IMPROVEMENTS NEEDED IN THE MOTOR CARRIER SAFETY STATUS MEASUREMENT SYSTEM

Federal Motor Carrier Safety Administration

Report Number: MH-2004-034
Date Issued: February 13, 2004

Memorandum

U.S. Department of Transportation

Office of the Secretary
of Transportation

Office of Inspector General

Subject:	ACTION: Audit Report on Improvements Needed in the Motor Carrier Safety Status Measurement System Federal Motor Carrier Safety Administration MH-2004-034	Date:	February 13, 2004
From:	Alexis M. Stefani Principal Assistant Inspector General for Auditing and Evaluation	Reply to Attn. of:	JA-40
To:	Federal Motor Carrier Safety Administrator		

This report presents the results of our audit of the Motor Carrier Safety Status Measurement System (SafeStat). An executive summary of the report follows this memorandum.

Our objectives were to determine whether the:

- SafeStat model is valid and whether the scores calculated are consistent with the model's design.

- data used by SafeStat are complete, consistent, accurate, and timely.

- data quality control systems are adequate to ensure information quality for intended uses.

We found that SafeStat generally calculated scores consistent with its design, and a 1998 study supported the model's validity. However, the model needs to be revalidated because changes have occurred since the earlier study, and more sophisticated analysis, not previously conducted, would optimize the model's effectiveness. Moreover, we found material weaknesses in the SafeStat data reported by states and motor carriers and with the Federal Motor Carrier Safety Administration's (FMCSA) processes for correcting and disclosing data problems. Consequently, while SafeStat is sufficient for internal use, its continued public dissemination and external use require prompt corrective action. Improvements in the model are important, but getting better data is essential.

A draft of this report was provided to FMCSA on December 10, 2003. In its comments, FMCSA agreed with our concerns for improving data quality and cited a number of improvements already implemented or ongoing to address the recommendations in the report. The improvements reported included:

- hiring a contractor to conduct a new study to revalidate the SafeStat model;

- implementing an improved system for tracking public challenges to the accuracy of SafeStat data;

- providing SafeStat users with comprehensive information on data limitations;

- assigning staff to review monthly state reports that address state data quality issues and to work with the states to resolve them;

- establishing goals for completeness, accuracy, and timeliness of data; and

- making state grant funding contingent on participation in certain data quality programs.

In commenting on the findings in the draft report, FMCSA did not agree with all of our assertions as to the impact of data quality problems on SafeStat. Specifically, FMCSA commented that the language in the draft report overstated the problem of out-of-date census data on SafeStat. FMCSA also disagreed with any implication in the report that some motor carriers who are categorized by SafeStat as high risk, may be categorized as high-risk carriers only because of the existing data problems.

We appreciate FMCSA's positive response to our recommendations and have revised the final report to recognize corrective actions that have been taken or that are ongoing. We do not agree that the language in the draft report overstated the problem with out-of-date census data, and we have provided additional information on the issue in this final report.

On the question of whether some carriers may be categorized as high-risk only due to the existing data quality problems, we agree with FMCSA that data quality problems are more likely to make a high-risk carrier look good. However, we continue to maintain that the opposite situation can also occur. Because SafeStat scoring involves a relative ranking of one carrier against another, *missing data may place a lower-risk carrier in a deficient category because data for a higher-risk carrier is not included in the calculation.* Missing crash data were most significant with six states failing to report any crashes for the 6 months analyzed. Nationwide, estimates for the underreporting of large trucks involved in crashes varied in magnitude with some states underreporting by 60 percent or more and other states underreporting by less than 20 percent.

The existing data quality problems should not prevent FMCSA from using SafeStat as an internal decisionmaking tool. However, while the data used for SafeStat calculations are sufficient for internal purposes, if public dissemination of SafeStat results is to continue, the data must meet higher standards for completeness, accuracy, and timeliness.

We request that within 30 days FMCSA provide clarifications and target completion dates for several planned actions, as noted in the attached report. In instances where we are in agreement on the corrective actions and target completion dates are provided, the recommendations are considered resolved subject to the follow-up provisions of Department of Transportation Order 8000.1C.

We appreciate the courtesies and cooperation of representatives from FMCSA, the Volpe Transportation Systems Center, state government offices, and motor carrier companies during this audit. If you have any questions concerning this report, please call me at (202) 366-1992 or Debra Ritt, Assistant Inspector General for Surface and Maritime Programs, at (202) 493-0331.

Attachment

#

cc: National Highway Traffic Safety Administrator

Executive Summary

Improvements
Needed in the Motor Carrier
Safety Status Measurement System
Federal Motor Carrier Safety Administration

Report No. MH-2004-034   **February 13, 2004**

INTRODUCTION

Safety investigators from the Federal Motor Carrier Safety Administration (FMCSA) conduct compliance reviews to ensure motor carriers are following safety regulations. However, FMCSA is able to assess less than 2 percent of about 650,000 active interstate motor carriers each year. To help select the motor carriers targeted for compliance reviews, FMCSA uses the Motor Carrier Safety Status Measurement System (SafeStat), an automated system for ranking motor carriers. The Volpe National Transportation Systems Center (Volpe) developed SafeStat in the mid-1990s.

The public also has access to SafeStat. SafeStat results have been available to the public via the Internet since 1999. Thus, motor carriers, shippers, insurers, and other Government users have free access to SafeStat information when making business decisions.

How SafeStat Works. When sufficient information from FMCSA's database is available, SafeStat assigns motor carriers a percentile ranking between 0 and 100 (with 100 being the worst) in one or more of the following four safety evaluation areas: accidents (crashes), driver, vehicle, and safety management. Carriers ranked in the 75th percentile or greater are considered deficient in an area. Carriers deficient in two or more areas are assigned an overall SafeStat score of 150 to 550, based on the weighted accumulation of the percentile rankings. The relative weights assigned to each area are accidents 2.0; driver 1.5; vehicle 1.0; and safety management 1.0.

Where SafeStat Data Come From. Key FMCSA data sources for calculating SafeStat rankings are:

- periodic census reports that motor carriers submit to FMCSA, which include information on the number of drivers and vehicles (expressed as power units) used by the motor carrier;

- police accident reports on crashes involving commercial vehicles;

- state reports on vehicle and driver violations found during roadside safety inspections, which include serious moving traffic violations such as speeding; and

- FMCSA compliance review and enforcement records on motor carriers.

Uses of SafeStat. FMCSA uses SafeStat to identify and prioritize high-risk motor carriers for compliance reviews.[1] FMCSA also uses SafeStat to generate warning letters advising carriers that continued performance problems may result in compliance reviews and potential state vehicle registration sanctions. Public uses for SafeStat include providing information to individuals making contract award or acquisition decisions and allowing carriers to assess their own safety strengths and weaknesses.

Breakdown of SafeStat Results. Of 645,551 active interstate carriers on record[2], in SafeStat, 26 percent or about 170,000 had sufficient data[3] to compute a value for one or more of the four safety evaluation areas. Of the 170,000 carriers, 7,821 were considered deficient; to the point they were classified as the riskiest carriers and placed in categories A, B, or C. Another 34,844 carriers, somewhat less deficient, were placed in categories D, E, F, or G. FMCSA field offices are directed to concentrate their compliance review efforts on category A and B carriers, which are considered the highest risk.

OBJECTIVES AND SCOPE

This audit responds to a request from Representative Thomas E. Petri, Chairman of the House Transportation and Infrastructure Highways and Transit Subcommittee. Concerned about data issues as well as the SafeStat methodology, Chairman Petri requested that we review the reliability, validity, and objectivity of SafeStat. Specifically, our objectives were to determine whether the:

---

[1] Federal and state safety investigators perform compliance reviews by examining motor carrier operations to determine whether motor carriers and their drivers meet safety requirements. Based on the results of a compliance review, FMCSA assigns carriers safety ratings of satisfactory, conditional, or unsatisfactory.

[2] The breakdown of SafeStat results is based on January 2003 data. As of August 2003, 665,646 active motor carriers were on record in SafeStat.

[3] One example of a data sufficiency rule is that carriers must have at least three inspections on record within the most recent 30 months to receive a value in the driver or vehicle areas.

- SafeStat model is valid and whether the scores calculated are consistent with the model's design.

- data SafeStat uses are complete, timely, consistent, and accurate.

- data quality control systems are adequate to ensure information quality for intended uses.

In conducting the audit, we used data obtained from Department of Transportation organizations, selected motor carriers, and state offices. Our work included an assessment of information from a sample of crash and inspection reports obtained in 10 states.

RESULTS

The SafeStat model generally calculated scores consistent with its design, and a 1998 study supported the model's validity. However, the model needs to be revalidated because changes have occurred since the earlier study, and more sophisticated analysis, not previously conducted, would optimize the model's effectiveness. Moreover, we found material weaknesses in the SafeStat data reported by states and motor carriers and with FMCSA's processes for correcting and disclosing data problems. Consequently, while SafeStat is sufficient for internal use, its continued public dissemination and external use require prompt corrective action. Improvements in the model are important, but getting better data is essential. Examples of problems with FMCSA's database include:

- The lack of updated census data for 42 percent of the 643,909 active carriers, and for 31 percent of the 170,623 carriers that had percentile values calculated in SafeStat for one or more of the four safety evaluation areas. (Census data include the current number of vehicles and drivers used by the companies.)

- Recorded values of zero vehicles for approximately 11 percent of active carriers and zero drivers for 15 percent of active carriers even though many of these same carriers also had crashes or inspections reported against vehicles or drivers.

- Missed reports for an estimated one-third of the large trucks involved in accidents annually, including 37,000 crashes involving interstate carriers.

- Late reporting of 20 percent or about 19,000 of the crashes in Fiscal Year 2002 by 6 or more months after the crashes occurred.

- Underreporting of serious moving traffic violations (mainly speeding) identified during roadside inspections. In one state visited, an estimated 29,000 serious moving traffic violations went uncounted over a 3-year period.

In addition, based on our tests of FMCSA's database, we estimated that 13 percent of the 21,000 crashes and 7 percent of the more than 1 million inspection transactions occurring in our 6-month sample period contained carrier identification errors, such as failure to identify a carrier associated with the violation, or in a smaller number of instances, identifying the wrong carrier.

Problems with inaccurate data are compounded because no effective system is in place now to facilitate the correction of errors in data reporting. Requests received by FMCSA for state data correction are forwarded to the states where the incidents occurred, and there is no system in place to centrally track the resolution of these requests.

We recognize that FMCSA relies on state officials to generate and process much of the safety event data used in SafeStat and that this presents many challenges for obtaining complete, timely, and accurate data. However, as a result of weaknesses in the data reported by states, SafeStat rankings are geographically biased against carriers operating in states that provide more complete data, while weaknesses in the data provided by carriers can produce errors in SafeStat calculations and cause some high-risk carriers to be missed.

Thus, while SafeStat is sufficient for targeting compliance reviews and considered valuable by internal users, its continued public dissemination and external use require prompt and complete action to improve the model and improve the quality of the data used. Because carrier safety data and the model's rankings are publicly disclosed, a higher standard of quality must be met to ensure fairness to motor carriers who may lose business or be placed at a competitive disadvantage by inaccurate SafeStat results. FMCSA will need to demonstrate timely improvements if it is to continue to publicly disclose carrier results across all SafeStat categories.

Need to Revalidate the Model. FMCSA must act to revalidate the SafeStat model because changes have occurred since a 1998 study that supported the model's validity. These changes include the addition of serious moving traffic violations to SafeStat calculations, revised weighting of fatal accidents, and altered methods for calculating the number of vehicles used by motor carriers. In addition, while the general approach used in the 1998 study was acceptable, more sophisticated analysis, not previously conducted, would optimize the model's effectiveness. Such analysis could determine the degree to which the weightings used produced internally consistent results across all categories of carriers and

demonstrate the degree to which the weighting reflected the views of experts. It could also test whether changes in the model would yield better predictive results.

Any future analysis should be conducted so that the results can be subject to independent confirmation. We could not independently verify the 1998 study because the data used were not available.

Data Quality Weaknesses. FMCSA should take action to improve data quality because significant problems exist with the data motor carriers and the states provide, and these data problems limit SafeStat's effectiveness and introduce bias into the ranking process. The most serious problems lie with outdated census data on the number of vehicles and drivers used by motor carriers, incomplete crash reporting from the states, and late or incomplete inspection reporting, particularly reports on serious moving traffic violations. Specifically:

- *Poor Carrier Census Data*. SafeStat calculations use census data that motor carriers are required to complete and update every 2 years. SafeStat cannot effectively rank carriers without accurate and complete census information. In the worst case, carriers incorrectly shown on census records as having zero vehicles can have crashes, including fatalities, without it negatively impacting their SafeStat score. We found that:

    - About 272,000 or 42 percent of the active interstate motor carriers had not met the congressionally mandated requirement to update census data every 2 years, as of January 2003. This includes 23,919 motor carriers on record without an update since coming on file in 1974. It also includes 4,086 carriers identified in SafeStat as having 2 or more accidents.

    - Approximately, 71,000 (11 percent) of the 643,909 active interstate carriers were on record in January 2003 as having "zero" power units and about 98,000 carriers or 15 percent of the 643,909 carriers were on record as having zero drivers. This included 15,136 carriers who had at least one inspection on record between October 2001 and September 2002. Getting good data on drivers and power units is important as these data are used frequently in SafeStat calculations. In January 2003, one or both of these data elements were used to calculate the SafeStat score for at least 74 percent of the category A carriers, the highest risk SafeStat category.

- *Poor Crash Data.* SafeStat calculations attach the greatest weight to crashes, because crash history is considered to be associated with future crash risks, but large gaps exist in how completely and timely states report crash data. For example:

- 6 of 51 states (including 2 of the 10 states we visited) did not report any crashes in the 6-month period we reviewed. The six states were the District of Columbia, Florida, New Hampshire, New Mexico, Pennsylvania, and Vermont. These six states constitute approximately 11 percent of the total commercial vehicle miles traveled in 2002 and are home to an equivalent percentage of the active, interstate motor carriers in the nation.

- 1 of the 10 states we reviewed reported only fatal crashes to FMCSA but failed to report 5,816 non-fatal large truck crashes from 2001 to the FMCSA database.

- Meanwhile other states reported most crashes within 90 days.

Such data variations across geographic locations indicate that FMCSA may fail to target certain high-risk carriers simply because of where they operate.

- *Poor Data on Moving Traffic Violations.* State reports on roadside inspections are supposed to include data on serious moving traffic violations, but often they do not. In addition, large state-to-state variations exist in the reporting of traffic violations, which introduce a degree of geographic bias in the ranking system. For example, California reported only 115 serious moving violations to the FMCSA database in FY 2001 compared to Indiana, which reported about 25,000. Although moving traffic violations are weighted less in SafeStat calculations than are crashes, these violations have been associated with higher crash rates. Thus, underreporting in this area reduces the effectiveness of SafeStat.

- *Inaccurate Data Records.* Even when crash and inspection data reach the FMCSA database, our review showed that errors or omissions occur during the process that could influence SafeStat scores. Based on a sample of FMCSA records, we estimate that errors occurred in approximately 13 percent (2,851 of 21,225) of the crashes and 7 percent (76,521 of 1.02 million) of inspection transactions on interstate carriers during a 6-month sample period. A critical mistake that occurred in an estimated 11 percent (9,484 of 79,372) of the errors was that the wrong carrier was held accountable for a SafeStat-related violation. We cannot estimate the impact of such errors on SafeStat category rankings, but reasons for errors included the misinterpretation of reporting rules by local officials and data entry errors.

Improving Systems for Disclosing and Correcting Poor Data. Given the state of data quality, FMCSA should improve processes for disclosing and correcting data problems. Areas requiring specific improvement include:

- *Insufficient Disclosure of Weaknesses in Publicly Disseminated Data.* To mitigate problems posed by public dissemination of inaccurate data, FMCSA should ensure public users of SafeStat are aware of any known data quality weaknesses. In the past, the SafeStat Internet site pointed to problems with data provided by carriers, but did not disclose reporting problems with states. Following the issuance of our draft report, FMCSA placed a more comprehensive disclaimer on the Internet site, and promised to provide state-specific information in the future.

- *Ineffective Systems for Facilitating Data Correction by the States.* Motor carriers cannot easily correct inaccurate crash and inspection information in SafeStat because carriers must deal with individual states to obtain corrective actions. In addition, FMCSA maintains no data on the adjudication or timeliness of correction challenges and does not enforce existing requirements. FMCSA standards call for the correction of data inaccuracies within 7 days, but we have no evidence of this being enforced.

  Since we initiated our audit, FMCSA and Volpe have developed a prototype for a data quality tracking system to centrally accept and track data accuracy challenges that motor carriers make to the states. However, timely action is needed to complete testing and field the system. FMCSA expects to have the system operational by the end of February 2004. With SafeStat scores being calculated and released monthly, versus the earlier practice of releasing scores semiannually, a properly operating data correction process will enable errors to be corrected on a more timely basis.

- *Insufficient Reviews of Data Quality.* We found that states had limited controls to ensure the quality of data submitted to FMCSA. To catch many of the data problems we found, quality reviews should include assessments of source documentation to make sure data transmitted to FMCSA and used in SafeStat calculations are accurate. Compounding the lack of state controls, FMCSA did not routinely conduct its own oversight reviews targeted at control weaknesses in the states. One area ripe for such a review would be identifying and correcting reporting problems associated with serious moving traffic violations observed in conjunction with roadside inspections.

- *Ineffective Use of Sanctions and Incentives.* Overall, FMCSA has not effectively used existing sanctions and incentives to promote better data reporting by states and motor carriers. FMCSA has not imposed sanctions, such as the withholding of basic Motor Carrier Safety Assistance Program (MCSAP) grant funds from states for failing to correct data quality problems. Also, current MCSAP incentive grant formulas, while useful in theory, are not sufficient. As implemented, only the timeliness of data is factored into the

incentive calculations, while completeness and accuracy of the data are ignored. Following the issuance of our draft report, FMCSA informed the states that certain grant funding would be contingent on participation in an improved system for tracking public challenges to the accuracy of SafeStat data.

- *Slow Implementation of Joint Programs.* Although the Motor Carrier Safety Improvement Act of 1999 directed FMCSA and the National Highway Traffic Safety Administration (NHTSA) to take action to improve the collection and analysis of state crash reporting, results have been slow to materialize. The agencies did not finalize a memorandum of understanding on the basic approach until February 2003. Encouragingly, since then the agencies have initiated a plan designed to bring about immediate improvement in the completeness of crash reporting for targeted states. Planned actions include reviews of how targeted states collect, store, and analyze crash data. Rapid implementation of the planned actions is needed and the plans should be extended to cover all states.

Correcting data quality problems are critical to ensure that SafeStat more effectively targets high-risk carriers for compliance reviews. However, quick action is even more important given the continued public dissemination of SafeStat results. Department of Transportation (DOT) guidelines issued in October 2002 underscore the need to apply stricter standards to "influential" information, such as SafeStat data, that has a "clear and substantial impact on important public policies or private sector decisions." To implement this higher standard, and thus ensure the continued public dissemination of SafeStat results, FMCSA officials will need to implement an overall data program. The program should include minimum standards for data completeness, timeliness, and accuracy and make sure that those standards are met.

Summary of Recommendations. Although we support SafeStat's use as an internal risk management tool, continued public disclosure of the information requires significant and timely actions to improve the system. Strong efforts are needed to obtain better data from the states on crashes and inspections. Funding to improve data reporting has been and continues to be provided to the states and FMCSA should act to ensure that the funding brings about the desired results. We have made the following recommendations to improve SafeStat.

First, FMCSA should revalidate the SafeStat model using a more sophisticated analysis and solicit public comment on model changes.

Second, to mitigate the impact of inaccuracies or incomplete data on public users of SafeStat, FMCSA should make available to all states, within 3 months of issuance of this report, an improved system for facilitating the correction of data

inaccuracies and the tracking of corrective actions. The agency should also expand the caution statements on the use of SafeStat recently placed on the Internet to include state-specific information on data quality problems.

Third, FMCSA should detail the components of an overall data quality improvement program that:

- Addresses longstanding issues associated with motor carrier census data by imposing fines on carriers that fail to provide updated carrier census information despite repeated opportunities to do so.

- Sets minimum standards for the quality of SafeStat data consistent with the Department's data quality guidelines. At a minimum, the standards should address completeness, accuracy, and timeliness of data.

- Accomplishes actions planned, in conjunction with NHTSA, for improving the completeness and timeliness of state-reported crashes.

- Enhances the depth, frequency, and type of FMCSA state data quality reviews and monitoring, and ensures state plans address data quality.

- Modifies FMCSA guidance and funding decisions so that MCSAP incentive grant awards are based, in part, on each state's implementation of guidelines established to provide accurate, complete, and timely safety event data.

MANAGEMENT COMMENTS AND OFFICE OF INSPECTOR GENERAL RESPONSE

We provided FMCSA a draft of this report December 10, 2003. In its comments, FMCSA agreed with our concerns for improving data quality and cited a number of improvements already implemented or ongoing to address the recommendations in the report. The improvements reported included:

- hiring a contractor to conduct a new study to revalidate the SafeStat model;

- implementing an improved system for tracking public challenges to the accuracy of SafeStat data;

- providing SafeStat users with comprehensive information on data limitations;

- assigning staff to review monthly state reports that address state data quality issues and to work with the states to resolve them;

- establishing goals for completeness, accuracy, and timeliness of data; and

- making state grant funding contingent on participation in certain data quality programs.

In commenting on the findings in the draft report, FMCSA did not agree with all of our assertions regarding the impact of data quality problems on SafeStat. Specifically, FMCSA commented that the language in the draft report overstated the problem of out-of-date census data on SafeStat. FMCSA also disagreed with any implication in the report that some motor carriers SafeStat categorized as high risk, may only be categorized as high-risk carriers because of the existing data problems.

The full text of a matrix provided by FMCSA detailing corrective actions that have been taken or that are ongoing is provided in the Appendix along with FMCSA's substantive comments on the draft report's findings.

We appreciate FMCSA's positive response to our recommendations and have revised the final report to indicate the number of corrective actions that have been taken or are underway. Although we did not agree that the problem with out-of-date census data was overstated in the draft report, we provided additional information on the issue in this final report to ensure the full context of the problem is presented.

On the question of whether some carriers may be categorized as high risk only due to the existing data quality problems, we agree with FMCSA that data quality problems are more likely to make a high-risk carrier look good. However, we continue to maintain that the opposite situation can also occur. Because SafeStat scoring involves a relative ranking of one carrier against another, missing data may place a carrier in a deficient category because data for a higher-risk carrier is not included in the calculation. Missing crash data were most significant with six states failing to report any crashes for the 6 months analyzed. Nationwide, estimates for the underreporting of large trucks involved in crashes varied in magnitude with some states underreporting by 60 percent or more and other states underreporting by less than 20 percent.

The existing data quality problems should not prevent FMCSA from using SafeStat as an internal decisionmaking tool. However, while the data used for SafeStat calculations are sufficient for internal purposes, if public dissemination of SafeStat results is to continue, the data must meet a higher standard.

Although FMCSA comments were generally responsive, we are requesting some additional information. Specifically, we request that FMCSA clarify whether its commitment to work with states to resolve data issues will include carrying out our recommendation to ensure that state plans address data quality. We also request that FMCSA clarify whether the review of source documentation proposed

will be a one-time or an ongoing task. FMCSA also needs to provide target dates for several corrective actions, as noted in the body of the report.

TABLE OF CONTENTS

INTRODUCTION

Background

The Federal Motor Carrier Safety Administration's (FMCSA) primary mission is to prevent commercial motor vehicle related fatalities and injuries. As of January 2003, FMCSA had approximately 643,000 active, interstate motor carriers on record. The Motor Carrier Safety Status Measurement System (SafeStat) is an automated system that measures the relative safety of motor carriers that operate in interstate commerce and haul hazardous materials. FMCSA uses SafeStat to identify and prioritize high-risk motor carriers for compliance reviews.[1] FMCSA also uses SafeStat to generate warning letters advising carriers that continued performance problems may result in compliance reviews and potential state vehicle registration sanctions. SafeStat is also part of a decision aid used by some state safety inspectors for selecting vehicles for inspection. The Volpe National Transportation Systems Center (Volpe) in Cambridge, Massachusetts, developed SafeStat for FMCSA in the mid-1990s.

In SafeStat, motor carriers are evaluated in four areas: accidents (crashes), drivers, vehicles, and safety management. Data sources for these areas are state-reported crashes, roadside safety inspections of vehicles and drivers, serious moving traffic violations, crashes and safety violations found during compliance reviews, enforcement cases brought against motor carriers by FMCSA, and motor carrier identification reports (census forms). Table 1 shows the data sources related to each evaluation area.

Table 1. SafeStat Safety Evaluation Areas and Data Sources			
Accident	**Driver**	**Vehicle**	**Safety Management**
State Reported Crashes, Compliance Reviews, and Census Information	Driver Roadside Inspections (including moving violations) Census Information, and Compliance Reviews	Vehicle Roadside Inspections and Compliance Reviews	Compliance Reviews and Closed Enforcement Cases

---

[1] Federal and state safety investigators perform compliance reviews by examining motor carrier operations to determine whether motor carriers and their drivers meet safety requirements. Based on the results of a compliance review, FMCSA assigns carriers safety ratings of satisfactory, conditional, or unsatisfactory.

Within each area, SafeStat categorizes carriers into peer groupings based on the number of safety events for which data are available, such as the number of crashes (see Table 2). SafeStat assigns weights to more recent and more serious safety events, such as crashes involving fatalities or injuries, and adjusts computed values using factors such as the number of power units[2] operated by the carriers. Based on computed values, carriers are ranked against other carriers within their peer group and are assigned a percentile. Percentiles are expressed from 0 to 100, with 100 being the worst.

Table 2. Peer Groupings for Accident Evaluation Area		
Peer Group Number	Number of Crashes	Percentile Ranking
0	0	0
1	1	Not Ranked[*]
2	2 - 3	0 - 100
3	4 - 8	0 - 100
4	9 - 20	0 - 100
5	21 - 88	0 - 100
6	89+	0 - 100

* SafeStat procedures state that carriers with only one crash are not considered to have sufficient data to calculate a score and are not ranked in SafeStat.

SafeStat assigns percentiles to a carrier in areas where the minimum number of safety event data are available. The minimum values include at least two crashes and at least three roadside inspections in the previous 30 months. Only carriers with percentiles of 75 or greater in two or more areas receive a SafeStat score. The maximum possible score is 550, computed based on weights for each of the four evaluation areas. The accident and driver areas received the highest weights. FMCSA's field offices are directed to concentrate on conducting compliance reviews of motor carriers with SafeStat scores in the top two categories—A (a score equal to or greater than 350) and B (a score equal to or greater than 225 but less than 350). Additionally, categories D, E, F, and G include carriers assigned a deficient percentile value in only one area (see Table 3).

Table 3. SafeStat Category Scoring	
Category	SafeStat Score/Percentile Range
A	> 350 to < 550
B	> 225 to < 350
C	> 150 to <225
D	Accident Area > 75 to 100 Percent
E	Vehicle Area > 75 to 100 Percent
F	Driver Area > 75 to 100 Percent
G	Safety Area > 75 to 100 Percent

In January 2003, SafeStat scored 7,821 (or 1.21 percent) motor carriers out of a possible 645,551 motor carriers. Another 34,844 carriers, somewhat less deficient, were placed in categories D, E, F, or G. FMCSA field offices are directed to concentrate on category A and B carriers first when conducting compliance reviews, as these are considered to be the highest risk. However, not all motor carriers had sufficient data to compute a value for one or more of the four safety evaluation areas. For example, in January 2003, about 170,000 (or 25 percent) carriers had percentile values calculated for one or more evaluation areas. SafeStat information, including a carrier's score,

---

[2] The number of power units is defined by the total number of trucks, tractors, hazardous material tank trucks, motor coaches, school buses, minibuses/vans and limousines that a carrier owns or leases.

percentile, value and rating in each area, and information on specific safety events used in the calculation, such as crashes and inspections, is available to the public on an Internet site maintained by Volpe. The Internet site contains specific carrier information, such as the reported number of vehicles and drivers, address, and any violations found in the latest FMCSA compliance review. The site also includes event-specific information, such as the date, time, and location of an accident; the number of injuries and fatalities involved; the age and condition of the driver; the vehicle license plate and identification number; and information on any hazardous materials released.

The Internet site has been publicly available since December 1999 and accessed as many as 80,000 times a month. In January 2003, SafeStat began making monthly, instead of semiannual, runs and postings to the Internet site. Public access to data allows carriers and the firms involved with carriers, such as shippers, insurers, and lessors, to use SafeStat information when making business decisions. Consequently, the reliability of SafeStat information is important because it can have an economic impact on motor carriers.

Objectives, Scope, and Methodology

This audit responds to a request from Representative Thomas E. Petri, Chairman of the House Transportation and Infrastructure Highways and Transit Subcommittee. Chairman Petri's request stated, ". . .while the goal of SafeStat is correct, there are continuing concerns regarding data issues as well as the validity and objectivity of the SafeStat methodology." Our overall objective was to determine whether SafeStat reliably identifies high-risk carriers. Specifically, we determined whether the:

- SafeStat model is valid and whether the scores calculated are consistent with the model's design.

- data SafeStat uses are complete, timely, consistent, and accurate.

- data quality control systems are adequate to ensure information quality for intended uses.

To form conclusions on the quality of data used in SafeStat and to make nationwide projections, we used data obtained from DOT organizations, selected motor carriers, and state offices. We conducted two-stage statistical sampling in which we selected 10 states for review, then selected crash and inspection reports on interstate carriers for examination from those 10 states.

We also used additional reports and data generated by FMCSA, the National Highway Traffic Safety Administration (NHTSA), and selected motor carriers from FY 1999 through FY 2002 in our analysis of data consistency and trends.

We evaluated the validity of SafeStat by discussing the basis for the model with its developers at Volpe, reviewing relevant documentation related to SafeStat's original development, and consulting with technical experts on the requirements for demonstrating model validity and on the statistical tools used to develop valid and unbiased models. Additionally, to determine whether scores calculated and posted on the Internet were consistent with the SafeStat algorithm, we recalculated scores assigned to a random sample of 65 carriers during the January 26, 2003 SafeStat run. Our efforts to confirm the model's validity and its effectiveness in putting potential unsafe carriers in proper categories were limited because data from 1993 to 1996 used in a 1998 study of the model's effectiveness were not available for our independent assessment of the simulated scores assigned to carriers. We did assess the methodology employed in the 1998 study.

States Visited Based on Statistical Sampling	
California	Montana
Georgia	New Mexico
Illinois	Pennsylvania
Indiana	Texas
Louisiana	Washington

To determine whether FMCSA's data quality control systems were adequate for their intended uses, we reviewed guidelines issued by the Department of Transportation (DOT) on the dissemination of data and compared the guidelines to systems in place or planned by FMCSA. We also observed quality control systems at the states visited and assessed selected areas, such as data entry controls. A list of activities visited or contacted is in Exhibit A.

The audit was conducted from November 2002 through January 2004 in accordance with Government Auditing Standards prescribed by the Comptroller General of the United States and included such tests of internal controls as were considered necessary. Further details on the audit methodology are in Exhibit B.

Prior Audit Coverage

One Office of Inspector General audit in 1999 and two General Accounting Office (GAO) audits, noted below, have addressed topics related to SafeStat. Details on the prior audits are in Exhibit C.

OIG Report: *Motor Carrier Safety Program, Federal Highway Administration,* Report Number TR-1999-091, April 26, 1999.

GAO Report: *Truck Safety: Motor Carriers Office Hampered by Limited Information On Causes of Crashes and Other Data Problems*, RCED-99-182, June 29, 1999.

GAO Report: *Commercial Motor Carriers: DOT is Shifting to Performance– Based Standards to Assess Whether Carriers Operate Safely*, RCED-98-8, November 3, 1997.

FINDINGS AND RECOMMENDATIONS

Finding A. More Can Be Done to Confirm the Model's Validity and Maximize Its Effectiveness

We verified that the scores produced by SafeStat generally are consistent with the model's design; and views of current users support the validity of the model. A 1998 study also supported the model's validity and an assessment of compliance review results shows SafeStat's ability to identify some high-risk carriers. However, changes have occurred since the 1998 study, and our attempts to confirm the results from the 1998 study were limited because data used in the study were not available. While the general approach used in the study was acceptable, a more sophisticated analysis, not previously conducted, would optimize the model's effectiveness. Moreover, further analysis, which expands on the previous study, should be done to determine the degree to which variations in the model, such as different treatment of exposure and weighting factors, would yield better results.

SafeStat Scores Generally Consistent with the Algorithm

Our recalculation of SafeStat scores across evaluation areas using randomly selected carriers showed that computer generated scores posted on the Internet were, with one exception, consistent with the algorithm used to calculate the scores. We determined this by recalculating carriers' scores in 13 different peer groups and 2 evaluation areas using SafeStat's stated methodology.

Although SafeStat's calculations generally were performed as intended, we noted one exception. The January 26, 2003 SafeStat run showed that the calculation for determining average power units over several time periods did not work correctly for 71 of 17,260 motor carriers. This error related to carriers with one or more power units in the current period and zero power units in prior periods. The error occurred because insufficient computer programming was done when SafeStat data were converted to a new database in September 2002.

As a result of this programming error, no SafeStat value could be computed in the accident area for the 71 carriers because of the incorrect computation of the average number of power units, which is used in the calculation of the accident value. Using the SafeStat model, Volpe staff estimated that 7 of the 71 carriers should have been categorized as A or B, the categories reserved for the highest risk carriers on SafeStat's January 2003 list. Volpe and FMCSA officials took

action to correct this programming error when we brought it to their attention in February 2003.

Current Users' Views and Assessment of Compliance Review Results Support Use of SafeStat

Government users of SafeStat generally view it as an effective tool. In the 10 states reviewed, both Federal and state motor carrier officials expressed confidence in SafeStat's ability to identify high-risk carriers and considered it a significant improvement over a prior system for selecting carriers for compliance reviews. Furthermore, representatives of a committee from the trucking industry, who expressed serious concerns about the SafeStat model, also agreed the model was an advance over an earlier system that identified carriers for compliance reviews.

Results of compliance reviews support the views of users as to SafeStat's ability to identify high-risk carriers. About 34 percent of the 3,815 carriers placed in the top three risk categories of SafeStat in March 2002 subsequently had compliance reviews and received less than satisfactory ratings. The higher the SafeStat score the greater was the likelihood that a carrier received a less than satisfactory compliance review.[3] Carriers that SafeStat rated in the highest category (A) had less than satisfactory ratings 40.4 percent of the time while those in the next highest risk category (B) had less than satisfactory ratings 36.1 percent of the time. Carriers in the lowest risk category (C) had less than satisfactory ratings 27.4 percent of the time. The data show SafeStat's ability to identify some high-risk carriers and support SafeStat's use as an internal decisionmaking tool.

A Range of User Views Were Considered in Model Development Although Certain Standard Statistical Tools Were not Used

Historical data on the model show that during its development in the 1990s, SafeStat concepts and methodology were based on expert opinions of stakeholders, solicited by the model's developers during a series of meetings. The participants in those meetings were individuals from state and Federal governments, trucking associations, insurance companies, shippers and consumer safety groups.

---

[3] The preferred analysis would be to compare compliance reviews for carriers selected by SafeStat against a group of carriers selected for compliance reviews at random. However, no large group of compliance reviews done at random was available. The vast majority of compliance reviews that are not done because of SafeStat ratings are conducted based on complaints or reported violations, or done at the request of carriers who have already received a compliance review and requested a follow-up review.

During these meetings stakeholders listed the most important safety fitness criteria and were asked to rank order the choices. In many of their lists, accidents, vehicle maintenance, and driver conduct were included as important safety criteria. Top ranked measures cited included the preventable accident rate, equipment out-of-service rate and driver out-of-service rate. Stakeholders also listed safety management as a lesser issue. Additionally, they expressed concern about data quality, suggesting that a sampling program to test accident data and inspection reports would be useful.

Although SafeStat's original development effort considered the views of industry stakeholders, such views have not been solicited for subsequent changes, such as those affecting how power unit numbers are calculated. FMCSA needs to ensure that industry and other members of the public are provided the opportunity to comment on proposed changes to SafeStat. Comments on model changes could be received using an on-line feedback system now in place on the Internet site.

While FMCSA consulted users during model development, such views are not a substitute for rigorous testing during model development. During the development of the SafeStat model, its developers did not use statistical tools to validate the weights in the SafeStat model, even though this is a standard practice in model development. Volpe's analysis of the weighting and normalization schemes in SafeStat has been limited to determining how changes affect the stability of the model. Such analysis determines whether changes in one safety event cause a dramatic change in a carrier's score, and thus cause an unstable model.

During meetings held during the model's development, FMCSA and stakeholders discussed the general use of weighting and normalization[4] factors, but they did not discuss specific formulas. An expert we consulted opined that, while the use of rank order choices at such stakeholders meeting is a recognized tool for determining the relative strength of measures, this information should be used in conjunction with other statistical tools to determine the relative importance or value of the indicators developed for the SafeStat model.

More rigorous testing could determine, for example, whether the weighting assigned to crashes produced internally consistent results across all categories of carriers. It could also assess the consistency of rankings under various raw scores and the impact of incomplete and untimely data. In the absence of such tests conducted during the model's development, our expert concluded that an approach used by Volpe in a 1998 study would serve as an alternative method for assessing

---

[4] Normalization is a method of adjusting safety scores so that carriers of dissimilar sizes can be compared on a more equal basis.

the model; however, improvements should be made in the approach used previously.

1998 Study Shows High-Risk Carriers Identified, but Work Needed to Maximize Model's Effectiveness

In the 1998 study, Volpe examined SafeStat's crash rate prediction capability as a "bottom line" test to confirm that SafeStat identified high-risk carriers. According to Volpe, the study was conducted using carriers' historical data from October 1993 through April 1996 to calculate simulated SafeStat scores.[5] Such an attempt to measure a model's forecast accuracy is an acceptable approach to model testing used in other fields. It was also a feasible approach at the time because Volpe had historical safety data available for analysis that was uninfluenced by decisions on compliance review assignments.

Volpe analyzed carrier crashes for the 18 months subsequent to April 1996 for three groups—those having the highest safety risk, those with a lesser safety risk, and those with sufficient data, under SafeStat criteria, but not considered a risk.[6] The number of crashes attributed to each carrier in the period after April 1996 was weighted based on severity and time factors. In addition, weighted crash values were converted into rates using the number of power units per carrier.

The published results in the effectiveness study supported the validity of SafeStat. Specifically, the weighted crash rate for the riskiest carriers was 169 percent higher than the rate for the least risky. Further, carriers that SafeStat identified in the two groups with the highest safety risk also had an 85-percent higher crash rate in the monitoring period than the carriers that SafeStat did not identify as having a high risk.

This analysis is convincing, but it needs to be updated to assess the current effectiveness of SafeStat for two reasons. First, changes have been made to SafeStat including the addition of serious moving traffic violations, revised weighting of fatal accidents, and altered methods of calculating average power units. Secondly, we could not validate the calculations made in the 1998 study because the original data used in the study were not available.

In addition, while the general approach used in the 1998 study was acceptable, more sophisticated analysis, not previously conducted, would optimize the model's effectiveness. Such analysis could determine the degree to which the weightings used produced internally consistent results across all categories of

---

[5] Simulated scores were used because SafeStat was not implemented until 1995.

[6] Some carriers did not have sufficient data under SafeStat criteria, such as at least two crashes, and in these cases no SafeStat score was calculated.

carriers and demonstrate the degree to which the weightings reflected the views of experts. It could also test whether changes in the model would yield better predictive results.

For example, such tests of the models predictive ability could address concerns raised by the trucking industry about whether power units or vehicle miles traveled are a better normalization factor for adjusting scores in the accident area. Both factors are used in SafeStat to some degree but power units are used most often.[7] Some motor carriers have argued that the use of vehicle miles traveled are a better measure of exposure to crashes; thus, those computations should be used to a greater degree for the accident measure calculations involving state reported crashes.

Similarly, such tests could address whether the use of team drivers should be a factor in SafeStat scoring. A motor carrier suggested that any power unit that is operated by a driving team should be counted differently from power units operated by a single driver. This is the practice for a system used in Canada for assigning safety ratings. In that system, the calculation of fleet size is adjusted upwards if vehicles are driven by teams. Such adjustments impact the calculation of the number of accidents per vehicle, and therefore, could also change SafeStat rankings in the accident area.

The revised analysis could also address the impact of relevant information on data quality or availability. Volpe's 1998 study assumed that state-reported crashes represented an unbiased sample of crash data, although acknowledging that some crashes were missing from the database. Given the geographic bias we found in the reporting of crashes, as discussed in Finding B, this assumption is questionable. If the data quality issues noted cannot be readily corrected, then the model should be modified, where possible, to reduce the chances for missing high-risk carriers because of bias in the reporting of certain data. For example, FMCSA could change the criteria used to evaluate the relative safety risk of carriers who predominately operate in locations where data are not complete. FMCSA already adjusts the deficiency threshold for hazardous material carriers from the 75th to the 70th percentile based on the desire to provide greater scrutiny for this type of carrier. The same concept could be used with regard to states or regions with data reporting problems.

Additional analysis can also point to the value of obtaining more complete data. For instance, expanded use of vehicle miles traveled as an exposure indicator, even if theoretically desirable, would be impractical at this time because 524,182

---

[7] Vehicle miles traveled are used in SafeStat for calculating the recordable accident indicator. This indicator is one of two calculations that may be used for determining a carrier ranking in the accident area. The other indicator, used more often, is based on power units.

or about 81 percent of active, interstate motor carriers, as of January 8, 2003 reported no data to the FMCSA database on vehicle miles traveled. Further assessment of the model could determine the value of increasing efforts to collect the data.

FMCSA officials reported that an independent analysis contractor has been hired to revalidate the SafeStat model. Additionally, FMCSA's FY 2005 Budget Estimate includes funding to test the assumption that SafeStat scores are a predictor of crash involvement for all commercial carriers.

RECOMMENDATIONS

We recommend that the Federal Motor Carrier Safety Administrator:

A1. Initiate an effectiveness study of the current model that tests the model's key parameters, assesses possible model adjustments to account for missing data, and evaluates whether the use of vehicle miles traveled or adjustments for team drivers would improve SafeStat calculations. The results of the study should be subjected to independent review by a party outside of Volpe.

A2. Establish processes for soliciting public comment on proposed changes in SafeStat calculations, to include those changes, if any, resulting from the revised effectiveness study.

MANAGEMENT COMMENTS AND OFFICE OF INSPECTOR GENERAL RESPONSE

We provided a draft of this report to FMCSA on December 10, 2003. In its comments, FMCSA cited a number of improvements already implemented or ongoing to address the recommendations in the report. FMCSA comments on Finding A and our response to those comments are summarized below. The full text of a matrix that FMCSA provided detailing actions planned or underway, along with significant FMCSA comments on the draft report findings, are provided in the Appendix to this report.

In response to Recommendation A1, FMCSA stated that an independent contractor has been hired to revalidate the SafeStat model, and the study should begin in early 2004. In response to Recommendation A2, FMCSA stated that it believes an effective public comment process is already in place on the SafeStat Internet site and that the majority of the changes to SafeStat have originated from industry and other stakeholder feedback. FMCSA will explore the feasibility of

using the existing feedback mechanism to solicit public comment on proposed changes.

We consider the planned actions to be responsive to the recommendations; however, target completion dates are needed for completing the revalidation study and for determining the feasibility of changes in the existing feedback mechanism on the Internet that would allow the solicitation of public comment on proposed changes. We request that FMCSA provide a written response with this information within 30 days.

Finding B. FMCSA Should Improve Key SafeStat Data Reported by States and Motor Carriers

We found that compliance reviews and enforcement data generally were reflected in SafeStat calculations in an accurate and timely manner; but problems occurred with data derived from motor carrier census forms and state-reported crashes and roadside inspections.   Problems included out-of-date and unreliable carrier information, incomplete data on crashes and serious moving traffic violations, untimely reporting of crash and inspection data, and errors in the crash and inspection reporting process that prevented the proper inclusion of some safety events in SafeStat scores.

When motor carriers and the states provide insufficient data it creates an unknown degree of bias in SafeStat's ranking of motor carriers and limits SafeStat's effectiveness as a tool for identifying high-risk carriers.   As a result, FMCSA cannot be sure it is focusing its resources for compliance reviews on carriers with the highest risk.   Moreover, as data quality varies significantly by state and by region, FMCSA may fail to target certain high-risk carriers because of where they operate.   Additionally, missing, incomplete, or untimely safety event data may cause public Internet users, who rely on specific rankings, to make incorrect decisions.

> **Data Sources For SafeStat**
>
> - **Federal safety violations and crashes from over 11,000 annual compliance reviews**
>
> - **Federal safety violations from closed enforcement cases**
>
> - **Census information from 643,000 motor carriers**
>
> - **State-reported crashes- 100,000 a year**
>
> - **State-reported inspections- 2.5 million a year with 240,000 serious moving traffic violations**

Data were deficient because effective control systems and quality standards were not in place at the Federal and state levels to provide reasonable assurance of data quality and to correct data inaccuracies. FMCSA has undertaken a number of actions to address data quality issues and to respond to the DOT Data Quality Guidelines issued in October 2002.   However, timely and complete corrective actions are needed to improve carrier reporting of current census data and to obtain more complete and timely data from states on crashes and inspections. Data improvements are important given the influential nature of SafeStat information and the stricter treatment of such publicly disseminated data called for in the October 2002 guidelines.

Motor Carriers Not Providing Current and Reliable Census Data Used in SafeStat

The Motor Carrier Safety Improvement Act (MCSIA) of 1999 required interstate motor carriers to periodically update their identification report (census forms) filed with FMCSA, but many still do not. Under FMCSA rules, motor carriers that were active on December 2000, should have filed an updated census report no later than December 31, 2002. However, according to FMCSA's Motor Carrier Management Information System,[8] as of January 2003, approximately 272,000 (or 42 percent) of the 643,909 active carriers had not done so. By our calculation this included 31 percent of about 170,000 carriers that had percentile values calculated in SafeStat for one or more of the four safety evaluation areas. For example, there were:

- 23,919 motor carriers, still active as of January 2003, that had not updated their records since 1974, such as one carrier with 45 roadside inspections on record and 6 accidents—4 with injuries—involving 6 different power units;

- 2,932 (17 percent) of 17,514 motor carriers with 2 or more crashes over a recent 30-month period; and

- 6,279 (20 percent) of 30,901 motor carriers with 3 or more serious moving traffic violations over the same 30 months.

In addition to not updating census information, carriers erroneously claimed that "zero" drivers or power units were in use. Overall:

- About 98,000 carriers or 15 percent of the 643,909 active carriers were on record as having zero drivers. This number includes about 26,000 carriers that updated the information, but still provided unreliable driver data, and 15,136 carriers with at least one inspection on record from October 2001 through September 2002, even though the company had no drivers.

- In January 2003, approximately 71,000 (11 percent) of the 643,909 active interstate carriers were on record as having "zero" power units—an improvement over levels FMCSA reported in 1999. At that time 24 percent of the carriers showed similar data. The reporting of "zero" power units still occurs in more than 1 in 10 motor carriers.

Accurate census data are important because information on power units and the number of drivers a carrier has are used for deriving SafeStat scores for a significant number of high-risk carriers. For example:

---

[8] The Motor Carrier Management Information System is the name given to the database managed by FMCSA from which data used for SafeStat calculations are drawn. The data analyzed were current as of January 2003.

- In January 2003, each of the 506 carriers in the highest risk category (A) were rated deficient in the accident area, the highest weighted evaluation area. For 376 (or 74 percent) of these carriers, the accident area calculation was based on *power units.*[9] An alternative measure—vehicle miles traveled—was used in the remaining 26 percent.

- Of the category A carriers ranked by SafeStat in January 2003, about 11 percent used data on the number of drivers in the calculations.

Inaccurate and unreliable census information can skew SafeStat calculations, particularly when data show zero drivers or power units, because SafeStat calculations are invalidated when zero values are present.[10] As a result, motor carriers can show crashes on record, including fatalities, but receive no SafeStat value in the accident area. For example, our analysis of the January 26, 2003 SafeStat run identified 391 carriers that had two or more crashes, but for which an accident area value was not calculated because the power unit data were recorded as a zero or a null (blank) value. One of the carriers had 21 crashes, but received no SafeStat value.

Biennial Update Requirements Not Enforced. FMCSA has promulgated rules, established in November 2000, to require carriers to update their census forms on a biennial basis, but has not sufficiently enforced them. Although FMCSA's regulation states that fines of up to $500 for each offense can be levied for failure to update carrier census forms, we found only one recorded case of enforcement of a fine related to a census form. FMCSA has undertaken a letter campaign to persuade carriers to comply with the reporting requirements. However, FMCSA management reports that would identify carriers, by state, with zero power units, and thus promote better enforcement of reporting requirements, were discontinued when a new FMCSA database was implemented in September 2002.

In our opinion, FMCSA must initiate an active program to correct the census data problem, including enforcing fines or penalties for failure to comply with biennial requirements to update carrier census information and targeting carriers reporting zero power units or drivers for special attention. Such efforts can improve the quality of data in SafeStat and reduce bias imposed by insufficient census data. Stronger actions could also include initiating a more vigorous letter enforcement campaign and, if carriers ignore repeat reminders, setting default values to one for zero values reported for vehicle and driver safety data. This would increase the likelihood of carriers receiving a high SafeStat score and a subsequent compliance review, and thus promote fuller compliance. But it would also introduce possibly

---

[9] The number of power units is defined by the total number of trucks, tractors, hazardous material tank trucks, motor coaches, school buses, minibuses/vans, and limousines that are owned or leased by a carrier.

[10] Because the number of drivers and power units is used as a divisor in the calculations, values cannot be computed using "zero" or blank data.

erroneous data into the system. An assessment of non-responding carriers would be needed to determine if this action was warranted.

State-Reported Crash Data are Incomplete and Untimely

Adding to the data problems created by insufficient census data is significant crash information also missing from the states. Overall, FMCSA's own estimates are that data on one in three large trucks involved in a Federal reportable[11] crash are missing from FMCSA's Motor Carrier Management Information System. This estimate, using FY 2001 data, is based on a probability-based nationally representative sample of crashes, as well as actual reports provided to FMCSA and to the Fatality Analysis Reporting System in NHTSA. In our estimate[12] 37,000 large trucks missing from the crash reports are involved in interstate commerce.

Our analysis of data from all states for the July 1 through December 31, 2002 period,[13] and our visits to 10 states showed significant gaps in crash reporting by the states. For example:

- 6 of 51 states did not report any crashes to FMCSA in the 6-month period. This included 2 of the 10 states we visited.[14] New Mexico had not uploaded any crashes for the previous 6 months while Pennsylvania had an estimated 18,000 crash reports waiting to be entered into its data system from the previous year.

- In another state we visited, although fatal large truck crashes were reported to FMCSA, 5,816 non-fatal large truck crashes from 2001 had not yet been uploaded to the FMCSA database as of January 2003. State officials planned to enter that data in the future.

Our review of crash data reporting showed variations in underreporting across states for all crashes involving commercial vehicles, and reporting problems specific to fatal crashes.

Variations in Underreporting of Crashes. The underreporting of crashes is widespread but varies in severity across states. Figure 1 shows underreporting estimates by states in 2001. States shown in red <u>did not report</u> at least 60 out of every 100 large trucks involved in reportable crashes, while states shown in green <u>reported</u> at least 80 percent. Averaging the variation by region shows that if a carrier operated in the southern region, there was about a 50-percent chance that a

---

[11] Federal reportable crashes are those with a fatality, an injury, or a tow away because of disabling damage to a vehicle involved in the crash.
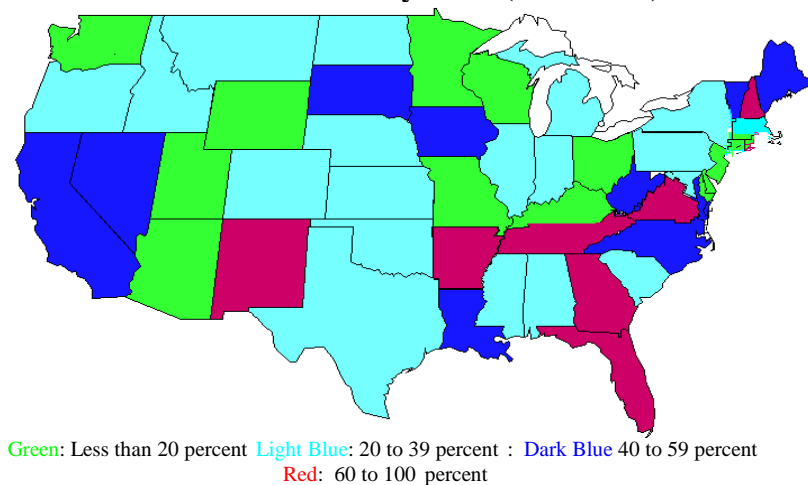
[12] This estimate is based on reported levels of crashes involving intrastate and interstate carriers as of January 2003.

[13] Data as of January 8, 2003.

[14] The six states not reporting any crashes in the last 6 months of 2002 were the District of Columbia, Florida, New Hampshire, New Mexico, Pennsylvania, and Vermont.

state did not report a large truck involved in a crash versus a 20-percent chance if a carrier operated in the mid-west region. Thus, based on the data for 2001, there is a greater chance in the southern region that a large truck involved in a crash will not be reported to FMCSA than in the mid-west region.

Figure 1. Estimate of Underreporting of Large Trucks Involved in Crashes by State (2001 Data)

Green: Less than 20 percent Light Blue: 20 to 39 percent : Dark Blue 40 to 59 percent
Red: 60 to 100 percent

Underreporting of Fatal Crashes. Fatal crashes alone also indicate significant underreporting and variations across states. Using November 2002 data, FMCSA identified 1,031 of 4,853 fatal large truck and bus crashes in NHTSA's database that were missing from FMCSA's database used for SafeStat. The largest number of crashes not matched was in the southern and western regions with 444 in the southern region and 239 in the western region. In addition, 425 or 41 percent of the unmatched crashes occurred in four states—California, Florida, Tennessee, and Texas.

Problems with Timeliness of Crash Reporting. Significant delays in the reporting of crashes will influence SafeStat calculations as crashes occurring within the most recent 6 months are weighted more severely than those from the previous 7 to 30 months. FMCSA has established a 90-day standard for crash reporting, but the agency's quarterly reports show difficulties in meeting this standard. The time it takes to add or upload a crash report to the FMCSA database has varied from 165 days in 1999 to 106 days in 2002. Our analysis of crashes on record for FY 2002 showed that about 18,000 or 20 percent of crashes that were reported in the 12-month period were entered into the database 6 months or more after the crash occurred.

In addition, average figures across the nation do not provide a complete picture because variations in the reporting times across the states are also a problem. Our analysis of a 4-year period (FY 1999 through FY 2002) indicates the variation in

reporting across states and shows how improving specific problem states could affect overall averages.

- As of November 2002, crash reports in FY 2002 were uploaded on an overall average of 103 days. However, 4 states averaged over 200 days while 20 states averaged less than 75 days.

- Decreasing the average reporting time for crashes in 6 to 8 states with slow reporting times could improve the overall reporting time by 22 percent.

If data were consistently late across all states, the impact on SafeStat would be minimal because each carrier's SafeStat calculations would be equally impacted, no matter where the carrier operated. Geographical bias is introduced in the rankings to the degree to which untimely data varies from place to place. Some variation would be expected, but a statistical analysis shows high variation in reporting times across states. The reporting variation across states or regions can be expressed statistically through the use of the coefficient of variation, a basic measure of dispersion, expressed as a number between 1 and 100. A standard or normal distribution has a value of 25.[15] For crashes we analyzed, the coefficient of variation over the 4-year period evaluated was more than 70, indicating that the data had a high variation. (For more details on this analysis, see Exhibit B.) Given the gaps and delays in the state reporting of crash information and the variations across states, we concluded that SafeStat calculations regarding crashes will be impacted by the location of a carrier's operations.

Impact of Incomplete Crash Reporting on SafeStat. How much impact could one or two missing crashes make in a SafeStat score? Although we cannot specify the impact of "missing data," we can help visualize the potential impact by addressing how adding only a small number of crashes to a carrier record can influence SafeStat peer group and percentile calculations. Table 4 shows the number of reported crashes by peer group in January 2003 and the weighted[16] accident value that corresponds to the 75th percentile cutoff point, the point at which a carrier is considered a high risk in a particular area.[17]

---

[15] The coefficient of variation is obtained by dividing the standard deviation by the mean value of the data set, then multiplying that result by 100.

[16] Crashes are weighted by age and severity and whether hazardous materials are released.

[17] The value for the 75th percentile changes for each SafeStat run and across evaluation areas.

Table 4. Variations in Selected Accidents Reported Per Peer Group		
Peer Group Number	Number of Crashes Reported	Approximate Cutoff Point of 75th Percentile
2	2 to 3	.971
3	4 to 8	.569
4	9 to 20	.408

As shown, only one crash separated the peer groups, but the cutoff point for the 75th percentile in the peer groups varied widely. Thus, the peer group that a carrier is placed in can affect the carrier's SafeStat score. The cutoff point for peer group 2 is .971 or about 70 percent higher than the cutoff point of .569 for peer group 3. Only one "missing crash" can move a carrier to a different peer group and result in differential treatment for the same crash rate. For example, a carrier in group 3 with 4 reported crashes and a weighted accident rate of .569 is considered deficient. The same carrier in group 2 with only 3 reported crashes, has to reach a rate of .971 before being considered deficient.

The January 2003 SafeStat run showed 445 carriers in peer group 2 that had 3 crashes and weighted crash rates above .569, but below .971. These carriers were not above the cutoff point for peer group 2; therefore, they would not be considered high-risk in that evaluation area. However, 18 of the same carriers were domiciled in Pennsylvania, Florida, or New Mexico—states that reported "zero" crashes in the last 6 months of 2002. If any of the carriers had crashes in those states during the last 6 months of 2002, and if the crashes had been reported, the carriers probably would have been placed in peer group 3, where their weighted crash rate of above .569 would have placed them above the 75th percentile and in the high risk category.

Even within peer groups, additional crashes will increase a carriers rankings. For example, if we hypothetically increased by 20 percent the number of crashes for 5 carriers that were part of the 65 carriers randomly selected within the different peer groups, the percentile score of each carrier would increase from 2 to 10 percentile points. Additionally, one carrier would move from a 56 to a 67 percentile.

Such a jump in rankings would not affect FMCSA's decision concerning a compliance review, because the cutoff point is 75, and two areas must be deficient. However, we found that users of SafeStat outside FMCSA may make decisions based on values below 75 in a single evaluation area. For example, a carrier scoring 67 versus 56 in the accident area would no longer be eligible to participate in a Department of Energy program, because its published criteria would deem a carrier unacceptable if it scored higher than 64 in the accident evaluation area.

Given the incident of missing data, it is reasonable to conclude that a public business making a decision based on SafeStat data, particularly if it relies solely on the accident area, will obtain biased information depending on where the carrier operates.

The effect of missing crashes on carrier rankings for specific geographic areas of operation can be inferred from looking at changes in the rankings of groups of carriers when crash reporting problems occur. As of January 2003, our review showed that no large truck crashes in Pennsylvania had been reported to FMCSA since December 2001. Our analysis of SafeStat scores for the period showed a 73-percent decline in the number of Pennsylvania-domiciled carriers with deficient values only in the accident area. At the same time, the number of Pennsylvania domiciled carriers in categories A and B, which are affected by both accident and non-accident safety data, declined by 44 and 35 percent, respectively. Nationwide, the number of carriers in the accident only category declined by less than 1 percent. We attribute this disparity not to the emergence of fewer high-risk carriers from Pennsylvania, but to the underreporting of crashes in that state.

Such underreporting impacts more than just one state. Since SafeStat scoring involves a relative ranking of one carrier against another based on available data, missing data may place a carrier in a deficient category because data for a higher-risk carrier is not included in the calculation. Additionally, for specific geographic areas or states where underreporting of crashes is more severe, SafeStat may fail to identify certain high-risk carriers simply because of where those carriers operate. In addition, since carrier rankings for the accident area are publicly disseminated via the Internet, some carriers may be excluded from consideration as high-safety risks by the public users because they operate in states with low crash reporting.

Efforts to Improve Crash Reporting. FMCSA has recognized the problem with states' incomplete reporting of crashes. Officials attributed the problem to the fact that FMCSA must rely on numerous local state jurisdictions to submit crash reports. Additional problems are the states' noncompliance with reporting requirements and resource issues at the state level. Although our 1999 audit recommended standardization in crash reporting and procedures, our current review showed that differences persist in the treatment of items, such as how federally-reportable crashes are recorded and what definitions are specified for commercial vehicles. MCSIA took note of data collection problems regarding crashes and directed FMCSA, in cooperation with NHTSA, to administer an improved data program.

In implementing the MCSIA requirement, the two agencies exhibited different viewpoints over what should be included in the improved data program. NHTSA argued for the creation of an entirely new data system involving electronic coding

of crashes with all states eventually participating, while FMCSA argued that the focus should be on improving the current system and the data collection process.

Given these differing viewpoints, the data collection improvement efforts carried out by the two agencies advanced slowly. Joint initiatives have included establishing pilot programs in seven states and comparing fatal accidents involving commercial vehicles recorded in each agency's database to identify incomplete information. However, the two agencies did not execute an agreement that focused joint efforts on improving the existing FMCSA database for commercial crashes until February 2003. The memorandum of understanding between the two agencies has been followed up with a more specific project plan that establishes specific short-term and long-term goals for the improvement of commercial vehicle crash reporting in targeted states. In our opinion, because of the common interest of NHTSA and FMCSA in obtaining complete and accurate crash data, the two agencies should jointly address crash data issues as much as possible.

State-Reported Inspection Reports are Also Untimely and Serious Moving Traffic Violations are Underreported

Although the rate of state inspection reporting was better than crash reporting, we found it was also a problem. According to FMCSA reports, states on average reported inspections in 35 days in 2002, versus the 21-day standard. More significant, however, was state underreporting of serious moving traffic violations[18] observed in conjunction with roadside inspections. Approximately 240,000 such violations are reported nationwide each year in the FMCSA database (the majority of which are for speeding). However, as with crashes, the states reported these occurrences incompletely and inconsistently. For example:

- California reported only 115 serious moving violations to the FMCSA database in Fiscal Year 2001 even though the state ranked number one in the nation for the number of commercial vehicle miles traveled and the number of roadside inspections. Conversely, Indiana, a state with 2.7 times fewer commercial vehicle miles traveled and 8.6 times fewer inspections, reported about 25,000 moving violations to FMCSA's database in the same period.

- Pennsylvania reported 60 times more serious moving traffic violations than Illinois, although it conducts fewer inspections annually than Illinois and logs about the same number of commercial vehicle miles traveled.

---

[18] To evaluate the driver, SafeStat uses the following serious moving traffic violations: (1) failure to obey traffic control device; (2) following too closely; (3) improper lane change; (4) improper passing; (5) reckless driving; (6) speeding; (7) improper turn; (8) failure to yield right of way; (9) use or possession of drugs; and (10) use or possession of alcohol. SafeStat uses only those traffic violations that are indicated on the inspection reports completed by law enforcement personnel in conjunction with roadside inspections.

Overall, 12 states reported serious moving traffic violations associated with roadside inspections at a rate of 0.2 violations per million commercial vehicle miles traveled while 13 other states reported rates more than 7 times higher.

Figure 2 depicts the wide discrepancy in the reporting of serious moving violations per million commercial vehicle miles traveled per year across the states. As shown, although the average rate per million miles is **.84** (see dotted line), the rate ranges from virtually **0** (less than **.01**) to **3.87**.

Figure 2. Rate of Reporting for Serious Moving Traffic Violations Varies Across the States (2001 Data)
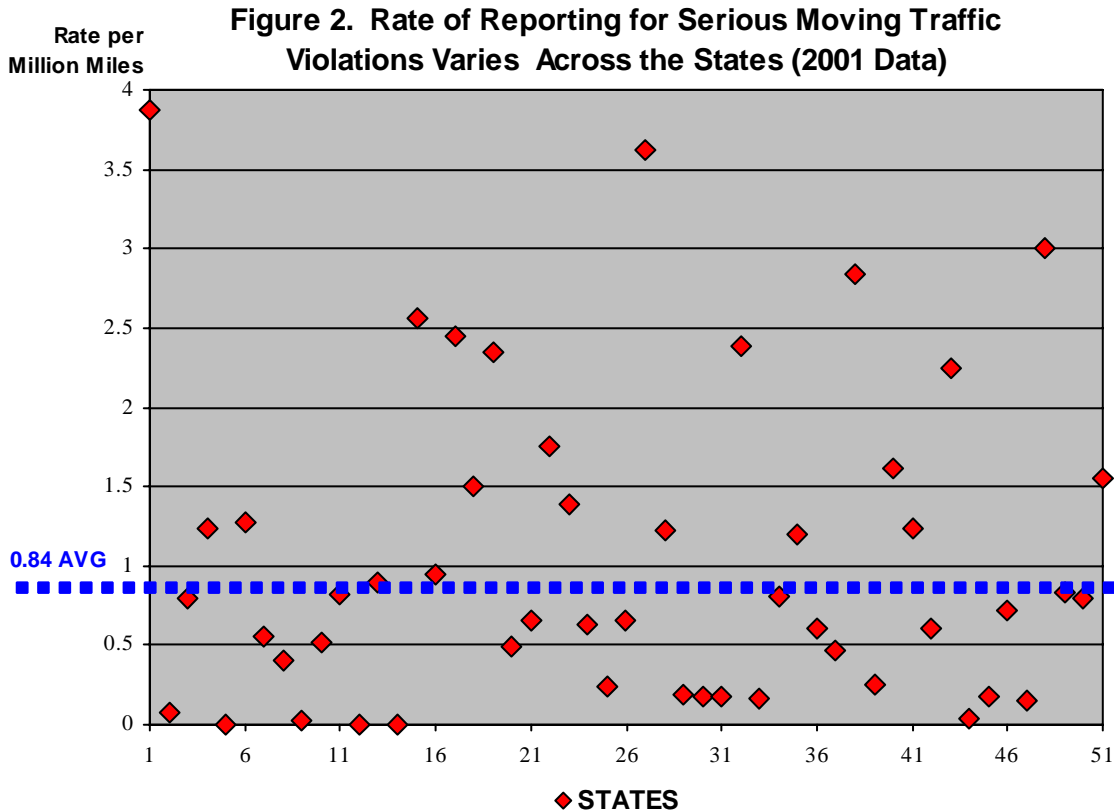
The reasons for underreporting varied. In Illinois, our comparison of moving violations cited on inspection reports with data on the same moving violations maintained in the FMCSA database showed that FMCSA's database did not correctly reflect the violation codes cited by the Illinois law enforcement officers. For example, a state police officer cited a specific speeding violation in his report, but FMCSA listed it as a "general violation," and it was not included in the SafeStat calculation. This occurred numerous times because of the improper conversion of data from the state database to the Federal database. FMCSA identified the problem only after we checked the source documentation as part of our review. In California, the low number of serious moving traffic violations was attributed, in part, to the forms officers used in performing traffic enforcement

inspections. In this case, the state forms, which were not automated, did not include a specific place to record a Federal violation code, such as speeding.

Underreporting of moving traffic violations has less effect on SafeStat than crash data, but it is still important. In January 2003, 74 percent of the highest-risk carriers (category A) used state-reported crash data in the SafeStat calculation, while only 11 percent partially used moving vehicle data to calculate the SafeStat score. However, underreporting of serious moving traffic violations represents a missed opportunity for targeting high-risk behavior, because a relationship has been established between moving traffic violations and higher crash rates. In the near term, FMCSA's targeted review of moving traffic violations could reveal reasons behind the disparities across states and promote corrective action. In the long term, devising a means to directly collect citation information and incorporating this information into SafeStat would be a significant enhancement to the system.

Although timeliness of inspection reporting was better than with crash reports, it was also a problem given the variation across the states. Based on FMCSA's quarterly reports, the time it takes to add or upload an inspection report to the FMCSA database has varied from 32 days in 1999 to 35 days in 2002. However, average figures across the nation are misleading because our analysis indicated variations in the states' reporting, which may influence SafeStat rank calculations. Specifically, our analysis of a 4-year period (FY 1999 through FY 2002) showed that:

- The overall average for inspections was 28 days as of November 2002. However, four states exceeded 60 days.

- Decreasing the average reporting time for inspections in three to four states with slow reporting times could improve the overall time by more than 20 percent.

As with crashes, the reporting variation across states or regions can be expressed statistically through the use of the coefficient of variation—a measure of how much data values in a group vary from the average, expressed as a value between 0 and 100. For inspections, the coefficient of variation over the 4-year period evaluated was more than 90, indicating that the data had an extreme variation. (For more details on this analysis, see Exhibit B.) Given the variations across states in the degree of and in the timeliness of reporting inspection violations, we concluded that SafeStat calculations could be influenced by where a carrier operates.

Sampled Crash and Inspection Records Showed Material Errors Affecting SafeStat

Complete and timely data are important, but accuracy is just as important to ensure that the correct carrier is associated with the correct safety event data. To assess the accuracy of data, we reviewed statistically selected crashes and inspections contained in FMCSA records, for the period from July 1 through December 31, 2002.[19] Based on our review, we made estimates regarding the proportion of crash or inspection transactions that were not correctly reflected in SafeStat. We counted as errors only those that could have a material effect on SafeStat. For example, a mistake in the number of injuries reported would not be considered material because the number of injuries is not a weighting factor in SafeStat—just whether or not an injury or fatality occurs. We are 95-percent confident that the estimates are correct plus and minus a margin of error.[20] Our review showed that material errors occurred that could influence SafeStat calculations. Specifically:

- Under the criteria established, we estimated that errors occurred in approximately 13 percent of the crash transactions and 7 percent of the inspection transactions reported on interstate carriers. For the 6-month period, we estimate that 2,851 of 21,225 crashes and 76,521 of the 1.02 million inspections reported contained material errors that could influence SafeStat.

- Out of an estimated 79,372 incorrectly reported safety events, we estimated that in 69,888 crashes or inspections reported, SafeStat did not hold the carrier accountable for the safety event and in 9,484 crashes or inspections reported, SafeStat erroneously held a carrier responsible.

Thus, about 89 percent of the errors excluded a safety event from a carrier's record that would have been used to calculate SafeStat scores and rankings. For example, a crash in Louisiana involving three injuries was not included in the carrier's SafeStat calculation because, although the police report included a description of the injuries, the data were not correctly entered into the database.

---

[19] Crash and inspection data were selected from FMCSA records as of January 8, 2003.

[20] We estimate a 13-percent error rate for crashes with a margin of error of plus and minus 6 percent; and we estimate a 7-percent error rate for inspections with a margin of error of plus and minus 4 percent.

The remaining 11 percent of the estimated errors were of a different type. In these cases, a mistake resulted in a carrier being erroneously charged with a SafeStat related violation. For example, a roadside incident in Indiana did not qualify as a crash under the Federal definition, but it was erroneously reported as one because the definition was incorrectly applied. We found

> **"…about 89 percent of the errors excluded a safety event from a carrier's record that would have been used to calculate SafeStat scores and rankings."…**
>
> **"[In] the remaining 11 percent of the errors…the mistake resulted in a carrier being erroneously charged with a SafeStat related violation."**

these data problems when we compared FMCSA records with source documents in the states or when we requested validation of selected crash records from motor carriers.

While the smallest in percentage terms, the most critical problems identified in the sample from an individual carrier's perspective are instances where carriers are incorrectly held accountable for SafeStat related violations. This can lead to time and expense to correct the problem and the possibility of a compliance review, even if such action would not have occurred if the transaction had been accurately recorded.

Errors due to "missing data" in which a carrier was not held accountable in SafeStat have a less obvious impact on individual carriers. However, these errors influence the entire system as one carrier's scores may be higher than they should be because missing data prevents carriers with more serious safety performance data from gaining a higher relative ranking. Overall, the results indicate that it is more likely that unsafe carriers would be identified as safe rather than safe carriers being identified as unsafe. We did not attempt to determine whether these error rates were consistent for different segments of the industry. Thus, the rates identified may vary significantly among truckload, less-than-truckload, or household goods carriers, or vary depending on the scope of carrier operations. Additional errors that we could not identify on inspections could also exist given the methodology we used. (See Exhibit B for details.)

Improving Data Quality Control Systems

Data deficiencies related to missing crash information, underreported moving traffic violations, untimely reporting, and inaccuracies in crash and inspection records exist because effective control systems and quality standards were not in place at the Federal and state levels to provide reasonable assurance of data quality. FMCSA has completed a number of actions with motor carriers and the

states to address data quality issues. These include edit checks, reviews of state MCSAP programs, incentive grants, and the addressing of data issues in the annual commercial vehicle safety plans that the states submit to FMCSA. However, actions in these areas can be improved.

Edit Checks. FMCSA used edit checks for identifying missing, invalid, duplicate, inconsistent, or otherwise erroneous data; but current checks do not go far enough and have not always worked properly. The current edit checks include identification of incomplete fields and entries outside of acceptable ranges, and match DOT numbers with carrier names. More sophisticated checks for logic and reasonableness that would detect data anomalies between reporting periods were not in place, although plans were underway for developing such checks.

Oversight Reviews. Even when perfectly implemented, computer edit checks of records cannot identify all the problems that relate to improper entry from source documents, such as those we found in our analysis of moving traffic violations and crashes. To find these problems, FMCSA will need to conduct or have the states conduct quality reviews of controls and source documentation.

FMCSA is required to conduct regular reviews of state programs that, in part, address data quality; however, it is not conducting the reviews regularly, and neither are the states. FMCSA's state offices are required to conduct program reviews at least every 3 years to address the states' implementation of grants provided through MCSAP. The reviews are to include grant expenditures, data quality and timeliness, and program effectiveness. Only 3 of the 10 states we reviewed had completed the required MCSAP reviews.

Given the data quality issues we identified, the frequency and depth of FMCSA's reviews of state MCSAP programs should be increased. In addition, FMCSA's reviews should be modified so that they regularly and uniformly address data quality issues and include tests to identify improper data entry from source documents. Alternatively, occasional focused reviews could be done so that nationwide issues can be addressed in a timely manner. A prime candidate for this review would be problems with the reporting of serious moving traffic violations. Based on our results, the inclusion of procedures comparing the accuracy of data to source documents would be a useful element in these reviews. Such reviews by FMCSA could also encourage better state reviews. Of the 10 states we reviewed, 5 did not conduct quality reviews of safety event data transmitted to FMCSA.

Incentive Grants. To promote data quality, FMCSA established incentive grants to states. These grants, however, are too limited in scope. An FMCSA rulemaking published in March 2000 recognizes that the collection of complete, accurate, and timely accident data is vital to reducing fatalities and accidents. As such, the rulemaking provides for incentive grants to the states for meeting upload

requirements established for crash and inspection reports. However, the focus of those incentive grants is on timeliness alone, not on other aspects of quality, such as accuracy and completeness. Based on the accuracy and completeness problems we have identified, we opine that incentives should also be developed to promote improvement in these areas. Where needed, changes should be made to FMCSA guidelines or policy to accomplish this as well as appropriate use of high priority MCSAP funding for data quality purposes.

State Plans. An additional mechanism for promoting improved data quality, which has not been fully utilized, is the annual commercial vehicle safety plan that the states submit to FMCSA under MCSAP. Neither the rulemaking nor FMCSA guidance requires states to submit plans with specific actions and milestones to improve the completeness and consistency of state reported crashes and other safety data. In our view, taking action in this area would improve the accuracy of SafeStat data and thus increase public confidence when making business decisions.

FMCSA's Plan in Response to DOT Guidelines is a Positive Step, but is not Sufficient

In response to DOT Information Dissemination Quality Guidelines, issued on October 1, 2002, FMCSA has prepared several proposals for improving the quality of safety data. Issued in response to the Office of Management and Budget guidelines directing agencies to "substantiate the quality of the information it has disseminated," DOT's guidelines have been cited as a benchmark standard for data quality assurance and recommended as a template for evaluating the data quality programs of other agencies.[21]

FMCSA has developed a draft plan to implement the DOT guidelines. In addition, FMCSA has presented and developed proposals for improved data correction and monitoring, recognizing that SafeStat scores fall into a category of information labeled "influential data."[22] Such data require stricter quality standards because public and private decisionmakers use the data. Although the draft plan is a positive step, FMCSA should add elements to the plan addressing data quality standards, correction procedures, missing data, and disclosure of data issues to fully implement DOT's guidelines regarding information quality.

Data Quality Standards. FMCSA's plan requires that methods be investigated and developed to improve the timeliness, completeness, and accuracy of safety

---

[21] Center for Regulatory Effectiveness, March 2003.
[22] Influential in this context means that the agency can reasonably determine that dissemination of the information will have or does have a clear and substantial impact on important public policies or important private sector decisions.

data. However, the plan does not address requirements for completeness, timeliness, and accuracy although such data requirements are included in the DOT guidelines. For example, the DOT guidelines cite an illustrative data requirement for the Fatality Analysis Reporting System (FARS), a database within NHTSA that contains fatality information associated with motor vehicle crashes. The example includes a timeliness standard of 3 months, a standard error rate of 6 percent, and a coverage rate of 90 percent. The DOT guidelines point out that decisions on the data requirements should be based on how the information will be used in decisionmaking. In our opinion, a similar data requirement for FMCSA crashes is needed, at a minimum, to address coverage rates for crash reports, the timeliness of reports, and the reliability or precision of the data. Following preliminary discussions on our audit findings, FMCSA has moved to develop standards and has indicated its adoption of several quality standards specific to crash data.

Improving Data Correction Procedures to Allow Timely Correction of Errors. Based on the October 2002 guidelines, DOT has established a mechanism to receive and monitor requests for correction of information. The system is part of DOT's Docket Management System, and thus may be used for all DOT data. As of August 2003, 87 of 92 requests for data correction involved SafeStat data items. Although the Docket Management System provides a central location for carriers to transmit requests for correction, the system has limited ability to directly correct errors. FMCSA's response to carriers with data problems related to state-reported crashes and inspections is that carriers must deal with individual states.

FMCSA recognizes a need to do more to facilitate the correction of errors in SafeStat data. To accomplish this, an improved, Internet-based system for tracking data correction requests is under development, with a prototype undergoing testing and expected to be available to the states by February 2004. This system is designed to automatically provide correction requests to state offices and to track the status of requests through resolution. Our audit work supports the establishment of such an Internet-based system as a means of facilitating correction of inaccuracies in state reported data. Although we found that errors resulting in a carrier being erroneously charged with a SafeStat-related violation were infrequent, such mistakes do occur and implementing a system for correcting this information as soon as possible is important.

FMCSA standards issued in 1996 call for the correction of data inaccuracies within 7 days, but we have no evidence of FMCSA enforcing this standard at the state level. Based on our survey of state officials, the number of current requests for corrections being handled by the states cannot be tracked reliably. Although all states received requests to correct inaccurate data, 80 percent did not track the

number of requests received, and 84 percent did not track the number of requests that resulted in data being changed by the state.

Address the Impact of Missing Data. The DOT guidelines note the need to regularly evaluate "bias due to missing data or coverage bias." As discussed previously, examples of this type of data problem include (1) states having no reported crashes, (2) states in particular regions of the country underreporting crashes, (3) significant variations in the reporting of serious moving violations, (4) substantial numbers of carriers reporting power unit data that are incomplete, and (5) failure to report updated information. FMCSA's plan should address what degree to which missing data can impact SafeStat calculations. Given public disclosure of SafeStat data and the potential competitive disadvantage that can result if rankings are influenced because certain states are not reporting safety events, addressing bias is important.

Disclosure of Data Quality. Openness is a key element of the DOT guidelines on dissemination of information. The DOT guidelines favor analysis which allows qualified members of the public to independently re-analyze data. This openness includes both the reporting of information sources and the limitations of the actual reporting. The Internet site discloses the potential negative impact on SafeStat scores when carriers do not provide up-to-date information on the motor carrier census forms. However, the SafeStat Internet site did not provide information on the known limitations in safety event data on crashes and inspections that states report. Increased visibility and awareness of limitations in state reported data would allow users to make informed choices about the use of the data and perhaps encourage individual states to improve the quality of the data they use for SafeStat. Openness would also be encouraged by allowing industry and the public to comment on proposed changes in SafeStat, which could be useful in minimizing unexpected results from system changes. Following the issuance of our draft report, FMCSA placed a more comprehensive disclaimer on the Internet site, and promised to add information on the timeliness, accuracy, and completeness of data from each state.

Conclusion

Although we support SafeStat's continued use as an internal risk management tool, the types and magnitude of data problems we found argue for immediate and effective action to correct data problems. The most serious concern is the continued public dissemination of motor carrier rankings for the accident evaluation area given the incompleteness of crash data.

We recognize that FMCSA relies on state officials to generate and process much of the safety event data used in SafeStat and that this presents many challenges for

obtaining complete, timely, consistent, and accurate data. However, improvements are needed to increase the quality and reliability of the data SafeStat uses, which will improve SafeStat's effectiveness as an internal risk management tool and permit continued public dissemination of SafeStat information.

RECOMMENDATIONS

We recommend that the Federal Motor Carrier Safety Administrator:

B1. Take immediate action to mitigate the impact of inaccurate or incomplete data on public users of SafeStat by:

a. Making available to all states an improved system for facilitating the correction of data inaccuracies and the tracking of corrective actions within 3 months following the issuance of this report.

b. Disclosing data problems, including variations in state reporting, to public users of SafeStat.

B2. As expeditiously as possible, establish an overall data quality improvement program that:

a. Addresses longstanding issues associated with motor carrier census data by imposing fines on carriers that fail to provide updated carrier census information despite repeated opportunities to do so.

b. Sets minimum standards for the quality of SafeStat data consistent with the Department's data quality guidelines. At a minimum, the standards should address completeness, accuracy, and timeliness of data.

c. Accomplishes actions planned, in conjunction with NHTSA, for improving the completeness and timeliness of state-reported crashes.

d. Enhances the depth, frequency, and type of FMCSA state data quality reviews and monitoring, and ensures state plans address data quality.

e. Modifies FMCSA guidance and funding decisions so that MCSAP incentive grant awards are based, in part, on each state's implementation of guidelines established to provide accurate, complete, and timely safety event data.

MANAGEMENT COMMENTS AND OFFICE OF INSPECTOR GENERAL RESPONSE

In responding to a draft of this report, FMCSA agreed with our concerns for improving data quality and cited a number of improvements already implemented or ongoing to address the recommendations in the report. FMCSA did not agree with all of our assertions regarding the impact of data quality problems on SafeStat. FMCSA's substantive comments on the report's findings are provided in the Appendix along with the full text of a matrix FMCSA provided on January 14, 2004, detailing actions planned or underway to address each report recommendation.

- **FMCSA Comments.** FMCSA stated that the language in the draft report overstated the problem of out-of-date census data on SafeStat, and that it would be preferable to present the data problems in terms of how they affect SafeStat rankings. Our report states that 42 percent of the 643,909 active carriers have not updated their census data. In contrast, FMCSA stated that over 80 percent of the carriers with sufficient data for SafeStat to evaluate have updated their data. FMCSA also noted that since the implementation of a new Motor Carrier Management Information System (MCMIS) in September 2002, carriers cannot obtain new DOT numbers with census data showing zero drivers or vehicles.

- **OIG Response.** We do not agree that the language in the draft report overstated the problem of out-of-date census data; but we provided additional information in this final report to present the full context of the problem. According to our calculation, of the approximately 170,000 carriers that had percentile values calculated in SafeStat for one or more of the four safety evaluation areas as of January 2003, 31 percent had not updated census data in the required 2-year period. In addition, 17 percent of motor carriers with 2 or more crashes over a 30-month period had not updated their census data. In our opinion, even if one is concerned only with the 17 percent of carriers with 2 or more *reported crashes*, the magnitude of out-of-date census data shows that it is a significant problem. Moreover, given our findings on crashes that are *not reported*, we are also concerned about the census reports from carriers without 2 or more reported crashes. On the issue of improvements in controls since September 2002, our analysis of the FMCSA database shows that as of January 2003, more than 800 carriers receiving DOT numbers in the 4 months since the introduction of the new MCMIS still showed zero total drivers and about 1,600 showed zero vehicles.

- **FMCSA Comments.** FMCSA took issue with our statement that "data problems limit SafeStat's effectiveness and introduce bias into the ranking

process." Although agreeing that data quality problems could limit SafeStat's effectiveness to identify all high-risk carriers, officials objected to the implication that "some identified high-risk carriers may not actually be high-risk because of data problems." FMCSA stated that its information indicates the data problems are much more likely to make a poor performing carrier look good.

- **OIG Response.** We agree with FMCSA that data quality problems are more likely to make a high-risk carrier look good. However, we continue to maintain that the opposite situation can also occur. Since SafeStat scoring involves a relative ranking of one carrier against another, missing data may place a carrier in a deficient category because data for a higher-risk carrier is not included in the calculation. Missing crash data was most significant with six states failing to report any crashes for the 6 months analyzed. Nationwide, estimates for the underreporting of large trucks involved in crashes varied in magnitude with some states underreporting 60 percent or more and other states estimated to be underreporting of less than 20 percent. In addition, our review of sampled crash and inspection records estimated that 11 percent of the estimated errors found resulted in a carrier being erroneously charged with a SafeStat violation. The report's discussion of different thresholds for high-risk carriers within different peer groups (Finding B) illustrates how a single erroneously charged crash could affect a carrier's ranking.

FMCSA also provided specific comments on Recommendations B1 and B2, detailing a number of corrective actions planned or ongoing to expeditiously establish an overall data improvement program. FMCSA comments and OIG responses are summarized below.

- **Recommendation B1a.** FMCSA plans to implement a new electronic system for filing concerns about Federal and state data released to the public by February 29, 2004. This system will simplify the existing process and provide a mechanism for FMCSA, the states and the public to track data challenges.

- **Recommendation B1b.** FMCSA plans to add clear and comprehensive guidance to Internet users as to the limitations of SafeStat data and provide information on each state's data quality. On January 22, 2004, FMCSA posted its initial statement on data limitations on the Internet and stated that this guidance would be enhanced later by adding information on the timeliness, accuracy, and completeness of data from each state.

- **Recommendation B2a**. FMCSA reported that stronger notices to carriers on filing biennial updates to census data started on January 1, 2004. It will also develop a plan for addressing missing respondents.

- **Recommendation B2b.** FMCSA will develop a comprehensive data quality plan during FY 2004. The plan will include contracting for a sample verification of source documents; identifying additional edit checks; conducting training sessions on data quality issues; assigning staff to monitor state reports; and reviewing analysis on biennial update rates. It will also set quality goals for timeliness, completeness, and accuracy as part of the Commercial Vehicle Analysis Reporting System (CVARS) program.

- **Recommendation B2c**. FMCSA reported providing CVARS grants to 22 states in 2003 to improve crash data reporting and expressed plans to expand state participation in CVARS.

- **Recommendation B2d.** FMCSA will assign a person to review the monthly state reports on timeliness, non-match, and traffic enforcement reports and to work with states to resolve the issues.

- **Recommendation B2e**. FMCSA offered an alternative to our recommendation that incentive grants be based on each state's implementation of data quality guidelines. FMCSA plans to make the granting of high priority MCSAP grants to states contingent on a state's participation in the electronic tracking system for data challenges.

We consider the FMCSA comments on Finding B to be positive and constructive. However, we request that FMCSA clarify whether its commitment to work with states to resolve data issues (proposed in response to Recommendation B2d) will include carrying out our recommendation to ensure that state plans address data quality. We also request clarification on whether its proposed action for Recommendation B2b will be a one-time or an ongoing task. In our opinion, such reviews should be conducted periodically to ensure data integrity.

Additionally, we request specific target dates for these actions:

- Enhancing the recently instituted guidance to Internet users as to the limitation of SafeStat data by including information on the timeliness, accuracy, and completeness of data from each state (Recommendation B1b),

- Developing a plan for obtaining missing census updates from carriers who have not responded to earlier requests for this information (Recommendation B2a),

- Including more states within CVARS (Recommendation B2c), and

- Assigning staff to review monthly state reports on timeliness, non-match, and traffic enforcement reports and to working with states to resolve the issues (Recommendation B2d).

EXHIBIT A. ACTIVITIES VISITED OR CONTACTED

SITE VISITS CONDUCTED

Federal Motor Carrier Safety Administration (FMCSA)

FMCSA Headquarters, Washington, DC

FMCSA State Division Offices in Virginia, Wisconsin, California, Georgia, Indiana, Illinois, Louisiana, Montana, Pennsylvania, New Mexico, Texas, and Washington

FMCSA Eastern Regional Office, Baltimore, Maryland

Research and Special Projects Administration, John A. Volpe National Transportation Systems Center, Cambridge, Massachusetts

National Highway Traffic Safety Administration (NHTSA)

State Officials

Department of California Highway Patrol, Information Processing Group

Georgia Department of Motor Vehicle Safety Enforcement Unit

Illinois State Police Division of Operations Commercial Vehicle Unit

Indiana State Police Commercial Vehicle Enforcement

Louisiana State Police Transportation and Environment Safety Section

Montana Motor Vehicle Inspection Bureau, Montana Highway Patrol

New Mexico Department of Public Safety, Motor Transportation Division

Pennsylvania Department of Transportation, Bureau of Maintenance and Operations

Texas Department of Transportation, State Patrol Motor Carrier and Inspection Services

Virginia Department of Transportation, State Police Motor Carrier Safety

Wisconsin Department of Transportation

Washington State Patrol Motor Carrier and Inspection Services

OTHER CONTACTS

American Trucking Association

Representatives of Selected Motor Carrier and Insurance Industry

Motor Carrier Safety Assistance Program Points of Contact at States

EXHIBIT B. AUDIT METHODOLOGY

Details on the methodology we used to assess the SafeStat model, data quality issues, and data quality control systems are provided below.

SAFESTAT MODEL

We reviewed the validity of SafeStat by discussing the basis for the model with its developers at the John A. Volpe National Transportation Systems Center (Volpe) in Cambridge, Massachusetts. We reviewed relevant documentation related to SafeStat's original development, including records on meetings held with stakeholders and a 1998 Effectiveness Study conducted by Volpe staff. We consulted with technical experts on the general practices used during model development and analyzed available data on the characteristics of motor carriers identified as high risk using SafeStat. Our efforts to confirm the model's validity and its effectiveness in putting potential unsafe carriers in proper categories were limited because data used in the 1998 study of the model's effectiveness were not available for our independent assessment.

Additionally, we determined whether values calculated and posted on the Internet were consistent with the Model's design. To conduct this test we selected a random sample of 65 carriers that had received a SafeStat value in at least 1 of 4 (out of a possible 10) safety measures. Then, using Microsoft Excel software we re-created the calculations for the four measures as an independent confirmation of their accuracy. We selected the following four SafeStat measures for testing.

- Accident Involvement

- Recordable Accident Rate

- Driver Review

- Driver Inspections

We discussed the controls that were in place at Volpe and FMCSA to verify that the calculations done during each SafeStat run were consistent with the model's design. We confirmed the information provided by examining the programming tools used to perform these validations.

DATA QUALITY

To estimate the occurrence of data quality problems in SafeStat and to make nationwide projections, we used data obtained from DOT organizations, selected motor carriers, and state offices. We conducted two-stage statistical sampling in which we first selected 10 states for review and then selected for examination crash and inspection reports from those 10 states covering the period July 1, through December 31, 2002. Details on the selection of states, the estimates of material errors in crash and inspection reports, and other analyses conducted are discussed below.

State and Sample Selection. To select the 10 states, we first used FY 2001 Fatality Analysis Reporting System data obtained from NHTSA to rank the states by the number of large truck-related fatalities recorded. The 51 states were then stratified into groups with stratum 1 containing the most fatalities and stratum 4 the least. Sample states were then selected at random from each stratum with a larger proportion selected from states with the greatest number of fatalities.

In each of the 10 states selected, a sample of crash and inspection records for a 6-month period was extracted from Motor Carrier Management Information System records as of January 8, 2003. The Army Audit Agency's Statistical Sampling for Auditors software, version 6.3, was used to estimate sample sizes that allow for estimates with 95-percent confidence. From among the sample records, carriers identified as interstate carriers were selected for review. A sample of 392 crash records and 400 inspection records were evaluated. In 2 of the 10 states selected, Pennsylvania and New Mexico, no crash records were available for the sample period, so samples were selected from earlier periods and imputed to the sample period.

Review of Sample Items. During our review of the sample, two general types of material errors were identified:

- Type 1 errors, where a carrier was incorrectly reported as:

  (a) either being involved in a crash or

  (b) held accountable for an out-of-service violation or a serious moving traffic violation.

- Type 2 errors, where a carrier was involved in a crash or had incurred an out-of-service violation, or serious moving traffic violation, but this information was not properly recorded in SafeStat.

After we identified errors of reporting and non-reporting, we passed the results to a statistical consultant for analysis and projection. The statistical consultant used

analysis formulas[1] consistent with the sample design used to perform the analyses, such as, stratified primary stage weights and secondary stage weights. Stratified primary stage weights involved the number of states in the population by stratum. Secondary stage weights involved the use of the total number of crashes or inspections performed in those states on interstate carriers for the 6- month period of the study. The resulting estimates and projections included both the proportion and number of errors for 6 months worth of crashes and inspections. Each estimate is accompanied by the statistical margin of error due to sampling.

With a 95-percent confidence level, we estimated that 13.4 percent of all crashes reported nationwide are reported incorrectly. The margin of error for that estimate is plus and minus 6 percent. Type 1 errors are estimated to occur about 1.4 percent of the time, with a statistical margin of error of plus and minus 1.4 percent. Type 2 errors are estimated to occur about 12 percent of the time, with a statistical margin of error of 5.2 percent. (Given the nature of the calculation, the statistical margin of error figures do not add to the overall margin of error.) The statistical estimates for errors in the reporting of crashes are in Table B1.

Table B1. Estimates of Errors in the Reporting of Crashes	
Error Type	**Projection of Error (by percentage)**
Overall errors	13.4 percent plus and minus 6 percent
Type 1 errors	1.4 percent plus and minus 1.4 percent
Type 2 errors	12 percent plus and minus 5.2 percent

With a 95-percent confidence level, we estimated that 7.5 percent of all inspection results reported nationwide are reported incorrectly. The margin of error on that estimate is plus and minus 4 percent. Type 1 errors are estimated to occur about 0.9 percent of the time, with a statistical margin of error of plus and minus 0.9 percent. Type 2 errors are estimated to occur about 6.6 percent of the time, with a

---

[1] "Elementary Survey Sampling", 5th edition; Scheaffer, Mendenhall and Ott; Duxbury Press (International Thompson), 2000. - chapter 9 for the two-stage sampling methodology and chapter 4 for the stratification methodology.

statistical margin of error of plus and minus 3.8 percent. (Note: Given the nature of the calculations, the statistical margin of error figures do not add to the overall margin of error.) The statistical estimates for errors in the reporting of inspection results are in Table B2.

Table B2. Estimates of Errors in the Reporting of Inspections	
Error Type	**Projection of Error (by percentage)**
Overall errors	7.5 percent plus and minus 4.0 percent
Type 1 errors	0.9 percent plus and minus 0.9 percent
Type 2 errors	6.6 percent plus and minus 3.8 percent

Although the various error estimates were calculated from data statistically sampled from only 10 of 51 states, the sampling process employed two stages with stratified random selection made during the first stage and simple random selection made during the second stage, which produced a valid estimation across all 51 states.

Our analysis does not permit us to estimate the degree to which errors have affected carriers' SafeStat rankings. SafeStat-related violations attributed to the wrong carrier will increase the calculated measure and possibly increase the overall SafeStat score of that carrier. The carrier responsible for the violation will experience a lower calculated measure and possibly a lower overall SafeStat score. To determine the possible effect these errors may have on the overall score of both carriers requires a recalculation of scores using corrected data. Before the recalculation, someone in FMCSA would have to identify carriers and manually correct erroneous data in FMCSA's motor carrier database. We did not pursue further analysis to estimate the impact to rankings for the errors identified in our sample, because we, along with Volpe officials, determined it would not be possible to attribute estimated errors to a specific carrier or carriers.

For inspections, we relied on information in the inspection report to determine whether the out-of-service violation was valid. Thus, if a report incorrectly identified a leasing company as the motor carrier inspected, but the matching DOT

number was provided, we would not detect the error unless some additional information on the inspection report led us to detect the problem. FMCSA has estimated that as many as 20 percent of reports involving leasing arrangements provide inaccurate DOT numbers, but it did not have documentation to substantiate this estimate. This area should receive additional scrutiny to assess the degree to which leasing arrangements contribute to improperly assigned safety events or the degree to which certain sectors may have experienced error rates above those found in our sample and the degree to which public disseminated data is affected by these errors.

Sample of Compliance Review and Enforcement Data. To test the degree to which information on compliance reviews and closed enforcement cases were correctly reflected in the database used for SafeStat calculations, we selected a sample of compliance reviews from each of the 10 states visited and compared the information with violations and enforcement actions from the database.

Analysis of Motor Carrier Management Information System Data. We obtained and independently analyzed data from the Motor Carrier Management Information System as of January 8, 2003. The data included information on state-reported crashes and inspections conducted since October 1, 2001, and operational data on active motor carriers reflected in the census files.

Identification of Crash and Inspection Issues. We also discussed state-specific crash and inspection reporting issues with state and FMCSA officials at the 10 states visited. To obtain additional records on crashes, we sent letters to 132 motor carriers requesting validation of selected information in FMCSA records on crashes reported in the 10 sampled states. We obtained additional information on crashes and inspections conducted from FY 1999 to FY 2002 from FMCSA, NHTSA, and states not sampled. The reports and data assisted us in the analysis of data consistency and trends.

Analysis of Variation. Because SafeStat is a relative ranking system, problems with untimely data would not seriously impact the rankings if the data were consistently late across all states. Each carrier's ranking would just be relatively higher. However, geographic bias is introduced in the rankings to the degree that untimely data varies from place to place. We analyzed variations in the timeliness of reported crash and inspection information through the use of the coefficient of variation, a basic measure of dispersion. The coefficient of variation is obtained by dividing the standard deviation by the mean value of the data set, then multiplying that result by 100.

This single number provides an indicator of the consistency of the data. When the value is 25, this indicates a normal or standard distribution for the variation. When the value is 60 or 70, the variation is high, suggesting biased outliers in the

data. Table B3 shows how timely states were in reporting inspection and crash data to FMCSA from FY 1999 through FY 2002.

Table B3. Variability in Timeliness of Data Reporting Across States (FY 1999 through FY 2002)		
Indicator	Coefficient of +Variation	Interpretation
Timeliness of Reporting Crashes	76.58	High Variation
Timeliness of Reporting Inspections	90.59	Extreme Variation

The variation across states for timeliness is high or extremely high, for each of the indicators above. The variability within regions[2] is less than across states. This indicates the impact of the variations may be less for carriers that operate within one geographic region. Conversely, the impact for carriers operating across regions could be greater. However, we did not attempt to estimate the impact of reporting variation based on patterns of operation.

DATA QUALITY CONTROL SYSTEMS

To determine whether FMCSA data quality control systems were adequate for their intended uses, we reviewed guidelines issued by DOT on the dissemination of data and compared the guidelines to systems in place or planned by FMCSA. We also observed quality control systems in the states we visited and assessed selected areas, such as data entry controls, correction of erroneous data, NHTSA and FMCSA cooperation to improve data, standardization of reporting, and concerns over public availability of data.

AUDIT PERIOD AND ASSESSMENT OF COMPUTER-GENERATED DATA

The audit was conducted from November 2002 through January 2004 in accordance with Government Auditing Standards prescribed by the Comptroller General of the United States and included tests of internal controls as were considered necessary.

---

[2] For this analysis, a region refers to areas sharing common attributes. A typical region consists of about four to six contiguous states.

In the conduct of this audit, we used computer-generated data from SafeStat system developers at Volpe, the FMCSA Motor Carrier Management Information System, the 10 states visited, and NHTSA. We did not assess the general and application controls for each of the automated systems. We conducted selected tests of the data that included checks to ensure the data provided from FMCSA's database were complete and data verification procedures were in place at Volpe. In our opinion, when the data are viewed in the context with other available evidence, the results and conclusions in the report are valid.

EXHIBIT C. PRIOR AUDIT COVERAGE

One Office of Inspector General (OIG) audit, issued in 1999, and two General Accounting Office (GAO) audits, issued in 1999 and 1997, have addressed topics related to SafeStat.

OIG Report: *Motor Carrier Safety Program, Federal Highway Administration*, Report Number TR-1999-091, April 26, 1999. We found that the existing organization, Office of Motor Carrier (OMC), was not sufficiently effective in ensuring that motor carriers complied with safety regulations. The existing enforcement program did not adequately deter noncompliance. Additionally, SafeStat could not target all carriers with the worst records, because OMC's database was incomplete and inaccurate, and data input was not timely. The report made the following recommendations for improving the data used to regulate the motor carrier industry:

- Require applicants requesting operating authority to provide the number of commercial vehicles they operate and the number of drivers they employ and, require all motor carriers to periodically update this information.

- Revise the grant formula and provide incentives through Motor Carrier Safety Assistance Program (MCSAP) grants for states to provide accurate, complete, and timely commercial vehicle crash data, vehicle and driver inspection reports, and traffic violation data.

- Within a reasonable notification period, such as 1 year, withhold MCSAP grant funds from those states that continue to report inaccurate, incomplete, and untimely commercial vehicle crash data, vehicle and driver inspection data, and traffic violation data.

- Initiate a program to train local enforcement agencies in reporting crash roadside inspection data including associated traffic violations.

- Standardize OMC and NHTSA crash data requirements, crash data collection, and reports.

- Obtain and analyze crash causes and fault data as a result of comprehensive crash evaluations to identify safety improvements.

GAO Report: *Truck Safety: Motor Carrier Office Hampered by Limited Information On Causes of Crashes and Other Data Problems*, RCED-99-182,

Exhibit C. Prior Audit Coverage   44

June 29, 1999. GAO found that OMC has not been effective in reducing fatalities resulting from crashes involving large trucks because it knows too little about the causes of crashes or the factors that contribute to them. The report noted that OMC had not corrected longstanding problems with the information it uses, such as information that identifies high-risk carriers. In addition, no reliable nationwide data existed on the causes of crashes involving large trucks.

GAO Report: *Commercial Motor Carriers: DOT is Shifting to Performance–Based Standards to Assess Whether Carriers Operate Safely*, RCED-98-8, November 3, 1997. GAO found that the new SafeStat system was designed to better identify problem carriers, but it was dependent on the states to provide complete, accurate, and timely data on recordable accidents and the results of roadside inspections and compliance reviews. GAO observed that some states lacked adequate data, particularly for accidents. Such gaps in the reported data, according to the report, can change a carrier's score, thus affecting SafeStat's reliability. GAO recommended the development of alternative approaches in states with inadequate data, but the Department disagreed because the development of separate processes for different states would be ineffective and impractical.

Exhibit C. Prior Audit Coverage 45

EXHIBIT D. MAJOR CONTRIBUTORS TO THIS REPORT

THE FOLLOWING INDIVIDUALS CONTRIBUTED TO THIS REPORT.

Name	Title
Barbara M. Cobble	Program Director
Joseph W. Come	Project Manager/Program Director
James W. Bess	Senior Management and Program Analyst
David W. Brown	Senior Management and Program Analyst
Alvin B. Schenkelberg	Senior Auditor
John M. Hannon	Senior Management and Program Analyst
Richard Hatcher	Auditor
Calvin L. Moore	Management and Program Analyst
Christopher T. Brothers	Management and Program Analyst
Constance B. Hardy	Management and Program Analyst
Petra Rose	Statistician
William E. Savage	Information Technology Specialist
Dr. Francis M. Ponti	Consultant
Harriet E. Lambert	Writer/Editor
Janice Alger	Senior Auditor
Wayne White	Auditor

APPENDIX. MANAGEMENT COMMENTS

A draft of this report was provided to FMCSA on December 10, 2003. In its response, FMCSA provided technical and editorial comments on the draft report as well as substantive comments on the report's findings. The substantive comments by FMCSA on the draft report's findings are provided verbatim below. (FMCSA's specific references to the draft report are shown in italics prior to its comments.) The Appendix also includes the full text of a matrix provided by FMCSA on January 14, 2004 detailing actions planned or underway to address each report recommendation. Where appropriate we made modifications in this final report to reflect FMCSA's comments.

Executive Summary, page ii: The report states: "SafeStat is also used by some states to link carrier safety performance to vehicle registration, and as a decision aid by State commercial vehicle safety inspectors to select vehicles for inspection. Public uses for SafeStat include providing information to individuals making contract award or acquisition decisions and monitoring the performance of drivers within a company."

Comment: While SafeStat was developed and tested during the Performance and Registration Information Systems Management (PRISM) pilot, SafeStat itself is not used to "link carrier performance to vehicle registration" sanctions. We are concerned that some State motor carrier associations have the mistaken impression that PRISM States will take registration action on a carrier's International Registration Plan plates, based solely on its SafeStat score. As such, some State motor carrier associations have been reluctant to support PRISM. We request that the language in the report be clarified to avoid these potential misconceptions.

Comment: The statement ". . . and monitoring the performance of drivers within a company" might give the impression that individual driver performance data (i.e. driver names, etc.) is available online to the public. Because that is not the case, we request that the report language be clarified.

Executive Summary, page iii: the first bullet under "Results in Brief" highlights the apparent fact that 42 percent of the 643,000 active carriers have not updated their census data.

Comment: While we share your concern for out-of-date data, we believe that the language in the report tends to overstate the problem—in relation to SafeStat. Therefore, it would be preferable to present the data problems in terms of how they affect SafeStat rankings. Please note that, for carriers that have generated data sufficient to be evaluated by SafeStat, over 80 percent have updated data.

Comment: On the issue of zero vehicles/zero drivers, since the implementation of "new MCMIS" (September 2002), carriers cannot obtain new DOT numbers with entries of zero vehicles or zero drivers. Also, the Bienniel Update process allows us to identify and correct errors on existing carrier data, and we continue to explore additional ways to improve the accuracy of these records.

Executive Summary, page v: Under "Data Quality Weaknesses" the report states: "FMCSA should take action to improve data quality because significant problems exist with the data motor carriers and the states provide, and these data problems limit SafeStat's effectiveness and introduce bias into the ranking process.

Comment: We share your concern that data quality problems could limit SafeStat's effectiveness to identify all high-risk carriers. However, we believe that the existing language in the report could be misconstrued to imply that some identified high-risk carriers may not actually be high-risk because of the data problems. We cannot agree with this implication and we request that the language in the report be clarified.

Executive Summary, page vii: Under the bullet: "Ineffective Systems for Facilitating Data Correction by the States." The report states: "Accuracy challenges will be increasingly important because SafeStat scores are now calculated and released monthly, versus the earlier practice of releasing scores semiannually."

Comment: We are concerned that this implies that monthly update of SafeStat scores makes an improved data correction process more urgent. Rather, we believe monthly update of SafeStat scores should be recognized as a data quality improvement. Now, when a carrier updates its census data and gets an inaccuracy corrected, its SafeStat score will show an improvement in 30 days or less, instead of up to 6 months later.

Executive Summary, page viii: In the paragraph preceding the "summary of recommendations," the report states: "Correcting data quality problems are critical to ensure the effective targeting by SafeStat of high-risk carriers for compliance reviews."

Comment: Based on information in other parts of the report that conclude SafeStat is effective for targeting high-risk carriers for compliance reviews, we suggest the report language be slightly modified as follows: "Correcting data quality problems are critical to ensure <u>more comprehensive targeting</u> by SafeStat of high-risk carriers for compliance reviews."

Page 8, second paragraph: The report states: "Although SafeStat's original development effort considered the views of industry stakeholders, such views have

not been solicited for subsequent changes, such as those affecting how power unit numbers are calculated."

Comment: The FMCSA's Analysis & Information online site that displays SafeStat data to the public has a feedback mechanism whereby the industry and public can offer suggestions on improvements. In fact, the majority of SafeStat changes, including the change related to how power units are calculated, have originated from industry and other stakeholder feedback.

Page 12, middle paragraph: The report states: "When motor carriers and the States provide insufficient data it creates an unknown degree of bias in SafeStat's ranking of motor carriers and limits SafeStat's effectiveness as a tool for identifying the highest risk carriers."

Comment: While we share your concern for improving data quality, we did not note any evidence in the report to substantiate the assertion that the data problems limit SafeStat's effectiveness as a tool for identifying the "highest risk carriers." We request that the statement be modified to read, ". . . limit SafeStat's effectiveness as a tool for identifying all high risk carriers."

Pages 13-14: Language regarding crash data issues and census data issues.

Comment: Please see our comment on Executive Summary, page v. In addition, a long-term solution is the above-noted PRISM program, which will eventually help increase data update frequency for active carriers that register through the International Registration Plan (IRP). Through PRISM, carriers are forced to update their MCS-150 data as part of their license plate renewal process unless it has been updated in the last year. Thirty-two states now have grant agreements to implement PRISM.

Page 24, last paragraph before the paragraph titled "Improving Data Quality Control Systems," the report states: "Errors due to "missing data" in which a carrier was not held accountable in SafeStat have a less obvious impact on individual carriers. However, these errors influence the entire system as one carrier's scores may be higher than they should be because missing data prevents carriers with more serious safety performance data from gaining a higher relative ranking. Overall, the results indicate that it is more likely that safe carriers would be identified as unsafe rather than unsafe carriers being identified as safe."

Comment: While we share your concern for improved data quality, we do not note any evidence in the report to substantiate the assertion that data problems make good carriers look bad. Rather, our information indicates that the data problems are much more likely to make a poor performing carrier look good.

[The following comment, not referenced to any specific draft report findings or recommendation, was also included in FMCSA's comments.]

In December 2003, FMCSA's ASPEN software was modified to collect better moving violation data. These changes greatly facilitate enforcement personnel's ability to select specific moving violation codes from a categorized list. Moreover, FMCSA recently provided a grant to the State of California to begin using the ASPEN software.

[January 14, 2004 Matrix from FMCSA Addressing Draft Report Recommendations]

OIG Recommendations	FMCSA Position
A1. Initiate an effectiveness study of the current model that tests the model's key parameters, assesses possible model adjustments to account for missing data, and evaluates whether the use of vehicle miles traveled or adjustments for team drivers would improve SafeStat calculations. The results of the study should be subjected to independent review by a party outside of Volpe.	Already planned. FMCSA has an independent analysis contractor on board to revalidate the SafeStat model. The study should begin in early 2004.
A2. Establish processes for soliciting public comment on proposed changes in SafeStat calculations, to include those changes, if any, resulting from the revised effectiveness study.	We believe FMCSA has an effective public comment process. FMCSA's Analysis & Information online site that displays SafeStat data to the public has a feedback mechanism whereby the industry and public can offer suggestions on improvements. In fact, the majority of SafeStat changes have originated from industry and other stakeholder feedback. FMCSA will explore the feasibility of using the existing feedback mechanism on A&I to solicit public comment on proposed changes.

B1. a. Take immediate action to mitigate the impact of inaccurate or incomplete data on public users of SafeStat by making available to all States an improved system for facilitating the correction of data inaccuracies and the tracking of corrective actions within 3 months following the issuance of this report.	Already underway and in progress. In late January 2004, FMCSA will implement the DataQs system, an electronic means for filing concerns about Federal and State data released to the public by FMCSA. Through this system, data concerns are automatically forwarded to the appropriate office for resolution. This system will simplify the existing process and provide a mechanism for FMCSA, the States, and the public to track data challenges. In addition, monthly update of SafeStat scores should be recognized as a data quality improvement. Now, when a carrier updates its census data and gets an inaccuracy corrected, its SafeStat score will show an improvement in 30 days or less, instead of up to 6 months later.
B1. b. Take immediate action to mitigate the impact of inaccurate or incomplete data on public users of SafeStat by disclosing data problems, including variations in State reporting, to public users of SafeStat.	FMCSA will add clear and comprehensive guidance for web users as to the limitations of the data by the end of January 2004. We will further enhance this guidance by adding information on each State's timeliness, accuracy, and completeness.

B2. a. As expeditiously as possible, establish an overall data quality improvement program that addresses longstanding issues associated with motor carrier census data by imposing fines on carriers that fail to provide updated carrier census information despite repeated opportunities to do so.	Already in progress. Beginning January 1, 2004, all reminder notices sent to motor carriers about filing their biennial update will contain the following statement: "Even if the information has not changed or your company is no longer in business (or no longer operating commercial vehicles in interstate commerce), we need you to file this update. Please note that, under 49 CFR 390.19(e), failure to file a new and accurate Form MCS-150 may result in the imposition of a civil penalty of up to $550 for each offense. Each day the violation continues constitutes a separate offense, and the total penalty for all offenses related to a single violation may reach $5,500." In January 2003, to help improve the response rate of the bienniel update, FMCSA provided an easy way for carriers to update their census data online. Currently, 40 percent of carriers are providing updates in this manner. We also developed a plan for addressing the missing respondents and are weighing implementation options. In addition, FMCSA's New Entrant program meets with each new motor carrier within the first 18 months of operations and conducts a safety audit. In this audit, census data is verified and updated. If the new entrant fails to pass the safety audit, the USDOT number is revoked.

B2. b. As expeditiously as possible, establish an overall data quality improvement program that sets minimum standards for the quality of SafeStat data consistent with the Department's data quality guidelines. At a minimum, the standards should address completeness, accuracy, and timeliness of data.	While we agree with OIG's concern for data quality, we cannot agree with all of your assertions as to the impact of data problems on SafeStat. FMCSA is undertaking a continuous data quality process, including developing a comprehensive data quality plan in FY 2004. Some of the planned components of this process include: (1) awarding a contract to conduct a sample verification of source documents (inspections and crashes) against what is uploaded; (2) coordinating with OIG to identify additional edit checks, (3) conducting training sessions at FMCSA's IT Workshop that specifically address data quality issues, (4) assigning a person in FMCSA to review the monthly timeliness, non-match, and traffic enforcement reports by State to identify problems and work with the States to resolve them, and (5) review recently completed analysis on improving the biennial update response rate. FMCSA has set data quality goals for timeliness, completeness and accuracy of crash data as part of the Commercial Vehicle Analysis Reporting System (CVARS) program.
B2. c. As expeditiously as possible, establish an overall data quality improvement program that accomplishes actions planned, in conjunction with NHTSA, for improving the completeness and timeliness of state-reported crashes.	Already underway. In early Summer 2003, FMCSA provided grants to 22 States under the Commercial Vehicle Analysis Reporting System (CVARS) program, to improve their crash data reporting. We recently found that 11 of these States had shown a marked increase in the completeness of crash data reporting. We continue to include more States in CVARS, and have established goals for accuracy, timeliness, and completeness. Through training and data system improvements, we expect the States to make significant strides in the quality of crash data reporting.

B2. d. As expeditiously as possible, establish an overall data quality improvement program that enhances the depth, frequency, and type of FMCSA State data quality reviews and monitoring, and ensures State plans address data quality.	While we agree with OIG's concern for data quality, we cannot agree with all of your assertions as to the impact of data problems on SafeStat. As noted above, part of our continuous data quality process will involve assigning a person in FMCSA to review the monthly timeliness, non-match, and traffic enforcement reports by State to identify problems and work with the States to resolve them.
B2. e. As expeditiously as possible, establish an overall data quality improvement program that modifies FMCSA guidance and funding decisions so that MCSAP incentive grant awards are based, in part, on each State's implementation of guidelines established to provide accurate, complete, and timely safety event data.	While we agree with OIG's concern for data quality, we cannot agree with all of your assertions as to the impact of data problems on SafeStat. In addition, we cannot agree with the specific approach of penalizing States for data problems, because we believe that withholding funds may lead to a reduced level of enforcement. In the alternative, we believe that providing incentive funding to States will assist them in improving data quality reporting processes. In fact, receiving high priority MCSAP grants will be tied to participation in the DataQs system. This information was sent out in mid-January to each MCSAP agency.

Textual Translation of Figures in the Report

A text only description of figures in the report is provided below.

Estimate of Underreporting of Large Trucks Involved in Crashes by State (2001 Data)

Figure 1, located on page 17 of the Report, provides a map showing estimates of the underreporting of large trucks involved in crashes by state (based on 2001 data). The map indicates the following.

In the states of Arkansas, District of Columbia, Florida, Georgia, New Hampshire, New Mexico, Rhode Island, Tennessee, and Virginia, it is estimated that 60 percent or more of large trucks involved in crashes will not be reported to the Federal Motor Carrier Safety Administration (FMCSA).

In the states of California, Iowa, Louisiana, Maine, Nevada, North Carolina, South Dakota, Vermont and West Virginia, it is estimated that between 40 and 59 percent of large trucks involved in crashes will not be reported.

In the states of Alabama, Colorado, Idaho, Illinois, Indiana, Kansas, Maryland, Massachusetts, Michigan, Mississippi, Montana, Oregon, North Dakota, Nebraska, New York, Oklahoma, Pennsylvania, South Carolina and Texas, it is estimated that between 20 and 39 percent of large trucks involved in crashes will not be reported.

In the states of Arizona, Delaware, Kentucky, Minnesota, Missouri, New Jersey, Ohio, Rhode Island, Utah, Washington, Wyoming, and Wisconsin, it is estimated that less than 20 percent of large trucks involved in crashes will not be reported.

The overall estimates of 50-percent for the Southern and 20-percent for the Midwestern regions consist of the following states. The Southern region includes Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, and Texas. The Midwestern region includes Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio, and Wisconsin.

Rate of Reporting for Serious Moving Traffic Violations Varies Across the States (2001 Data)

The information below provides the textual translation for Figure 2, located on page 22 of the Report. The information on variances in reporting for serious moving traffic violations provided in Figure 2 is the following.

The chart shows that the average number of serious moving traffic violations per million commercial vehicle miles reported by each state is .84.

The chart indicates there is a wide discrepancy in the reporting of serious moving traffic violations per million commercial vehicle miles number reported from state to state. The chart divides the states into groups alphabetically with the first group representing the first six states Alabama, Alaska, Arizona, Arkansas, California, and Colorado. The states in this group reported at rates of 3.87, .07, .8, 1.24, 0, and 1.28 violations per million commercial vehicle miles as compared to the average of .84.

The second group, states 7 through 11 reported at rates of .56, .41, .03, .52 and .82 as compared to the average of .84.

The third group, states 12 through 16 reported rates of 0, .89, 0, 2.56 and .95 as compared to the average of .84.

The fourth group, states 17 through 21 reported at rates of 2.45, 1.5, 2.35, .49, and .66 as compared to the average of .84.

The fifth group, states 22 through 26 reported at rates of 1.75, 1.39, .63, .24, and .65 as compared to the average of .84.

The sixth group, states 27 through 31 reported at rates of 3.62, 1.23,.19, .18 and .18 as compared to the average of .84.

The seventh group, states 32 through 36, reported at rates of 2.39, .17, .81, 1.2, and .6 as compared to the average of .84.

The eighth group, states 37 through 41, reported at rates of .47, 2.84, .25, 1.62, and 1.24 as compared to the average of .84.

The ninth group, states 42 thru 46, reported at rates of .6, 2.25, .04, .18, and .72 as compared to the average of .84.

The tenth group, states 47 thru 51, reported at rates of .15, 3, .83, .79, and 1.55 as compared to the average of .84.