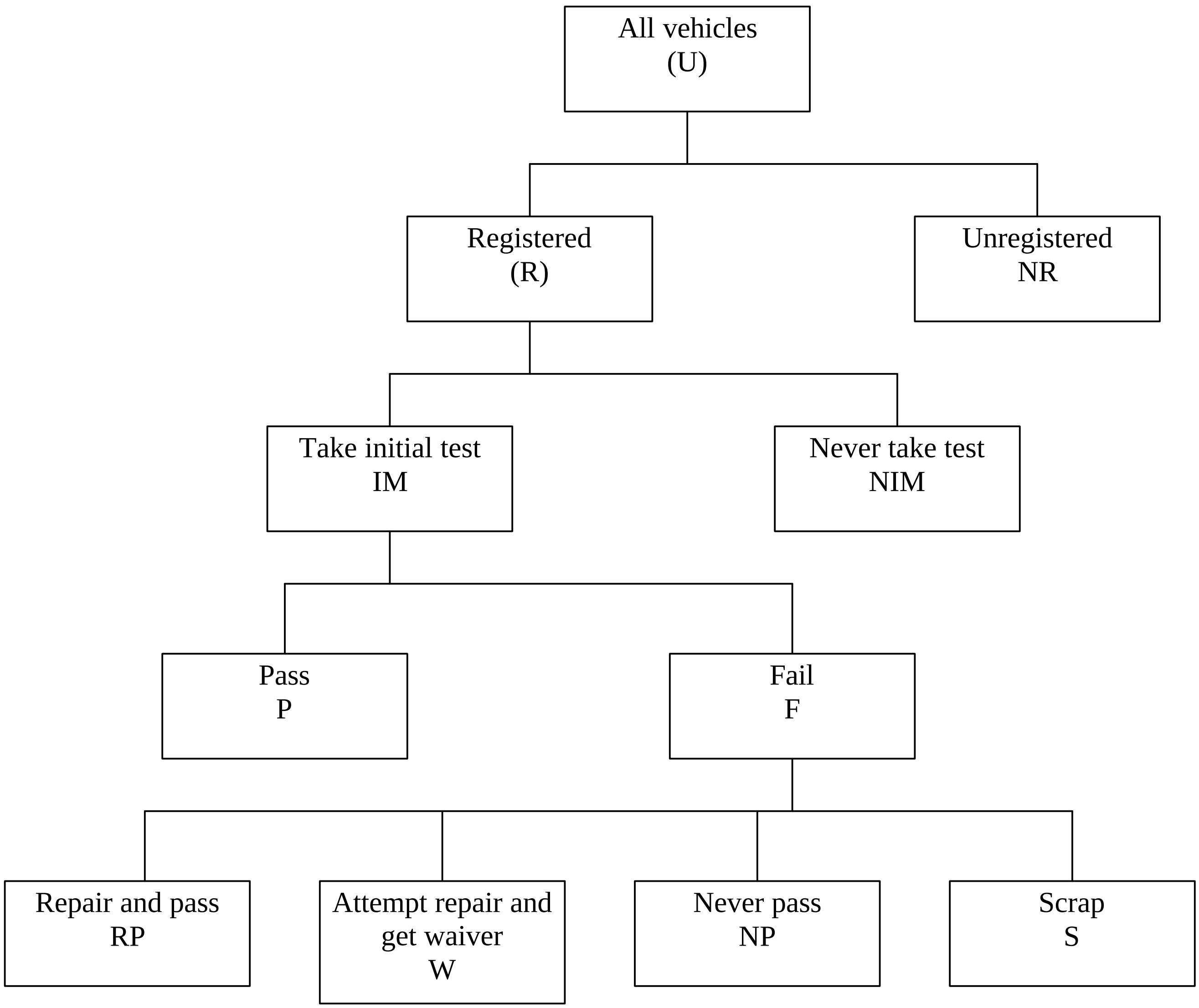
[14] Several on-road evaluations of I/M programs conducted prior to the introduction of the IM240 Test have suggested that emission reductions attributable to I/M were being overestimated (Scherer and Kittelson, 1994, Scherer 1996, Lawson et al. 1990). Because of the newness of Enhanced I/M programs using the IM240 test, there has been only one on-road evaluation that we are aware of. The Colorado Enhanced I/M program, which began in January 1995, was evaluated about a year later using data collected on the streets of Denver by remote sensing (Stedman et al. 1997). While emission reduction estimates made using the I/M retest data were 23, 18 and –2 percent for HC, CO and NOx, respectively, Stedman et al. found no emission reduction for HC and NOx and 4 to 7 percent reductions for CO.

In any case the noncompliance rate, using the symbols defined in Figure 1, is

$$\text{Noncompliance rate} = \frac{NR + NIM + NP}{U} \qquad (5)$$

It can be difficult for states to estimate each component of the numerator of. States may not have a good estimate of the number of unregistered vehicles, and registration records often do not fully reflect vehicles that have been scrapped or which are removed from the area. Even for registered vehicles, the data links between motor vehicle agencies and environmental agencies are not available to determine which have passed I/M, and vehicles with incomplete repairs are difficult to track. Finding the vehicles that are no longer registered in the area but continue to be driven there may be most difficult of all, since remote sensing technology is the only current way to identify such cars.

Figure 2. Classification of Vehicles Subject to I/M



CHAPTER 4

COMPARISON OF MOBILE ASSUMPTIONS ABOUT I/M TO EVIDENCE FROM OTHER DATA SOURCES

There are many aspects of modeling I/M in the TECH and MOBILE Models where the behavior of motorists or mechanics may be important. We identify three such areas--repair effectiveness, compliance and tampering--and examine in some detail how these are handled in either the TECH or MOBILE Models. In all of these areas, the underlying assumptions in the Model are not evident to the user, but are often imbedded in the program code. We have tried to discern where and how the various assumptions are made, and when possible, describe the sources of the underlying data on which the assumptions are based. Then, because evidence from actual I/M programs reflect all aspects of the I/M process including the underlying behavior of motorists and mechanics, we summarize actual data from the Arizona I/M program and other repair studies and compare them to the EPA results. Examining and evaluating data from a number of sources will shed light on the validity of the assumptions and the accuracy of the MOBILE Model results.

We first take an in-depth look at what is assumed about repair in the Model. We look at what the Model assumes about repair effectiveness and at the EPA dataset on which the assumptions are based. We then compare repair effectiveness assumed in the Model to the evidence from on-going I/M programs and various repair studies. We also look at some important issues that may have been left out of the repair effectiveness component of the Model, such as the cost of repair and joint pollution issues. We conclude the repair section with a summary regression analysis, in which we look at the determinants of repair effectiveness across the different datasets.

Next, we look at the assumptions in MOBILE about compliance with I/M. There is growing evidence that motorists have a number of ways of not repairing their failing vehicles. We look at how compliance is handled in the Model and compare that to evidence from the Arizona I/M program. Finally, we look at the assumptions about how tampering affects fleet emissions within the Model. We then present some preliminary evidence about tampering from Arizona.

4.1. Repair Effectiveness

How well failing vehicles get repaired is one of the most important determinants of I/M effectiveness. Repair effectiveness is a critical component of how TECH models I/M programs, but little is known about this part of the Model since the assumptions about repair effectiveness are built into the code of the program itself.[15] Repair effectiveness assumptions are not observable to the user and certainly not possible for the user to modify easily. Here we examine how the repair component of the TECH model works, what the assumptions about repair are, and at the underlying data used to make these assumptions. We then compare assumptions about repair effectiveness in the Model to evidence about repair from the real world by looking at a number of different datasets.

4.1.1. EPA Repair Dataset

The evidence upon which the repair component is based is from data collected on repair from several EPA labs, which we refer to as the EPA repair dataset. There are 266 vehicles in this dataset, all from the Emission Factor testing program, which has been recruiting vehicles for testing and repair from 1977 to the present. The vehicles included in the larger dataset are supposed to reflect a representative sample of in-use vehicles.[16] Most of the vehicles in the repair dataset were chosen because they had FTP scores above some level, although in recent years IM240 scores were used for flagging vehicles. There were a limited number of vehicles repaired each year, however, so not all vehicles which met the test criteria were repaired and included in the dataset. And, the cutpoints or criteria used to recruit vehicles varied from year to year. Hence, it is not at all clear that the repair dataset is a random sample of high emitting vehicles.

The repairs were done over a period of years (1977 to 1993) at contractor labs in Hammond, Indiana or Ann Arbor, Michigan and at the National Fuel Emission Laboratory in Ann Arbor, Michigan. Mechanics were directed to perform only those repairs that had been identified by a thorough visual and functional inspection of each vehicle after the FTP test. They were directed to repair everything they could from the list of what was wrong, and they did not target particular cutpoints or check emissions levels during repairs. They also did not pay attention to the potential cost of repair, except that they were told not to repair certain parts such as the catalyst. There are several reasons for this. First, new catalysts often take several hundred miles of driving before their emissions are representative of after-repair emissions, so after repair emissions would not be easy to measure. Second, catalysts are the most expensive repair, and the costs may have exceeded the waiver limit in most states at the time the repair work was being done.[17] The repair dataset contains information on pre- and post-repair FTP and IM240 emissions and is used to estimate changes in HC and CO emissions resulting from vehicle repairs (266 vehicles). An additional set of vehicles with

---

[15] Data are included in the file DOTS.INP in TECH. Emissions before and after repair are specified for three pollutants, for different I/M program types (IM240, idle and two speed idle at different cutpoints) for different classes of vehicles (LDV, LDT) and for each emitter group (super, very high, high and normal) is fed into the program as input data.

[16] There are no rentals, commercial vehicles or self-selected vehicles in the sample.

[17] However, data from Arizona indicates that catalyst repair or replacement occurred in about 7% of the cars that got repaired in 1995 or 1996. The costs of this type of repair appears to be under $200 for many cars. To the extent that some of these are catalyst replacements, of repair under the higher waiver limits in place in states today.

high NOx emissions was added to the dataset to estimate the effect of repairs on NOx emissions (27 additional NOx vehicles).

Estimation of Repair Effectiveness in the EPA Repair Dataset. In estimating repair effectiveness, the impact of repairs on HC/CO emissions was analyzed independently from the impact on NOx emissions.[18] Estimation of the repair effectiveness of an I/M program having a specific HC/CO cutpoint utilized the subset of the 266 vehicles with initial IM240 emissions above the given cutpoint. The impact of repairs was then estimated through a comparison of pre- and post-repair FTP emissions of those vehicles which had failed the IM240 test. In the case of vehicles having post-repair IM240 emissions below the given cutpoint, the observed pre-and post-repair FTP emissions were used as the measure of repair effectiveness.

For vehicles having IM240 emissions after repair that were above the cutpoint, EPA adjusted the emissions downward to reflect the fact that vehicles are required to pass the I/M test. The adjustment is complicated by the fact that repair effectiveness is measured by comparison of the before and after FTP readings, so an after repair FTP reading that is consistent with the adjusted IM240 was needed. The way the EPA adjusts the IM240 reading and makes the conversion to an FTP score is described in Appendix C. In the EPA dataset we are working with,[19] 46% of the vehicles had IM240 post repair readings above the standards (or cutpoints) for at least one pollutant and had to have their emissions adjusted downward.

Once before and after repair (observed or adjusted) FTP emissions were determined, vehicles were separated into four emitter groups (Normal, High, Very High, Super) based on initial FTP emissions. For each category, mean initial and final FTP emissions were determined. These values provided four points showing pre- and post-repair emissions which are used to create a piecewise linear function from which post-repair emissions can be determined from a pre-repair emission measurement. Figure 3 gives an example of the function for post-83 light duty gasoline vehicles in an IM240 test with cutpoints of 0.8 HC and 15 CO.

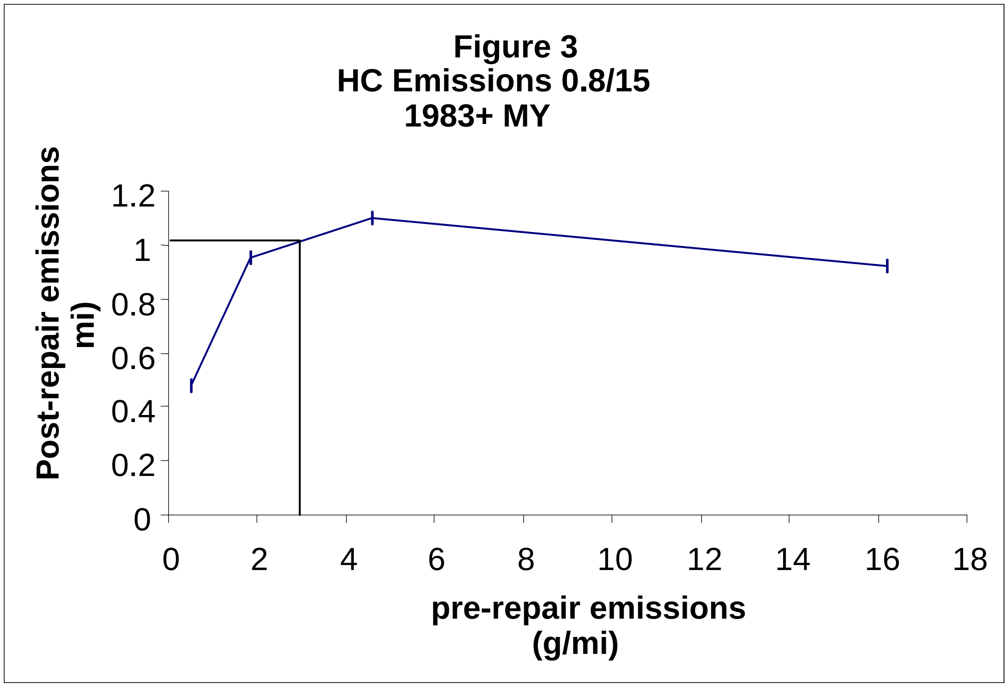
In modeling repair effectiveness within the TECH model, it is assumed that repairs are not made exactly to the cutpoint in question, but rather are made with a margin of error. Vehicles are repaired to a level below the cutpoint. To do this, repair effects corresponding to tighter cutpoints are used in modeling the effect of any given cutpoint. For example, the repair effects corresponding to .8/15 HC/CO cutpoints (such as that shown in Figure 3) are used to model the effects of a program that actually has looser 1.2 HC/20 CO cutpoints. Vehicles in this program are then assumed to be repaired as shown in Figure 3. For example, for a 1.2 g/mile HC cutpoint, super emitters are assumed to have 16.2 grams/mile HC before repair, and they are all repaired (assuming a 100% identification rate) to .92 grams per mile,

---

[18] Repairs that reduce HC or CO emissions through increases in engine efficiency can result in increases in NOx emissions. The model accounts for this fact by including NOx disbenefits for I/M tests that do not measure NOx emissions. Overall, the disbenefits are small.

[19] The data we obtained from the EPA corresponds to .8/15 HC/CO cutpoints. In that dataset, there were 32 vehicles with adjustments to both final HC and final CO, 42 with adjustments to final HC only, 3 with adjustments to final CO only, and 189 which were repaired to the standard and needed no adjustment.

as shown in Figure 3. The point is that this type of functional relationship between pre-repair and post-repair effectiveness is what is used in the model to determine I/M effectiveness. How well this relationship depicts actual repair effectiveness has not been confirmed through empirical analysis.



There are a number of other potential issues with the repair effectiveness modeling and parameter assumptions described above.

- Repair effectiveness is based on EPA lab results, and as a result may be optimistic compared to what mechanics in the field can achieve. We examine this possibility in more detail below.
- The repair effectiveness parameters for the IM240 test are based on a relatively small dataset - there are only 266 vehicles in total. These parameters purport to reflect variation in model year, technology type, cutpoint, etc. However, looking at the actual data in the repair effectiveness block data file in TECH shows that there is no variation due to technology type or model year. The only variation is by cutpoint or I/M test regime. This is probably due to the paucity of data on repair effects in the EPA dataset. In addition, the variation in repair effectiveness among I/M test regimes does not seem to be based on much empirical evidence. In fact, the repair effectiveness of the 2500 idle test is based on repairs of only 36 vehicles

(cars and trucks). Finally, it is not clear that the 266 vehicle IM240 data set sheds much light on repair effectiveness under different cutpoints. Mechanics were not required to repair the vehicles to specific cutpoints, so it is not clear that this data can be used to infer what would happen in the field under tighter cutpoints.

- Joint pollution reduction impacts may need to be examined more carefully in the Model. The EPA dataset assumes HC and CO reductions occur jointly, and NOx reductions are independent (except for small NOx disbenefits associated with HC/CO reductions from I/M programs that do not measure NOx)**.** Repair may be more difficult in some cases than others. For example, when all three pollutant levels are high, and reductions have to be made in all three, there is some evidence that it is more difficult to make reductions. In any case the evidence about joint reductions are complex and may need to be accounted for in repair effectiveness (more discussion below).
- The repair component in the TECH model assumes that all vehicles will be repaired to the standard, regardless of repair costs. Repair effectiveness depends only on I/M cutpoints and vehicle type. However, there is evidence from programs in operation that not all vehicles are repaired to their designated cutpoint. There is also little empirical evidence linking more stringent I/M cutpoints to actual improved repair effectiveness. Comparison of IM240 programs in different states may soon allow some conclusions to be drawn about this.
- The model assumes no relationship between current repairs and the probability of repair in the future. The nature of this relationship, in fact, is not an issue that the current TECH model can address since it is not a dynamic model. Implicitly, the assumption in the Model is that even after vehicles have been repaired, they have the same probability of becoming excess emitters again as any other similar type vehicle in the fleet.

4.1.2. Comparison of EPA Repair Data and Modeling to Evidence from other Repair Data

It is important to compare the EPA repair dataset and assumptions to evidence from other studies, especially those that represent conditions in the real world. This will allow us to determine how well the MOBILE Model is capturing what occurs in practice and to examine the causes of any differences. Here we focus on a comparison of the repair effectiveness from the EPA repair dataset and data from other studies.

Until recently there was very little evidence from I/M programs about repair effectiveness. There are now a handful of empirical studies that have focused on vehicle repair. In addition there is the growing body of data from State IM240 programs, some of which have repair data. Here we examine the evidence from the California Pilot Project done in 1994, and two earlier studies, the Sun Oil Company Study, and the California I/M Review Committee study, and we have obtained data from Arizona's IM240 program which includes information on what was repaired for failing cars and the cost of repairs. We describe each of these studies below and they are summarized in Table 1 and Figure 4 (and in Appendices D and E).

Arizona I/M240 Program. The State of Arizona has had an I/M program in place which includes an IM240 test for post 1980 vehicles since 1995. We have obtained data on all vehicle tests from this program for January 1, 1995 through May 1996. Like many states, Arizona's test protocol allows for vehicles to "fast pass" in less than the full 240 seconds of the test. However, Arizona does require a random sample of 2% of the fleet to complete the entire test. We have selected the failing vehicles from the 2% random sample of vehicles which have full tests to analyze repair issues in the Arizona program. Arizona's program is a centralized program that is registration enforced, and with a $450 waiver for vehicles newer than a 1979 model year.

Sun Oil Company study. The study was undertaken by Sun Company, Inc. in 1992 in order to determine the feasibility of earning emission credits through the repair of high emitting vehicles. Vehicles identified as having high emissions through the use of remote sensing devices were requested to participate in the program. Participating vehicles having emissions in excess of 1997 EPA-mandated standards as measured through an IM240 test were repaired until emissions were reduced to the standards or costs exceeded $450. Only one attempt at repairs was permitted. Repairs were performed at a select group of Sunoco Ultra Service Centers where participating dealers had been given special training in diagnosing and repairing faulty emissions control systems. The dataset contains information on pre- and post-repair IM240 emissions.

California I/M Review Committee. California conducted an undercover study of 1100 cars that were recruited in California in 1992 as the central part of an assessment of the performance of California's "Smog Check" program. This assessment was done by the California I/M Review Committee, which was set up by the legislature to study the State's troubled I/M program. The program attempted to recruit a large sample of vehicles in use, but there has been some dispute over how random the final sample was (Aroesty et al., 1993). The vehicles were given an initial FTP test, and those that failed the FTP were sent out to a sample of Smog Check stations in Southern California as if they were cars out to get their required Smog Check certificates. These "undercover" cars were given emission tests by the (presumably) unsuspecting Smog Check stations and if failing, were repaired and retested. The cars were then given a post-repair FTP test. Of approximately 1100 vehicles originally included in the program, the analysis described here is of 681 vehicles for which repairs were attempted and a second FTP completed.[20]

[20]We are indebted to Douglas Lawson of the Desert Research Institute for providing us with this data set. For more information see Lawson (1993).

Table 1. Comparison of EPA Repair Dataset to Other Evidence about Vehicle Emission Repair							
	EPA FTP [1]	**EPA IM240 [1]**	**Arizona IM240**	**CA Pilot Project FTP [2]**	**CA Pilot Project IM240 [2]**	**Sun Oil IM240**	**CA I/M Review Committee FTP**
Number of observations	266	266	5909	199	201	155	681
Years repair took place	?	?	1995-96	1994	1994	1992	1992
Model years repaired	1981-90	1981-90	1981-96	1968-93	1968-93	1969-92	1966-89
Average Model Year	1985	1985	1986	1983	1983	1984	1979
HC Emissions (g/mile)							
Pre-repair	3.13	1.84	2.69	3.34	3.03	4.74	4.95
Post-repair	1.24	0.69	1.7	1.65	1.33	1.56	3.71
Change in emissions	1.88	1.15	1	1.69	1.7	3.18	1.24
Average improvement	60.4%	62.5%	36.8%	50.6%	56.1%	67.1%	25.1%
CO Emissions (g/mile)							
Pre-repair	44.76	32.98	40.35	35.88	32.1	68.31	48.44
Post-repair	12.69	9.69	25.65	20.82	17.47	17.69	44.43
Change in emissions	32.04	23.26	14.7	15.06	14.63	50.62	7.01
Average improvement	71.6%	70.6%	36.4%	42.0%	45.6%	74.1%	8.3%
NOx Emissions (g/mile)							
Pre-repair	-	-	3.14	2.05	2.5	2.83	2.13
Post-repair	-	-	2.24	1.23	1.38	2	1.89
Change in emissions		-	0.9	0.82	1.11	0.82	0.24
Average improvement	-	-	28.7%	40.0%	44.8%	29.3%	11.3%
Average cost	N/A	N/A	$172.85	$305.50	$305.50[4]	$336.00	$90.00
Average number of retests	N/A	N/A	1.52	1.35	1.35[5]	N/A	N/A

[1] Data sample used to estimate changes in HC/CO emissions resulting from repairs in MOBILE. 27 vehicles with high NOx emissions were added to this dataset to estimate the effect of vehicle repairs on NOx emissions.

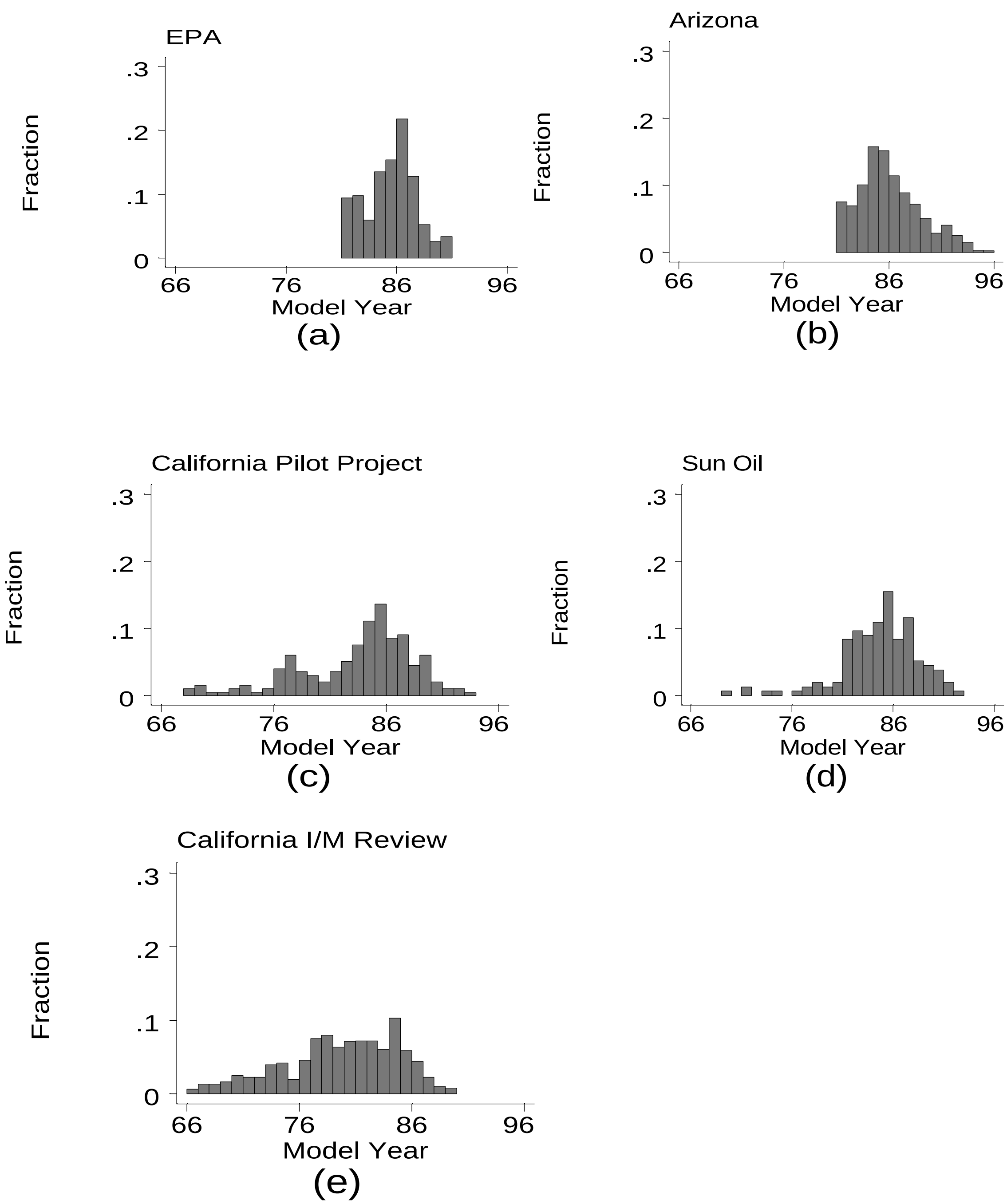
[2] Pre- and post-repair emissions as recorded at CARB's Haagen-Smit Laboratory. Emissions recorded after extended CARB repairs were not included in mean post-repair emissions.

[3] Costs are comprised of repair costs ($170.04) and tampering costs ($2.82).

[4] Costs are comprised of repair costs ($215.25) and tampering costs ($90.25) occurring at Clayton facility. Average cost based on 96 vehicles in the IM240 bin.

[5] Repair round at Clayton facility.

Figure 4. Distribution of Vehicles by Model Year, Repair Datasets



California Pilot Project. This study was conducted by CARB between June and December 1994 in order to compare the effectiveness of IM240 and ASM testing procedures in identifying excess emissions and reductions in excess emissions as a result of vehicle repairs. Vehicles were required to participate in the study through California Senate Bill 2018. Initial tests were performed at CARB and failing vehicles were assigned to IM240 or ASM bins. Repairs were performed by BAR employed licensed Smog Check mechanics at two repair bays established at Clayton Industries in Sacramento for the purpose of the study. The mechanics received training on ASM and IM240 testing and on vehicle diagnosis and were asked to repair vehicles until emissions were below the standard or repair costs exceeded $500. Multiple rounds of repairs and testing are included in the dataset. The dataset we use in the analysis here contains FTP and IM240 emissions for those vehicles that failed the IM240 test.

We want to compare MOBILE repair effectiveness assumptions against the evidence about repair effectiveness from these other studies. All of the datasets are made up of vehicles that are found to be in need of repair because their emissions are higher than some specified emission rate; all are made up of failing vehicles as defined in the respective programs. The EPA dataset we are using is the raw data from the 266 vehicles repaired in EPA labs--we are not using the adjusted data.

Among the datasets, there are some important differences in which model years vehicles are repaired, as shown in Figure 4. The EPA repair dataset and the Arizona data set include only post 1980 vehicles. Arizona only performs IM240 tests on post -1980 vehicles, continuing to give the idle test to pre-1981 vehicles. EPA handles pre-1981 vehicles separately in the MOBILE model. In contrast, the other three datasets--the California Pilot Project, the Sun Oil Study, and the California I/M Review Committee Study all have substantial numbers of older cars. Another important difference is that the Pilot project was done in Sacramento, California where there was no I/M program in place, and the Arizona and California I/M Review Committee evaluations were of regions where an I/M program had been in place for a period of time. The TECH Model structure is such that the emissions reductions are from a no I/M case to an I/M case. It should, therefore, be more similar to the Pilot Project results.

Table 1 provides summary statistics of all of the repair datasets. The Arizona program provides by far the largest number of observations with 5,909 failing vehicles from the random sample of cars during 1995 and the first half of 1996. The other datasets range from several hundred to 681 in the California I/M Review Committee study. Before comparing the results from the different datasets it is important to note some differences between them.

The cutpoints are different for each study, and it is unclear what effect this will have on the results. In some cases, mechanics were aware of a specific target cutpoint for repair, in other cases they were not. The cutpoints for each study are summarized in Table E-1 of Appendix E.

The directions to repair technicians were different in each case. In the California I/M Review Committee study and in Arizona, mechanics are aware that their clients are trying to pass an emissions test. In the EPA analysis, it is not clear that mechanics were given identical instructions through time (the data were collected over a period of time). In the case of the Sun Oil Company and the EPA repairs, only one round of repairs was allowed. In the Pilot Project and the Arizona IM240 program, multiple rounds of repair are allowed. In the undercover car study done for the California I/M review committee, each vehicle received a second FTP test in CARB labs after obtaining a Smog Check certificate from the Smog Check station.

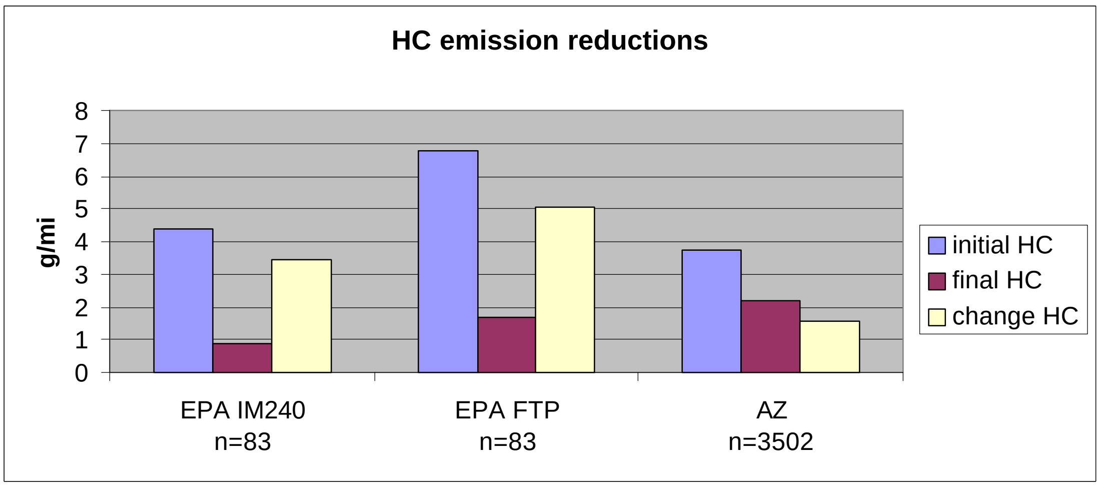
Table 1 summarizes the repair effectiveness from the various programs. EPA estimates appear to be optimistic. The percentage reduction for HC and CO are the smallest for the two datasets from actual I/M programs: the Arizona IM240 and the California I/M Review Committee study. The Arizona program provides perhaps the best real world benchmark to which the EPA dataset should be compared. For both HC and CO, the Arizona program has emission reductions that are about half of that predicted from the EPA repair dataset. In addition, the IM240 readings for HC and CO in Arizona are above those from the EPA data.[21]

The results of the Pilot Project look more similar to the EPA results for HC and CO reductions than does the Arizona data. The California I/M Review Committee study of the Smog Check program makes that program look the least effective, though the limit on what had to spent to repair a car was the lowest, at less than $100. The average cost of repair for that program was only $90, whereas for Arizona it was almost double that. The average costs for the laboratory programs was higher, at about $300 per vehicle. However, even in these programs there was a limit on what mechanics were supposed to spend to bring a vehicle into compliance - about $450 or repair per vehicle. To make full repairs might have been even more expensive.

Figures 5 and 6 focus on HC and CO emissions reductions as a result of repair from several of the datasets. Figure 5 compares repair effectiveness in the EPA dataset to the results in Arizona. The EPA datasets show much larger emission reduction compared to what is happening in Arizona. In addition, recall that the MOBILE Model adjusts the FTP results farther downward from the levels we observe here to build in the result that vehicles are repaired to pass at whatever cutpoints are in effect. Figure 6 compares the EPA and Pilot Project results. The repair effectiveness is quite similar to the EPA dataset results.

[21] The NOx readings from the EPA dataset were incomplete and so we do not show NOx changes in Table 1. Twenty-seven additional vehicles that failed NOx were added to the dataset to determine NOx reductions for the MOBILE model, but we did not have that full dataset.

Figure 5. Comparison of Emission Reductions: EPA Repair data set vs. Arizona IM240 Program



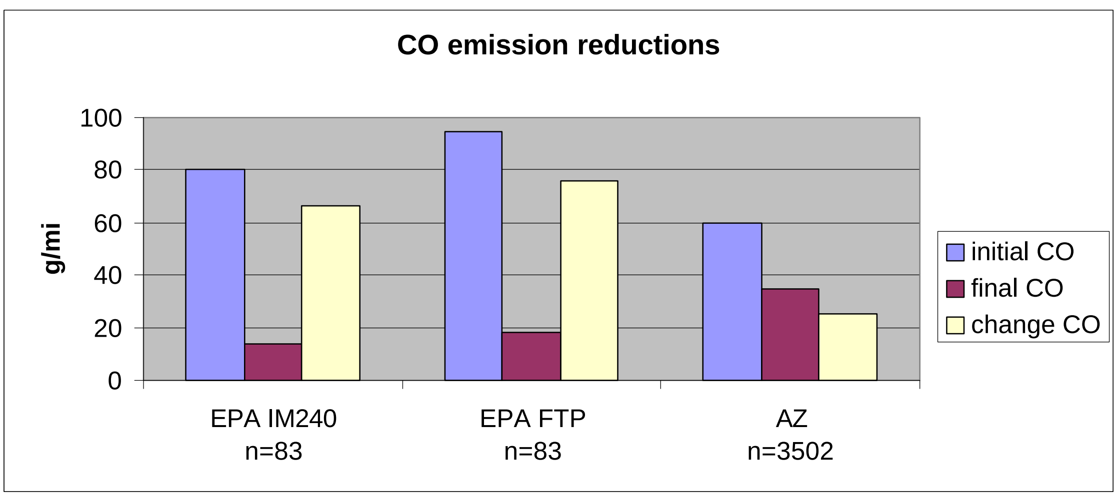
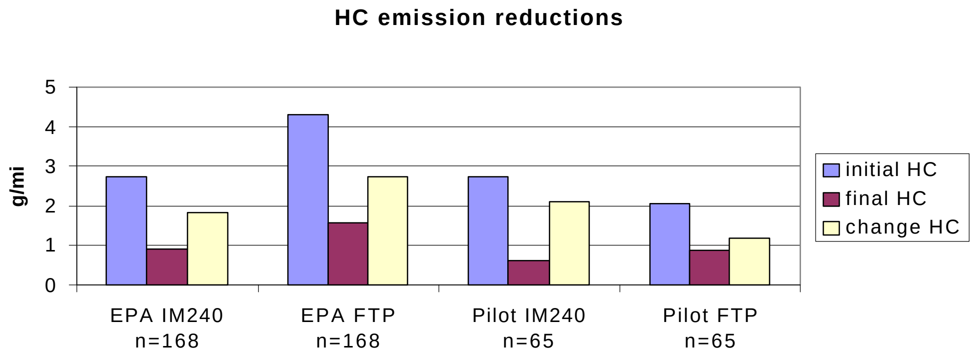
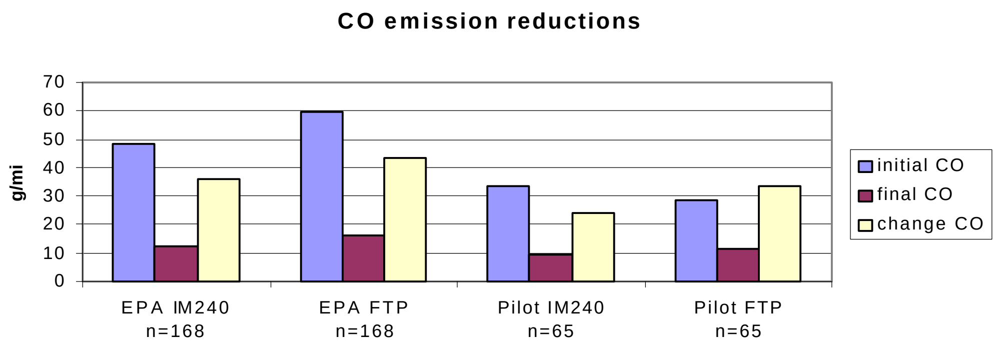


Figure 6 Average HC Emissions of Vehicles Receiving Repairs with Initial Test Emissions above Pilot Cutpoints for HC/CO/Nox





4.1.3. Emitter Group Categories and Repair Effectiveness in the TECH Model

A critical aspect of repair effectiveness in the TECH model is how the emissions distribution of the fleet is handled. This aspect of the model is essentially where the emissions "deterioration rates" (how emissions change as vehicles age) are determined. The emissions profile of the vehicle fleet and how that profile changes are, it seems to us, the central element of how an I/M program works. Here, we examine how the TECH model handles this issue and compare the Model assumptions to what we observe about the emissions distribution in the Arizona fleet.

How repair of high emitting vehicles changes the fleet emissions characteristics, and whether and how long those emission reductions last would seem to be important for determining the effectiveness of I/M. The current TECH model makes strong assumptions about repair effectiveness of each emitter group, and assumes emitter group emissions and repair effectiveness are the same regardless of earlier I/M testing.

Repairs may be more expensive or more difficult with certain emitter categories -- this fact may influence motorist and mechanic behavior. It may also suggest that certain policies are more effective or cost effective (e.g. fixing high, compared to very high or super) emitting vehicles may be very expensive and produce only small emission reduction; or repair of some super emitting vehicles may be so expensive that either vehicles are scrapped, changing the emission distribution, or emission reduction to the standard will only occur with some type of repair subsidy).

In the current TECH model, the emissions distributions among supers, very highs, highs and normals is encoded within the model. The user is not aware what that distribution is, and there is no ability or incentive to test those distributions in the real world.

We attempt here to compare the repair effectiveness assumptions for different emitter group categories. We want to compare the assumptions imbedded within the TECH model to what we observe in the Arizona I/M program.

This is difficult because the emitter categories from the EPA model are based on FTP results, and only IM240 results are available from Arizona. In the TECH model vehicles are split into emitter groups for each vehicle type and standard. The splits between groups were originally based on vehicle emissions levels relative to the standard the car was certified to.[22]

The divisions between emitter groups used in the TECH model are based on FTP readings as defined below. Figure 7 shows how the divisions between emitter categories changes as vehicles age. The first listing in the key shows the designation for HC/CO (such as high, very high, etc); the second listing is for the NOx level which is either high or low.

---

[22] Normals: less than or equal to 2 times the standard for HC and less than or equal to 3 times the standard for CO; Highs: between 2 and 4 times the standard for HC and/or between 3 and 4 times the standard for CO; Very Highs: greater than 4 times the standard for HC but less than or equal to 10 g/mi and/or greater than 4 times the standard for CO but less than or equal to 150 g/mile CO; Supers: greater than 10 g/mile HC and or greater than 150 g/mile CO.

High: greater than 0.82 g/mi HC or 10.2 g/mi CO but less than Very High
Very High: greater than 1.64 g/mi HC or 13.6 g/mi CO but less than Super
Super: greater than 10.0 g/mi HC or 150.0 g/mi CO

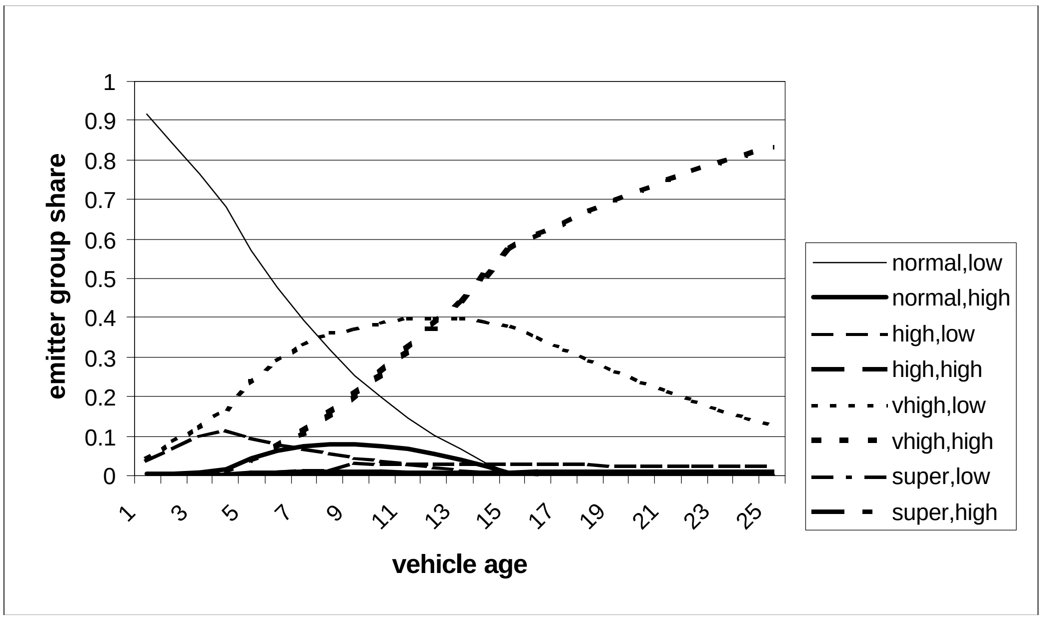


Figure 7. Emitter Group Share by Vehicle Age Assumed by EPA for Closed-Loop Multi Port Fuel Injected LDGV

We want to compare the EPA assumptions about how emissions are reduced under I/M for each emitter group and compare that to what is observed in Arizona and in the California Pilot Project. We do this for HC/CO definitions of emitter categories since our dataset for NOx is incomplete. The Pilot Project has FTP readings and emitter groups can be split by the same definition as used for the EPA data. However, only IM240 results are available in Arizona. Because the IM240 is one component of the FTP, test results from the two tests do not produce identical results. The problem is in determining where the emitter groups should be split. We use two different methods to split the Arizona dataset. The third column of Table 2 shows the results from the Arizona program when the EPA definitions of the emitter groups are used and applied without change to the IM240 results; a super emitter has IM240 emissions above 10 grams per mile HC or above 150 grams per mile CO. The second column shows an alternative method using regression analysis to convert the splits between the emitter groups to IM240 equivalent readings. Linear regressions of IM240 emissions on FTP emissions for each pollutant produce a fairly good fit (R^2=0.85. See Appendix F for sample regressions). Column 2 uses these regression results to predict IM240 levels that correspond to the divisions between emitter groups. We then apply these divisions to the Arizona data and determine the number of vehicles and the emissions rates corresponding to the different emitter groups.

Table 2. Emitter Groups and Repair Effectiveness: MOBILE Assumptions Compared to Evidence from Arizona and California Pilot Project[4]

	EPA[1] (FTP) cutpoint=1.2/20	Arizona[2] (IM240 Adjusted)	Arizona[3] (IM240)	Pilot Project[3] (FTP)
Super Emitters				
Number of Vehicles	20	544	219	15
% of Vehicles	12.58%	9.21%	3.71%	7.58%
HC-Before Repair (g/m)	16.2	7.25	8.11	21.51
HC-After Repair (g/m)	0.92	3.31	3.44	7.73
CO-Before Repair (g/m)	194.69	119.69	157.04	106.17
CO-After Repair (g/m)	8.3	56.11	63.67	77.96
Very High Emitters				
Number of Vehicles	55	5008	4161	105
% of Vehicles	34.59%	84.75%	70.42%	53.03%
HC-Before Repair (g/m)	4.58	2.37	3.11	2.74
HC-After Repair (g/m)	1.1	1.61	1.96	1.53
CO-Before Repair (g/m)	78.19	34.46	46.5	46.69
CO-After Repair (g/m)	14.1	23.76	29.85	22.14
High Emitters				
Number of Vehicles	79	296	760	38
% of Vehicles	49.69%	5.01%	12.86%	19.19%
HC-Before Repair (g/m)	1.84	0.36	1.12	0.95
HC-After Repair (g/m)	0.95	0.5	0.92	0.91
CO-Before Repair (g/m)	24.62	2.4	9.21	10.4
CO-After Repair (g/m)	11.18	6.28	10.13	10.48

1 Repair effectiveness data used by the TECH model. After repair values reflect adjusted, rather than actual, after repair FTP emissions. Before and after-repair emissions correspond to .8/15 cutpoints in order to simulate repairs beyond cutpoint.

2 Emitter group cutpoints based on IM240-FTP regressions contained in Appendix F. Average emission values do not include vehicles not having retests.

3 Data divided into emitter groups using separating emissions level used by EPA.

4 % s do not add to 100% because normal vehicles were not included here.

Table 2 shows the results from comparing the three datasets, EPA Arizona and the Pilot Project. By either method of splitting the Arizona data into emitter groups, the EPA and Arizona fleets appear to have very different distributions. The EPA dataset appears to have a slightly higher number of Super emitters, many fewer Very High emitters and many more High emitters compared to the Arizona dataset. In addition, the EPA data show Super and Very High emitters to have higher before repair emissions and lower after repair emissions compared to the Arizona data. The Pilot Project data seems to show a distribution that lies roughly in between the EPA dataset and the Arizona data. The Pilot Project data shows much higher before and after repair emissions compared to the EPA data. It is important to note, also, that the high emitter category shows very little change in emissions with repairs in both the Arizona program and the Pilot Project. In some cases (for CO in particular) emissions actually increase after repair.

A comparison of the distribution of emitter groups and emitter group emissions between MOBILE and real world data such as that from Arizona is difficult for a number of reasons. It is quite difficult to extract from the Model what the distribution of emitters groups is and the associated emission levels: the model is not designed to make this comparison so most of the information is imbedded as input data in the program code. There is, in addition, the problem of comparing the distribution of IM240 readings from an I/M program to MOBILE's definition of emitter groups by FTP levels.

4.1.4. The Cost of Repair

Repair costs might not be expected to be part of an inventory model of vehicle emissions. However, costs are likely to play a role in the response of either motorists, mechanics or state regulatory authorities, or all three to requirements to keep pollution controls working, particularly as cars age. Economic theory predicts that costs would matter, even under a regulatory regime like I/M in which there is no choice over how many repairs is supposed to be done. The higher are the costs of compliance, the more likely motorists will be to do the minimum, or incomplete repairs, or take their vehicles outside the system and drive illegally.

Original estimates of the costs of I/M were that repair costs would be relatively low. EPA predicted that they would be about $90 on average. As Table 1 shows, however, the costs of repair evidence from the other repair studies repair costs to be much higher, as high as $330 in the Sun Oil Company study. For the Arizona I/M program, costs of shop repaired vehicles are about $173.13 for the first 5 months of 1996. However, many vehicles were having some difficulty passing the test so these costs are not the cost of achieving compliance. Table 3 provides some evidence about the number of retests in the Arizona program. The costs of compliance will include all costs, including driver costs and repair costs, for as many rounds of repair as it takes to fix the car. This could be a significant cost for a small number of cars.

Table 3. Number of Retests for Failing Vehicles Arizona I/M Program, January 1995 - May 1996		
	Frequency	Percent of Total
0	1410	19.3
1	4106	56.1
2	1088	14.9
3	427	5.8
4	155	2.1
5 or more	133	1.8
Total	7319	

Table 3 shows the percentage of cars that have trouble passing the test, and the frequency of retesting for the 2% random sample of cars in the Arizona dataset. Almost 20% of vehicles take only one test, and over half only have to take one retest. However, almost 6% have to take 3 retests, and one vehicle actually had 17 retests. This kind of retesting can add dramatically to the cost of the I/M program, and could have some impact on behavior. Table 4 shows the pattern of emission rates as vehicles are retested. NOx emissions occasionally go up, indicating the mechanics may have been trying to fix the other pollutants (HC and CO), and succeeded in making NOx worse. This table also shows that there is not a gradual decline in emissions for these difficult to fix vehicles. Instead, emissions appear to improve dramatically on the last round of testing.

4.1.5. <u>Joint Pollution Issue</u>

The ease of effectiveness of repair may depend on which combination of pollutants is high. This joint pollution problem complicates modeling repair effectiveness. The MOBILE model simplifies the joint emission issue in a number of ways. In the repair dataset used to determine the basis for repair effectiveness, vehicles are placed into emitter groups on the basis of emissions of any of the three pollutants. For example, vehicles will be considered supers if they exceed certain levels of either HC, CO or both. However, for running the model, supers are considered to have high emissions of both HC and CO, and repair reduces both. NOx is treated independently from CO and HC.

The evidence from the various datasets about how many vehicles fail for each of the different pollutants or combination of pollutants is mixed. Table 5 shows the possible categories of failure by pollutant for three datasets, the EPA repair dataset, the Arizona IM240 dataset and the California Pilot Project. In the EPA dataset, the largest number of vehicles fail both HC and CO. In the Arizona dataset, by far the largest number fail for NOx alone, and this is with a fairly lax NOx standard. The Arizona data are also different from the EPA data in that almost the same proportion of vehicles fail HC alone as fail both HC and CO, and many fewer fail CO alone compared to those that fail HC alone.

Table 4. Emissions By Repair Round
Arizona Random Sample of Failing Vehicles
Jan. 1995-May 1996

	Maximum number of retests	
	Three	Six
	(n=427)	(n=29)
HC Emissions (g/mi)		
Initial test	3.12	3.04
1st retest	2.81	2.54
2nd retest	2.56	2.75
3rd retest	1.88	2.35
4th retest	-	2.28
5th retest	-	2.16
6th retest	-	1.16
CO Emissions (g/mi)		
Initial test	48.92	55.6
1st retest	43.75	47.27
2nd retest	39.98	47.78
3rd retest	29.62	56.56
4th retest	-	43.66
5th retest	-	43.68
6th retest	-	22.46
NOx Emissions (g/mi)		
Initial test	3.15	3.02
1st retest	2.97	3.08
2nd retest	2.88	2.88
3rd retest	2.17	2.28
4th retest	-	2.84
5th retest	-	2.36
6th retest	-	1.61

Table 5. Average Reductions in Pollutant Levels from Repair, by Pollutant Failure (g/mi)

Initial Failure	N	% of Total	HC	CO	NOx
EPA IM240 Data					
HC	19	7.2	0.49	2.52	-0.33
CO	16	6.0	0.38	18.8	-0.35
NOx	20	7.5	-0.03	0.12	1.1
HC,CO	84	31.7	3.07	62.7	-0.55
HC,NOx	15	5.7	0.03	-0.36	1.1
CO,NOx	3	1.1	0.35	8.18	1.37
None	86	32.5	-0.04	0.05	-0.32
All	22	8.3	1.56	23.8	0.42
Total	265	100.0			
Arizona IM240 Data					
HC	1088	18.4	1.57	7.38	0.31
CO	408	6.9	0.64	29.33	-0.34
NOx	2174	36.8	0.19	-0.79	2.07
HC,CO	1291	21.8	2.08	50.37	-0.48
HC,NOx	533	9.0	1.18	2.56	1.93
CO,NOx	23	0.4	0.27	16.03	1.04
None	279	4.7	-0.01	-1.1	0.11
All	113	1.9	1.68	18.77	1.54
Total	5909	100.0			
California Pilot Project CARB IM240 Data					
HC	22	11.1	1.28	5.37	0.02
CO	10	5.1	0.63	25.6	0.15
NOx	33	16.7	0.38	-0.91	2.17
HC,CO	46	23.2	2.76	46.99	0.34
HC,NOx	34	17.2	2.99	1.3	2.18
CO,NOx	1	0.5	0.32	12.72	4.71
None	36	18.2	0.22	-0.31	0.33
All	16	8.1	3.27	19.36	2.55
Total	198				

Note: EPA cutpoints (in grams per mile) used to determine failures based on initial IM240 emissions are: 0.8 HC, 15.0 CO and 2.0 NOx. See Table E-1 of Appendix E for cutpoints used in Arizona and California Pilot Project datasets.

Table 5 also shows the ancillary pollution reductions that result from repair of failing vehicles. There are some consistent results across all of the datasets. If a vehicle fails for either HC or CO alone, there will be some improvement in the other pollutant. The EPA dataset indicates that for vehicles failing either HC, CO or both, repair will make NOx worse. For vehicles failing only NOx, all datasets show very minor effects on either HC or CO.

The joint pollution issue is complex, but it may be important to separate out different types of failures for better forecasts of repair effectiveness.

4.1.6. Regression Analysis of Repair Effectiveness

We can look at the simultaneous influence of many of these factors on repair effectiveness using regression analysis on repair data. In this section we examine the EPA, the Arizona IM240 program and the California pilot program datasets. Results are shown in Tables 6,7 and 8 for the three pollutants HC, CO and NOx. In each case, the dependent variable is the reduction in pollutant emissions. Each column represents one equation, with the different variables affecting emission reduction on the left-hand side. Standard errors are given in parentheses under each coefficient, and the coefficients whose t-statistics are significantly different from zero are shown with asterisks (see table footnotes).

The first two equations (columns) give results for the EPA repair data in terms of the measured FTP emissions and measured IM240 emissions. The independent variables include the three initial pollutant levels, plus dummy variables for fuel injection and, in a carburetor system, whether it is closed-loop. In other specifications (not shown) we examined the effect of model year group but found that due to multicollinearity we could not have model year variables and the engine technology variables in the same equation.

The third column contains results from the Arizona IM240 program, with emissions measured by the IM240 test. Data on engine technology were unavailable, but we did include dummy variables for model years 83-90 and 91 or later. The coefficients show emission reduction effectiveness relative to the left out age category of vehicles, the 81-82 vintage. The Arizona regression also includes two repair cost variables: repair costs and tampering costs.

The last column shows results for the California pilot program, in which emission results are measured by the FTP test. Besides the three pollutants, the independent variables include the two types of cost, plus the vehicle's odometer reading.

The important results from these regressions are:

- By far the most important variable explaining the effect of emission repair on emission reductions is the initial level of that pollutant. However, initial levels of HC or CO do not appear to significantly affect the emission improvements in the other. NOx, however, is negative and significant for both HC and CO, indicating an inverse relationship between NOx on the one hand and HC and CO on the other. Hence, if repairs must reduce NOx as well as HC (or CO), the repair is more difficult and likely to be less effective.

Table 6. Results of Repair Regressions: Determinants of HC Emission Improvements

	EPA (FTP)	EPA IM240	Arizona (IM240)	CA Pilot (FTP)
HC initial	0.964** (0.036)	0.947** (0.021)	0.592** (0.011)	0.710** (0.046)
CO initial	-2.01E-03 (3.20E-03)	5.82E-04 (1.39E-03)	-4.92E-04 (5.95E-04)	0.017 (0.010)
NOx initial	-0.582** (0.110)	-0.189** (0.033)	-0.048** (0.009)	0.197 (0.203)
Fuel injection	0.554* (0.248)	0.277** (0.089)		1.118* (0.527)
Closed loop	0.051 (0.434)	2.51E-03 (0.157)		
Tampering repair cost (×100)			-0.035 (0.09)	-0.050 (0.12)
Total repair cost (× 100)			0.080** (0.010)	0.17 (0.14)
Model year 83-90			0.321** (0.047)	
Model year 91 or later			0.745** (0.072)	
Odometer				-0.0054 (0.0054)
Constant	-0.678 (0.442)	-0.516** (0.158)	-0.811** (0.064)	-1.901* (0.859)
N	265	265	5909	95
R^2	0.8794	0.9522	0.4943	0.8478
Root MSE	1.8056	.6524	1.2617	2.3031

Notes:

Standard errors in parentheses.
* Significant at the 5% level (2-tailed test)
** Significant at the 1 percent level (2-tailed test)

Table 7. Results of Repair Regressions: Determinants of CO Emission Improvements

	EPA (FTP)	EPA IM240	Arizona (IM240)	CA Pilot (FTP)
HC initial	0.195 (0.439)	0.314 (0.370)	-1.253 (0.215)	-1.178 (0.316)
CO initial	0.938** (0.039)	0.940** (0.024)	0.660** (0.012)	0.970** (0.069)
NOx initial	-5.173** (1.351)	-1.636** (0.583)	-0.315 (0.188)	-1.144 (1.380)
Fuel injection	7.968* (3.036)	6.630** (1.564)		6.978* (3.588)
Closed loop	1.337 (5.315)	-1.699 (2.754)		
Tampering repair cost (×100)			-0.026 (0.019)	-2.53E-03 (7.94E-03)
Total repair cost (×100)			0.011** (2.01E-03)	0.013 (0.010)
Model year 83-90			6.708** (0.953)	
Model year 91 or later			13.378** (1.457)	
Odometer				0.022 (0.037)
Constant	-9.997 (5.419)	-8.450** (2.767)	-14.843** (1.290)	-18.135** (5.849)
N	265	265	5909	95
R^2	0.8649	0.9400	0.5218	0.8277
Root MSE	22.134	11.441	25.472	15.674

Notes:
Standard errors in parentheses.
* Significant at the 5% level (2-tailed test)
** Significant at the 1 percent level (2-tailed test)

Table 8. Results of Repair Regressions: Determinants of NOx Emission Improvements				
	EPA (FTP)	EPA IM240	Arizona (IM240)	CA Pilot (FTP)
HC initial	-0.030 (0.017)	-0.062* (0.030)	-0.081** (0.011)	-0.023 (0.013)
CO initial	-1.89E-03 (1.50E-03)	-7.44E-04 (1.94E-03)	-0.003** (5.92E-04)	-2.47E-03 (2.86E-03)
NOx initial	0.423** (0.052)	0.538** (0.047)	0.566** (9.28E-03)	0.783** (0.057)
Fuel injection	0.259* (0.116)	0.343** (0.125)		0.044 (1.49E-01)
Closed loop	-0.022 (0.204)	-0.017 (0.220)		
Tampering repair cost (×100)			1.11E-03 (9.44E-04)	3.05E-04 (3.30E-04)
Total repair cost (× 100)			6.27E-04** (9.90E-05)	1.16E-03* (4.21E-04)
Model year 83-90			-0.096* (0.047)	
Model year 91 or later			0.159* (0.072)	
Odometer				-4.27E-03 (1.54E-03)
Constant	-0.610** (0.208)	-0.925** (0.221)	-0.528** (0.064)	-0.543* (0.243)
N	265	265	5909	95
R^2	0.2992	0.3960	0.5269	0.8171
Root MSE	0.8486	0.91275	1.2547	.652

Notes:
Standard errors in parentheses.
* Significant at the 5% level (2-tailed test)
** Significant at the 1 percent level (2-tailed test)

- For each pollutant the size of the coefficient on the initial level is a good indicator of the effectiveness of repair. The repairs in the EPA dataset are much more effective than those in Arizona; 95 percent of the incremental gram of HC emissions are removed by repair, compared to only 60 percent in Arizona.
- Repair costs, when the data are available, have a consistently positive and significant effect on emission reductions, which is as we would expect. The more money spent on repair, the greater the emission reduction, but though statistically significant the effect is not very large. An additional $100 spent on repair results in HC emission reductions of 0.8 g/mi, CO emission reductions of 1.1 g/mi, and NOx reductions of 0.8 g/mi.[23] The lack of sensitivity to cost may arise partly from measurement error on the cost variable, but we believe that even with perfect cost data the emission reductions would not be very much more sensitive to cost.
- We suspect that either older cars or cars with certain technologies may be more difficult to repair. We include dummy variables for model year in the regressions in columns (1) and (2) for the Arizona data. For each model year subgroup (83-90) and (90+) the comparison is to the base model year group 1981-82. We find that both model years have lower after repair emissions, other things the same, and the coefficients are significant. The coefficients on the 91+ vehicles are about double the value of those on the 83-90 model years. It is difficult to separate out several issues here. Are the model year variables simply picking up a vehicle age effect (that vehicles are more difficult to repair as they get older), or is there are technology effect that has little to do with age? Fuel injection technology was almost nonexistent in the base model year (81-82) and slowly came to penetrate the market through the 1980s. By the 1990+ model year most of the vehicle were fuel injection.

Repair Effectiveness by Component. We attempt to shed additional light on repair effectiveness by examining the effect on emission reductions of individual repairs. These data come are self-reported by the motorist, who hands in a form completed (ostensibly) by the person who repaired the vehicle when he returns to the test lane for the retest.

If the impact of model year on repair effectiveness were just a matter of vehicle aging, then we would expect that with each component, repair effectiveness would decline with earlier model years. Table 9 shows results of repair effectiveness in the Arizona program for two model year groupings. The first group is the 1981-88 model years and the second are the 1989-95 model years. We regressed dummy variables representing whether different components were repaired on emissions changes for each of the two groups. As shown, the

---

[23] Costs are missing for many of the observations. These missing costs are largely associated with warranty situations and follow-ups on repairs that were unsuccessful the first time. We have assumed zero costs for these observations. When we restrict the sample to observations containing complete cost information, results are quite similar.

effectiveness of repairs on specific components is not systematically related to vehicle age. We plan to explore this issue more in future work.

Table 9. Repair Effectiveness by Model Year Group

Component Repaired	HC reduction		CO reduction		NOx Reduction	
	1981-88	1989-95	1981-88	1989-95	1981-88	1989-95
Dwell	.21	1.00	23.72	68.09	4.19	2.16
Air intake	.86	1.51	-6.13	23.37	-.74	1.17
Vacuum	3.90	1.43	81.43	17.83	2.39	2.87
Af mix	1.82	3.84	63.43	75.58	0.17	2.52
Idle speed	3.53	1.82	35.01	21.55	1.56	3.11
Spark plug	1.97	1.82	35.61	19.82	1.12	1.52
Carburetor	2.59	3.93	64.46	83.35	3.90	2.75
Oxygen sensor	1.90	4.03	45.25	114.05	2.00	3.42
Sensors	3.25	2.72	86.28	105.30	.24	1.08
EGR valve	0.07	0.44	10.00	8.93	4.50	2.97
Corr cat	4.51	3.19	60.73	37.92	4.21	5.06
air inject	3.08	1.92	78.92	81.33	5.39	2.68
Electric	4.22	3.04	33.22	51.97	2.36	2.68

To summarize what we have learned about repair effectiveness, it appears that the MOBILE model is very optimistic about how effective repairs will be compared to other datasets. In particular, evidence from I/M programs in practice shows repair effectiveness to be substantially lower than the EPA forecast. We find some evidence that cost matter in emission reductions achieved through repair.

4.2. I/M Compliance

Compliance with an I/M program is another critical determinant of the effectiveness of an I/M program. Here we examine how the MOBILE model incorporates compliance assumptions and we attempt to compare those assumptions to the evidence from the ongoing Arizona program.

4.2.1. MOBILE Model Incorporation of Compliance

The MOBILE model defines compliance as those registered vehicles that successfully complete an I/M testing cycle (including waivers). Considering that vehicles may be unregistered or registered elsewhere, this definition is a little vague. In terms of Figure 1, apparently the noncompliance rate in MOBILE is NIM / R, rather than the expression in Equation (4). The model assumes that the vehicles that are tested eventually pass or get a waiver, i.e. $NP = 0$.

A simple adjustment is made to the I/M credits within MOBILE to capture the basics of I/M compliance. The MOBILE user inputs a percentage compliance parameter that is the proportion of all registered vehicles complying with I/M, which remains constant over time if the program is run for a number of years.[24] This percentage is converted through an adjustment factor which is then used to reduce the "credits" or "benefits" of the I/M program.

Table 10 shows the sensitivity of the Model results to variation in the compliance parameter. The Model was run for a biennial IM240 I/M test with cutpoints of 1.2 grams/mile for HC, 20 grams/mile for CO and 3 grams/mile for NOx. (The other assumptions for the Model run are given in the Appendix.) The table shows that for all vehicles, a 50% compliance rate reduces I/M benefits by about 60% for each pollutant. In general, when the compliance rate is *x* percent, the emission reductions remaining are somewhat less than *x*. Table 10 gives results for light-duty gasoline vehicles, but the results for other vehicle categories are similar.

Table 10. MOBILE Emission Factors and Compliance Rate Average for 1995 (g/mi)

	No I/M	98%	95%	90%	75%	50%
LDGV						
HC (g/mi)	2.722	2.23	2.26	2.30	2.39	2.52
HC (pct of benefit remaining)	0%	96%	90%	83%	64%	40%
CO (g/mi)	21.95	15.04	15.46	15.979	17.337	19.113
CO (pct of benefit remaining)	0%	96%	90%	83%	64%	39%
NOx (g/mi)	1.595	1.40	1.41	1.419	1.457	1.506
NOx (pct of benefit remaining)	0%	97%	91%	84%	66%	43%

The compliance adjustment factor attempts to account for the fact that the dirtier cars may be less likely to comply with I/M. This relationship is imbedded in the Fortran code[25] in the MOBILE Model. From Table 10, if the compliance rate is 98%, the implied failure rate of those two percent of non-complying vehicles is assumed to be twice as high as the population of registered vehicles, so the benefits lost to the I/M program are twice as high (about 4%) as the non-compliance rate (2%). Or, if the compliance rate is 95% (5% don't comply), the benefits lost from the I/M program are about 9-10%. At much lower compliance rates, however, the 2:1 ratio no longer holds.

There are several problems with the MOBILE assumptions about compliance rates. From equation (4) above there are two types of non-complying cars--those that never get

[24] I/M-related information is inputted by the user in the One-Time Data Section of the MOBILE input file. Thus, if a single input file contains multiple scenarios over a range of years, the same compliance rate would be applied to all years. Multiple input files could be used to model a scenario where the compliance rate varied over time.

[25] See fortran file ENFORC.FOR for this function.

tested and those that get tested and don't get repaired. There is some evidence that many cars fall in the latter category and do not successfully get repaired to the standard (they may get scrapped, sold outside the area or get driven illegally).[26] The failing cars that do not get repaired and continue to driven in the area may be the dirtiest of the failing cars, so their non-compliance will directly reduce the emissions reductions from an I/M program. Non-compliance among these cars may have much bigger impacts on I/M benefits than Table 10 suggests. We use a simple, if somewhat extreme example to illustrate the point. If there is a 10% failure rate among all vehicles but half of the failing vehicles do not get repaired, then a 5% non-compliance rate (5% of the entire fleet) results in a 50% loss of emission benefits from I/M. From Table 10, MOBILE would assume a loss of about 10% of I/M benefits. The truth for many programs probably lies somewhere in between, but it seems likely that cars that are non-complying would be mostly failing vehicles, and of those they might be the dirtiest and most difficult to fix. We examine the evidence from Arizona below to shed more light on these issues.

Non-compliance rates are likely to depend on the type of I/M regime in place, and to vary over time as adjustments are made to the requirements of a program. An I/M program that is more difficult to pass, like the IM240, is likely to result in more non-compliance. And, over time, after the vehicles that have the most difficult time complying either get scrapped or move outside the region, compliance rates may improve. In MOBILE currently, there is a compliance rate that remains in effect for all forecast years, although it is possible that the user could input different compliance rates and run each year separately.

4.2.2. Evidence from Arizona

Using the dataset of failing vehicles from the random sample of cars in Arizona, we can examine what happens to failing vehicles over a period of time. We first selected from the 2% random sample the vehicles which failed the I/M test between January 1, 1995 and July 1, 1995. We want to compare the vehicles that eventually passed the test with those that never did.[27] We follow the vehicles up to May 1996 (over a year later for most) and separate them into two groups depending on whether they have or have not passed the test as of that date.

Results are shown in Table 11. In this table we show the number of light-duty vehicles in each of the two groups, further separated into cars and trucks. For each pollutant Table 11 shows the average results of the initial and final emission tests (a percent of the vehicles were only tested once, so the initial and final tests are the same), as well as the initial and final results for the vehicles that initially fail for that pollutant. For example, 58 percent of the cars that failed the initial test exceeded the HC cutpoint (row 7), and those cars had average HC emissions of 3.71 g/mi (row 8).

---

[26] See discussion below about the evidence from the Arizona program.

[27] Some of these vehicles may have been moved out of state, and other may have been scrapped.

	Table 11. Preliminary Evidence of Compliance in Arizona									
		All failing vehicles[a]			Vehicles that pass[b]			Vehicles that never pass[c]		
		Cars	Trucks	Total	Cars	Trucks	Total	Cars	Trucks	Total
1	Sample size	2136	799	2935	985	455	1440	1151	344	1495
2	Percent of all				46.1%	56.9%	49.1%	53.9%	43.1%	50.9%
3	HC									
4	Initial emissions, g/mi	2.65	3.89	2.99	2.18	3.66	2.65	3.06	4.19	3.32
5	Final emissions, g/mi	1.96	2.64	2.15	0.95	1.76	1.20	2.84	3.80	3.06
6	Percent passing/failing HC	58.0%	53.3%	56.7%	48.4%	47.0%	48.0%	66.2%	61.6%	65.2%
7	Initial emissions, g/mi	3.71	5.55	4.17	3.31	5.49	3.98	3.96	5.61	4.32
8	Final emissions, g/mi	2.62	3.57	2.87	1.12	2.16	1.44	3.57	4.99	3.88
9	CO									
10	Initial emissions, g/mi	40.42	57.8	45.15	34.49	54.01	40.66	45.5	62.8	49.48
11	Final emissions, g/mi	29.49	41.92	32.87	13.91	28.25	18.44	42.83	59.99	46.78
12	Percent passing/failing CO	36.3%	25.9%	33.5%	31.9%	23.3%	29.2%	40.1%	29.4%	37.6%
13	Initial emissions, g/mi	79.67	136.76	91.7	73.04	136.81	89.14	84.18	136.7	93.62
14	Final emissions, g/mi	50.59	84.36	57.71	17.25	47.94	25	73.31	122.58	82.16
15	NOx									
16	Initial emissions, g/mi	2.93	3.72	3.14	2.88	3.77	3.16	2.97	3.65	3.12
17	Final emissions, g/mi	2.29	2.93	2.46	1.62	2.53	1.91	2.86	3.45	3.00
18	Percent passing/failing NOx	50.8%	43.2%	48.8%	53.2%	46.6%	51.1%	48.8%	38.7%	46.5%
19	Initial emissions, g/mi	4.47	5.68	4.76	4.27	5.57	4.65	4.66	5.85	4.89
20	Final emissions, g/mi	3.10	3.76	3.25	1.85	2.66	2.08	4.25	5.50	4.49

Notes:
[a]All failing vehicles initially tested prior to July 1, 1995
[b]All failing vehicles initially tested prior to July 1, 1995 and having emissions below cutpoints on final test
[c]All failing vehicles initially tested prior to July 1, 1995 and having emissions above cutpoints on final test

The most interesting observations from the table are as follows:

- The vehicles that eventually pass are repaired well below the relevant standards. The final average HC emission rate among the cars in the sample, for instance, is 0.95 g/mi, compared to a cutpoint of 2 g/mi for 1990 and earlier model years and 1.2 g/mi for post 1990 model years.
- Over half the vehicles, however, never pass. As shown, of the 2,935 cars and trucks in this sample, 1,495 never appeared for a retest or never had a passing test. (Some of these cars may have waivers, but unfortunately, we have not yet been able to obtain the waiver status of cars in the Arizona program).
- The vehicles that never passed were dirtier to begin with compared to the average of all failing cars, at least for HC and CO. And clearly, the cars that did not pass had higher final-test emissions. The most likely interpretation is that these vehicles are harder to repair.
- These vehicles also don't show much improvement between the initial and final test. Among the cars that never pass, for example, HC, CO and NOx improve on average by 7, 6 and 4 percent, respectively; in the vehicles that eventually pass the improvement is 56, 60 and 44 percent.
- The "never pass" group includes relatively more cars than trucks. This is most likely because the test cutpoints for trucks are much more lax than they are for cars.
- For trucks, there are smaller differences in initial emissions between the two groups. In fact, initial NOx emissions in the "never pass group are actually higher than in the other group. About half of the trucks of the group that do not get repaired were tested once and then never retested.

4.2.3. Summary of Compliance

Again, we have found it difficult to make a comparison of compliance issues in the real world (in Arizona) to what is assumed in the MOBILE Model. For the compliance issue, MOBILE makes assumptions about the emissions impact of non-complying vehicles. The underlying assumption is that the failure rate is about twice the average in the fleet, but this is difficult to test. We do know the emissions rates of failing vehicles in real world programs, but we don't know those rates for MOBILE. The problem is made more complicated by the fact that there are several types of non-complying vehicles. Some non-complying vehicles do not go through the text regime, and their is no mechanism currently to test their emissions at all (remote sensing might be one possibility).

It appears that the MOBILE assumptions about the impact of non-compliance are optimistic. The non-compliance rate is out of the entire fleet, yet the vehicles that are most likely to be non-compliant may be predominantly the failing vehicles, and from the evidence from Arizona, they may be the dirtiest of the failing vehicles.

The Arizona data suggests that the extent of non-compliance may be quite large. The data also suggest that of the failing vehicles there are some that are getting fixed and another large group that appear to be difficult to repair to the standards (and the standards are very lax in Arizona currently). At a minimum, all of this suggests that determination of the compliance rate may be quite complicated and that there should be more guidance for states about how to calculate it. States may need to have more coordination between states environmental and motor vehicle agencies to track the registration, movement and eventual disposal of vehicles. Certainly, allowing states to take a high default value of compliance (98%) should be reviewed.

4.3. Tampering

Tampering with emission control or emission-related equipment is one of the most obvious ways motorists' behavior can affect vehicle emissions. Whether it is by disconnecting the catalyst or losing the gas cap, motorist tampering of pollution control equipment has been found to be a serious problem, especially in early studies of I/M programs.[28] As a result, the MOBILE Model takes tampering seriously, and includes tampering rates for a range of different emissions control components, and includes tampering mitigation in I/M programs. In the Model, there are baseline tampering rates for eight emissions-related components that can be tampered with, and then there are both deterrent assumptions associated with I/M and anti-tampering programs that can be initiated as part of an I/M program.

Most of the basis for the MOBILE assumptions about tampering come from a number of tampering surveys EPA conducted in the 1980s and early 1990s. These surveys have been discontinued because tampering is perceived to be a less important problem for more recent model year vehicles because it is both more difficult to tamper and because the potential performance benefits are much lower than they used to be. In this section on tampering, we first examine the EPA assumptions about tampering in the MOBILE Model; we then examine some limited empirical evidence from Arizona and try to draw possible comparisons to the EPA assumptions.

4.3.1. MOBILE Model Assumptions about Tampering

There are several ways tampering behavior is included in the MOBILE model.[29] In the absence of any emission control program, there is assumed to be some base amount of tampering. However, tampering can be reduced by I/M programs in two different ways. The

---

[28] CAL study and GAO study

[29] We have tried to cover all places where tampering is included in the Model but its inclusion is complex, and not well documented

first is a reduction in tampering due to a deterrent effect which occurs just as a result of the presence of an I/M program: because motorists and mechanics know there is an I/M program in place, they will be less likely to tamper. Second, I/M testing can include a separate anti-tampering program that specifically checks for the impact of certain types of tampering and reduces it. The data on which the tampering parameters are based come from a number of EPA's tampering surveys, including the last ones done in the early 1990s.

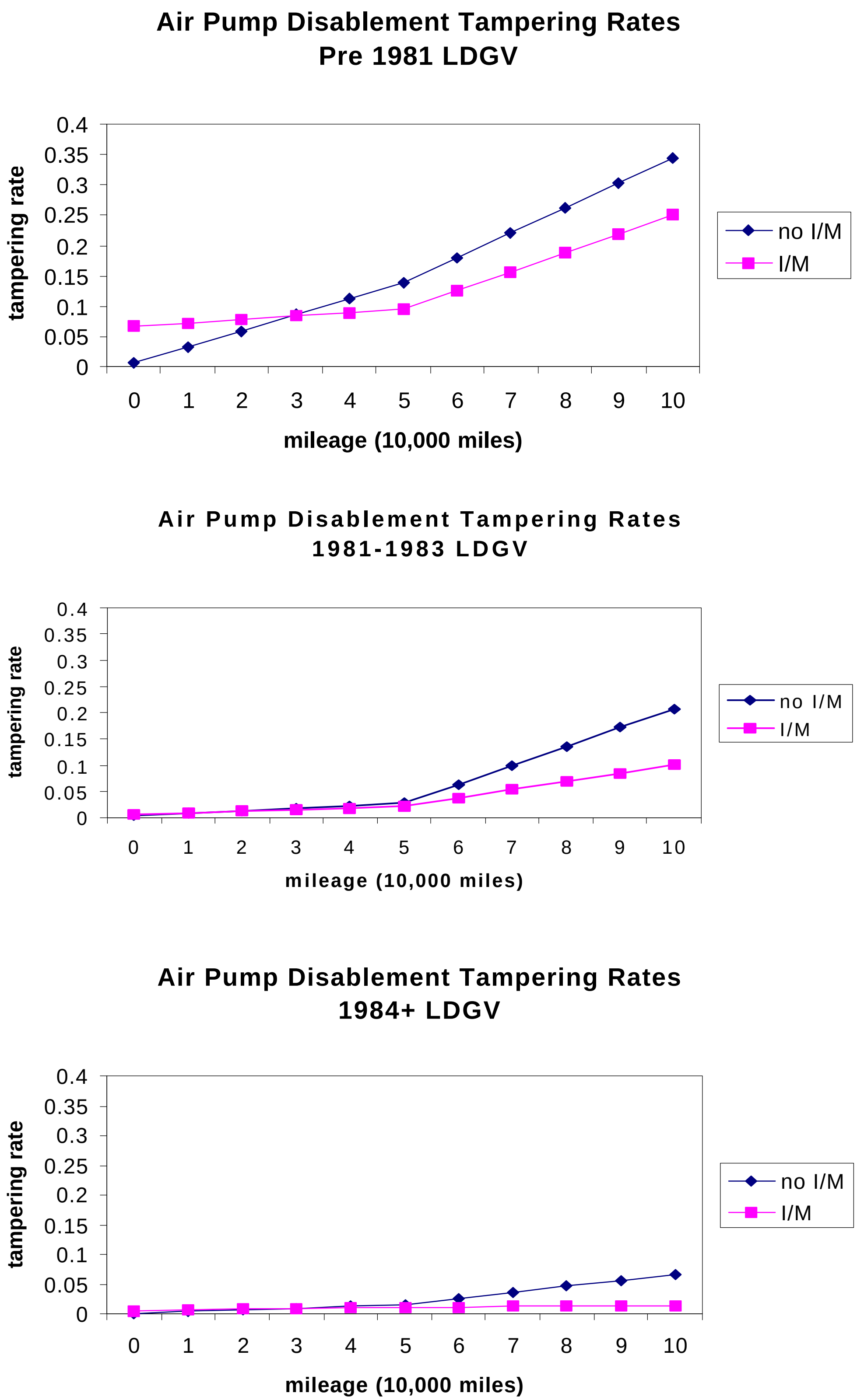
Tampering is defined within the MOBILE model as the malfunctioning of one or more of a set of emission control devices. The Model definition of tampering includes malfunctions that occur as a vehicle ages, in addition to those caused by deliberate disablement. Tampered vehicles are those that can be identified through visual inspection as having a malfunction to one of the following devices: air pump disablement, catalyst removal, overall misfueling, fuel inlet restrictor disablement, exhaust gas recirculation (EGR) system disablement, evaporative control system disablement, positive crankcase ventilation (PCV) system disablement, and missing gas caps.

Tampering is treated separately in the MOBILE Model from the determination of base vehicle emissions. The base emissions rates are supposed to represent non-tampered vehicles; tampering effects are included through tampering incidence rates and emissions offsets for the tampered vehicles, whose product is added to the base emission factors. The incidence rates and offsets are determined separately for vehicles equipped with different control technologies, and model year groups (pre-1981, 1981-1983, 1984+). MOBILE users can utilize default tampering incidence rates or provide alternate tampering rates in the MOBILE input file. The offsets (or the amount emissions rise when, for example, the catalyst has been tampered with) are fixed within the Model and cannot be changed by the user.

The tampering incidence rates are defined in terms of zero mile and deterioration rates. In addition, since tampering is assumed to be lower in a region having an I/M program, separate tampering rates are needed for the time periods before and after the implementation of an I/M program. Examples of the MOBILE's default tampering incidence rates for Air Pump Disablement are shown in Figure 8 for both the I/M and no I/M cases. The tampering rates are percentage of the fleet that has had tampering. The figure indicates that tampering incidence rates are projected to be significantly lower for later model year vehicles. Below, we examine the incidence of tampering failures in the Arizona IM240 I/M test.

In addition, MOBILE allows I/M programs to have separate anti-tampering credits. The existence of specific anti-tampering programs is assumed to decrease the effect of tampering on emission rates. The effect of anti-tampering programs on emissions reductions depends on vehicle type, inspection type (inspection only or inspection and repair); the frequency of inspection (annual or biennial); the compliance rate; and the inspections performed (i.e., air pump inspection, catalyst inspection, fuel inlet restrictor inspection, etc.). As in the case of I/M programs, the relationship between the compliance rate and the emission reductions allowed for anti-tampering program is assumed to be nonlinear: that is, the non-complying vehicles are assumed to be relatively more polluting and more likely to be tampered with than average cars in the fleet.

Figure 8. Air Pump Disablement Tampering Rates



4.3.2. The Arizona Program

The Arizona IM240 data provides some evidence of the incidence of tampering in an on-going IM240 program. Here we focus on the incidence of tampering by model year and not on the emission rates of tampered vehicles or the effect of anti-tampering programs.[30] Table 12 shows the summary of tampering by model year in Arizona for the 2 percent random sample for 1995 and the first half of 1996. Column A shows the share of failing vehicles for which mechanic surveys indicated some tampering repairs had to be made (this includes repairs to any of the different tampering possibilities including air pump disablement, etc., see Appendix G for the survey form filled out by mechanics as part of the Arizona program). There does not seem to be much variation or trend in tampering incidence by model year; there is no evidence of the dramatic decrease in the incidence for the more recent model years assumed by MOBILE.[31]

In the California Pilot Project, tampering incidence was found to be between 1 percent and 18% of the fleet, depending on component. Of the vehicles that failed the IM240 test in the California Pilot Project, almost 40 percent were found to have some type of tampering. This is much higher than what we found for Arizona. However, vehicles in the Arizona program had been subject to an anti-tampering program for a number of years. Also, the Pilot Project vehicles were older, on average, than the vehicles inspected in Arizona as part of the IM240 or those recruited at the EPA labs.

Table 12 also shows the share of vehicles in the Arizona program for which positive costs were reported for tampering related repairs. The tampering costs, on average, were quite small and often there were no costs reported.

4.3.3. Tampering Summary

Tampering is likely to depend on a number of factors, including the performance benefits of tampering, the costs avoided if complete repairs do not have to be made, and the costs of tampering, including any penalty if tampering is discovered. These costs and benefits have likely changed over time with changing vehicle technology and changing test regimes. The MOBILE Model does attempt break out tampering incidence rates by technology and tampering type. But there is no way for the Model account for repair costs in the determination of tampering, nor does it compare the Model assumptions to evidence in on-going I/M programs. The higher the costs of repair, the more likely motorists will be to tamper. In addition, high penalties for tampering would serve as a deterrent to tamper. It would seem to be important to collect tampering data from the states, and to use that to compare to the Model assumptions. States like Arizona do collect at least some of that data, as we report above.

---

[30] With the 2 percent random sample and spotty reporting of tampering, our dataset is small for examining emissions by tampering type. We will look at this issue more carefully in the future when we use the entire dataset.

[31] Note that the Model tampering rates are given as a percentage of the fleet, and we show the Arizona tampering rates as a percentage of the failing vehicles.

Table 12. Incidence of Tampering in Arizona Among Vehicles Failing Inspection, 1/95 - 5/96			
Model Year	Sample size	Frequency of Tampering	
		Tampering repairs performed (% of failing vehicles)	Vehicles with nonzero repair cost (% of failing vehicles)
1981	590	8.64	2.54
1982	512	6.64	1.95
1983	761	4.86	1.71
1984	1152	7.47	3.30
1985	1102	5.72	2.54
1986	838	6.32	2.63
1987	657	5.48	2.59
1988	507	6.71	3.35
1989	372	5.38	1.88
1990	214	4.21	2.80
1991	277	5.42	1.81
1992	182	6.04	2.20
1993	109	7.34	4.59
1994	28	17.86	10.71
1995	10	30.00	10.00
1996	6	50.00	33.34
All years	7317	6.39	2.64

Notes: Costs were not reported for all tampering related repairs.

CHAPTER 5

SENSITIVITY ANALYSES

To examine the relative importance of different parameters in MOBILE we have conducted several sensitivity analyses. The first looks at variations in different parameters affecting I/M test regimes. The second is a more general sensitivity analysis on a variety of both technical and behavioral parameters to see the relative impact of each.

5.1. Sensitivity on I/M Test Regimes

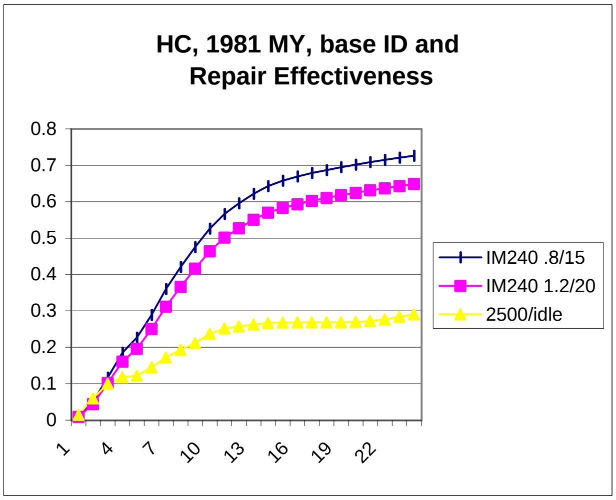
MOBILE forecasts emissions credits for a number of different types of idle tests and two different dynamometer tests. Table 13 shows the variation in emissions assumed by the Model under the range of different I/M test regimes (other assumptions of the IM240 test are shown in Appendix H). The IM240 test regime gets almost double the emissions reductions of any of the idle tests for HC and CO, and does much better than the ASM test for all the pollutants.

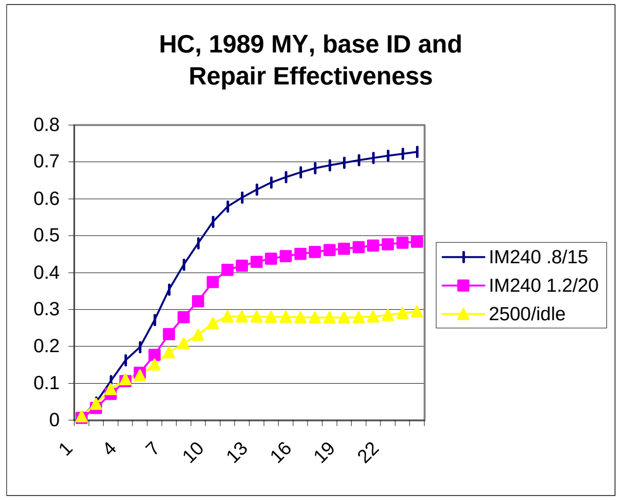
Table 13. MOBILE Emission Factors Compared to No I/M by I/M Test Regime LDGV, Year 1995

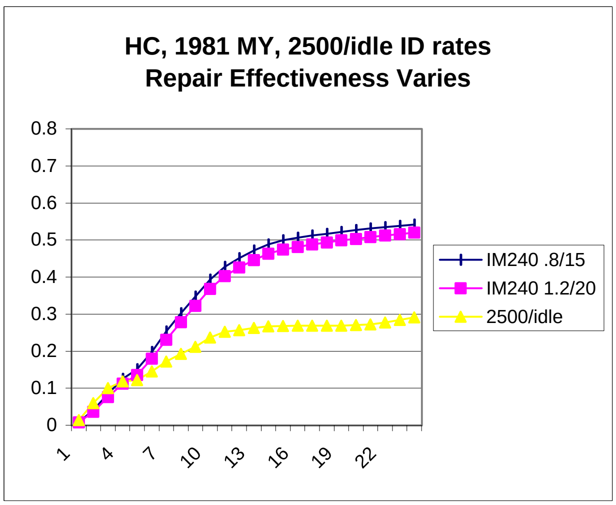
		CO	HC	NOx
no I/M program	g/mi	21.95	2.722	1.595
Idle	g/mi	18.68	2.482	1.583
	% reduction	14.9%	8.8%	0.8%
2500/idle	g/mi	17.184	2.415	1.585
	% reduction	21.7%	11.3%	0.6%
Loaded idle	g/mi	17.526	2.416	1.585
	% reduction	20.2%	11.2%	0.6%
IM240 (1.2/20/3)	g/mi	14.761	2.21	1.386
	% reduction	32.8%	18.8%	13.1%
ASM 2 Mode	g/mi	14.584	2.214	1.282
(2525/5015)	% reduction	33.6%	18.7%	19.6%

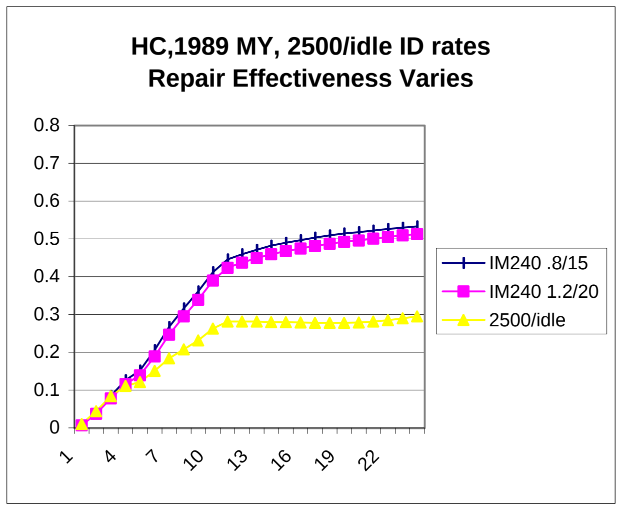
Figure 9 sheds light on what underlies the large difference between the IM240 and the 2500 idle test in the MOBILE model, and illustrates the importance of the parameters of the TECH model. We show in Figure 9 the credits or the percent reduction in emissions for two model year vehicles, from each of three I/M test regimes: the first is the 2500 idle test, the

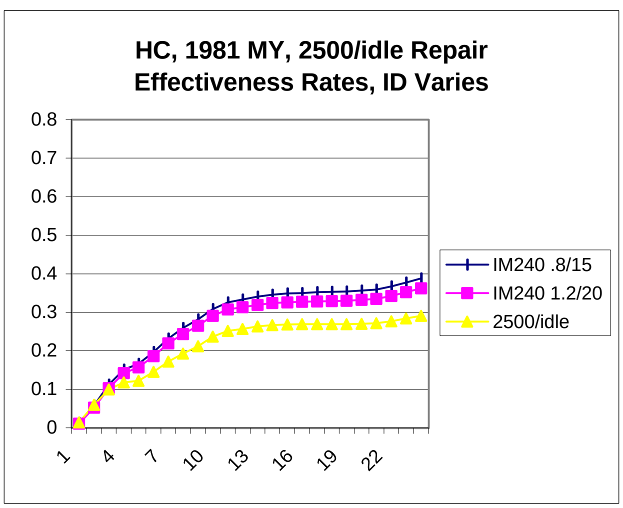
Figure 9
I/M Credits as a Function of Vehicle Age

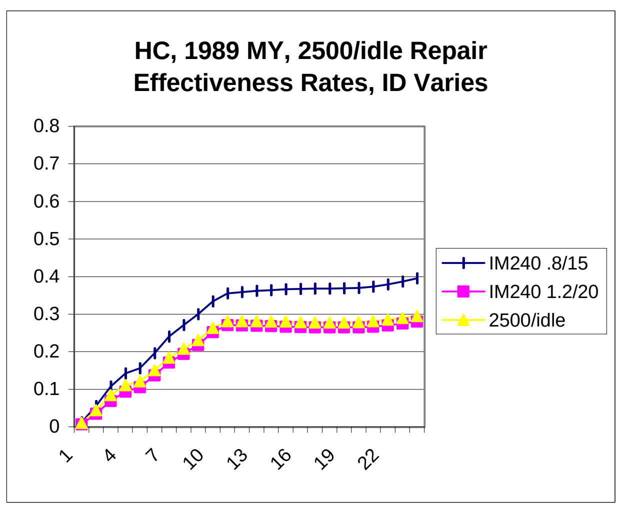












second is the IM240 with loose cutpoints (1.2HC/20CO), and the last the IM240 with tighter cutpoints (.8HC/15CO). The two parameters of the model that contribute to the effectiveness of the I/M test are the identification rates and the repair effectiveness. Identification rates in the Model vary by technology group (multiport fuel injected, carbureted, etc.), by model year group, by emitter group and by I/M test type. We want to isolate the impact of test type so we look at two different model years, which capture differences in technology and model year group. The 1981 model year includes mostly carbureted vehicles, while the 1989 model year includes the more recent generation of emission control equipment and mostly fuel injected technology.

The difference between the IM240 test (either cutpoint) and the 2500 idle test is dramatic for the 1981 vehicle in Figure 9. The less strict IM240 test in the 1989 vehicles is quite different from the stricter IM240 test, and lies in between that test and the 2500 idle. How much of this difference is accounted for by identification rates and how much by repair effectiveness? The second and third set of graphs in Figure 9 isolate the impact of each for the two model year vehicles. The second set of graphs show the credits that would result if the identification rates were held at the 2500 idle test level. It is clear that for both model years, repair effectiveness alone is quite a bit higher for the IM240 and there is not much difference in repair effectiveness across the two IM240 cutpoints. The third set of graphs in Figure 9 hold the repair effectiveness constant at levels assumed for the 2500 idle test, and allow only the identification rates to vary across tests. Here there are some differences between the model year vehicles. The identification rates for the two IM240 are similar and somewhat greater than the 2500 idle tests. For the 1989 vehicle, the identification rates for the 2500 idle test and the IM240 with the loose cutpoint are similar and significantly lower than that of the IM240 with the stricter cutpoint.

To summarize, the IM240 test regime obtains much higher emissions reduction credits than alternative test regimes. This outcome, however, is based on a relatively small sample of cars -- 36 for the 2500 idle test and 266 for the IM240 test. It would seem to be important to test this result from evidence from the field as that evidence becomes available. In addition, the identification rates are shown to vary with the IM240 test cutpoint, particularly for the more recent model years. This could have important implications for tightening cutpoints in the future in IM240 programs. The costs and potential for emission reduction from tighter cutpoints will need to weighed to determine if higher cutpoints are cost effective.

5.2. Sensitivity Analysis on Model Parameters

A general sensitivity analysis can shed light on the parameter changes that have most impact on the final emission factors in the Model. We focus the sensitivity analysis on the model components that we have discussed in this report or that may have important behavioral aspects. We examine various I/M test regimes: we compare the baseline IM240 centralized test to the same test with a tighter cutpoint, to a 2500 idle test, to a no I/M case, and to a decentralized I/M program. We then compare the base case to variation in the identification rate of the I/M test, to repair effectiveness that approximates that found in

Arizona, and to a lower compliance rate. The next set of sensitivities is on the assumed speed of travel; and finally we examine a range of assumptions about the age distribution of the vehicle fleet. Figure 10 illustrates the impact of various parameter changes on the base emission factors for NOx, CO and HC.

The base emission scenario for Figure 10 is for a centralized IM240 test with cutpoints set at 1.2 grams/mile HC, 20 grams/mile CO and 3 grams/mile NOx. All of the baseline assumptions are listed in Appendix H.

The first four sets of bar graphs in Figure 10 show the sensitivity of the Model to variation in assumptions about the type of I/M program. The tighter cutpoint on the IM240 test reduces emissions. HC emissions are reduced by about 7%, CO emissions by 12% and NOx emissions by about 10%. Compared to no I/M, the 2500 idle test is quite effective for reducing CO, and reduces HC or NOx by about 50%. Decentralized I/M increases CO the most dramatically compared to a centralized program.

The identification rates and repair effectiveness sensitivities are done only for HC and CO because NOx is handled separately in the Model.[32] The Model results are somewhat sensitive to changes in the identification rate, but they are more sensitive to the repair effectiveness and compliance rate assumptions examined here. The repair effectiveness assumption examines changes to the repair component of the Model to reflect emissions before and after repair based on repair evidence in Arizona. The resulting percent change in CO emissions seems particularly large. The compliance rate change is to use the compliance rate implied by evidence from the Arizona program: 34% of failing vehicles did not comply (see compliance section above), which is about 4% of the entire fleet (overall failure rate of 12%). This compliance assumption (shown as 96% compliance in the fleet) has a very small impact on emissions and as described above in the compliance section is likely to dramatically underestimate the impact on emissions reduced from this amount of non-compliance. This is because the focus should be on non-compliance of failing vehicles since this where the emissions reductions in an I/M program will come from. For comparison, we also show an overall 66% compliance rate for the entire fleet, which would more closely reflect what is happening in Arizona. This has quite a large impact on emissions.

The next parameter changes are for average speed. The model is has been found by other studies to be very sensitive to changes in speed of travel. We examine changes of plus or minus 4 miles per hour from the base speed of 19 miles per hour. We consider these to be large changes in average speed for an urban area. For example, to increase average speed from 19 to 23 miles per hour would require either a large investment in new highways or quite restrictive congestion measures. The Model is fairly sensitive to these changes, especially for HC.

Finally, we examine changes in the vehicle registration distribution. This represents an important set of parameters in our analysis because so many policies can influence vehicle holdings, and thus the age distribution of the fleet. Here we include two comparisons to the

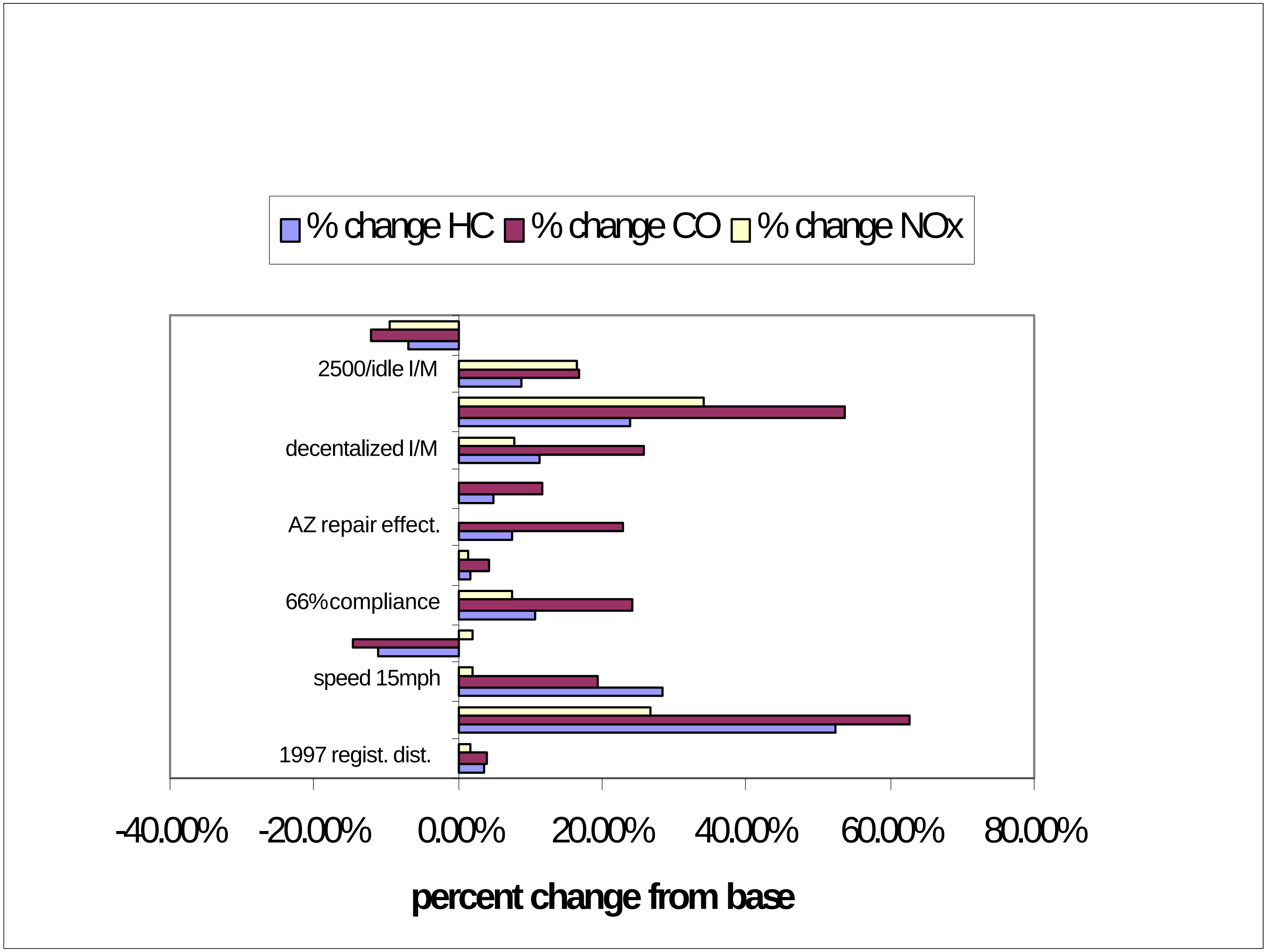
---

[32] We were not fully clear about where and how the NOx reductions were taken in the Model, so we did not include NOx in this analysis.

base case which is the 1990 national fleet distribution. This distribution, which is shown in Figure J-1 of Appendix J, is used as the default in the Model if the user does not input an alternative distribution. For forecast years after 1990, clearly this fleet should be aged using the best available assumptions about scrappage rates and new cars sales. The EPA has designed a program to do just that, and we use that program to forecast the fleet to 1997. The aging of the fleet seems to make little difference in the emission forecast as shown in Figure 10. We also examine the impact on emission from using a fleet distribution that is supposed to reflect the California fleet in 1997.[33] This distribution, which is shown in Figure J-2 of Appendix J along with the others used in Figure 10, has a larger proportion of older cars than the national distribution. Using this distribution has a very large impact on emissions of all three pollutants. Much of this impact occurs through the pre1981 vehicles, which we have not examined in detail in this report. Mobile handles the pre 1981 vehicles entirely separately from the post 1980 vehicles (the latter have been the focus of the this report).

[33] This distribution was obtained from the California Air Resource Board.

Figure 10. MOBILE Sensitivity Analysis: 1997 LDGV



CHAPTER 6

CONCLUSIONS AND NEXT STEPS

The MOBILE emission factor model is capable of dividing the vehicle fleet into subcategories based on a very detailed set of vehicle characteristics, producing emission factor estimates for each subcategory under a wide variety of policy assumptions and operating conditions. It first calculates "base" emission rates, which assume a baseline set of operating conditions, and then uses "correction factors" to adjust those base emission rates to specific conditions found at a particular time. The most important adjustments include those for average vehicle speed and average ambient temperature.

MOBILE has proved to be a reasonably accurate and useful model -- at least since the revision that produced MOBILE5 -- when used for its original purpose, namely the estimation of emission inventories for mobile sources. For this purpose the Model takes as given the distribution of vehicles in the vehicle fleet, their technological attributes, and the characteristics of their use, and produces estimates of vehicle emission rates under various conditions. However, MOBILE is now used for other purposes, purposes for which it was not originally designed and is less well suited, including policy analysis and estimation of emission credits for SIP revisions in nonattainment areas. In particular, the I/M component of the model that is used to estimate emission reductions resulting from different types of I/M programs requires the Model to be able to predict the response to I/M regulations over a period of time. This would be a difficult task for any model to do with precision, and is particularly difficult for a static, inventory model like MOBILE.

We have identified several problems that may interfere the ability of the MOBILE model to produce useful results when used for policy assessments. We group these problems in three areas: model structure, model validation and model calibration.

Model structure. Many of the Model's parameters that are considered to be fixed by technology are in fact endogenous and behaviorally determined at least to some degree. However accurate the technical data used by MOBILE, it appears not to account adequately for behavior likely to occur in real world situations. We find that model results are sensitive to changes in behavioral assumptions, especially in matters concerning (i) the age structure of the vehicle fleet and (ii) various aspects of I/M programs. Furthermore, we find that there has been little attention to behavioral responses by motorists, mechanics and others, and as a result there is little empirical information on potentially important variables describing how individuals may respond to changes in mobile source policy.

Model calibration. Efforts at model calibration have been hampered, in our view, by the limits placed by EPA on the what constitutes acceptable data. This is especially true for the parts of the Model used to predict the effectiveness of I/M programs. For the measurement of repair effectiveness, for example, data collection efforts have been limited to studies of repair in EPA or contractor laboratories, without comparison to data available about repair from other analyses. Furthermore, EPA's repair studies may also give misleading indications of what can be expected from vehicle repair because the repairs were conducted in a highly artificial laboratory situation. In addition, for the purpose of setting the basic parameters of the model, the EPA will only consider emission data from FTP tests. As a result, the data sets that have been used by EPA to determine the basic input parameters of the Model are strikingly limited in both their size and scope -- that is, the number of observations on which some important assumptions of the Model are based are often quite small, and the definition of what data are considered acceptable is limited. While FTP tests are perhaps the most precise and replicable of available emission tests (and certainly the most expensive), the FTP data sets suffer from a selection bias that may, in the end, be a more serious problem than the measurement errors of less sophisticated emission tests.

Model validation. MOBILE was not developed and is not used with the idea of making it testable against evidence from the real world. There have been some broad attempts at validating the Model emissions outcomes through speciation studies from ambient air quality models, and by comparison on average emissions in the Model to emissions from tunnel studies. With few exceptions, moreover, the empirical studies have attempted to test the overall results of the Model; we are aware of only a few studies that have attempted to compare to real-world outcomes the predicted results of the I/M components. While aggregate emission measurement is the "gold standard," there are in fact a number of other ways the Model could be compared to real world results. This is especially true of the I/M components of the Model, which could easily be modified to report out implications that could be compared to real-world outcomes.

We now elaborate on these points and suggest some alternatives for addressing them, proceeding in reverse order from the above.

6.1. Model Validation

Currently, the only way the Model is checked against reality is through comparison of final grams per mile estimates from the model to other estimates of vehicle emissions, most often from tunnel studies. To our knowledge, MOBILE's predictions regarding the effectiveness of I/M have only rarely been tested empirically.

With model validation limited to the results of ambient measurements and tunnel studies, there is always the prospect of calibrating the model by adding an adjustment factor at the end of the process. This may have pragmatic value, in that a model calibrated by a fudge factor may be able simply by trend extrapolation to predict changes in emissions in the short term, but it is almost certain to be "right" for the wrong reasons. This is important because without a more fundamental understanding of the structure of the problem, the model is unlikely to predict the effect of changes in policy, the more so as policies become further removed from the status quo.

Nonetheless, there are many other ways the Model could be checked against evidence from the field, especially in the I/M component of the Model. Within the TECH Model, there are implicit failure rates under different assumptions; there are implicit numbers of vehicles in different emitter categories; there are numbers of vehicles repaired to different levels, etc. The implications of the assumptions of the Model could all be made explicit, and the model checked against evidence from on-going real world programs. If this is difficult to ask of all states, then at least this could done on a pilot program basis. Such checks could suggest revisions to the Model, or isolate the importance of behavioral responses relative some of the more technical issues.

For example, it is reasonable to expect that in newly-established I/M programs, if they are successful, failure rates will begin at a high level and decline until a steady state is reached after a few years. The rate of decline will depend on the interplay between the ability to find and repair high-emitting vehicles and the durability of repairs.[34] However, the TECH model does not report failure rates, and in fact failure rates cannot easily be determined since identification rates are defined in terms of excess emissions rather than vehicles.

In a related issue, the embedding of all of the assumptions in the program code so that the user is not aware of them or has no control over many aspects seems to divert attention away from possible effective policies. In the repair component, repair effectiveness is built into the Model, regardless of mechanic training, enforcement of repair requirements, or the cost of repair. There is no incentive for state programs to look for ways to measure or improve repair, since the repair assumptions are pre-determined and imbedded in the program code. In general, there has been very little policy focus on repair, in part for this reason.

6.2. Model Calibration

Repair

EPA's Emission Factors data base consists of vehicles selected from the general population and tested (and sometimes repaired) in EPA laboratories or those of its contractors. The agency has also considered FTP test data submitted by vehicle manufacturers or collected as part of the Auto-Oil program. These vehicles and emission test results are used to define both the identification rates of various emission tests and the effectiveness of repair. The repair data set is a small subset of the Emission Factors data base, consisting of only 266 vehicles. Since this data set is used to set repair effectiveness by emission test type, fuel system type, model year, and age, after subclassifications the estimates of some parameters in the model are based on observations of fewer than ten vehicles.

The repair effectiveness factors were based on efforts to repair these vehicles in a laboratory setting by EPA mechanics or those of its contractors, not by ordinary mechanics in real-world situations. In addition, when even this kind of repair did not bring a vehicle into compliance with the emission cutpoints, it was assumed that further repair would do the job, and the vehicle's emissions after repair were assumed to be below the cutpoints even though that result could not be achieved in the laboratory.

---

[34] This is the result that obtains in the RFF model (McConnell and Harrington, 1992, 1994).

Our comparison of the repair effectiveness in the Model based on these results, with the repair effectiveness achieved in a number of non-EPA projects and in the Arizona Enhanced I/M program suggests that the Model's assumptions about repair are very optimistic. There are two major reasons why the current Model results may be optimistic: first, Model results are based on laboratory data to the exclusion of information provided by real-world data. Differences between the laboratory results and data from on-going programs could help to identify either weaknesses in an I/M program or advantages of laboratory repairs. In addition, there is currently no accounting for mechanic quality, or mechanisms to account for mechanic learning. In fact, the use of laboratory mechanics seems to base repair effectiveness on results after mechanic learning has been completed.

Second, the results from the lab are overridden when laboratory repairs did not bring the vehicle's emissions below the cutpoints. This additional adjustment seems to be particularly unrealistic. Although it might represent some ideal program, it is important to at least attempt to compare actual programs to this ideal. The evidence from on-going I/M programs is that all vehicles are not getting repaired, but we are only beginning to find this out. No one knows why these repairs are not getting made in current I/M programs, in part because states have felt no need to investigate repair performance since the reductions for repair are all taken within the Model. As we observe in Chapter 3, though, there are some good reasons to question whether emission repair would ever achieve the ideal outcome even in a competitive environment.

As a result, we believe forecasts of repair effectiveness to be overestimates, and that this element of the model needs to be revisited. We have discussed in the text a number of other data sets that could be useful in making this reassessment. In particular, the results that are beginning to emerge from the Enhanced I/M programs in Arizona, Colorado, Wisconsin and Ohio provide a wealth of information about the cost, frequency and effectiveness of various types of repairs. These data sets may not consist of FTP tests but they have other attributes that the EPA repair data set does not have and that may be of greater long-term value. They are very large; they consist of real-world repairs; they also contain data about repair cost. These data could be used to validate or at least identify problems with the laboratory data.

This can be expanded into a more general point about the value of other kinds of emission data. Every vehicle emission data base has strengths and weaknesses. The advantages of FTP data are that FTP tests are more precise and more replicable tests of vehicle emissions. But the FTP data has disadvantages as well, the main one being a selection bias due to the ways in which vehicles are recruited. Other vehicle emission data sets may consist of observations with larger measurement errors, but may have smaller errors overall because they are less afflicted by selection bias. Both provide useful information in the policy process.

Compliance

The modeling of the compliance rate is also of critical importance for assessing I/M programs. Again, there were a number of issues that arose in our review of compliance rates:

It is complex to define the compliance rate, and even more difficult for states to measure it. We believe the Model must address compliance more formally, and define several measures of it.

The estimation of compliance appears to be on the wrong variable in the current Model. Compliance rate defined as a percent of the entire registered fleet would seem to underestimate the importance of failure for causing vehicles to be non-compliant. It seems to us that the large majority of vehicles that are non-compliant are those that fail the test, and do not get successfully repaired. The fraction of these vehicles in Arizona seems to be quite large.

The compliance rate should be modeled as function of the cost of repair. Therefore, compliance would be a function of the strictness of an I/M program or test type, or of the existence of a waiver limit on repair.

6.3. Treatment of Behavioral Responses to I/M and Other Policies

Analysis of policies that affect the relative cost of holding vehicles, and therefore influence motorists decisions to repair, scrap or replace existing vehicles are difficult if not impossible to fully assess within a static emissions inventory model like MOBILE. To some degree this observation extends to any policy that changes costs, either directly or indirectly, and results in a feedback change in behavior. For example, implementing a stricter I/M program may induce more vehicles to be scrapped. The type of I/M program has a feedback effect on the age of the fleet. There are, indeed, few policies of any consequence that do not affect relative prices or the incentives motorists or mechanics face in some way. Fortunately, the separation of base emission rates from the emission correction factors offers the prospect of making modifications or additions to the model. The flexibility imparted by this separation means that behavioral considerations can *in principle* be introduced into parts of the model without affecting other parts. Since most of the behavioral responses we have been concerned with affect the base emission rates, we will focus on the TECH model. There are two possible approaches, depending on whether behavioral and dynamic considerations are added partially and in an *ad hoc* way, or whether a comprehensive modeling approach is taken.

<u>Partial approaches</u>

It would be possible to add behavioral feedback effects into the base emissions rates as they are calculated in the TECH Model. For example, with more stringent I/M policies or higher waiver rates, more vehicles may be scrapped or sold outside the region and the age distribution of the vehicle fleet may change. There are a number of studies examining scrappage response to used vehicle repair cost that could be drawn on to develop such a response. Scrappage elasticities could be drawn from such literature without the need to develop an economic model that fully reflects household decisions involving vehicle ownership. The partial effects of various policies on fleet composition could be fed back to MOBILE to use as input for the calculation of emission rates for the following year, giving the overall model a recursive structure.

There are extant models that do this. For example, Energy and Environmental Analysis, Inc. (EEA) has developed a model that includes a policy analysis component as an extension to the TECH model that can be linked to MOBILE. The EEA model is an extension of MOBILE but includes endogenous scrappage through a separate subprogram to recalculate the fleet age distribution in response to certain policy changes (EEA, 1994). Vehicle scrappage or retirement impacts are particularly important to account for in estimating emissions, because they have the largest impact on the older end of the age distribution, where emissions are likely to be highest. The EEA model also allows for analysis of the impact of emissions fees and the resulting impact on vehicle scrapping and vehicle use. There are, however, some problems with the way the behavioral component has been included in this model. For example, the EEA model's current assumption about induced scrappage is that all scrapped cars will be replaced by new cars, even though motorists have other options, including driving fewer miles or driving existing vehicles more intensively. This assumption finesses one of the most difficult and contentious issues surrounding induced scrappage programs.

As a second example we have developed a model at RFF (see Harrington, McConnell and Alberini, 1996a, 1996b) that compares the existing command-and-control policy to economic incentive approaches to I/M policy. The distinguishing feature of the model is its explicit propagation of the uncertainties involved in emission measurement and vehicle repair, showing how the magnitude of those uncertainties can affect the relative attractiveness of various policies. The model has also been used to examine the distributional consequences of in-use emission policies and to demonstrate how various cost-sharing schemes can affect those distributions (Harrington and Walls, 1995). The principal weakness of the model is that, as in the EEA model, the modeling of vehicle holdings decisions is *ad hoc* and incomplete.

A third example refers to an earlier RFF model that has only limited behavioral content (Harrington and McConnell, 1992, 1994) but explicitly allows for dynamic changes in the vehicle fleet in response to I/M policies. This model could easily be integrated into the existing MOBILE structure, but it would require a certain amount of elaboration to be compatible with the level of existing technical subcategorization in the model.

We also should mention an earlier model developed by the EPA, and one possible approach to the problems of the I/M components would be to revert to this earlier I/M modeling approach, which apparently was used by EPA in MOBILE2, before development of the current I/M credit calculation found in the TECH model. Our knowledge of this model is limited to a brief examination of a staff report (Rutherford 1982) describing the model. The Rutherford model does appear to have some features that we have identified as important and useful, including an explicit treatment of uncertainty, an attempt to base model parameters on identification, deterioration and repair data from real-world I/M programs. Apparently it was felt that unavailability of sufficient data to calibrate such a model prevented its immediate use. Perhaps the demands on OMS to produce immediate results, together with a lack of research funds for the purpose, prevented further development of this model or the data needed to support it.

Comprehensive approaches

The second approach to deal with the endogenous aspects of both vehicle fleet structure and I/M is to turn the problem inside out, making the core model an integrated economic model of vehicle ownership and use. MOBILE would then be used at the end as a database of emission factors used to compute base emission rates. This approach has been used by E. Deakin and G. Harvey (1996), who have developed an economic model of transportation demand that allows them to examine demand related policies such as congestion and pollution fees. This model uses a vehicle choice model based on data from California.

We would like to mention two other modeling efforts that could be used to produce dynamic estimates of fleet emissions. Maureen Sevigny (1995) has also developed an econometric model of vehicle use which can be used to forecast the impact of various policies on vehicles miles traveled. Emissions rates can be combined with the vehicle use estimates to determine emissions changes resulting from different policies.

The other modeling effort involves ongoing work at RFF. We have developed an integrated model of vehicle ownership and use, using data from the 1990 Nationwide Personal Transportation Survey in a conditional logit framework. This model has been used to estimate the consequences of fuel taxation (Krupnick, Walls and Hood 1993) and is now being used to examine the relationship between population density and the demand for travel (Walls, Krupnick and Harrington, 1997). This model could be combined with the I/M repair and scrappage model discussed above to produce an integrated dynamic stock turnover model that explicitly incorporates in-use emission programs such as I/M.

In sum, although there are no finished I/M models ready at this time to be substituted for the existing I/M part of MOBILE, there are several models that could be introduced for dealing with the behavioral issues we raise. Some of these new approaches involve relatively modest changes to the underlying structure and others involve more fundamental change. All would treat behavioral responses more explicitly than they are being treated now. The main weakness of all these models, and perhaps the most important barrier to their adoption, is their need for new data on which to base the behavioral parts of the model. Nonetheless, it would still be useful for EPA to begin to consider models of this sort, because more data are now available on the performance of I/M and on vehicle repair than ever before.

REFERENCES

Alberini, Anna, Winston Harrington, and Virginia McConnell. 1996. "Fleet Turnover and Old Car Scrap Policies" Discussion Paper 96-31, Department of Economics, University of Colorado.

Bishop, Gary A., Donald H. Stedman and Lowell Ashbaugh. 1996. "Motor Vehicle Emissions Variability." *Journal of the Air & Waste Management Association* v. 46 pp. 667-675.

Brzezinski, David J. 1985. *Tech IV Credit Model: Estimates for Emission Factors and Inspection and Maintenance Credits for 1981 and Later Vehicles for MOBILE3*. U.S. Environmental Protection Agency Technical Report, EPA-AA-IMG-85-6.

Cadle, Steven H., Mark Carlock, Kevin Cullen, Robert A. Gorse, Kenneth T. Knapp and Douglas R. Lawson. 1994. "Real-World Vehicle Emissions: A Summary of the Fourth Annual CRC-APRAC On-Road Vehicle Emissions Workshop," *Journal of the Air & Waste Management Association*, 44, 1180-1187.

Cadle, Steven H., Mark Carlock, Richard E. Gibbs, Kenneth T. Knapp, Alan C. Lloyd and William R. Pierson. 1991. "CRC-APRAC Vehicle Emissions Modeling Workshop," *Journal of the Air & Waste Management Association*, 41, 817-820.

Cadle, Steven H., Robert A. Gorse and Douglas R. Lawson. 1993. "Real-World Vehicle Emissions: A Summary of the Third Annual CRC-APRAC On-Road Vehicle Emissions Workshop," *Journal of the Air & Waste Management Association*, 43, 1084-1090.

Cebula, Francis J. 1994. *Report on the Sunoco Emissions Systems Repair Program*, prepared for Sun Oil Co., Philadelphia.

Day, J. Paul and Louis S. Socha. 1992. "Technique for the Analysis of FTP Emissions," SAE #920724.

Deakin, Elizabeth and Greig Harvey. 1996. "Transportation Pricing Strategies for California: An Assessment of congestion, Emissions, Energy and Equity Impacts," prepared for California Air Resources Board, Sacramento (November).

Decker, Stephen, John Suhrbier, Krista Rhoades, Herb Weinblat, Garry Brooks, Edmund Dickson. 1996. *Use of Locality-Specific Transportation Data for the Development of Mobile Source Emission Inventories*. Report prepared for Greg Janssen, U.S. Environmental Protection Agency, Office of Mobile Sources.

Energy and Environmental Analysis, Inc. 1994. "Highway Vehicle Emission Fees: Summary of Policy considerations and Analysis of Selected Implementation Scenarios." Arlington, Virginia, (September).

Guensler, Randall. 1993. "Data Needs for Evolving Motor Vehicle Emission Modeling Approaches."

Harrington, Winston. 1997. "Fuel Economy and Motor Vehicle Emissions," *Journal of Environmental Economics and Management*, v. 33 No. 3 (July 1997).

Harrington, Winston, Virginia D. McConnell. 1994. "Modeling In-Use Vehicle Emissions and the Effects of Inspection and Maintenance Programs," *Journal of the Air and Waste Management*, v. 44 (June 1994).

Harrington, Winston, Virginia D. McConnell, and Anna Alberini. Forthcoming. "Economic Incentive Policies under Uncertainty: The Case of Vehicle Emission Fees," in Ken Small and Roberto Roson, eds., *Environment and Transport in Economic Modelling* (Amsterdam: Kluwer Academic Publishing).

Harrington, Winston and Margaret A. Walls. 1996. "Controlling Vehicle Emissions: The Efficiency and Equity of Taxes versus Standards," Presented at the annual Meetings of the American Economic Association (January).

Heiken, Jeremy G., Barbara S. Austin, Alison K. Pollack, Dana L. Coe, Douglas S. Eisinger and Lyle Chinkin. 1996. *Methodology for Gathering Locality-Specific Emission Inventory Data*. Report prepared for Greg Janssen, U.S. Environmental Protection Agency, Office of Mobile Sources.

Kelly, Nelson A. and Peter J. Groblicki. 1993. "Real-World Emissions from a Modern Production Vehicle Driven in Los Angeles," *Journal of the Air & Waste Management Association*, 43, 1351-1357.

Lawson, Douglas R., P. J. Groblicki, D. H. Stedman, G. A. Bishop, and P. L. Guenther. 1990. "Emissions from In-Use Motor Vehicles in Los Angeles: A Pilot Study of Remote Sensing and the Inspection and Maintenance Program," *Journal of the Air and Waste Management Association* v. 40 No. 8, p. 1096-1105.

McConnell, Virginia D. and Winston Harrington. 1992. "The Cost-Effectiveness of Enhanced Motor Vehicle Inspection and Maintenance Programs," RFF Discussion Paper QE92-18, May 1992.

Patel, Dilip, Steven Magbuhat, Jeff Long and Mark Carlock. 1996. *Comparison of the IM240 and ASM Tests in CARB's I&M Pilot Program*. California Environmental Protection Agency Air Resources Board.

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

Scherrer, Huel C. 1996. "Methodological Flaws in I/M Program Self-Evaluation Studies." Testimony to the State of Colorado, Denver, CO 1996.

Scherrer, Huel C. and D. B. Kittelson. 1994. "I/M Effectiveness as Directly Measured by Ambient CO Data," SAE Technical Paper No. 940302.

Schwartz, Joel. 1995. "An Analysis of the USEPA's 50 Percent Discount for Decentralized I/M Programs," report prepared for the California Inspection and Maintenance Review Committee (February).

Sevigny, Maureen. 1997. *Taxing automobile Emissions for Pollution Control*, (Cheltenham, UK: Edward Elgar Publishers).

Shih, Celia. 1990. *MOBILE4 Sensitivity Analysis*. U.S. Environmental Protection Agency Technical Report, EPA-AA-TEB-90-02.

Sierra Research, Inc. 1994. *Evaluation of MOBILE Vehicle Emission Model*.

Stedman, D. H., G. A. Bishop, P. Aldrete, and R. S. Slott. 1997. "On-Road Evaluation of an Automobile Emission Test Program," *Environmental Science and Technology*, vol. 31, pp. 927-931.

Stephenson, A. R. and G. B. Dresser. 1994. *State-of-the-Practice Report on Mobile Source Emissions*.

Transportation Research Board, National Research Council. 1997. *Clean Air and Highway Transportation, Mandates: Challenges and Research Opportunities*.

U.S. Environmental Protection Agency. 1992. *Procedures for Emission Inventory Preparation, Volume IV: Mobile Sources*. EPA-450/4-81-026d (Revised).

U.S. Environmental Protection Agency. 1993. *Motor Vehicle Tampering Survey, 1990*. EPA-420-R-93-001.

U.S. Environmental Protection Agency. 1994. *Dynamic Registration Preprocessor*. MOBILE5 Information Sheet #4.

U.S. Environmental Protection Agency. 1994. *User's Guide to MOBILE5 (Mobile Source Emission Factor Model).* EPA-AA-AQAB-94-01.

U.S. General Accounting Office. 1990. *Air Pollution; Reliability of EPA's Mobile Source Emission Model Could Be Improved.*

U.S. General Accounting Office. 1997. *Air Pollution; Limitations of EPA's Motor Vehicle Emissions Model and Plans to Address Them.* GAO/RCED-97-210, November.

Venigalla, Mohan M., Terry Miller and Arun Chatterjee. 1995. "Alternative Operating Mode Fractions to the F.T.P. Mode Mix for Mobile Source Emissions Modeling," Transportation Research Board 74th Annual Meeting.

Venigalla, Mohan M., Terry Miller and Arun Chatterjee. 1995. "Start Modes of Trips for Mobile Source Emissions Modeling," Transportation Research Board 74th Annual Meeting.

Walls, Margaret A., Winston Harrington, and Alan Krupnick. 1997. "Examining the Relationships Between Urban Density, Transit Availability, and Vehicle Travel: Results from a Nested Logit Model of Vehicle Choice," working paper, Resources for the Future, Washington, D.C.

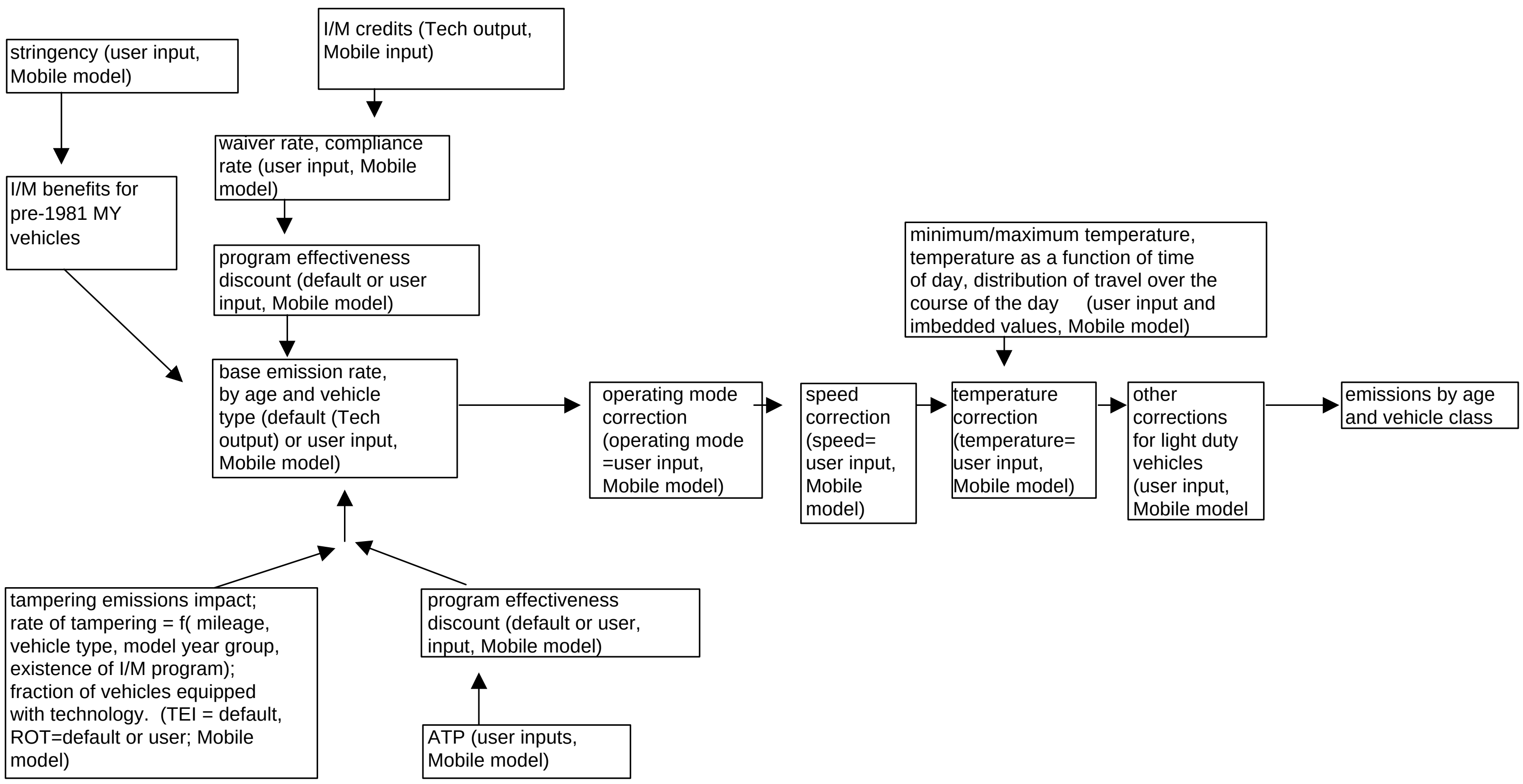
Walls, Margaret A., Alan J. Krupnick, and H. Carter Hood. 1993. *The Distributional And Environmental Implications Of An Increase In The Federal Gasoline Tax*, Discussion Paper 93-24, Resources for the Future, Washington, D.C.

Walls, Margaret A., Alan J. Krupnick, and H. Carter Hood. 1995. *Estimating the Demand for Vehicle-Miles-Traveled Using Household Survey Data: Results from the 1990 Nationwide Personal Transportation Survey*, Discussion Paper 93-25, Resources for the Future, Washington, D.C.

Appendix A

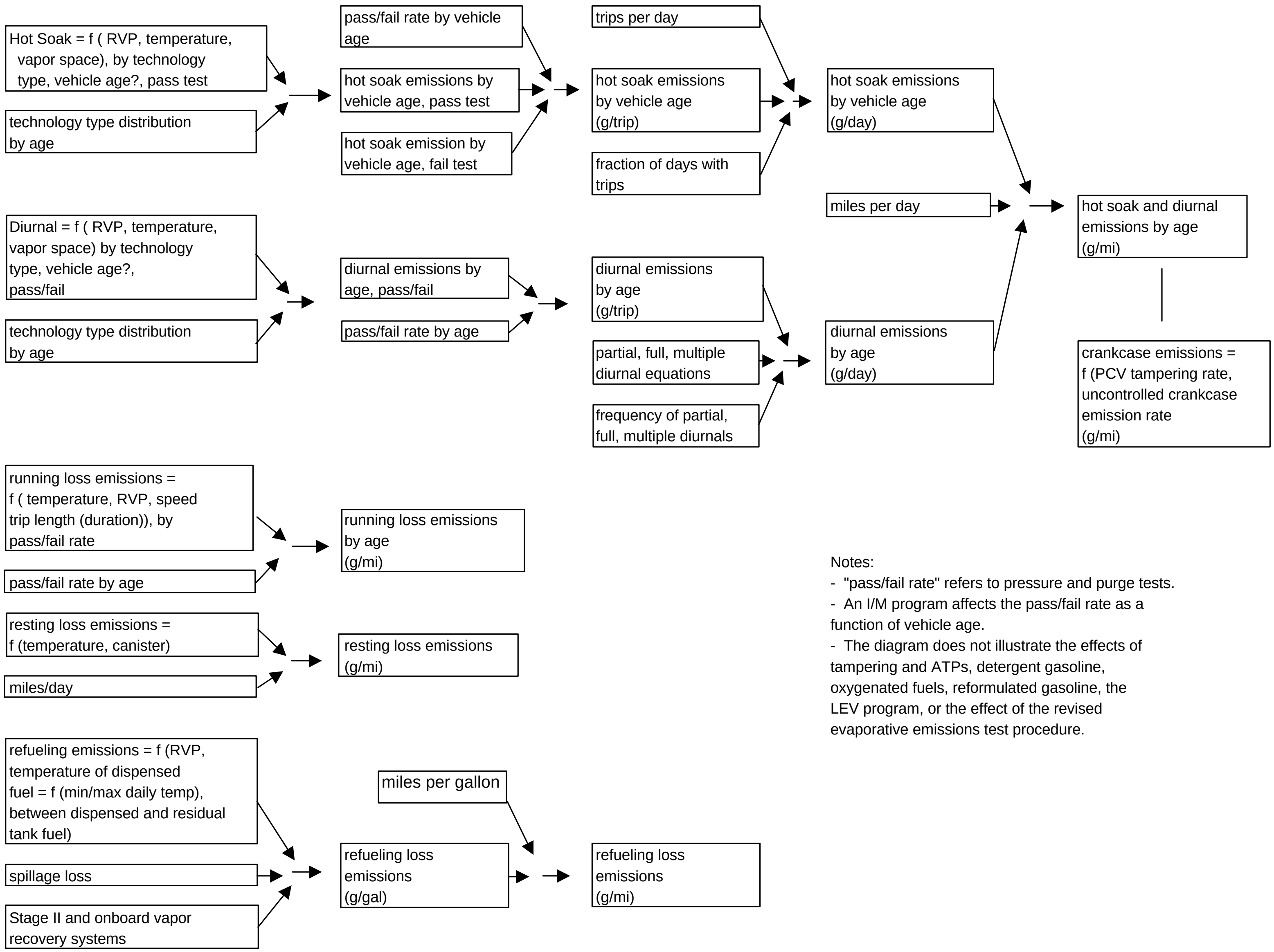
MOBILE and TECH Flowcharts

Figure A-1
Exhaust Emissions by Age and Vehicle Class



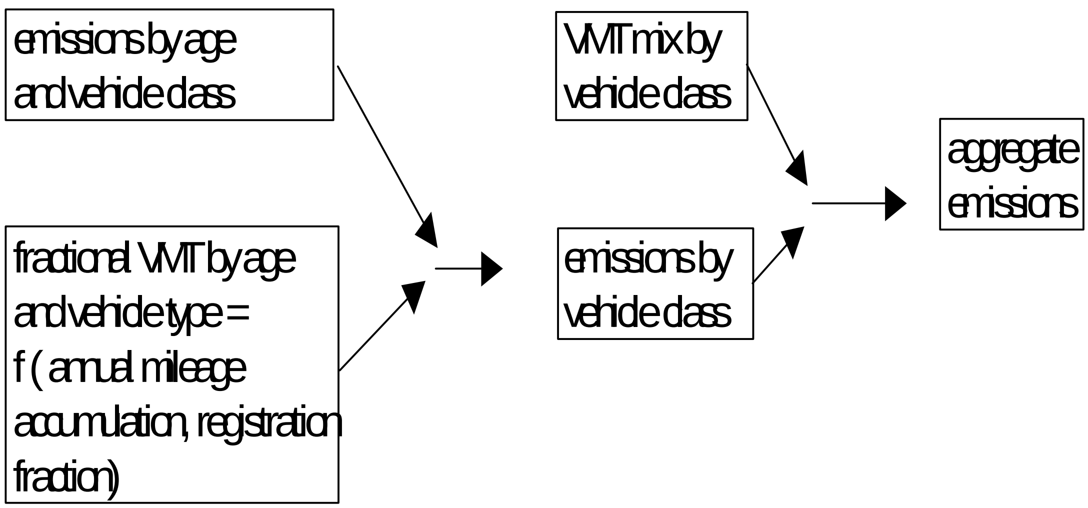
Notes:
- The "stringency" input is used to determine I/M benefits for LDGV before MY1981 and LDGT prior to MY 1984. The Tech model is used to estimate I/M credits for 1981-1993 MY LDGV.
- The diagram does not illustrate the effects of detergent gasoline, oxygenated fuels, reformulated gasoline, or the LEV program.

Figure A-2
Evaporative Emissions



Notes:
- "pass/fail rate" refers to pressure and purge tests.
- An I/M program affects the pass/fail rate as a function of vehicle age.
- The diagram does not illustrate the effects of tampering and ATPs, detergent gasoline, oxygenated fuels, reformulated gasoline, the LEV program, or the effect of the revised evaporative emissions test procedure.

Figure A-3
Aggregation of Emissions (Mobile Model)



Appendix B
Fuel Economy and Emissions

The results presented in this Appendix are intended to follow up on previous research on fuel economy and emissions described in Harrington (1997), which challenged the notion that fuel economy no longer affects emissions of motor vehicles because all vehicles of the same type must meet the same emission standards in grams per mile. found that the

On the contrary, it was found that the only vehicles for which fuel economy does not affect emissions are new vehicles. Among older vehicles those with better fuel economy have lower emissions, and the difference grows as vehicles age. The difference was attributed to the manner in which vehicle emissions deteriorate as vehicles age.[1] We found that the presumed independence of fuel economy and emissions depended on the assumption, made by OMS and built into the MOBILE model, that emission rates deteriorate on an average gram-per-mile basis, regardless of fuel economy. (This was referred to as the GPM hypothesis.) An alternative hypothesis was proposed, namely that that emission rates deteriorate on a grams per gallon basis (the GPG hypothesis). The alternative was found to fit the HC and CO emission data much better.

The work described below extends the previous work in three ways:

Physical mechanisms by which fuel economy may or may not affect emissions in emission-constrained vehicles are spelled out, through the use of a numerical example. In particular, we suggest that GPM deterioration is consistent with the gradual increase in engine-out emissions, whereas GPG deterioration is consistent with the steady loss of catalyst efficiency. Since malfunctions and equipment wear can affect both engine-out emissions and catalyst performance, it might be useful to have a statistical model that allows both GPM and GPG deterioration.

A new data set, consisting of IM240 tests conducted in 1995 on 1980-1993 model-year light-duty vehicles[2] in the Colorado Enhanced I/M program, is used. The data set used in the earlier paper contained emission measurements made in 1991 on 1978-1990 model-year vehicles. Cars of more recent vintage, so the argument goes, are much more reliable and have lower emissions, so these results of the earlier paper may not apply, or may apply only in attenuated form. In addition, the emission measurements in the earlier paper were made by using remote sensing (RSD) on vehicles on the highway. Using RSD measurements to estimate average emission rates in grams per mile is controversial.

[1] At present EPA uses cumulative vehicle mileage rather than chronological age as a measure of vehicle wear. While mileage is no doubt a superior measure in principle, vehicle odometer readings are often not available (as in the current case) and not very reliable even when available. To our knowledge no empirical study using real-world data has demonstrated that odometer measurements provide better predictions of emission deterioration than vehicle age.
[2] The data sets consists only of vehicles up to an including the 1993 model year, because of difficulties obtaining fuel economy data on 1994 and 1995 vehicles.

It is considered by many knowledgeable observers to be unacceptable for this purpose because of large measurement errors and the difficulty of inferring anything about average emissions from a split-second RSD measurement. In addition, the RSD data set we used contained no measurements of NOx emissions. (Current RSD technology is capable of measuring NOx.)

An alternative statistical model is used to test the hypotheses. In the previous paper two distinct hypotheses about emission deterioration were tested against one another. The result was that both hypotheses were rejected in formal statistical tests, although one fit the data much better than the other. These results were interpreted to mean that both the GPM and GPG deterioration occur, but that the latter predominates. Below we embed both the GPM and GPG hypotheses in a nonlinear model, in which one of the estimated coefficients can be interpreted as an indicator of the relative importance of each.

Two kinds of emission deterioration

Table B-1 contains information taken from detailed FTP tests conducted on two late-model (but not new) vehicles by Marc Ross and his colleagues at the University of Michigan. These vehicles are compared with average emissions of a pre-regulatory vehicle. The two 1994 vehicles have very similar tailpipe emissions in g/mi, reflecting adherence to the same emission standard. However, tailpipe emissions in terms of g/gal differ by nearly 50 percent, mostly reflecting differences in fuel economy. The opposite is true for engine-out emissions (i.e. pre-catalyst). Engine-out emissions are quite close in g/gal., reflecting generally similar engine emission characteristics and technology. But when engine-out emissions are expressed in g/mi, there is now a difference of 50 percent.

Compared to the 1968 (preregulatory) vehicle, tailpipe emissions (in g/mi) of modern vehicles are reduced by over 98 percent. Most of the reduction is in engine-out emissions, which are reduced by 75 percent. The catalyst pass fraction (*p)* is the percent of engine-out emissions that are not captured by the catalyst. As shown, *p* is 6.83 percent for the Cadillac and 10.7 percent for the Saturn. The better fuel economy of the latter means that it does not have to reduce emissions by as much to reach the emission standard.

One of the most important variables determining the catalyst pass fraction, is its surface area, relative to the rate of exhaust gas flow. The catalyst pass fraction approximately obeys a first-order exponential law:

$$p = \exp(-kA) \qquad \text{(B-1)}$$

where *A* is the specific area of the catalyst, or the ratio of surface area to flow rate, and *k* is a constant. The relative specific catalyst area in the table is the exponent *kA*, or $-\ln(p)$.

Now consider two examples of emission deterioration, the first of which is a tripling of engine-out emissions while the catalytic converter remains intact. As shown in the second section of Table B-1, this increase triples all emission rates for the Cadillac and Saturn. In other words, emission rates in grams per mile for these vehicles remain about

Table B-1 Two kinds of HC emission deterioration					
	HC emissions	**units**	**1968 car, 17mpg**	**1994 Cadillac Seville (17 mpg)**	**1994 Saturn (28 mpg)**
	Fuel economy	mpg	17	17	28
	Engine out emissions	(g/gal)	225	60	55
	Engine out emissions	(g/mi)	13.2	3.53	1.96
	Catalyst pass fraction	%	100%	6.83%	10.73%
	Relative specific catalyst area[a]			2.68	2.23
	Tailpipe emissions	(g/gal)	225	4.1	5.9
	Tailpipe emissions	(g/mi)	13.2	0.24	0.21
Emission deterioration: 200% increase in engine-out emissions					
	Engine outs	(g/gal)		180	165
	Engine out	(g/mi)		10.59	5.89
	Catalyst pass fraction	%		6.83%	10.73%
	Catalyst area (relative)			2.68	2.23
	Tailpipe	(g/gal)		12.2	17.7
	Tailpipe	(g/mi)		0.71	0.64
Emission deterioration: 50% loss of catalyst area					
	Engine out	(g/gal)		60	55
	Engine out	(g/mi)		3.53	1.96
	Catalyst pass fraction	%		26.1%	32.8%
	Catalyst area (relative)			1.34	1.12
	Tailpipe	(g/gal)		15.68	18.01
	Tailpipe	(g/mi)		0.92	0.64
Emission deterioration: 75% loss of catalyst area					
	Engine out	(g/gal)		60	55
	Engine out	(g/mi)		3.53	1.96
	Catalyst pass fraction	%		51.2%	57.1%
	Catalyst area (relative)			0.67	0.56
	Tailpipe	(g/gal)		30.7	31.4
	Tailpipe	(g/mi)		1.81	1.12
Source: Ross et al. 1995.					

the same, as shown in Table B-1. In the second example of emission deterioration, suppose no change to the engine-out emissions, but the catalyst becomes less effective by losing half its effective area. As shown in the table, the relative catalyst area is now half its former value, which in turn affects the catalyst pass fraction and the tailpipe emissions. Expressed in grams per mile, the tailpipe emissions are now much higher for the Cadillac. Loss of 75 percent of the catalyst only accentuates the effect.

Thus, it is plausible to expect that emission deterioration will obey either a GPM law, in which case emission deterioration in grams per mile is independent of fuel economy, or a GPG law, in which case emission deterioration is directly dependent on fuel economy.

Statistical models

The GPM and GPG hypotheses can be expressed in the following simple statistical models:

GPM model: $$m_t = m_0 + \beta_t + \varepsilon_t \qquad (2)$$

GPG model: $$m_t = m_0 + \gamma_t D^{-1} + \eta_t$$

where

m_t emissions of a vehicle *t* years old, in grams per mile

D certification fuel economy, mpg

β_j, γ_j emission deterioration coefficients after t years.

ε_t, η_t normally distributed and uncorrelated error terms, each with variance proportional to vehicle age.

Because we only have measurements taken over a very short time interval (several months), we cannot separate the effects of age from the effects of vintage in these estimates. The estimated coefficients β_t and γ_t, that is, refer to the emission deterioration of a one-year old 1994 vehicle, a two-year-old 1993 vehicle, etc. Separation of vintage from age effects requires a data set composed of emission tests taken at a sequence of times on vehicles of various ages. This would, for example, permit one to determine whether emissions deteriorate at a slower pace in newer cars (and hence whether the effect of fuel economy on emissions is being attenuated). At best, we can at best compare these results to the results of the earlier study on California.

In Harrington (1997) the two models in (2) were tested against one another in a classic hypothesis testing framework, using the Davidson-MacKinnon J test for nonnested hypotheses. In the J test, the predicted values under the null hypothesis are used as a regressor in the equation specifying the alternative hypothesis. The null hypothesis is rejected if its coefficient is significantly different from 1.

The possibility that mechanisms may be causing emissions to be deteriorating in terms of both GPG and GPM simultaneously suggests that it may be useful to investigate statistical models that do not require one to choose between the two hypotheses but allows both to contribute to emission deterioration. Below we present results for two such models, a linear model and a nonlinear model, and compare them to the “pure” linear GPG and GPM model results.

(i) Nested linear model:

$$m_t = m_0 + \beta_t + \gamma_t D^{-1} + \varepsilon_t \tag{3}$$

(ii) Nonlinear model:

$$m_t = m_0 + \beta_t D^{\lambda} + \varepsilon_t \tag{4}$$

The linear model thus contains the variables of both the submodels. As for the nonlinear model, when $\lambda = 0$, (4) reduces to the GPM model; when $\lambda = -1$, to the GPG model. The estimate of λ provides a measure of the relative success of each model in explaining the data. Unfortunately, it’s not that easy to disentangle the contributions of the two hypotheses in either model. For one thing, the independent variables are highly collinear. With large data sets it is possible to distinguish between the two, but the coefficients have large error variances, especially in the nested linear model. In the nonlinear model the problem is that when the estimate of the exponent λ gets too close to zero the estimation procedure fails to converge (Thus it is impossible to estimate the “pure” GPM model in this way). An additional source of difficulty in choosing among these models is the lack of an adequate underlying theory to guide the choice of specification.

Data

The models (3) and (4) were estimated for both cars and trucks for the pollutants HC, CO and NOx, using a dataset of IM240 test results collected in late 1995 and early 1996 by the fledgling Colorado Enhanced I/M Program. The data set consisted of full IM240 tests on randomly selected vehicles (i.e. no fast passes). We used a subset of the data consisting of about 24,500 cars and 9,500 trucks. These were the vehicles for which we were able to decode VIN numbers so as to assign fuel economy estimates. This is the same dataset that has been used by Stedman et al. (1997) to evaluate the performance of the Colorado Enhanced I/M program. One of the features of the Enhanced I/M program in Colorado is that in its first year it applied to odd-model-year vehicles plus only those even-year vehicles that were newly registered in the state. For our purposes that meant that the number of vehicles for even model years was small, so that standard errors were large for even-model-year coefficients.

Results

We first used partial F tests to test the GPG and GPM models against the full linear model (3). Results are shown in Table B-3. The F-statistics indicate that in all six cases both the GPG and the GPM models are significant at the 1 percent level. Neither model can be left out of the full model without a loss of explanatory power. However, a comparison of the sum of squared residuals between the full model and each of the partial models suggests a high degree of overlap in the variation explained by each model, which is consistent with the high collinearity among the coefficients of the full linear model. This causes a high degree of instability in the estimated coefficients.

Coefficient instability is illustrated in Table B-3 compares the estimated coefficients of the nested model for HC emissions from cars with those of the two submodels. A variable of the form "AGE*n*" in the table is a dummy variable for a vehicle of age *n* in 1996, while the "AG*n*dMPG" variables refer to the interaction of this dummy with the reciprocal of fuel economy. Comparison of the full model with either submodel shows, first, that the standard errors of the coefficients in the former are much larger. The coefficients themselves are also much more erratic. That is, in the submodels the coefficients increase with more regularity as the age of the vehicle increases. Similar results obtain for CO and NOx emissions as well as for emissions from trucks.

The nonlinear coefficients are shown in Table B-4 for car emissions and for truck HC. The nonlinear models of truck CO and truck NOx emissions did not converge. The coefficient l was between -1 and 0 in all four cases, indicating that the best fit in this specification lies somewhere between the GPG and GPM model. For car and truck HC and for car CO the value of l is large enough to suggest that the GPG deterioration can substantially affect emissions. The car NOx coefficient is small, though, suggesting that GPG deterioration is only a minor consideration.

Tables B-5 through B-10 use the coefficients from the four specifications (or three for truck CO and NOx) to produce estimates of average emission rates for a 20 mpg vehicle and a 40 mpg vehicle. For each class-vehicle combination, the results of the various specifications are broadly similar. The GPM model emission rates are identical, of course. For the GPG and nonlinear models, the difference in emission rates are small until the vehicle reaches middle age (about 8-9 years) and then the emission rates of the gas guzzler start to grow faster. The nested model estimates are more erratic, but generally similar. Fuel economy appears to make the largest difference in emissions of HC and CO, especially as the vehicle ages. The difference still exists for NOx emissions, but it is smaller and does not seem to be as limited to the oldest vehicles.

Table B-2 **Nested tests of GPG and GPM models**		
	Sum of squared residuals	F-statistic
HC – cars		
Full model	20,649	
GPG submodel	20,832	10.3
GPG submodel	20,761	15.6
CO – cars		
Full model	5,334,390	
GPM submodel	5,386,406	16.5
GPG submodel	5,380,956	17.2
NOx – cars		
Full model	9,031	
GPM submodel	9,175	12.9
GPG submodel	9,093	28.0
HC – trucks		
Full model	8,938	
GPM submodel	9,159	7.1
GPG submodel	9,031	16.8
CO – trucks		
Full model	2,283,446	
GPM submodel	2,401,239	5.1
GPG submodel	2,299,552	35.1
NOx – trucks		
Full model	5,022	
GPM submodel	5,088	15.7
GPG submodel	5,130	8.8

Table B-3
Estimated coefficients of the nested linear model and submodels:
Car HC emissions

	Full linear model		GPM model		GPG model	
	Coefficient	std err	Coefficient	std err	Coefficient	std err
AGE2	0.162	0.136	0.056	0.022		
AGE3	0.175	0.138	0.142	0.022		
AGE4	0.155	0.203	0.289	0.033		
AGE5	-0.046	0.156	0.305	0.025		
AGE6	0.062	0.236	0.417	0.039		
AGE7	0.654	0.175	0.481	0.027		
AGE8	0.305	0.288	0.665	0.046		
AGE9	0.745	0.176	0.8	0.030		
AGE10	1.818	0.288	1.232	0.049		
AGE11	1.113	0.188	1.487	0.033		
AGE12	1.226	0.309	1.871	0.061		
AGE13	0.293	0.241	2.012	0.045		
AGE14	-1.464	0.476	2.876	0.092		
AG1dMPG	1.719	2.758			-5.225	0.939
AG2dMPG	-0.823	1.702			-3.868	0.858
AG3dMPG	1.092	2.149			-2.198	1.039
AG4dMPG	5.822	4.693			1.979	1.242
AG5dMPG	11.825	2.901			2.535	1.067
AG6dMPG	12.139	5.783			5.720	1.383
AG7dMPG	-2.945	3.755			7.194	1.139
AG8dMPG	12.207	7.384			12.512	1.520
AG9dMPG	5.618	3.759			18.234	1.188
AG10dMPG	-15.154	7.639			28.352	1.652
AG11dMPG	12.578	4.132			34.970	1.227
AG12dMPG	20.432	8.015			45.934	1.892
AG13dMPG	49.250	5.700			49.244	1.457
AG14dMPG	119.819	12.298			73.859	2.562
_const	0.068	0.115	0.139	0.018	0.362	0.035

Table B-4
Nonlinear model results

	Car HC	Car CO	Car NOx	Truck HC
A	0.444	12.17	-1.280	-0.589
	(0.076)	(1.08)	(0.722)	(0.379)
b1	-3.671	-98.40	3.208	3.675
	(1.682)	(28.00)	(0.790)	(1.309)
b2	-2.962	-87.14	3.438	4.208
	(1.415)	(24.40)	(0.741)	(1.160)
b3	-2.216	-81.49	3.931	5.287
	(1.435)	(25.39)	(0.660)	(1.149)
b4	-0.159	-42.66	4.152	6.079
	(1.210)	(19.63)	(0.618)	(1.124)
b5	0.112	-52.88	4.457	6.808
	(1.060)	(20.06)	(0.556)	(1.044)
b6	1.668	-28.11	4.673	7.880
	(1.070)	(17.76)	(0.519)	(1.142)
b7	2.411	-20.11	4.879	9.585
	(0.926)	(15.49)	(0.479)	(1.217)
b8	5.034	17.50	5.264	11.472
	(1.220)	(14.71)	(0.414)	(1.595)
b9	7.858	64.41	5.308	11.287
	(1.569)	(16.59)	(0.405)	(1.512)
b10	12.842	148.53	5.652	16.437
	(2.676)	(32.63)	(0.355)	(2.739)
b11	16.228	184.85	5.806	18.945
	(3.350)	(38.88)	(0.331)	(3.262)
b12	21.673	255.12	6.044	22.488
	(4.631)	(54.34)	(0.308)	(4.139)
b13	23.395	296.65	5.896	25.011
	(4.953)	(62.14)	(0.321)	(4.897)
b14	35.554	412.09	6.202	24.375
	(7.760)	(87.27)	(0.298)	(4.661)
l	-0.786	-0.740	-0.214	-0.565
	(0.071)	(0.066)	(0.065)	(0.085)

Table B-5 Comparison of 4 sets of emission estimates, 20 vs. 40 mpg car Car HC (units: grams per mi)								
	nested		gpm only		gpg only		Nonlinear	
Age	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg
1	0.154	0.111	0.139	0.139	0.101	0.231	0.096	0.242
2	0.189	0.209	0.195	0.195	0.169	0.265	0.163	0.281
3	0.298	0.270	0.281	0.281	0.252	0.307	0.234	0.322
4	0.514	0.368	0.428	0.428	0.461	0.411	0.429	0.435
5	0.613	0.318	0.444	0.444	0.489	0.425	0.455	0.450
6	0.737	0.433	0.556	0.556	0.648	0.505	0.602	0.536
7	0.575	0.648	0.620	0.620	0.722	0.542	0.673	0.577
8	0.983	0.678	0.804	0.804	0.988	0.675	0.922	0.721
9	1.094	0.954	1.008	1.008	1.274	0.818	1.190	0.876
10	1.129	1.508	1.371	1.371	1.780	1.071	1.662	1.150
11	1.811	1.496	1.627	1.627	2.110	1.236	1.984	1.337
12	2.316	1.805	2.011	2.011	2.659	1.510	2.500	1.636
13	2.824	1.593	2.151	2.151	2.824	1.593	2.664	1.731
14	4.595	1.599	3.015	3.015	4.055	2.208	3.817	2.400

Table B-6 Comparison of 4 sets of emission estimates, 20 vs. 40 mpg car Car CO (units: g/mi)								
	nested		gpm only		gpg only		Nonlinear	
Age	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg
1	2.462	3.455	2.809	2.809	1.539	5.526	1.460	5.760
2	3.188	5.031	3.760	3.760	2.789	6.151	2.687	6.494
3	4.895	5.530	5.273	5.273	3.612	6.563	3.301	6.862
4	9.293	7.963	8.506	8.506	8.194	8.854	7.529	9.393
5	9.578	6.029	7.542	7.542	7.000	8.256	6.417	8.727
6	11.901	8.167	9.675	9.675	9.952	9.733	9.113	10.341
7	6.126	13.549	10.662	10.662	10.800	10.157	9.984	10.862
8	15.941	11.914	13.574	13.574	15.322	12.418	14.078	13.313
9	16.351	18.529	17.686	17.686	20.791	15.152	19.185	16.371
10	20.287	27.205	24.711	24.711	30.792	20.153	28.343	21.853
11	29.844	26.474	27.872	27.872	34.815	22.164	32.297	24.220
12	34.097	33.822	33.933	33.933	42.932	26.223	39.947	28.799
13	48.556	27.833	37.224	37.224	47.779	28.646	44.468	31.506
14	72.524	23.105	46.456	46.456	61.520	35.517	57.035	39.029

Table B-7 Comparison of 4 sets of emission estimates, 20 vs. 40 mpg car Car NOx (units: g/mi)								
	nested		gpm only		gpg only		Nonlinear	
Age	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg
1	0.374	0.276	0.340	0.340	0.334	0.366	0.411	0.178
2	0.511	0.363	0.465	0.465	0.483	0.441	0.532	0.283
3	0.843	0.495	0.636	0.636	0.747	0.572	0.792	0.507
4	0.994	0.572	0.745	0.745	0.900	0.649	0.909	0.607
5	1.076	0.771	0.901	0.901	1.101	0.750	1.070	0.746
6	1.175	0.878	0.998	0.998	1.246	0.822	1.183	0.844
7	1.328	0.946	1.095	1.095	1.396	0.897	1.292	0.938
8	1.468	1.167	1.291	1.291	1.646	1.022	1.495	1.113
9	1.518	1.164	1.301	1.301	1.681	1.040	1.518	1.133
10	1.856	1.229	1.455	1.455	1.943	1.170	1.699	1.289
11	1.853	1.340	1.553	1.553	2.010	1.204	1.781	1.359
12	2.002	1.431	1.661	1.661	2.172	1.285	1.906	1.467
13	1.764	1.492	1.615	1.615	2.029	1.214	1.828	1.400
14	1.795	1.761	1.777	1.777	2.203	1.301	1.990	1.539

Table B-8 Comparison of 4 sets of emission estimates, 20 vs. 40 mpg car Truck HC (units: g/mi)								
	nested		gpm only		gpg only		Nonlinear	
Age	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg
1	0.096	-0.047	0.123	0.123	0.123	0.113	0.088	-0.131
2	0.186	-0.026	0.230	0.230	0.221	0.162	0.186	-0.065
3	0.384	0.101	0.343	0.343	0.366	0.234	0.385	0.070
4	0.536	0.163	0.485	0.485	0.517	0.310	0.531	0.168
5	0.610	0.635	0.614	0.614	0.648	0.375	0.665	0.259
6	0.882	0.334	0.777	0.777	0.855	0.479	0.863	0.393
7	1.309	-0.014	1.062	1.062	1.194	0.648	1.177	0.605
8	1.435	1.195	1.375	1.375	1.540	0.821	1.525	0.840
9	1.430	1.027	1.310	1.310	1.503	0.802	1.491	0.817
10	2.115	2.063	2.095	2.095	2.538	1.320	2.440	1.459
11	2.859	1.939	2.605	2.605	2.952	1.527	2.902	1.771
12	3.641	1.655	3.250	3.250	3.619	1.861	3.555	2.213
13	3.882	2.849	3.485	3.485	4.183	2.143	4.019	2.527
14	3.750	3.860	3.764	3.764	3.885	1.993	3.902	2.448

Table B-9 Comparison of 4 sets of emission estimates, 20 vs. 40 mpg car Truck CO (units: g/mi)								
	nested		gpm only		gpg only		Nonlinear	
Age	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg
1	2.09	0.98	2.29	2.29	1.97	0.32	–	–
2	2.94	-0.07	3.56	3.56	3.12	0.89	–	–
3	5.93	3.50	5.58	5.58	6.08	2.37	–	–
4	8.39	5.33	7.97	7.97	8.61	3.63	–	–
5	9.90	7.79	9.58	9.58	10.38	4.52	–	–
6	13.39	5.83	11.93	11.93	13.36	6.01	–	–
7	20.30	-1.93	16.14	16.14	18.47	8.57	–	–
8	18.40	20.23	18.85	18.85	21.27	9.97	–	–
9	22.03	16.74	20.46	20.46	23.82	11.24	–	–
10	37.30	32.01	35.21	35.21	43.51	21.09	–	–
11	55.95	22.63	46.75	46.75	55.00	26.83	–	–
12	68.04	29.15	60.38	60.38	67.62	33.14	–	–
13	73.27	35.73	58.86	58.86	73.19	35.92	–	–
14	62.66	43.76	60.13	60.13	63.58	31.12	–	–

Table B-10 Comparison of 4 sets of emission estimates, 20 vs. 40 mpg car Truck NOx (units: g/mi)								
	nested		gpm only		gpg only		Nonlinear	
Age	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg	20 mpg	40 mpg
1	0.199	-0.090	0.252	0.252	0.287	0.424	–	–
2	0.391	-0.039	0.480	0.480	0.499	0.530	–	–
3	0.855	0.447	0.796	0.796	0.822	0.692	–	–
4	1.108	0.427	1.015	1.015	1.059	0.810	–	–
5	1.040	0.956	1.027	1.027	1.061	0.811	–	–
6	1.243	1.304	1.255	1.255	1.312	0.937	–	–
7	1.386	1.140	1.340	1.340	1.413	0.987	–	–
8	1.520	1.480	1.510	1.510	1.628	1.095	–	–
9	1.688	1.484	1.627	1.627	1.789	1.175	–	–
10	1.985	1.556	1.815	1.815	2.110	1.336	–	–
11	1.805	1.819	1.809	1.809	1.935	1.248	–	–
12	1.744	1.394	1.675	1.675	1.768	1.165	–	–
13	1.679	2.114	1.846	1.846	2.029	1.295	–	–
14	1.989	1.723	1.953	1.953	2.020	1.291	–	–

Appendix C
Adjustment of Post Repair Emissions for Repair Effectiveness in MOBILE

Vehicles in the EPA repair dataset that were not repaired to the IM240 emissions cutpoints had their emissions adjusted downward to meet the cutpoint. An adjustment was made to the observed post-repair FTP reading for these vehicles. Bag1 minus Bag2 FTP emissions were calculated in order to obtain a FTP-IM240 relation corresponding to zero IM240 emissions (point A). Assuming a linear relation between FTP and IM240 emissions, a function between FTP and IM240 emissions was estimated using this point and the observed post-repair IM240 and FTP emission levels (point B). From this line, FTP emissions associated with the IM240 cutpoint were estimated. Figure 1 shows the method used.

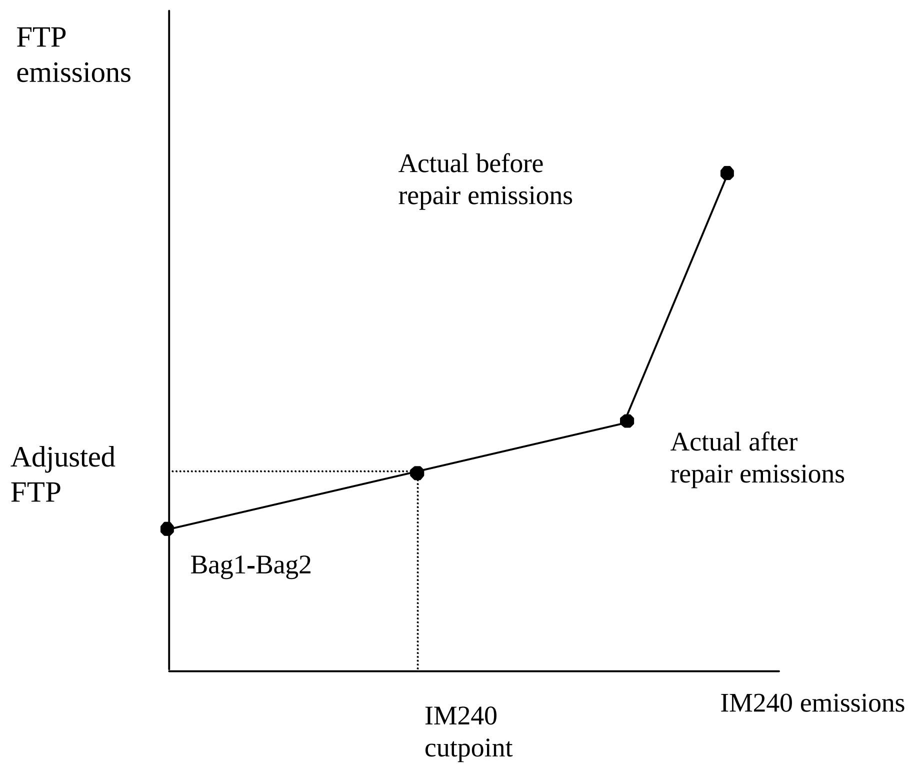


Figure 1. EPA method for adjusting after repair FTP emissions to meet the cutpoint (Used only if IM240 reading is above cutpoint)

Appendix D
Initial Emission Distributions

Figures D-1, D-2 and D-3 show the initial emission distributions of vehicles in the different datasets. For the EPA and California Pilot Project, both IM240 and FTP emissions measurements were made. For the Arizona IM240 program and the Sun Oil study, only the IM240 readings are available. The California I/M Review Committee study used only FTP. The Arizona program appears to have fewer extremely high super-emitting vehicles, but this may reflect the fact that the Arizona dataset is much larger and probably more representative of the fleet; it may include some "outliers" but there are so few that the share would be a fraction of 1%. Also, it is notable that when there is data on both the FTP and IM240 readings for the same vehicles (the EPA and California Pilot Project datasets) the emissions distribution results look quite similar.

Figures D-4, D-5,D-6 show the distribution of the emission reductions for each dataset. The average emission reductions are quite small and most of the repairs are tightly bunched just above zero. There are some very large emissions reductions, and many falling in the negative range.

Figure D-1
Distribution of Initial HC Emissions Across Vehicles, Repair Datasets

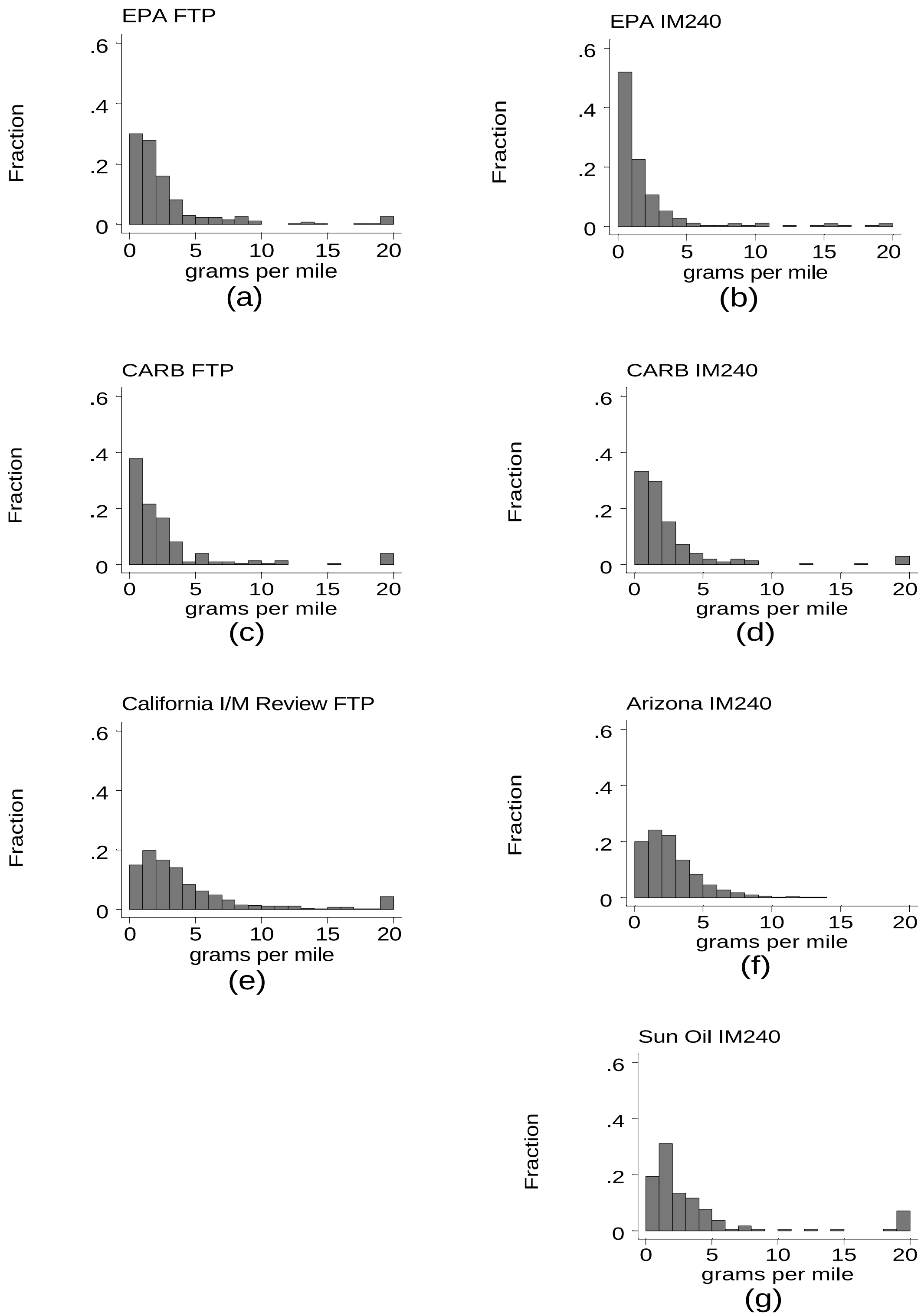


Figure D-2
Distribution of Initial CO Emissions Across Vehicles, Repair Datasets

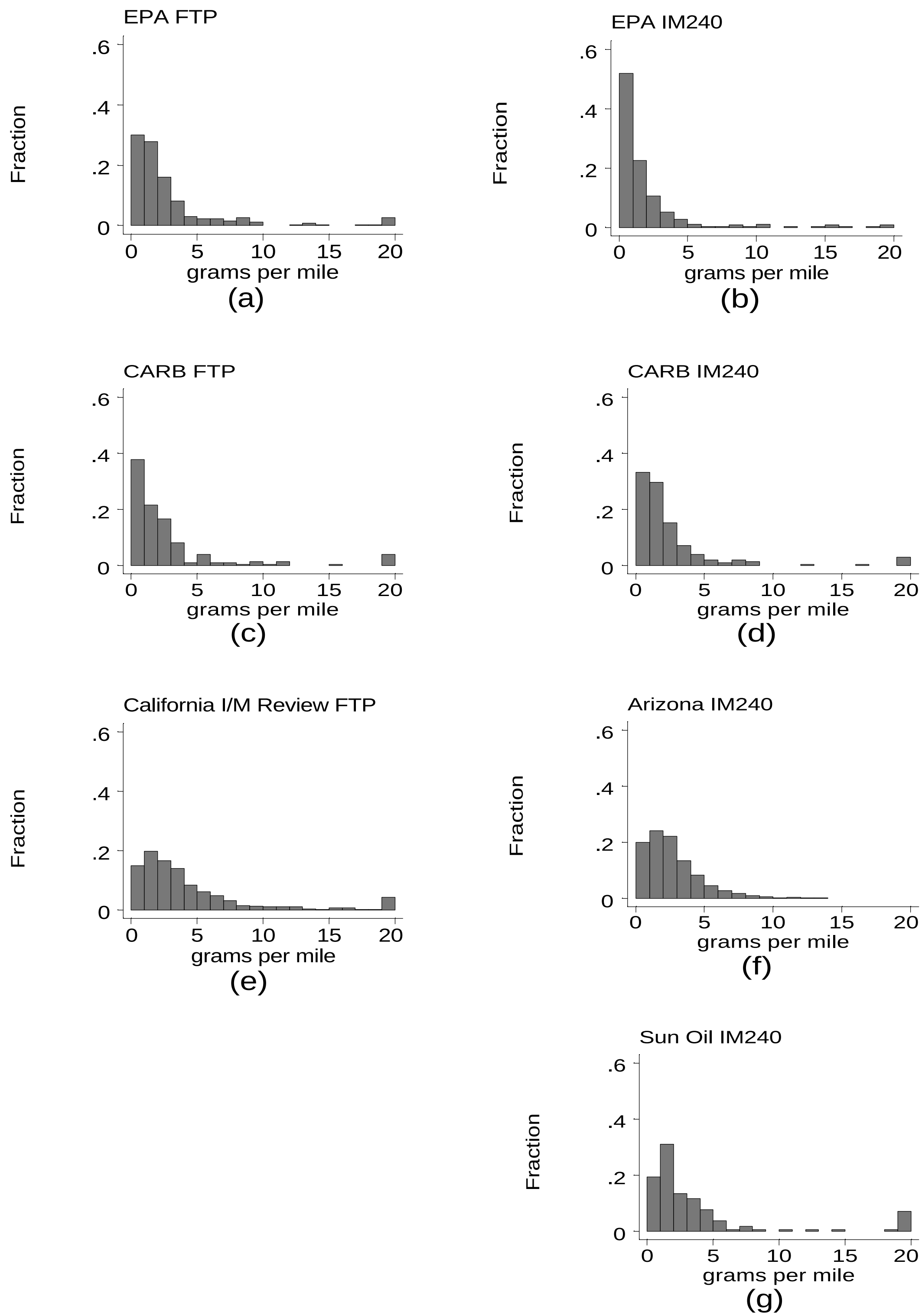


Figure D-3
Distribution of Initial NOx Emissions Across Vehicles, Repair Datasets

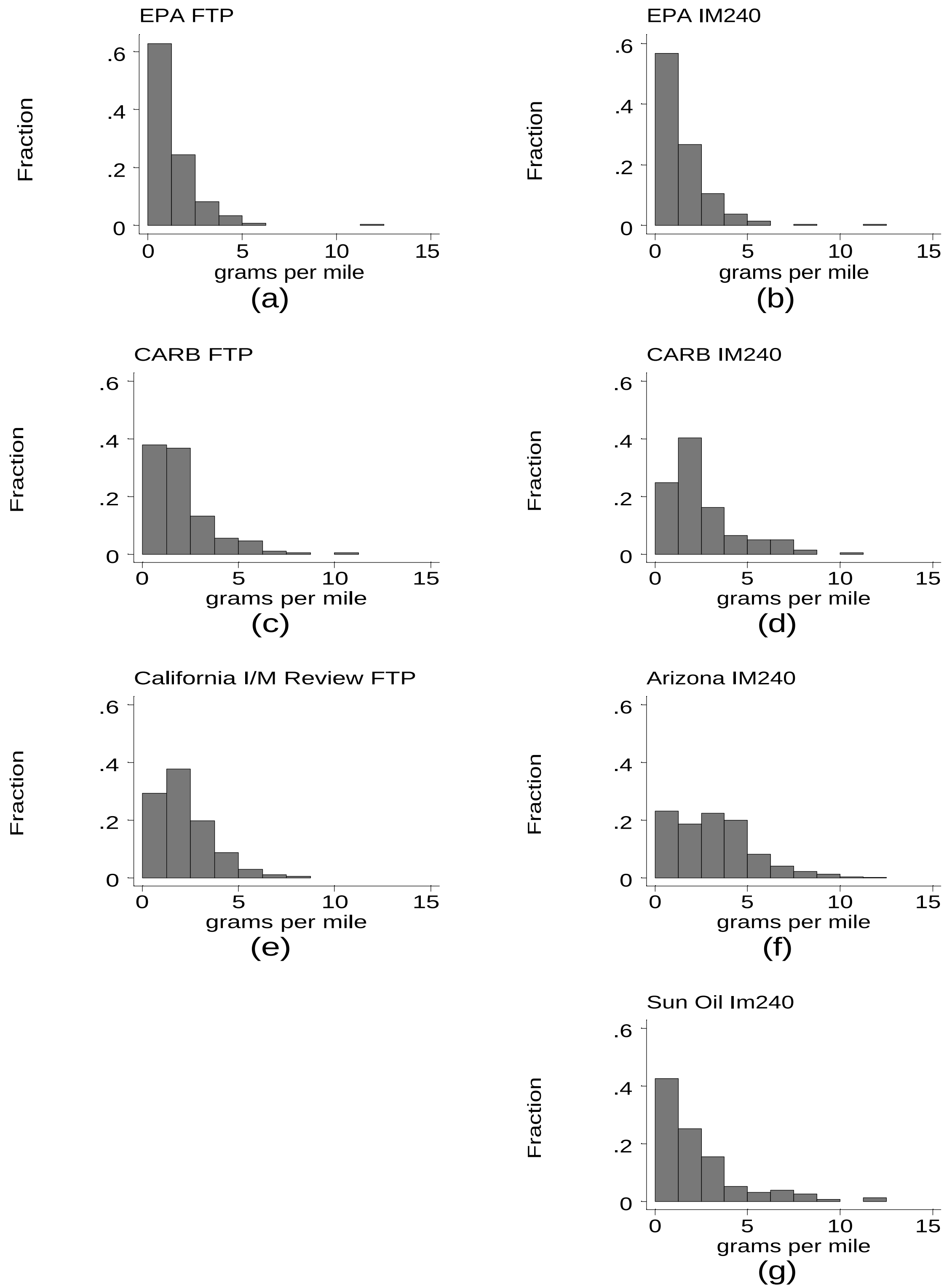


Figure D-4
Changes in HC Emissions as a Result of Repair, Repair Datasets

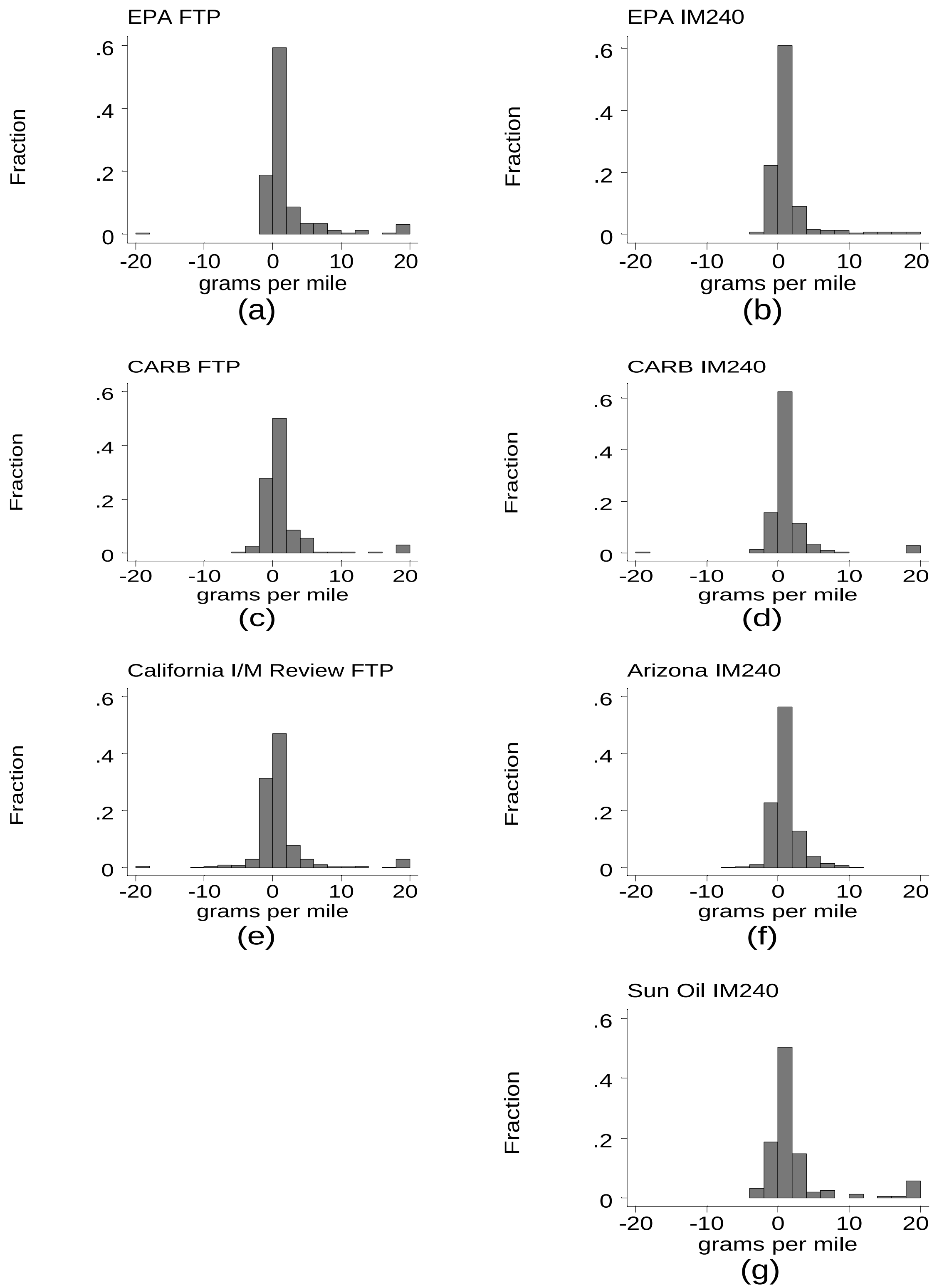


Figure D-5
Changes in CO Emissions as a Result of Repair, Repair Datasets

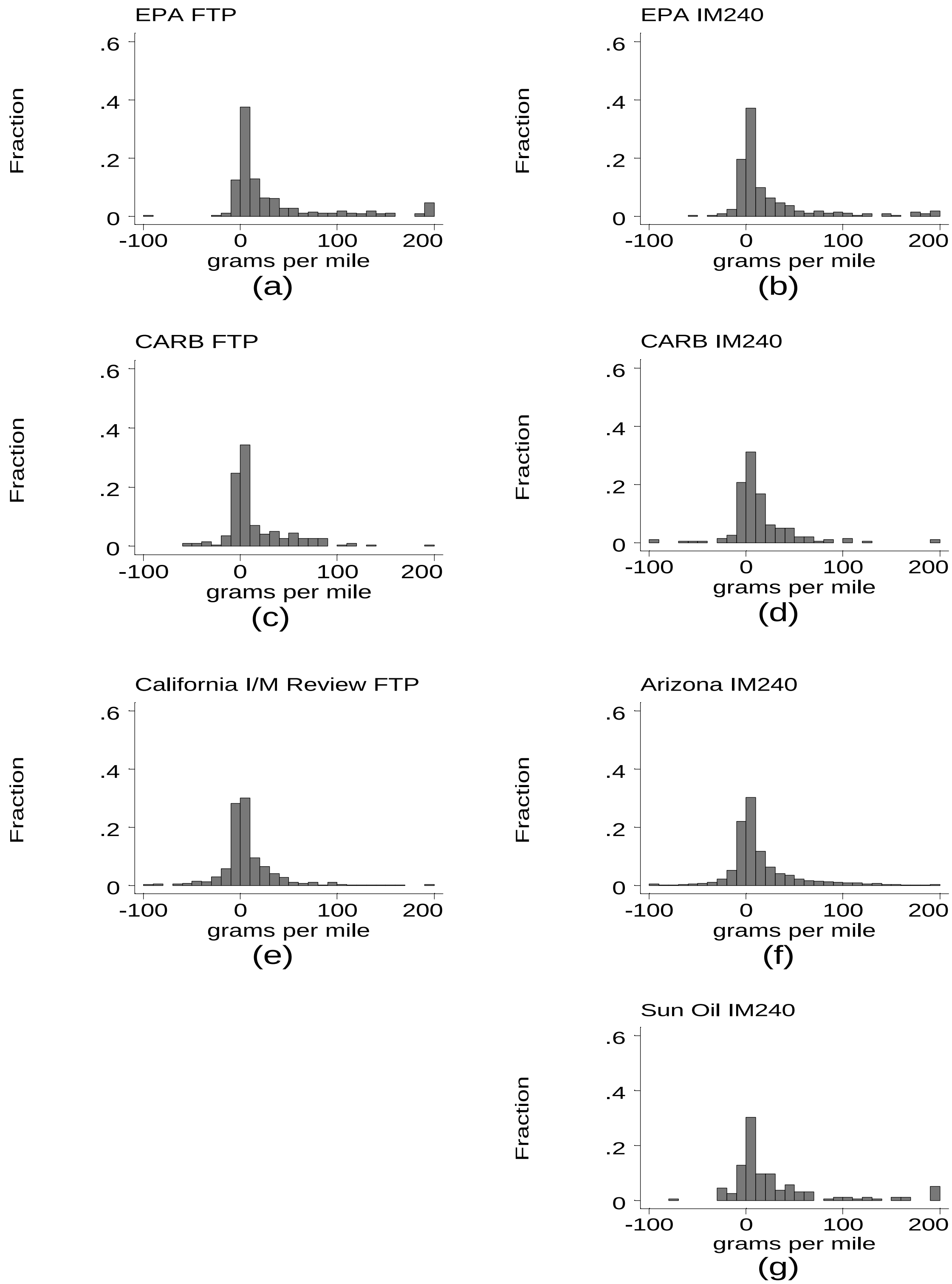
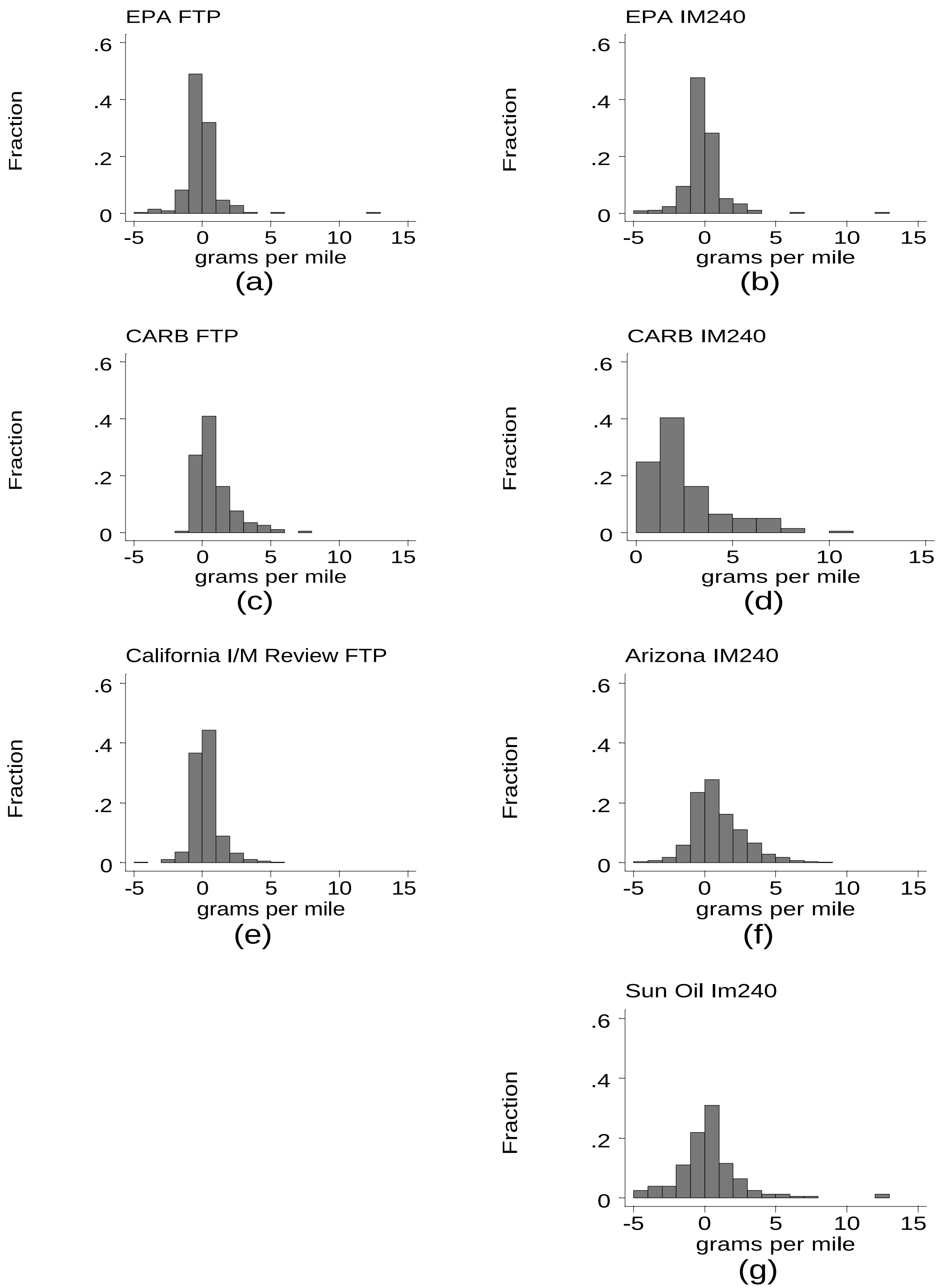


Figure D-6
Changes in NOx Emissions as a Result of Repair, Repair Datasets



Appendix E
Inspection and Maintenance Program Cutpoints grams per mile)

Arizona:

	HC	CO	NOx
LDV, 1981-82	2.0	60.0	3.0
LDV, 1983-90	2.0	30.0	3.0
LDV, 1991+	1.2	20.0	2.5
LDT1 and LDT2, 1979-83	7.5	100.0	7.0
LDT1 and LDT2, 1984-87	3.2	80.0	7.0
LDT1 and LDT2, 1988-90	3.2	80.0	3.5
LDT1 and LDT2, 1991+	2.4	60.0	3.0

EPA:

Vehicles included in the dataset used to estimate repair effectiveness were not selected on the basis of having failed a particular cutpoint. The MOBILE model can estimate the effects of an IM240 program having the following cutpoints for HC/CO and NOx :

HC/CO	NOx
0.6/10	1.5
0.6/12	2.0
0.6/15	2.5
0.8/15	3.0
1.2/20	no cutpoint
0.6/20	
0.8/20	

Any combination of the above cutpoints can be modeled. In addition, MOBILE can model an idle, loaded/idle or 2500/idle test with cutpoints of 1.2/220.

California Pilot Project:

	HC	CO	NOx
All Vehicles, 1981+	0.97	23.3	1.99
Passenger Cars and LDT, 1975-80	1.45	20.27	5.13
Passenger Cars and LDT, 1974-	5.83	65.63	5.67
Medium Duty Vehicles, 1980-	5.83	65.63	5.67

Sun Oil:

	HC	CO	NOx
1983+	0.8	15.0	2.0
1981-82	0.8	30.0	2.0
1980	0.8	30.0	4.0
1977-79	3.0	65.0	4.0
1968-76	3.0	65.0	6.0

Appendix F
Relationship between FTP and IM240 Tests: Regression Results

Table F-1
IM240-FTP Regressions
EPA Repair Dataset
Initial Emissions

Equation estimated is linear:
FTP emissions= a + b * IM240 emissions

Pollutant	R^2	b	S.E.	t
HC	.8827	1.561	.035	44.48
CO	.7701	1.073	.036	29.68

Table F-2
Equations to Predict FTP Emission Levels
From Indolene IM240 Results
(EPA Study)

HC (log fit)	N	X	a	b	R^2
1981-82	58	0.309	0.1382	1.0715	0.909
1981+ Oplp	24	0.315	0.1448	0.9654	0.879
1983+ Carb/Cllp/Air	73	0.195	0.0000	0.9745	0.905
1983+ TBI/Cllp	224	0.180	0.0000	0.9840	0.873
1983+ MPFI/Cllp	211	0.222	0.0000	0.9520	0.915
CO (log fit)	N	X	a	b	R^2
1981-82	58	2.140	0.0000	1.0040	0.943
1981+ Oplp	24	1.640	0.3090	0.6510	0.904
1983+ Carb/Cllp/Air	73	1.579	0.0000	0.9060	0.873
1983+ TBI/Cllp	224	1.541	-0.1386	1.0720	0.782
1983+ MPFI/Cllp	211	1.696	0.0000	0.8860	0.780
NOx (linear fit)	N		a	b	R^2
1981-82	58		0.2534	0.7737	0.825
1981+ Oplp	24		0.0000	0.9306	0.976
1983+ Carb/Cllp/Air	73		0.0000	0.8925	0.961
1983+ TBI/Cllp	224		0.0767	0.8234	0.901
1983+ MPFI/Cllp	266		0.1250	0.7730	0.825

log (base 10)equation: $\log(\text{FTP emissions} - X) = a + b\log(\text{IM240 emssions})$
where X = estimate of engine start emissions
linear equation: FTP emission $= a + b$IM240 emissions

Table F-3
Emitter Group Emission Levels

Levels Used by EPA in Differentiating Repair Effectiveness Data into Emitter Groups (FTP Emissions)

High: greater than 0.82 g/mi HC or 10.2 g/mi CO but less than Very High
Very High: greater than 1.64 g/mi HC or 13.6 g/mi Co but less than Super
Super: greater than 10.0 g/mi HC or 150.0 g/mi CO

Levels Estimated for IM240 Emissions Based on Coefficients from IM240-FTP Regressions in Table F-1 above (IM240 emissions- used for splitting Arizona data into emitter group categories when only IM240 data are available)

High: greater than 0.36 g/mi HC or .77 g/mi CO but less than Very High
Very High: greater than .89 g/mi HC or 3.94 g/mi Co but less than Super
Super: greater than 6.24 g/mi HC or 131.04 g/mi CO

Appendix G
Vehicle Repair Information Form from Arizona I/M Program

APR 08 '97 09:08AM SIERRA RESEARCH

VEHICLE REPAIR INFORMATION

To qualify for a retest you must complete this form and return within 60 days of the first test.
Please see the brochure "Failed Vehicle Information"
For repair information and diagnostic assistance.

EQUIPMENT REPAIR INFORMATION Check the appropriate boxes for those repairs performed on the vehicle.

- ☐ AIR INJECTION SYSTEM
- ☐ CATALYTIC CONVERTER
- ☐ INLET RESTRICTOR
- ☐ GAS CAP

MARICOPA COUNTY ONLY
- ☐ POSITIVE CRANKCASE VENTILATION (PCV)
- ☐ EVAPORATIVE EMISSIONS SYSTEM

EQUIPMENT REPAIR COST ☐☐☐ (TO THE NEAREST WHOLE DOLLAR)

PRESSURE & PURGE REPAIR INFORMATION - MARICOPA COUNTY ONLY Check the appropriate boxes for those repairs performed on the vehicle.

- ☐ CANISTER
- ☐ PURGE VALVE
- ☐ HOSES / FITTINGS
- ☐ GAS CAP
- ☐ CHECK VALVES
- ☐ OTHER ______

PRESSURE/PURGE REPAIR COST ☐☐☐ (TO THE NEAREST WHOLE DOLLAR)

EMISSIONS REPAIR INFORMATION Perform the following and make necessary repairs and adjustments in accordance with manufacturer's specifications and procedures. Check appropriate boxes for repairs performed.

A
- ☐ EMISSIONS FAILURE DIAGNOSIS
- ☐ DWELL / TIMING
- ☐ AIR INTAKE SYSTEM (Air filter, Choke/Cold Start System, EFE/Heat Riser, TAC)
- ☐ PCV SYSTEM
- ☐ VACUUM LEAKS
- ☐ AIR / FUEL MIXTURE
- ☐ IDLE SPEED

B
- ☐ PLUG WIRES
- ☐ SPARK PLUGS
- ☐ SPARK CONTROL (Coil/Coil Packs, Distributor, Pickup Coil, Module, Rotor Cap, Mechanical)

C
- ☐ CARBURETOR/FUEL INJECTION Clean / replace injectors or repair carburetor components (float setting, power valve, etc.)

D
- ☐ EVAPORATIVE EMISSIONS SYSTEM
- ☐ OXYGEN SENSOR
- ☐ THROTTLE POSITION SENSOR
- ☐ COOLANT TEMP. SENSOR
- ☐ AMBIENT SENSOR
- ☐ MAP/BARO SENSOR
- ☐ MASS AIR FLOW SENSOR
- ☐ RPM SENSOR
- ☐ ELECTRONIC CONTROL MODULE
- ☐ EGR SYSTEM
- ☐ AIR INJECTION SYSTEM
- ☐ CATALYTIC CONVERTER
- ☐ ELECTRICAL PROBLEMS (Connectors, broken wires, corrosion, battery, alternator)
- ☐ COOLANT SYSTEM
- ☐ FAULT CODES (List in comment section)

EMISSIONS REPAIR COST ☐☐☐ (TO THE NEAREST WHOLE DOLLAR)

DIESEL REPAIR INFORMATION Check the appropriate boxes for those repairs performed on the vehicle.

- ☐ AIR CLEANER
- ☐ FUEL INJECTION TIMING
- ☐ FUEL INJECTORS
- ☐ FUEL PUMP
- ☐ SET AIR/FUEL RATIO
- ☐ PUFF LIMITER/ANAROID VALVE
- ☐ EGR
- ☐ ELECTRONIC FUEL/ENGINE CONTROL
- ☐ FAULT CODES (List in comment section)

DIESEL REPAIR COST ☐☐☐ (TO THE NEAREST WHOLE DOLLAR)

TO BE FILLED OUT BY REPAIR FACILITY OR VEHICLE OWNER (Please print or use rubber stamp)

FACILITY PHONE NUMBER ☐☐☐-☐☐☐☐

- ☐ HIRED REPAIR FACILITY
- ☐ DO-IT-YOURSELF REPAIR

Person or Facility Performing Repairs

Street

City Zip

I certify under penalty of law that I have completed the above repairs solely in an attempt to bring this vehicle's emissions into compliance.

Mechanic's Name (Please Print)

Mechanic's Signature

Mechanic's Comments:

Appendix H
Sensitivity Analysis Assumptions

MOBILE Input File Used in Sensitivity Analysis Base Case

```
1         PROMPT
    base IM240 1.2/20/3 cutpoints
1         TAMFLG   - Default tampering rates.
1         SPDFLG     -One value of average speed for all vehicle types.
1         VMFLAG    - MOBILE VMT mix.
1         MYMRFG    - MOBILE accumulation rates and registration distributions
1         NEWFLG      - MOBILE exhaust emissions
2         IMFLAG - One I/M program.
1         ALHFLG    - no additional correction factors.
1         ATPFLG -  No ATP
1         RLFLAG  - refueling emissions reflect onboard VRS
1         LOCFLG - LAP record will appear in each scenario.
1         TEMFLG  - temperature calculated from max and min
6         OUTFMT - Spreadsheet output format
4         PRTFLG - Print exhaust HC, CO and NOx results.
1         IDLFLG   - no idle emission factors calculated
3         NMHFLG - Calculate emissions for volatile organic hydrocarbons.
3         HCFLAG - Print sum (excludes refueling) and component; evap in gms.
92 20 75 50 00 00 100 112 2222 4211 1.20 20.0 3.00              I/M 240 test
1 95 19.6 75.0 20.6 27.3 20.6 07                        Scenario description record
Year 1995        C  72.  92. 11.5 08.7 92 1 1 1 1      Local Area Parameter record
1 97 19.6 75.0 20.6 27.3 20.6 07                        Scenario description record
Year 1997        C  72.  92. 11.5 08.7 92 1 1 1 1      Local Area Parameter record
000000000000000000000000000000000000000000000000000000000000000000000
00000000000
```

Local Area Parameter Record, Scenario Description Record and I/M Record Used in Sensitivity Analysis
Base Case

Local Area Parameter Record
- fuel volatility class: C
- minimum temperature: 72
- maximum temperature: 92
- "period 1" RVP: 11.5
- "period 2" RVP: 8.7
- "period 2" RVP start year: 1992

Scenario Description Record
- region: 1 (low altitude)
- calendar year: 1997
- speed: 19.6
- ambient temperature: 75
- operating modes: 20.6, 27.3, 20.6
- month of evaluation: July

I/M Record
- start year: 1992
- stringency: 20%
- first and last model years covered by program: 1975, 2050
- pre-1981 model year waiver rate: 0%
- post-1981 model year waiver rate: 0%
- compliance rate: 100%
- inspection only
- biennial inspection
- LDGV, LDGT1, LDGT2, HDGV covered by program
- I/M type: IM240 with HC/CO/NOx cutpoints of 1.2/20/3 g/mi

Appendix J
Vehicle Registration Distributions