OMCHS SAFETY PROGRAM PERFORMANCE MEASURES:

ASSESSMENT OF INITIAL MODELS AND PLANS

FOR SECOND GENERATION MODELS

DRAFT

5-28-1999

Prepared for:
Federal Highway Administration
Office of Motor Carrier & Highway Safety
Analysis Division, HIA-20
400 Seventh Street SW
Washington, D.C. 20590

Prepared by:
John A. Volpe National Transportation Systems Center
Economic Analysis Division, DTS-42
Kendall Square
Cambridge, MA 02142

Table of Contents

Preface

This report has been prepared for the Federal Highway Administration's (FHWA) Office of Information Analysis, Analysis Division (HIA-20) by the John A. Volpe National Transportation Systems Center (the Volpe Center). It documents a project that was undertaken to identify the key elements that constitute the FHWA motor carrier safety program and to define a means to assess their effectiveness through the use of performance measures. The work addresses the requirements of the Government Performance and Results Act (GPRA) of 1993, which obligates federal agencies to measure the effectiveness of their programs as part of the budget cycle process.

Work on Office of Motor Carrier and Highway Safety (OMCHS) Program Performance Measures was initiated during FY93. In December 1994, a report titled "Office of Motor Carrier Safety Program - Performance Measurement" was prepared. That report provided a comprehensive breakdown of OMCHS programs and described about a dozen potential evaluation models. Based on OMCHS's review of that document, the Volpe Center revised the report in July 1998 to provided an evaluation model and approach to assess two key OMCHS programs: roadside inspections conducted by participating states under the Motor Carrier Safety Assistance Program (MCSAP) and on-site compliance reviews conducted by the OMCHS field offices. This report also provided preliminary results of these programs using 1996 data in these initial models. In September 1998, the OMCHS convened an expert panel to review the Volpe preliminary report on the initial roadside inspection and CR program performance measure models. Subsequent to the expert panel, the Volpe team produced additional findings concerning the initial models and their limitations related to future implementation.

The OMCHS has identified the need to measure the effectiveness of the third safety program, the OMCHS's traffic enforcement program. Also, there are plans for a fourth evaluation effort to provide a comprehensive assessment of the combined effect of the entire package of OMCHS programs.

This report summarizes the work that has been conducted to date on the initial program performance models, as well as the review panel comments and additional findings by the research team. This report also introduces second generation models for the compliance review and roadside inspection programs. Furthermore, this report provides an opportunity to present approaches that have been explored for evaluating the traffic enforcement program as well as for assessing the combined effect of all programs

The staff at the Volpe Center drew on their experiences working on OMCHS and other transportation safety related projects. These projects included the development of an improved process for motor carrier safety fitness determination and an analytical information system tool (SafeStat), which identifies and monitors unsafe motor carriers for the compliance review program and the Performance & Registration Information Systems Management (PRISM) project. Individuals who contributed the ideas and material presented in this report are:

Dale Sienicki of the FHWA's Office of Motor Carriers, HIA-20, and the OMC Project Manager;

Donald Wright, the Volpe Center Principle Investigator;

David Madsen of the Volpe Center;

Dr. Thomas M. Corsi, Logistics and Transportation, The Robert H. Smith School of Business, University of Maryland, College Park, Maryland.

I. Project Background

A. Objectives and Scope

Since the early 1980s, Congress has passed several acts that strengthened federal motor carrier safety regulations and led to Federal Highway Administration (FHWA) programs to enforce them. The Surface Transportation Assistance Act of 1982 established the Motor Carrier Safety Assistance Program (MCSAP), a grants-in-aid program to States to conduct roadside inspections and enforcement programs aimed at commercial motor vehicles. The 1984 Motor Carrier Safety Act directed the Department of Transportation (DOT) to establish safety fitness standards for carriers. In response to this legislation, the FHWA, in conjunction with the states, implemented the MCSAP to establish and fund the roadside inspection and enforcement program as well as the Safety Fitness Determination Process (SFDP) and rating system based on on-site safety audits called compliance reviews (CRs).

It is believed that a major benefit of these programs has been and will continue to be an improved level of safety in the operation of commercial motor vehicles. The John A. Volpe National Transportation Systems Center (the Volpe Center) has helped the FHWA's Office of Motor Carrier and Highway Safety (OMCHS) systematically evaluate the effectiveness of its motor carrier safety programs as called for in the requirements of the Government Performance and Results Act (GPRA) of 1993. The short-term objective of this work effort has been to measure the impact of the safety program activities on reducing crashes involving motor carriers.

In July 1998, the Volpe Center issued a report providing a methodology and approach to evaluate the OMCHS's roadside inspection and compliance review programs through two models.[1] In addition, some preliminary results for these programs were presented using 1996 data in these initial models. The OMCHS further identified the need to measure the effectiveness of the OMCHS's traffic enforcement program; methods to assess the effectiveness of this program are currently being explored. Finally, future plans for program performance effectiveness include a fourth evaluation effort to provide a comprehensive assessment of the combined effect of the entire package of the OMCHS programs. It is hypothesized that the combined effect model would assess the effects of the three separate programs on carriers with different levels of exposure to each.

In September 1998, the OMCHS convened an expert panel to review the Volpe preliminary report on the initial roadside inspection and CR program performance measure models. This panel issued a formal review of the Volpe models on September 11, 1998.[2] Subsequent to that report, the Volpe team produced additional findings concerning the initial models and their limitations related to future implementation.

The purpose of this report is to synthesize the work that has been conducted to date on the initial program performance models, and, in light of the review panel comments and additional findings by the research team, to introduce second generation models for the compliance review and roadside inspection programs. Furthermore, this report provides an opportunity to present approaches that have been explored for evaluating the traffic enforcement program as well as for assessing the combined effect of all programs.

---

[1] John A. Volpe National Transportation Systems Center, Economic Analysis Division, DTS-42, OMC Safety Program Performance Measures, July 1998 (Revised on December 18, 1998), prepared for Federal Highway Administration, Office of Motor Carriers.

[2] Robert M. Nicholson, Expert Panel Review of the Office of Motor Carrier's Performance Measure Model Development Activities, October 23, 1998, prepared for Federal Highway Administration, Office of Motor Carriers.

B. Summary of Initial Assessment Models

1. Compliance Review: Impact Assessment Model

The compliance review (CR) is perhaps the single greatest resource-consuming activity of the OMCHS. In addition to actually conducting CRs, the OMCHS invests in extensive analysis of the requirements of the Federal Motor Carrier Safety Regulations (FMCSRs), the design of the CR to assess safety performance and compliance with the FMCSRs, safety investigator training, prioritization methodologies and systems such as SafeStat[3] to determine who should receive CRs, and information systems to report and store the results of the CRs that are conducted. When performing CRs, OMCHS and state safety investigators spend many hours examining the safety records of individual motor carriers to assess their compliance and safety performance. The investigators also discuss their findings with the carriers' safety managers to improve understanding of the carriers’ safety programs. After the reviews are completed, the results are incorporated with other safety data in SafeStat to reassess carrier safety status and are used to assign overall safety ratings (i.e., satisfactory, conditional, unsatisfactory). In the instances where serious violations are discovered, enforcement cases are initiated and fines may be imposed. It is intended that through education, heightened safety regulation awareness, and enforcement effects of the CRs, carriers will improve the safety of their commercial vehicle operations and, ultimately, reduce their crash rates. Several thousand CRs are conducted each year. In 1996 alone, over 8,000 CRs were conducted on individual motor carriers by federal and state enforcement personnel.

The CR Impact Assessment Model was developed to determine the effectiveness of the CR program element. The model is based on the individual and cumulative "before and after" changes in safety performance of carriers that received CRs. The approach taken by the model is to analyze changes in a motor carrier's safety performance in a time period after an on-site compliance review in comparison to its safety performance prior to that review. This analytic model shows the direct impact of the performance of compliance reviews on carrier safety. The model uses data collected during CRs that include the carrier's recordable crashes (crashes involving injuries and/or fatalities and "towaways," in which an involved vehicle cannot leave the crash scene due to damage) and vehicle miles traveled (VMT) during the 12 months preceding the review. The results of CRs are stored in the OMCHS's Motor Carrier Management Information System (MCMIS), which is the source of the data used by the model. The model measures the collective changes in individual carrier crash rates between two successive reviews one year or more apart, effectively capturing the effect of the first review.

As indicated, the model uses crash rate information collected at the time of the CR. As part of the CR procedure, investigators are required to obtain the number of recordable crashes in which a carrier was involved over the past 12 months as well as the VMT by the carrier's fleet over the same 12 months. Therefore, crash rates (in the form of the number of recordable crashes per million VMT) for all carriers having received CRs can be calculated. Because the CR Impact Assessment Model determines the change in crash rates from before to after the CRs, it required not only pre-CR crash rates but also crash rates after the CRs. Therefore, the model only considers carriers with two CRs. The earlier (or initial) CR provides the pre-CR crash rate data and the subsequent (or follow-up) CR provides the post-CR crash rate data.

The first step performed by the model involves determining the percentage change in the combined crash rates of carriers prior to and after the CRs. Using CR data from MCMIS as of June 1997, the research team conducted the empirical analysis for carriers with two or more CRs between April 1993 and June 1997 that were conducted from 1 to 2 years apart. 1,738 carriers met these criteria. These carriers had a cumulative crash rate

---

[3] SafeStat (Safety Status Measurement System) is an automated, data-driven analysis system developed at the Volpe Center that is designed to incorporate on-road safety performance information and enforcement history with on-site compliance review information in order to measure the relative safety fitness of interstate motor carriers. Since April 1997, SafeStat has been used by OMCHS to identify and prioritize interstate motor carriers for CRs.

of 0.750 crashes per million VMT based on the initial CRs and a cumulative crash rate of 0.661 based on the follow-up CRs - a 12 percent reduction in crash rate. Therefore, the research team concluded that CRs do have a positive effect on carriers by lowering their crash rates and that the best estimate of that effect is a 12 percent reduction in crash rate on an annual basis during the initial year after the CRs.

The next step in building the model was to estimate the total number of crashes avoided that were attributed to all CRs, not just multiple CRs conducted from one to two years apart on the same carriers. This was done by applying the average reduction in crash rates (12%) to a baseline crash rate (the average crash rate for all carriers receiving CRs). This provided a crash-avoided rate per million VMT. It is believed that, although the effect of the CR on crash reduction diminishes over time, it is still present after one year. Therefore, the effects of the CR on crash reduction in ensuing years needed to be estimated as well. The model assumed CRs to effect crash rates for three years, declining in influence each year. The decision to reflect the influence of the CR over a three-year time frame was based on input from OMCHS field staff. They believed that the safety ratings generated from CRs became obsolete three years after they were performed due to the volatile nature of the industry (e.g., changes in personnel, equipment, and operations), which causes the effect of the review to "wear off." Therefore, the model limited the effect of the CR in reducing crashes to three years after the initial CR under the belief that, on average, the effect of the CR on carrier safety performance beyond this period of time is insignificant relative to other factors that influence safety performance. In order to reflect the full multi year effects, the exposure of the reviewed carriers was estimated in VMT for each of the three years following the CR. This VMT estimate involved adjusting for industry growth as well as for attrition.

Next, the benefits were calculated by applying an average cost per crash to the number of avoided crashes. The National Public Service Research Institute average weighted cost of truck crashes ($135,000) was used as the basis for benefits calculations. This figure was multiplied by the expected number of crashes avoided to estimate the benefits of the CR program.

The model was applied to all 8,111 carriers that received CRs in 1996. Over the following year, it was estimated that they would have traveled 24.7 billion miles. Applying a base crash rate of 0.70 recordable crashes per million VMT over these miles, produces an estimated 17,162 expected crashes. Of these crashes, 12% (or 2,040 crashes) were avoided due to the effects of the CR program in the first year. Using the aforementioned assumptions regarding VMT changes and crash rate reductions in each of the three years, the model estimated that 1,476 and 801 crashes would be avoided in the second and third years (1997 and 1998), respectively. The total number of crashes avoided in all three years was 4,317. By applying the estimated average crash cost to the 4,317 crashes avoided, the model estimated a benefit of approximately $580 million from the CRs administered in 1996. This represents an average of over $71,000 in cost savings per review.

2. Roadside Inspection: Safe-Miles Model

A vital component of the OMCHS effort to improve truck safety on our nation's highways is the roadside inspection program. As a result of funds provided by OMCHS through the MCSAP, states perform inspections of vehicles and drivers at fixed and mobile sites to ensure compliance with the FMCSRs, Hazardous Material Regulations (HMRs), and related state laws. These inspections have increased significantly throughout the late 1980s and early 1990s. Most recently, 4,500 full and part-time state personnel nationwide performed more than 2.0 million inspections in 1998.

Inspections performed under the MCSAP are conducted in accordance with the North American Standard (NAS) developed by the Commercial Vehicle Safety Alliance (CVSA) in cooperation with the FHWA. These standards establish national uniform inspection procedures and criteria for removing unsafe vehicles and drivers from the highway. Serious violations that are detected result in the vehicle/driver being placed "out-of-service" until the deficiency is remedied. The out-of-service violations encompassed within the NAS preclude operation of a commercial motor vehicle by its driver for a specified period of time or until a particular defect is corrected or condition met.

The Safe-Miles Model was developed to determine the effectiveness of the roadside inspection program. The Safe-Miles Model is based on the belief that the roadside inspection program has both direct and deterrent (or preventative) effects, each of which reduces crashes. The direct effects of the program are based on the corrections of out-of-service deficiencies which, if not corrected, would have increased the likelihood of crashes occurring. The deterrent effects of the program are based on the safety improvements made by carriers due to their awareness of the program and the potential consequences that the program can have on their operations.

a. Direct Effects

The Safe-Miles Model provides the necessary link between an actual roadside inspection and its safety benefits by defining a specific procedure for estimating the number of crashes avoided as a result of detecting and correcting these out-of-service conditions for each vehicle inspection that uncovers either a vehicle or a driver with an out-of-service condition. The inspection results in changing the future operation of the inspected vehicles/drivers from “unsafe” to “safe” conditions, thereby reducing the probability of their being involved in crashes.

The starting point is the detection during any roadside inspection of either a vehicle or a driver out-of-service condition (all levels of roadside inspections were included in the analysis). Direct benefit “credits” are accumulated when vehicles or drivers with out-of-service conditions are detected and taken out-of-service. Thus, the initial step is to determine the number of vehicles placed out-of-service for vehicle and for driver problems during a given time period. In cases where an inspection resulted in the detection of at least one vehicle defect and one driver problem, the vehicle is counted in both the driver and vehicle defect categories.

Once the inspections were reviewed and those involving vehicle and/or driver defects are identified, it is necessary to determine the length of time that the benefits from detecting and correcting vehicles/drivers defects will last into the future. Since the CVSA has a three-month period after a satisfactory inspection during which the vehicle is exempt from additional inspections, such a time period was considered appropriate for establishing a “safe” post-inspection period. There was no empirical basis governing the duration of the effects of finding out-of-service defects for drivers. In order to adopt a more conservative approach, the post-inspection safe period for corrected driver out-of-service defects was shortened to two months.

The model is based on the hypothesis that the inspection, detection, and correction of defects leads to a subsequent lowering of base crash rates, i.e., the crash rates that would have been obtained in the absence of the inspection program. Since this rate is not known, the overall rate at which crashes occurred in the general motor truck population was used. Based on 1994 data, the overall crash rate is set at 0.885 crashes per million miles.

The subsequent steps in the model development process involve an assessment of the probability that the particular vehicle or driver problem detected in an inspection would, in the event of a crash, have contributed to its occurrence or be identified as a causation factor in the crash. It is precisely these crashes that can be avoided due to the inspection program. Thus, if there are twenty crashes per twenty million vehicle miles, and ten percent of these crashes have vehicle defects as contributing factors, then two crashes will be avoided over the twenty million miles traveled if the defects detected during roadside inspections are corrected subsequent to them.

Results from a study at the University of Oregon show that approximately 4.6 percent of all commercial vehicle crashes involved truck mechanical defects (i.e., defects involving tires, steering, brakes, lighting devices, securing cargo, windshield, etc.) as a factor contributing to the crash.[4] It is believed that the 4.6 percent figure

---

[4] S.G. Miller, P.E. Montagne, S.U. Randhawa, and C.A. Bell, Out-of-Service Criteria for Commercial Vehicles: Evaluation of Accident Data in Relation to Vehicle Criteria, September 1996, Transportation Research Report, 96-6, Transportation Research Institute, Oregon State University, p. 58.

is a conservative estimate since crash reports underreport vehicle defects as contributing factors due to a lack of thorough post-crash investigations and limited mechanical knowledge of investigators.

Based on the directly linked factors (intoxication/possession of illegal substances, fatigue, other driver-factors) and the indirectly linked factors (i.e., inattention due to hours-of-service violations), the Volpe research team estimates that 5.7 percent of the crashes had driver-contributing factors that could be identified during a roadside driver inspection. Again, as with the estimate of crashes due to vehicle defects, this estimate is felt to be very conservative due to a lack of time for thorough post-crash investigations and evidence available to accident investigators. By applying the overall expected crash rate to the total number of “safe-miles” generated as a consequence of vehicle and driver out-of-service inspections, an estimated expected number of crashes (had there been no inspections) is developed for both totals. The National Public Service Research Institute average weighted cost of truck crashes ($135,000) was used as the basis for benefits calculations. This figure was multiplied by the expected number of crashes avoided to estimate the direct benefits.

b. Deterrent Effects

An equally important element of the roadside inspection program is the impact that the very existence of the program has upon managers of motor carriers and drivers. The realization that annually approximately 2 million roadside inspections of motor vehicles are conducted nationwide has led to permanent changes by motor carrier managers in their vehicle maintenance and inspection procedures and their driver qualifications and behavior. These indirect or deterrent impacts of the roadside inspection programs are assumed to be substantial, but much more difficult to estimate. For the Safe-Miles Model, these effects are estimated by assuming that they are partially reflected in a motor carrier’s awareness of the program as a function of the number of vehicle and/or driver inspections that it has had. If this assumption is true, the deterrent effect would cause carriers with a higher frequency of (that is greater exposure to) either driver or vehicle inspections to change their behavior and voluntarily improve their safety, resulting in lower driver and vehicle out-of-service rates.

By studying data on driver inspections conducted during 1995 and 1996, carriers with six or more driver roadside inspections showed a 25 percent improvement in their overall driver out-of-service rate versus carriers who have had only one driver inspection. In reviewing the vehicle roadside inspections, a similar improvement in vehicle out-of-service rate was observed. However, in the case of vehicle inspections, the improvement didn’t occur until carriers had more than 40 inspections over the two-year period. For these carriers, the improvement was 21 percent.

The Volpe Center researchers estimated total 1996 program benefits from the roadside inspection program at $86 million: $47 million in direct benefits and $39 million in deterrent benefits.

C. Recognized Limitations of Initial Models

The Volpe report recognized that both the Impact Assessment and the Safe-Miles Models have limitations. These limitations were discussed in the report and became a starting point for the evaluation of the models by the expert panel.

1. Impact Assessment Model

Several limitations associated with the initial CR Impact Assessment Model were recognized by the model developers. Specifically, further work needs to be done to empirically establish duration of impact of a CR and reduction of its impact over time. The model assumes a diminishing impact from a 12 percent reduction in the first year, to an 8 percent reduction in the second year, to a 4 percent reduction in the third year. While the decline assumption is intuitively reasonable, further empirical work could confirm or improve upon the rate of decline and the duration of the impact assumptions that are currently used in the model.

Second, the model is based upon results from carriers receiving two CRs at least twelve months apart within the study period. This stipulation was necessary in order to develop before and after crash rates for the study group of carriers. Only carriers with two separate CRs could provide the before and after crash rate base, since the two CRs were the instruments used to collect total crashes and VMT, the necessary ingredients for calculating crash rates. However, the model does not differentiate on the basis of whether carriers had multiple CRs within the study's time frame. Developing continuous, long-term carrier crash rate profiles, not dependent upon data collected during a CR, would enable researchers to compare before and after CR crash rates based on data from carriers with a single CR.

Third, the initial CR Impact Assessment Model measured improvement based on the subset of carriers who received two CRs 12-24 months apart (in step 1) and then applied this improvement to all carriers who received a CR during the study period (in step 2). There was an assumption that the subset of carriers in step 1 experienced the same reduction in crash rate as all carriers who received a CR. Carriers can get follow-up reviews for a variety of reasons -- e.g., the carrier required an enforcement case follow-up, a non-frivolous complaint was filed against the carrier, the carrier itself requested a new CR to improve its rating. Some of these reasons for a follow-up review may potentially make the subset of carriers with multiple CRs 12-24 months apart different from the rest of the carriers not receiving follow-up reviews. It would be preferable to base the reduction on a representative sample of all carriers experiencing reviews rather than just the subgroup subjected to second reviews.

Fourth, the model does not differentiate impact based on any characteristics of the carriers involved. It is important to realize that the objective of the Impact Assessment Model is to provide overall program evaluation. It should not be viewed as dictating that a one-size-fits-all approach is in the best interest of achieving overall maximum safety performance. Indeed, CR resources may be better spent on a particular size of carriers, while it may be more efficient to use other types of safety-enhancing treatments for carriers of another size. Model advancements should give consideration to the differing impact of CRs depending on selected carrier characteristics, such as overall size.

Fifth, the Impact Assessment Model uses a before and after approach with the CR itself as the control event used to establish the "before" and "after" period. However, there are clearly other "events" in the "before" and "after" time periods that may impact a carrier's safety performance (crash rate). These events may consist of roadside inspections, enforcement cases, etc. The planned work on combined effects of this and the other OMCHS safety program components will help to isolate the effect of the CR.

2. Safe-Miles Model

The Volpe report identified a number of limitations associated with the Safe-Miles Model. First, while there is some empirical evidence that three months is a reasonable time period to use for accumulating safe-miles from vehicle defects detected during a roadside inspection as a result of CVSA procedures and practices, there is no such evidence for establishing a driver safe-mile time period. Future empirical assessment of driver behavior after driver out-of-service violations may help establish the reliability of the two-month period used in the initial version of the model

Second, there is a fundamental assumption in the model that detected defects will be repaired. To the extent that out-of-service defects are addressed and corrected after the inspection, then the assumption that safe-miles are being accumulated for vehicles not undergoing repair is faulty. The model may need to build in a factor to adjust safe-mile totals accordingly.

Third, the Safe-Miles Model assumes that 4.6 percent of all crashes have vehicle defects that contribute to crash causation. Better empirical determination of the causation factor may yield the finding that more safe-miles are accumulated and crashes avoided through defect detection at roadside inspections. However, solution of this issue is not simple. There are relatively few studies available that provide accurate assessment of the extent to which vehicle defects contributed to crash causation. Such studies are time-consuming and, due to incompleteness and accuracy of accident investigation reports, their results will always be questionable.

The model notes that a certain percentage of crashes involve driver inattention, but only a small fraction of those are based on driver deficiencies that could have been detected in a roadside inspection. The major factor contributing to inattention that could be detected during an inspection is, of course, violations of driver hours-of-service regulations. The current version of the model may understate the percent of the inattention situations identified as contributing to the cause of the crashes that could be linked to hours-of-service violations identified at a roadside inspection.

The model identifies vehicle and driver out-of-service factors that contribute to accident causation. However, there is not direct proof that the absence of the out-of-service condition will prevent any crash in which that condition is identified as a contributing cause.

The deterrence portion of the model relies on measured changes in out-of-service rates of carriers that have had multiple inspections as a basis for calculating indirect effects from inspections. Actually, the model undoubtedly understates the deterrence effects since there is most likely an overall deterrent effect across all carriers, regardless of the number of inspections they receive. However, the overall improved preparation and compliance of drivers and vehicles, motivated by the presence of a roadside inspection program, is difficult to measure. Specialized studies may be required to determine if the influence of deterrence can be widened in the model to account for the more general deterrence phenomenon.

D. Review Panel Assessment of Model Limitations

The Review Panel in its report reached a general agreement that Volpe's efforts in developing both the Impact Assessment and the Safe-Miles Models represented "a good first attempt to create models to permit the evaluation of several key OMCHS programs." Indeed, the Panel noted that the "general forms of the two models that were developed are logical and straightforward."

Nevertheless, the Panel concluded that the models rest on assumptions that require certain kinds of data that are currently unavailable. "Due to the absence of appropriate, reliable, valid data, particularly truck crash data, truck crash causation data, and truck driver performance data, the researchers had to make several assumptions which they incorporated into their models. There is no reason to believe that these assumptions are correct and, if they are incorrect, there is no basis for knowing how they might be biased." In fact, the Panel's consensus opinion is that "until such time as data such as these (as listed above) are available, it will not be possible to make the effectiveness evaluations desired."

For example, a key assumption of the Impact Assessment Model concerns the generalizing of crash reduction rates that are determined from a subset of carriers that have had multiple CRs. Specifically, the Impact Assessment Model relies on data from companies with two CRs conducted 12-24 months apart to develop crash reduction percentages to apply to all carriers receiving CRs. The Panel believed that companies receiving multiple CRs are not the same as companies receiving a single one. As a result, they believed, there is risk in generalizing the experience of firms with multiple CRs to those with a single CR. In addition, the Panel was uncomfortable with the Impact Assessment Model's assumption that effects of the CR would diminish by a fixed percentage in the second and third years. The Panel concluded that there was simply no justification for the percentage reductions assumed in the model. However, member(s) of the Panel came up with the suggestion that determination of the CR's impact on crash rates would require researchers to draw a systematic sample of motor carriers whose experience with CRs would be tracked over time as a basis for estimating the impact of CRs.

Opinions were expressed that the answer to this question would require data collected specifically to answer the question.

Turning to the Safe-Miles Model, the Panel felt quite uncomfortable with the assumption that the safe period after detecting a driver defect could be set at 60 days. The Panel argued that there was no basis for that number. Furthermore, in its opinion, "drivers repeatedly violate hours-of-service regulations in order to stay in business, even after being caught and fined." It concluded that any assumption about driver behavior post detection with an out-of-service situation would require a "thorough study and understanding of all the issues involved."

The Panel was also generally uncomfortable with the notion that if factors (driver or vehicle) contribute to the cause of a crash, then the crash would not have occurred in the absence of the contributing factor as implied in the Safe-Miles Model. Furthermore, the figure used of 4.6 percent of the crashes with vehicle defects and 5.7 percent with driver defects, as basic assumptions of the Safe-Miles Model, could be significant underestimates or overestimates of the extent to which driver and vehicle factors contribute to the causes of crashes. The Panel felt that data were inadequate to make this judgement and that the figures used by the research team lack any scientific validity.

Overall, the Panel believed that an assessment of the roadside inspection program should focus on developing "reasonable assumptions about percentages of crashes related to defects that could be prevented by detection of vehicle and driver out-of-service defects." Furthermore, the analysis must recognize that individual defects are not separable, but must be viewed as interrelated impacts contributing to the likelihood of crash involvement. As presently constituted, the Safe-Miles Model does not account for these interrelated impacts, but treats each defect as a separate contributing factor.

In conclusion, the Panel felt that the research team should place "more emphasis on experimental research" and rely less on data that are already available, having been gathered for other purposes. Such controlled study designs are relevant for both the evaluation of the compliance review and the roadside inspection programs.

E. Additional Findings Related to Future Implementation

Further analysis has been performed on the CR Impact Assessment Model using existing data in an attempt to address some of the limitations of the initial model. One study was conducted to examine the effects of policy changes related to carrier CR selection (national implementation of SafeStat for prioritization) on the results of the CR Impact Model. Another study examined the results of the model using an alternative means of obtaining crash rates.

1. Effects of SafeStat Prioritization on the CR Impact Assessment Model

One of the assumptions of the initial CR Impact Assessment Model was that the subset of carriers with two CRs one to two years apart experienced the same reduction in crash rate as did all carriers that received a CR. Potential problems with this assumption are that there might be significant reasons why the crash rate experience of the subset of carriers with follow-up reviews may not be indicative of all carriers that had reviews, especially carriers with no follow-up reviews. One of those reasons might be the way carriers are selected for CRs.

Before April 1997, many of the carriers were selected for CRs based on the Selective Compliance and Enforcement (SCE) algorithm. The SCE algorithm was heavily based on non-safety-related carrier characteristics, such as carrier size and type of commodities hauled. Therefore, some carriers were selected for initial CRs and even follow-up CRs based on non-safety characteristics. In of April 1997, the OMCHS replaced the SCE algorithm with the safety performance-based algorithm, SafeStat, for selecting carriers for CRs. SafeStat evaluates and ranks carriers strictly by their safety records. Carriers with the worst safety status are recommended to receive CRs.

The use of SafeStat should improve the overall effectiveness of the CR program by targeting carriers with the worst safety problems, because these carriers have the most potential to gain from such a program. However, the effects of using SafeStat may paradoxically show a reduction in the CR program's effectiveness, according to the initial CR Impact Model. Specifically, carriers that receive a first CR and then significantly improve their safety performance will not be re-selected by SafeStat and, therefore, will not receive follow-up CRs so their crash rate reductions will not be measured and used by the model. On the other hand, carriers that received CRs and did not significantly improve their safety performance will be re-selected by SafeStat for subsequent CRs and their crash rate reductions will be measured and used by the model. Therefore, the crash reduction rate that is applied to carriers that receive only one CR is understated. Because the assumption that the crash reduction experience by carriers that received follow-up CRs is indicative of all carriers that received CRs is false, the initial model will understate the effectiveness of the CR program.

An analysis was performed to verify the above hypothesis. Using MCMIS data available as of March 1998, the model showed that there were 1,675 carriers that had initial CRs and follow-up CRs one to two years later between the dates of March 1993 and March 1997, which is before SafeStat was employed for prioritization purposes. These carriers experienced the same cumulative crash rate reduction of 12% as was used by the initial model. Following the introduction of SafeStat to prioritize carriers (after April 1997), 746 carriers had follow-up CRs. These carriers showed a nominal crash rate increase of 1% after the initial CR.

These results suggest that the introduction of a more efficient means of carrier selection for CRs invalidates the initial CR Impact Assessment Model. The assumption that the crash reduction of the subset of carriers that received follow-up CRs is indicative of the crash reduction of all carriers that received CRs is erroneous when using safety performance-based criteria for the selection of CRs. While these results negate further use of the initial CR Impact Assessment Model, they give guidance to the construct of an improved model and suggest the need to obtain new data that will more accurately reflect the true post-CR crash rate reduction for all carriers and reduce the potential for biased results.

2. CR Impact Assessment Based on State-Reported Crashes

Prior to December 1998, the CR Impact Assessment Model estimated the performance of the OMCHS's CR program using crash rate information collected at the time of the review. As part of the CR, recordable crashes that occurred over the previous 12 months and annual vehicle miles traveled (VMT) data are used to calculate a carrier's crash rate over the past year. Because the CR Impact Assessment Model determines the change in crash rates before and after a CR, it required not only pre-CR crash rates but also crash rates after the CR. Therefore, the model only considers carriers with two CRs. The earlier (or initial) CR provided the pre-CR crash rate data and the subsequent (or follow-up) CR provided the post-CR crash rate data. The CR Impact Assessment Model measured the collective changes in individual carriers' crash rates between two successive reviews one year or more apart to calculate the effects of the initial CRs, which showed a 12 percent reduction in crash rates.

One of the limitations of the above approach is that the model is restricted to carriers with 2 successive CRs. These carriers are only a small subset of the all of the carriers that have received CRs. The effects of the CR program on this subset of carriers may not be indicative on the effects of the CR program on the entire carrier population that received CRs. Also, the CR Impact Assessment Model does not account for other external influences on motor carrier safety that could change over the time-frame used in the model.

A way to address the limitations of the CR Impact Assessment Model is to use crash rate information that has a wider coverage of the carrier population. One such source of crash information is the crash data reported by the States [according to the National Governors' Association (NGA) standard] that reside in the crash file of the MCMIS. Another source for measuring exposure that is available for all interstate carriers is the power unit (PU) information from the OMCHS's Motor Carrier Census. These data do come with several problems. The state-reported crash data are not complete and often take a lengthy period of time to be entered into MCMIS. PU data are not updated often nor on a periodic basis. Despite these problems, these data provide crash rates (in state-reported crashes per 1,000 PUs) for all interstate carriers, which can not be obtained by using other currently available data.

The Volpe Center researchers re-visited the CR Impact study using MCMIS data available as of September 1998. The study considered the pre- and post-crash rates of carriers that received CRs within a specified one-year period. The study used the above-mentioned crash rate of the number of state-reported crashes per 1,000 PUs. The differences in pre-CR and post-CR crash rates were calculated and compared to a control group that did not receive CRs within the time-frame of this study.

Both the reviewed and control carrier populations were divided into three strata based on size (i.e., the number of PUs): (1) 1-5 PUs, (2) 6-20 PUs, and (3) 21 or more PUs. The strata were defined in such a manner that each stratum had roughly the same number of reviewed carriers. All comparisons in the changes of crash rates were made within these strata.

This new study imposed requirements on the reviewed carriers such that the carriers must have received a CR between 6/30/96 to 6/30/97 and that the carrier must have at least 1 PU. When a carrier had more than one review during the specified interval, the first review was used. There were 6,194 carriers that passed the above requirements.

The control group was selected such that a carrier: (1) was not reviewed since 9/1/93; (2) was operating since 1/1/95; (3) must have at least one power unit; and (4) must have at least one roadside inspection between 1996 and 1998. There were 116,504 carriers that were selected according to the four criteria above.[5]

Since the carriers in the control group did not have review dates, by definition, a pseudo-review date was assigned to each control carrier. This was accomplished by selecting an actual review date from the reviewed carrier data and assigning it, at random, to a control carrier within the same stratum.

The crash rates were calculated by obtaining the number of state-reported (reportable) crashes that occurred within a year prior to the CR/pseudo-CR date and a year after the CR/pseudo-CR date. The number of pre-CR and post-CR crashes were divided by the number of PUs to obtain a pre- and post- CR crash rate, respectively.

The final comparisons were done in three steps. The first step was to compute the difference between the collective pre-CR crash rate and the post-CR crash rate for each of the two groups (reviewed and control) within a stratum. The second step was to compute the difference between the reviewed group rate differential and the control group rate differential within the stratum. If the reviewed group differential is greater than the control within a stratum, the difference between the two groups would show how much more the carriers in the reviewed group improved their safety than did the carriers not receiving CRs. The third step is to weight the improvement of the reviewed carriers in each of the strata (should it exist) by the amount of exposure represented by reviewed carriers in each stratum. This is done by calculating the percentage of PUs that are operated by the reviewed carriers in each stratum.

These percentages of exposure are then by multiplied by the percent of improvement shown by the reviewed carriers in each stratum. The resulting totals are summed to obtain the overall improvement of the reviewed carriers. The results are shown in the accompanying tables.

---

[5] A further requirement was applied to discard cases from the study where a carrier had unreasonably high or low crash rate in either the pre-CR or post-CR time periods. The Poisson distribution was used to eliminate 157 reviewed carriers and 325 control carriers that had unreasonable crash rates.

The results of the study showed that reviewed carriers overall experienced a reduction of less than 2 percent in crash rates relative to the carriers that did not receive CRs. This 2 percent reduction in crash rates using this new approach fails to confirm the previous results of the CR Impact Assessment Model, which showed a reduction of 12 percent. While this new approach addresses some of the limitations of the previous study, it also has its own limitations such as the incompleteness of crash information and inaccuracies in exposure (PU) data. The results do confirm the need to develop a new means of obtaining complete and accurate crash rate information on all carriers that received CRs and on comparable control groups to better estimate the effectiveness of the CR program.

Group 1: 1-5 PUs

	# of Carriers	Pre-CR Crashes	Post-CR Crashes	# of PUs	Pre-CR Crash Rate	Post-CR Crash Rate	% Improvement
Reviewed	2124	311	234	5468	56.9	42.8	24.8%
Control	86047	4564	4429	182665	25.0	24.2	3.0%
						Difference:	21.8%

Group 2: 6-20 PUs

	# of Carriers	Pre-CR Crashes	Post-CR Crashes	# of PUs	Pre-CR Crash Rate	Post-CR Crash Rate	% Improvement
Reviewed	2238	1288	1098	25445	50.6	43.2	14.8%
Control	22900	5055	4695	231143	21.9	20.3	7.1%
						Difference:	7.6%

Group 3: 21+ PUs

	# of Carriers	Pre-CR Crashes	Post-CR Crashes	# of PUs	Pre-CR Crash Rate	Post-CR Crash Rate	% Improvement
Reviewed	1832	5089	5133	152708	33.3	33.6	-0.9%
Control	7557	9233	9012	520933	17.7	17.3	2.4%
						Difference:	-3.3%

Effectiveness of CR Program

Reviewed Carrier Group	Total of PUs in Groups	% of PUs In Groups (A)	% Improve. of Reviewed Grp (B)	(A) x (B)
(1) 1-5 PUs	5468	3.0%	21.8%	0.6%
(2) 6-20 PUs	25445	13.9%	7.6%	1.1%
(3) 21+ PUs	152708	83.2%	None	0.0%
			Overall Effectiveness of All Review Grps:	**1.7%**

This study also indicated a strong relationship between the size of the carrier and impact of CR program. The Group 1 carriers (1-5 PUs) with CRs cumulatively experienced a substantial reduction in crash rate of 22 percent beyond that of carriers with no review. The reviewed Group 2 carriers with 6-20 PUs had crash rate reduction of 8 percent. However, the largest reviewed carriers in Group 3 (21 or more PUs) showed a slight increase in crash rates of 3 percent. This study indicates that relationship of size and program effectiveness is important, and therefore, should be accounted for in future models that measure the effectiveness of the CR program.

F. Revised Approaches: Second Generation Models-General Strategy

In preparing its draft report on the initial models for the Panel in July 1998, the research team was aware of the limitations of the Impact Assessment and Safe-Miles Models and documented such limitations in the report. The OMCHS asked the Panel to address a series of questions developed from the deficiencies list. In response to these questions, the Panel advised that the research team re-evaluate the identified models on the basis of new experimental research being conducted as part of the effort to evaluate the OMCHS programs. The additional research findings, with respect to the Impact Assessment Model by the research team, have re-confirmed the need to design new models based on the collection of additional data with experimental design in mind. The remainder of this report is devoted to a discussion of proposed second generation models for the Roadside Inspections program and the Compliance Review program, as well as the strategic plans for the definition of initial models for Traffic Enforcement and the combined effects model.

II. Compliance Review Program: Second Generation Approach

As shown in Section I.E.1, the initial CR Impact Assessment Model will no longer be able to accurately measure program safety performance now that carriers are selected for CRs using performance-based carrier selection criteria, such as the SafeStat algorithm. Although the model remains essentially valid, alternative means of measuring changes in crash rate (different than basing the rate on carriers that receive two or more CRs one to two years apart as used in the initial model) need to be found. The analysis in Section I.E.2 employed another means of measuring changes in crash rate using currently available data. However, known problems with the data used in that study (e.g., incomplete reporting of crashes and the lack of updating of PU information throughout the study's time frame) undermine the acceptance of that approach. Therefore, it is proposed that an effort be undertaken to obtain new data that will allow an unbiased estimate of post-CR crash rates. A modified "second generation" Impact Assessment Model will be developed based on these new data.

A. Establish Cross-Sectional Follow-up Data Collection Approach

Given that the post-CR crash rates should not be estimated from crash data from CR follow-ups, as is the case in the initial model, a new effort in collecting crash rate information on a cross-sectional sample of carriers that have received CRs is proposed. The input of these new data into a "second generation" CR Impact Assessment Model should lead to better measurement of the CR program effectiveness. These new data with periodic updates can provide a basis for annual measurements of CR program effectiveness. These data also can be used in analyses to determine the influence of certain factors on the program's effectiveness. For example, these data can be used to determine the duration and extent of impact that CRs have on carrier crash rates. Additional analysis can be performed with these new data to determine the effects of multiple CRs. The results of such analyses will lead to improvements in the accuracy of the CR Impact Assessment Model, as well as to provide potentially valuable insight on improving the CR program.

B. Establish Control Group of Carriers with No CR

The collection of new data should not be limited to carriers with CRs. A control group of carriers, carriers with no recent CR, should be carefully matched to the study group, carriers with recent CRs, and followed in tandem. Using a control group will address several of the limitations of the initial model and will improve the accuracy of the second generation model's results.

Use of a well-matched control group will provide a tracking mechanism for the influence of external factors (such changes in the driving environment, influences of other safety programs, improvement in truck technologies, etc.). Therefore, change in the study group crash rate that is not mirrored by a change in the crash rate of the control group can be attributed to the influence of the CR.

Members of the control group will be selected so that carrier size and location will be comparable to the study group. It is noted that matching the carriers in the study group and the control group with respect to their safety status will be difficult. Many of the carriers receive reviews because they are identified as “at-risk” according to SafeStat. It is the current OMCHS policy to review all “at-risk” carriers. The control group will not have any “at-risk” carriers and will differ from the study group with respect to their safety status. Ideally, the control group should be as much like the experimental group as possible. One way to compensate for having no “at-risk” carriers in the control group is to select the control group of carriers with some known safety problems, but not enough poor performance safety data to make the carriers “at-risk”. These carriers will be chosen from the SafeStat Single SEA list.

C. Develop Data Collection Plan and Analysis Framework

The study will require the development of a detailed data collection and analysis plan. That plan would have to address the following issues:

- Sampling Procedure and the sample size required to meet study objectives. Ensure adequate number of carriers to obtain statistical reliability in both the CR group and the control group.
- Determine the types of data to be collected
- Determine methods of data gathering (mail, phone, interview)
- Check data quality
- Develop necessary model framework modification and associated analyses

III. Roadside Inspection Program: Second Generation Approach

A. Direct Effects: Intervention Model

The Safe-Miles Model used a two-step analysis process to evaluate the effectiveness of the roadside inspection program. First, the model recorded instances in which vehicles and/or drivers were taken out-of-service during a roadside inspection. The second step involved converting the subsequent travel by the out-of-service vehicles and drivers, once conditions were corrected, into “safe-miles” and, then, estimating avoided crashes during this “safe-miles” period.

This two-step approach involved significant drawbacks. First, each out-of-service condition was treated as a separate outcome with an independent impact on crash probability. The two-step Safe-Miles Model didn’t increase crash probability when more than one driver and/or vehicle out-of-service conditions existed. Second, the post-crash time period during which safe-miles were accumulated was difficult to establish with any degree of certainty. Third, the safe-miles approach, by focusing on individual out-of-service factors, had to relate those factors to crash occurrences--i.e., the model implied that the finding and correcting of a single out-of-service condition could avoid a crash.

With these significant weaknesses of the initial Safe-Miles Model in mind, the research team conceived a second generation model, the Intervention Model, to address these concerns and limitations. The Intervention Model, like the Safe-Miles Model, has both direct and deterrent components. In fact, the deterrent effect parallels the approach documented in the Safe-Miles discussion, although further refinements in this approach are contemplated as part of the second generation models. The Intervention Model, like the Safe-Miles Model, begins with the out-of-service inspection being the initiating event. However, in contrast to the Safe-Miles Model, the Intervention Model evaluates the total out-of-service inspection results. Thus, inspections with multiple driver and/or vehicle problems are treated differently than inspections with a single out-of-service condition.

The Intervention Model makes the assumption that observed deficiencies at the time of a roadside inspection can be converted into crash involvement probabilities based on the type and number of these deficiencies. The underlying assumption of the Intervention Model is that deficiencies found during out-of-service inspections vary in severity and, as a result, vary in their crash involvement probabilities. Yet, the model posits that these differences can be estimated and used as a basis for estimating crashes avoided.

The challenge of the Intervention Model is to develop crash involvement probabilities based on occurrences of multiple and/or single vehicle and/or driver out-of-service conditions. The model’s underlying assumption is that multiple out-of-service conditions have higher crash involvement probabilities than do out-of-service inspections with a single out-of-service condition. It is also recognized that there may be exceptions in that certain single out-of-service conditions (i.e., driving under the influence of alcohol) may have a higher crash involvement probability than an inspection with multiple out-of-service conditions.

In addition, the Intervention Model, in contrast to the Safe-Miles Model, calculates crash involvement probabilities directly from the results of the out-of-service inspection without the intervening conversion to safe-miles. Therefore this approach eliminates a potential source of error in the Safe-Miles Model associated with the estimation of the safe-miles traveled after the inspection and correction of out-of-service deficiencies. An example may illustrate the concept more fully. After investigation, it may be ascertained that an out-of-service inspection with a driver found to be under the influence of alcohol and exceeding hours-of-service regulations combined with a vehicle with multiple defects has a high accident probability. In fact, it is the objective of the Intervention Model to develop crash involvement probabilities covering specific categories of out-of-service inspections. In this instance, the Intervention Model may place a crash involvement probability of .5 on this circumstance. The interpretation of the crash involvement probability is that the model would estimate a crash avoided for every two incidents of drivers caught under the influence of alcohol, exceeding hours-of-service regulations and driving vehicles with multiple defects. In the Intervention Model, these out-of-service conditions are not associated with a given number of miles or time period following an inspection, thereby eliminating the need to estimate these factors which were required by the Safe-Miles Model.

1. Critical Assumptions of New Approach

Since it is simpler and more straightforward, the Intervention Model requires fewer assumptions than does the Safe-Miles Model. However, an additional challenge posed by this approach is the determination of crash probabilities for the out-of-service conditions cited during a roadside inspection. In particular, the Intervention Model requires crash involvement probabilities for the entire range of circumstances covered during an out-of-service inspection. In general, an out-of-service inspection with no driver problems and a single vehicle defect will have a very low crash involvement probability. Under these circumstances, according to the Intervention Model there would have to be many such inspections before a single crash would be avoided and thus credited to the program.

2. Data Requirements

The Intervention Model will base its crash probabilities on two factors. First, some estimates of the frequency that out-of-service conditions (single or multiple) occur among truck vehicle miles. Second, the frequency that out-of-service conditions (single or multiple) occur in or contribute to a crash.

The determination of out-of-service conditions occurring among all truck miles will be based on studies similar in nature to the National Fleet Safety Survey. This study provided some indication of the extent to which observed out-of-service conditions exist among truck vehicle miles on a random basis. Combining this knowledge with frequency of observed out-of-service conditions identified in crash situations will provide the basis for determining crash probabilities for each single or multiple out-of-service condition.

An example will illustrate the basic concept of the Intervention Model. If, for example, bad brakes are found in 20 percent of vehicles inspected during a random roadside inspection, then the assumption will be that 20 percent of vehicle miles are with vehicles having bad brakes. If ten percent of the crashes had bad brakes identified as a crash causation factor, then we can estimate the number of crashes with bad brakes as a factor. There might, for example, be 10,000 crashes in a single year with bad brakes as the cause. Based on the estimated miles driven with bad brakes and the number of crashes with bad brakes, an estimate can be developed for the number of crashes with bad brakes per million miles traveled. Based on an average trip length available from the Census of Transportation (Truck Inventory and Use Survey), an estimated number of crashes per trip or the crash probability given a particular defect can be derived. If this figure is 2% of the trips with crashes, then we estimate that for every one hundred out-of-service inspections with bad brakes, two crashes are avoided.

This example illustrates the challenge of the Intervention Model. First, all inspections must be classified and grouped based on their combinations of out-of-service conditions detected. Second, estimates must be developed regarding the percent of truck miles in which stipulated conditions exist. At present, the best source of data for this will be the National Fleet Safety Surveys. Third, vehicle/driver defects must be related to their identification as factors contributing to the cause of crashes.

The model will also benefit from results on a crash investigation project currently underway by the Analysis Division in OMCHS to identify and measure vehicle and driver factors that are related to crash causation.

B. Indirect Effects: Improved Approach

As with the Safe-Miles Model, the Intervention Model also includes an indirect or deterrent impact of the Roadside Inspection Program. In general, the approach taken for measuring an indirect impact as outlined in the

Safe-Miles Model is incorporated as part of the Intervention Model. However, as stated in the discussion of the Safe-Miles Model, the research team believes the approach presented for measuring indirect effects does not capture the full deterrent effect. This subsection explores the feasibility of measuring deterrence on a broader scale.

1. Examination of Opportunities to Document Wider Impact of Deterrence

The discussion of deterrence in a previous section focused on carrier exposure to roadside inspections of drivers and vehicles as measures of deterrent impact as manifested in lower out-of-service rates once inspections reached threshold levels. It is hypothesized that the roadside inspection program has a broader deterrent impact than the evidence presented above suggests. There should be a systematic exploration of new opportunities to document a wider deterrent impact.

2. Need for Broader Approach

One of the main challenges, however, is to compile evidence to support the hypothesis of a wider deterrent impact. There needs to be some documentation that as carriers became more aware of the roadside inspection program, they initiated management policy changes and that these changes had an impact on crash rates. However, such documentation will require data collection over time and in a systematic fashion. The research team believes that discussions with fleet managers and/or the examination of individual carrier records over time may provide a basis for data gathering to document a wider deterrent impact than previously measured. Records on carrier inspections and the results of those inspections would be gathered where available. Longitudinal data on crashes or safety performance while somewhat more problematic, would also be obtained and related, where possible, to inspections.

IV. Traffic Enforcement Options

A subset of all roadside inspections are conducted as a consequence of commercial motor vehicle (CMV) traffic enforcement actions funded through the OMCHS Motor Carrier Safety Assurance Program (MCSAP). This section presents an initial investigation into alternative approaches for measuring the impact of these traffic enforcement actions on overall carrier safety performance. There is a discussion of three approaches to evaluate the impact of the traffic enforcement program. These approaches are labeled as: carrier-based; highway-based; and driver-based. Each, in turn, has some limitations and problems. Having presented these issues, there is a discussion of the need for additional research to investigate the means of measuring the impact of the traffic enforcement program.

A. The Carrier-based Model

The Carrier-based model begins with the investigation of the relationship between the level of carrier exposure to traffic enforcement actions and its crash rates. The focus is an investigation of whether carriers with a higher frequency of traffic enforcement actions in a given time period have a lower crash rate in a subsequent time period as a consequence of changes in policy, practices, and behavior initiated due to the exposure to traffic enforcement actions directed at commercial motor vehicles.

The model dictates the collection of carrier-based data over time on crash rates and traffic enforcement actions initiated against a carrier's drivers. During a given time period (one calendar year, for example), carriers will be stratified on the basis of the exposure of their drivers to traffic enforcement actions. The strata will indicate low, medium, and high traffic enforcement actions. Carrier driver or vehicle population will be used to standardize the data. The crash rates one year prior to and one year subsequent to the calendar year analysis period will be examined to determine the impact of traffic enforcement. The hypothesis is that carriers in the high enforcement group will have a higher crash reduction than the carriers in the low enforcement exposure group.

The challenge of this model is formidable. It is difficult to establish a situation in which enforcement actions during the study period are isolated to assess their impact on a crash rate comparison between the time prior to and the time subsequent to the inspection activity. Indeed, there is difficulty in separating out factors during the study period that might influence the observed difference in carrier safety performance in the pre- and post-periods other than the observed differences in traffic enforcement actions. Thus, there must be accurate records for each carrier of all OMCHS-related program activity. In addition, there needs to be a control group of carriers with minimal exposure to OMCHS programs.

Preliminary investigations of this model, without sufficient experimental controls, have not yielded statistically significant results.[6] It is believed, however, that not much weight should be placed on these results since the studies were done post-facto and did not involve any systematic research design.

B. Driver-based Model

There are many similarities between the driver- and carrier-based approaches. As expected, the focus of the Driver-based model is the individual driver and the impact of traffic enforcement actions on his/her crash record.

---

[6] John A. Volpe National Transportation Systems Center, Analysis of Three Proposed Approaches to Developing a Model of the Effectiveness of Commercial Motor Vehicle Traffic Enforcement Roadside Inspections, September 1998, prepared for Federal Highway Administration, Office of Motor Carriers.

The research design would use the traffic enforcement record of a driver during a given time period as the study treatment period. Crash performance differences during the time period immediately prior to and immediately subsequent to the study period would serve as the basis of comparison. The underlying assumption would be that enforcement activity during the study/treatment period impacts crash rate performance differences that are observed from the time period before to the time period after the study/treatment period.

As in the Carrier-based model, the Driver-based model would have to control for other factors influencing driver behavior during the analysis period. These could include non-traffic enforcement related inspections, crashes, or other disciplinary actions taken against the driver by the carrier. It seems clear that the driver-based approach has some important limitations. For example, it is more straightforward to hypothesize that carriers are impacted by traffic enforcement actions and are capable of initiating safety-enhancing policies as a response. It is more difficult to view and measure enforcement impacts on individual drivers' behavior and their effect on safety in response to individual single traffic enforcement actions. Furthermore, it is quite difficult, on an individual driver basis, to get accurate information on safety performance differences between the time period before and the time period after the analysis treatment time period. Aggregating the results to the carrier or higher level poses additional problems. It is very likely that, on a driver-by-driver basis, most drivers will have no crashes, in either the time period before or the time period after the treatment. The actual occurrence of a crash for an individual driver is subjected to a random element. If experience across drivers can be aggregated to a carrier level, there is a better chance of being able to measure significant differences in crash rate experience. Since safety policies and programs are initiated and applied at the carrier level, it seems inappropriate to make the driver the preferred analysis unit.

C. Roadway Approach

The unit of analysis in this approach shifts from the carrier to the roadway. The underlying hypothesis of the model is that increased traffic enforcement activity in a given stretch of a roadway may result in reduced crashes. The approach requires the measurement of traffic enforcement activity during a specified period of analysis. Traffic enforcement activity may be measured in a variety of ways ranging from hours of patrol, manpower expenditures, traffic citations issued. These levels of activity would be compared with crash rates in the time period prior to the increased activity versus crash rates after the enhanced activity.

There are numerous problems associated with this approach. First, there are a number of external/environmental factors that might impact crash rates that are not impacted by the increased enforcement. There may be poorer weather conditions during the post-treatment period in comparison to the pre-treatment period that may skew results. Second, the increased enforcement actions may divert traffic to alternative routes which might experience higher crash rates as a consequence of the traffic diversions. Third, the observed crash rate impacts may be ephemeral and quickly dissipate if the enforcement activity is reduced and returns to normal levels.

The conclusion is that the Roadway-based model has some significant obstacles that prevent findings of fact regarding increased enforcement on crash reduction for the numerous reasons cited above. Further, the extrapolation of results from specific roadway measurements to all roadways, States, or the U.S. would require assumptions that could not be confirmed.

D. Selecting An Approach

Before recommending the adoption of any of these approaches or another approach, additional basic research needs to be conducted to obtain information and guidance from state enforcement agencies with their experience and knowledge of the ongoing activities in CMV traffic enforcement. Once this basic research is completed, a document titled, "Traffic Enforcement Effectiveness Measurement Options" will be written. This paper will provide a description of the traffic enforcement programs for CMVs at state and federal level, a summary of other traffic enforcement-related studies done in the past, and a thorough list of approaches to measure the effectiveness of CMV traffic enforcement. The Options Paper will outline the strengths and weaknesses associated with each approach.

V. Combined Effects

A. The Research Challenge

It is recognized that while the OMCHS program for addressing motor carrier safety consists of individual, separately identifiable components (i.e., roadside inspection, traffic enforcement, and compliance reviews), motor carrier safety performance is influenced by all program elements simultaneously. Indeed, it is a major challenge to isolate the individual effects of each of the OMCHS program elements as well as to separate program effects (separately or in combination) from other general factors impacting safety performance.

Motor carriers respond to a variety of outside influences in developing their safety programs. While specific changes may flow from an individual compliance review or a series of roadside inspections, there may be a combination of factors (based on OMCHS programs) influencing behavior as well as a host of external factors.

The research objective is to design an approach that will best address this complex set of influences in a way that will recognize contribution of each program element to overall safety improvement. A second objective will be to determine how the application of these separate "treatments" or programs can be applied in an optimal fashion in order to have the maximum "combined effect."

B. Model Design and Construction

The research team believes that a combined effects model of OMCHS program impact can be designed and serve as a basis for assessing the individual impact of each OMCHS program element. Such a combined effects model would require a stratification of the motor carrier population on the basis of exposure to and results from the set of OMCHS programs--i.e., roadside inspections, traffic enforcements, and compliance reviews. The sample design requires the identification of distinct sub-groups based on levels of exposure to each of the programs over time. Thus, one stratum may include firms with no compliance reviews, very few traffic enforcement actions, and a modest number of roadside inspections. While it is premature to *a priori* decide on the delineation of each of the strata, it is anticipated that it will be feasible to create a series of strata with sufficiently distinct attributes to allow an assessment of the impacts of individual component elements as well as the effects of combined treatments. With regard to the latter analysis, the combinations having the greatest combined effect will be sought. The intention is to track safety performance over time for each of the strata with the objective of determining the effects of OMCHS programs, individually or in combination. The expectation is that there will be enough individual strata to allow assessment of individual programs or combinations of programs. The goal will be to determine whether there are statistically significant safety performance differences depending upon exact exposure to OMCHS program elements. The identification of strata will include a stratum with minimum exposure to OMCHS programs--i.e., no compliance reviews, minimal traffic enforcement actions, and few roadside inspections. This stratum with minimal exposure would serve as a control group to measure the effects of non-OMCHS programs on safety performance.

C. Data Gathering/Data Analysis Plan

The research model described above requires the determination of a number of distinct strata based on level of exposure to the range of OMCHS programs. Taking one or two calendar years worth of data on exposure to OMCHS programs, distinct strata can be established. It is anticipated that one stratum will consist of firms with high exposure to roadside inspections; a second will consist of firms with high exposure to compliance reviews; and a third will consist of firms with high exposure to traffic enforcement activity. Additional strata will consist of firms with high exposure to all or several of the OMCHS programs. A final group will consist of firms with low exposure to all of the OMCHS programs. This latter group will serve as the control group in the study.

The focus of the study will be on an examination of the change in crash rates for each stratum between the calendar year prior to the activity measurement period and the calendar year subsequent to that period. The observed differences and the magnitude of those differences will provide the basis for measuring the impact of each individual program, by itself, or in combination with other programs.

VI. Project Plan Summary

In devising the initial approaches to determine the impact of OMCHS programs, the research team developed non-data collection-intensive approaches that focused on short-range answers to very difficult questions. While the team recognized that such approaches had serious shortcomings, there was a strong belief that with a consistent application of conservative assumptions, the answers developed would be indicative of program impact, while not necessarily capable of yielding precise program impacts.

Based on the experience gained from the development of the initial models and the input of the Expert Panel, however, it became clear that a "second generation" of models, requiring new data collection in a controlled, systematic research design, is required to develop defendable estimates of program effectiveness. These new data collections and analytical efforts constitute the plan for the next major phase of the project.

A. Compliance Review Program: Second Generation Model

Two new data acquisition efforts are required. First, in order to better estimate the impact of the CR program, a new effort is need to collect crash rate data from reviewed carriers over a period of time after the CR was conducted. This new post-CR data will be compared to the pre-CR crash rate data which were already collected during the CR. It is envisioned that OMC field will collect these new data. Second, crash data need to be obtained and analyzed for a control group of carriers not receiving CRs.

1. Data Collection on Carriers with CRs

Step 1. Select Carriers for Data Collection: Data collection should be limited to carriers that have had a CR that is approximately 12 months old. This will provide a similar duration of pre and post-CR crash rate information for the data analysis. Because reviews are performed on a continuous basis, the data collection on carriers with 12 month old reviews will also be collected on a continuous basis. The most preferable scenario for the data analysis is if all carriers that received a CR in a given year (or other period of time) have their crash data collected 12 months after the CR. This will provide complete coverage of all carriers directly impacted by the program. If this data collection process is deemed infeasible (e.g., due to cost and limited OMC resources), a sample of these carriers can be chosen for data collection. Such a sample will be stratified accounting for carrier size, location, type of operation, and CR history. The exact number of strata will be determined based on input from OMCHS.

Step 2. Determine Data to Collect: Fundamental crash rate data will be collected. At a minimum, this will include the number of VMT since the CR and number of recordable crashes. Other potentially useful data that can be collected include updated census data and more information on individual crashes. The extent of the data collected is subject to OMCHS resources.

Step 3. Determine Means of Collecting Data: It is assumed that OMCHS field staff will collect the data. Given the wide logistical distribution of OMCHS field offices and of motor carriers, it is imperative that data gathering and recording are performed in an efficient matter. The Volpe team will work with the OMCHS to develop an effective means of gathering and recording data that will assure data quality.

Step 4. Conduct Data Analysis: After the new data are collected, they will be incorporated into the Impact Assessment Model to determine if and/or how much improvement in safety was demonstrated by carriers that received reviews.

2. Data Collection on Control Group of Carriers (without CRs)

As previously mentioned, the use of a control group will be used to mitigate many of the limitations of the initial CR Impact Assessment Model. A control group of carriers with no CRs will be chosen based on the strata used in Phase 1 for the reviewed carriers. The same types of data collected for the reviewed carriers will be collected for the control group. 12 months after the initial data are collected, the crash rate data will be collected again. This two "snapshots" of data that are 12 months apart will be comparable to the pre and post-CR crash rate data from the reviewed carriers in Phase 1. The control carrier data will be integrated into the Impact Assessment Model to account for potential biases found in the initial model.

3. Longitudinal Study

The study outlined in Phases 1 and 2 can be repeated from year to year to observe the effectiveness of the CR program on carriers that received reviews on an annual basis. This can provide valuable insight as to how changes in the program during a particular year impact the effectiveness of the program.

Furthermore, additional crash rate data can be collected on a periodic basis on same carriers selected in Phases 1 and 2 over multiple years. This data can be use to estimate the duration and magnitude of the CR impact as well as the account for the effects of multiple CRs have on carriers.

B. Roadside Inspection Program: Intervention Model

Step 1. Profile Roadside Inspections to Determine Crash Involvement Probability Categories: The entire set of roadside inspections with vehicle and/or driver out-of-service conditions are sub-divided into crash reduction probability categories. The underlying assumption is that multiple out-of-service conditions (driver and vehicle) have high crash involvement potential (i.e., greater risk) than do out-of-service inspections with a single violation. It is, however, recognized that certain single out-of-service conditions alone (i.e., driving under the influence of alcohol) may have a high crash potential.

The objective of Step 1 is to develop definable categories of crash involvement probabilities based on roadside inspection results. Specific estimates of the crash involvement probabilities are developed in subsequent steps. The Step 1 output focuses on defining the crash involvement probability categories.

Step 2. Develop VMT/Trip Estimates for Crash Involvement Probability Categories: Once categories are defined in the initial step, estimates must be made for the number of miles accumulated on an annual basis by vehicles/drivers in each category. At present, the best identified source for random estimates of vehicle/driver operating conditions is the National Fleet Safety Survey. If the assumption is made that a profile of vehicle/driver conditions is representative of all vehicles, then data in this study can be used to estimate annual VMT/Trips for each crash involvement probability category identified in Step 1.

Thus, if the National Fleet Survey indicates that 1% of vehicles have drivers under the influence of alcohol as the sole out-of-service categories, then this 1% estimate can be used to estimate total annual truck VMT with drivers under the influence of alcohol. Using an estimate of average trip length from the Truck Inventory and Use Survey (TIUS), for example, the number of truck trips with drivers under the influence of alcohol can be estimated.

Step 3. Develop Crash Contribution Factors for Crash Involvement Probability Categories: This step provides a link between the vehicle inspections and truck crashes. The objective of Step 3 is to develop an estimate of the annual number of truck crashes in which conditions identified in each of the crash involvement probability categories were present. Thus, if one crash involvement probability category is vehicles with multiple defects, including bad brakes, Step 3 will provide an estimate of the number of truck crashes in which vehicles had multiple defects (including bad brakes) identified as crash contributing factors.

Step 3 will require extensive examination of presently ongoing crash causation studies in the expectation that some of previously identified concerns (i.e., underestimates of causation factors) are addressed.

Step 4. Estimate Crash Involvement Probabilities: This step integrates information from Steps 1-3 and produces estimates of crash reduction probabilities for each category identified. Thus, if there are 10,000 truck crashes on an annual basis corresponding to the crash involvement probability categories of vehicles with multiple defects, including bad brakes, and 300,000 annual truck trips in which vehicles had multiple defects, including bad brakes, then the crash involvement probability for this category would be .033. Therefore, it would be estimated that for every 1,000 inspections in which vehicles had multiple defects, including bad brakes, 33 crashes would be avoided.

C. Traffic Enforcement Program

Prior to developing a specific approach, additional research needs to be conducted with states that have done studies on commercial vehicle traffic enforcement and in particular done analysis on relating that information to crashes. The OMCHS has provided grants to states to perform such studies and to reflect those commercial vehicle traffic enforcement efforts in their performance-based state safety improvement plans. Using information gained from these state experiences, discussions with state and OMCHS traffic enforcement experts and further analysis of moving violation citation data reported with roadside inspections, an options paper will be prepared on the alternative approaches to measuring the effectiveness of traffic enforcement.

D. Combined Effects Model

The research objective of the combined effects model is to design an approach that will best address the total set of OMCHS program influences in a way that will attribute contribution of each program element to overall safety improvement. A second objective will be to determine if the application of these separate “treatments” or programs can be applied in an optimal fashion in order to have the maximum “combined effect.” A sample study is proposed.

Step 1: Determine Strata for Analysis: Establish number of strata for analysis. There will be a stratum for high exposure to each of the major programs(i.e., roadside inspections, carrier reviews, and traffic enforcement). In addition, there will be a stratum for carriers with high exposure to all three programs as well as strata for carriers with high exposure to all possible combinations of two programs. Finally, there will be a stratum of carriers with below average exposure to all three programs. This group will serve as the control group.

Step 2: Define "high exposure levels" for each program category: For each OMCHS program determine average levels of activity for various sized carriers. In addition, establish criteria for above average and below average activity levels. Use the above average cut-off as establishing the basis for defining the "high-exposure" carriers.

Step 3: Find carriers to fill each stratum: Based on the definitions developed in step 2, find carriers meeting the criteria for each stratum. The objective is to have three carrier size groups for each stratum and to have at least 30 carriers in each stratum. Thus, there will be a stratum for carriers with above average exposure to the roadside inspection program. In fact, there will be three such strata based on carrier size (i.e., small-sized, medium-sized, and large-sized groups).

Step 4: Collect complete OMCHS profiles for carriers in each stratum: Having determined the individual strata and the size groupings as well as the carriers meeting the established criteria, collect complete OMCHS profiles for all carriers. That is, if a carrier is in the strata for having high exposure to roadside inspections, collect complete information on its exposure to all OMCHS programs during that period of time.

Step 5: Determine Pre-and Post-crash Rates for all carriers selected in the sample: For all carriers included in the analysis, determine their crash rates during the year prior to the OMCHS exposure measurement period as well as the year subsequent to it.

Step 6: Calculate differences in crash rates and assess program impact: Based on crash rates, calculate differences in crash rates for each stratum (conduct the analysis separately for each carrier sized group). Compare observed differences in crash rates for each stratum against the performance of carriers in the control group.

Step 7: Observe patterns and combinations: Based on the differences in crash rates observed, draw inferences and estimates of program impact. Be prepared to evaluate strategies that might combine different levels of program exposure for maximum effectiveness.