PB98-127194

Driver/Carrier Data Relationship Project

Phase II Report

February, 1997

Prepared for:

United States Department of Transportation
Federal Highway Administration

by

AAMVAnet, Inc.
Keane Federal Systems

AAMVAnet inc.

KEANE Federal Systems, Inc.

For more information about this report, please contact:

Paul Alexander
Office of Motor Carriers
Federal Highway Administration
US Department of Transportation
400 Seventh Street, S.W.
Washington, D.C. 20590

Management Summary

This document is the Phase II report for the Driver/ Carrier Data Relationship Project.

The goal of Phase I was to determine if it could be demonstrated that there was a relationship between a carrier and the citations that were received by the carrier's drivers. Five States, California, Idaho, Indiana, Michigan, and North Dakota included carrier identification on their driver citations. Four of those States, Idaho, Indiana, Michigan, and North Dakota were able to provide copies of their driver citation files. Of these four States, two were also able to produce exposure data. Idaho provided ton-mileage tax data, and Indiana provided IRP mileage.

Based on the analysis of this data, Phase I clearly demonstrated that violations were not randomly distributed among carriers.

Phase II had three goals. The first was to validate the results of Phase I. The second was to determine if the difference in carrier violation rates could be related to a difference in safety performance. The third was to seek a way to use this information in a real environment to target potentially unsafe carriers.

For Phase II, both Indiana and Michigan provided additional driver violation and IRP mileage data. Further, FHWA provided data from its Motor Carrier Management Information System (MCMIS), including carrier crash history, SCE score and safety rating.

Each of the three project goals was achieved.

- Using the new data, the Work Group obtained results similar to the results from Phase 1. Clearly, driver violations were not randomly distributed.
- There was a demonstrable relationship between the violation rate of the carrier's drivers' and the crash history of the carrier. To some extent, the difference in driver violation rates was indicative of a difference in carrier safety performance.
- Using the carrier identification on the citation along with data that was available from the citation records, the Work Group developed a statistic that was capable of identifying potentially problem carriers for further review and scrutiny. The statistic is called Carrier's Drivers' Violation Rate (CDVR).

It has been demonstrated that there is public policy benefit in collecting data about the carrier on driver citations. This data can be used to identify potentially unsafe carriers. Further, there is benefit in collecting data about a driver's relationship to a carrier any time that the data is presented to law enforcement. Finally, there is an implication that there are a wide variety of possible next steps which could be taken to improve the data collection, and which could be used to target potentially unsafe motor carriers for further scrutiny.

Table of Contents

Chapter 1 - Introduction

A. Background

1. Safety Problems That Lead to the Project

The deregulation of the trucking industry in the late 1970s affected the economic regulation of the trucking industry. However, as a part of that effort, the role of the US Department of Transportation, and particularly the Office of Motor Carriers (OMC) , in the assurance of safety of the nation's highways, became even more critical.

OMC responded to the Reagan era efforts to return responsibility to the States with the Motor Carrier Safety Assistance Program (MCSAP). Through an emphasis primarily on vehicle safety, the MCSAP program has been a key factor in the overall nationwide increase in truck safety.

It remains part of the mission of the OMC to continually search for ways to improve truck safety on the nation's highways.

The Motor Carrier Act of 1991 specifically addresses issues related to motor carrier safety. The relationship between safety acts committed by a commercial vehicle driver and the carrier employing that driver are part of the title IV. A steering committee was formed which later created a Project Work Group. This Work Group organized and began this study.

2. Project Sponsors

This report has been prepared by AAMVAnet, Inc. under contract to the Federal Highway Administration (FHWA) for the Driver Citation/Carrier Data Relationship Project study. Mr. Paul Alexander is the FHWA Project Manager for this study.

This report is based on information provided by five Driver Citation/Carrier Data Relationship Project participant States; California, Idaho, Indiana, Michigan and North Dakota. This report also includes a statistical analysis that used information from FHWA's Motor Carrier Management Information System (MCMIS).

3. Organization of the Project

The Driver/Carrier project was organized in two phases. Phase I, which ran from 1993 through 1994, addressed the question of whether there was a difference in the rate that drivers receive citations based on the carrier that was employing the driver. Phase I findings determined that the data strongly suggested that there was a difference in the rate that drivers received citations, and that a piece of the difference was based on the carrier that was employing the driver.

There were two documents issued during Phase I. The first, titled the *Driver/Carrier Summary Analysis Report*, was produced in June 1994. The second, titled the *Driver/Carrier Statistical Analysis Report*, was produced in September 1994.

Phase II of the project started at the end of Phase I and ran from 1995 through 1996. Phase II had three project goals:

- Revalidate the Phase I results.
- Determine if the difference in carrier violation history is associated with a difference in safety performance, and
- Identify appropriate measures for use in identifying potential problem carriers.

This report represents the end of Phase II of the project.

4. Goals of The Report

The goal of this report is to address the three goals stated in the previous section. This report will discuss the overall findings and conclusions of the project research and analysis. The report will also suggest recommendations for moving forward. The various methods used for the research and analysis will be discussed as well as the States' participation and information. The conclusions show the benefit of collecting carrier information.

B. Description of Project Activities

1. Data Preparation and Description

For analysis, this project used citation, mileage, "snapshot" and crash data. Citation data for the Phase I analysis was provided by the States of California, Idaho, Indiana, Michigan and North Dakota.

Phase II used citation data from Indiana and Michigan. This data represented citations on which the motor carrier could be identified that were issued by the State Police during the time periods studied. Both Indiana and Michigan provided International Registrant Plan (IRP) mileage data. Snapshot information, which included current and immediate prior Safety Rating, SCE rating and Motor Carrier Safety Improvement Process (MCSIP) score was received from FHWA. In addition, FHWA provided crash data.

2. Summary of Tools and Methods Used

a. Software Used

The software used to manipulate the data collected was Microsoft Visual FoxPro. Visual FoxPro is a relational database package that can manipulate large amounts of data quickly. It also has built-in statistical functions and a report writer. Databases were built from the data received from the States. Additionally, special queries, calculations and reporting were performed on the data. Microsoft Excel and Borland Quattro Pro were also used for some of the statistical analysis.

The software chosen for use during the project did not drive the results of the project. Many software packages provide the same functionality.

b. Statistical Methods Used

The statistical methods used included standard deviations, Z-scores, Chi-squared statistics, coefficients of correlation, rankings and several statistically based distributions. Appendices B and C contain explanations of the statistical methods.

The project data was summarized into logical subsets to aid in the statistical analysis.

3. States' Participation

Idaho, Indiana, Michigan and North Dakota participated in Phase II of the project and formed the core of the Work Group. Indiana, Michigan and FHWA provided the data that was used in this phase. A number of Work Group meetings were held during the course of the project so that the States and the group as a whole could exchange ideas.

C. Organization of Report

The report is organized into a management summary and seven chapters. The chapters tend to be written as stand alone analyses. The reader does not need to complete the prior chapters in order to read any of the later chapters. The first section of the report is the Management Summary.

The other sections are:

- Chapter 1 - The Introduction. This chapter presents background information, objectives and project activities, as well as the organization of the report.

- Chapter 2 - Summary of Phase I Reports. This chapter summarizes the results of the reports developed in Phase I of the project.

- Chapter 3 - Validation of Phase I Conclusions. The Work Group performed the same analysis as was done in Phase I using the additional data received from Indiana and Michigan. Through the same rigorous statistical analysis that was performed previously, the results of Phase I were revalidated. This met the first goal of Phase II.

- Chapter 4 - Violation Rates and Safety Performance. The second goal of Phase II was to demonstrate that, to some degree, the difference in carrier's drivers' performance can be related to a difference in carrier safety fitness. This chapter presents the statistical analyses necessary to demonstrate that the difference in carrier performance is related, to some degree, to a difference in carrier safety performance.

- Chapter 5 - Alternative Statistical Indicators. The third goal of Phase II was to determine if there is any way, based on the data that was collected, to identify potential problem carriers. The Work Group identified a potentially feasible method called the Carrier's Drivers' Violation Rate (CDVR). This section presents a non-technical explanation of the approach that was used to develop the CDVR, and how it could be useful.

- Chapter 6 - Validating the Carrier's Driver's Violation Rate. This chapter includes the results of the process that was performed to validate the use of the Carrier's Drivers' Violation Rate. This validation includes both expert review and statistical analysis.

- Chapter 7 - Phase II Conclusions. This chapter includes a review of the conclusions from Chapters 3-6, as well as overall project conclusions.

This report also contains 5 appendices. These are,

- Appendix A - Glossary of Terms and Acronyms

- Appendix B - Statistical Methods used in Phase I

- Appendix C - Statistical Methods used in Phase II

- Appendix D - The States' Comments

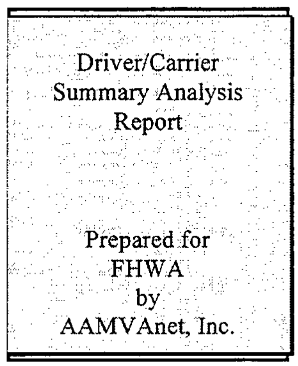
- Appendix E - Selected Reports and Specific Carriers Comments. This section will not to be released to the public.

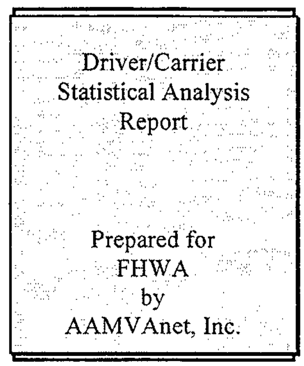
Lastly, the report contains an Acknowledgment section.

Section E has been packaged under separate cover, and is marked for Official Use only. It is the only portion of the report that contains information about specific carriers.

Chapter 2 - Summary of Phase I Reports

Two reports were produced in Phase I of the project. The first report, the *Driver/Carrier Summary Analysis Report,* was released in June, 1994. The second report, the *Driver/Carrier Statistical Analysis Report,* was released in September, 1994.





Phase I had two project goals.

- The first goal was to determine if a relationship existed between the violations that a driver received and the carrier that the driver was working for at the time of citation.
- The second goal was to demonstrate that the differences in carrier violation rates could not be explained by randomness.

The reports discussed the various tasks and activities that took place to complete Phase I of the project. The Phase I conclusions are also discussed.

Phase I concluded that there was a relationship between the violations that a driver receives and the carrier the driver was working for at the time of the citation. Violation rates greatly differed among carriers. The difference was more than exposure, it appeared that there were other factors.

A. The *Driver/Carrier Summary Analysis Report*

The Phase I *Driver/Carrier Summary Analysis Report* began with a Management Summary section. A summary of project events and the organization of the report was included in the Introduction Chapter. The report included a section to review the States' results. In addition to statistical observations and results, the participating States identified issues and considerations related to the data collection activity.

> **Management Summary**
>
> This report presents the Driver/Carrier Summary Analysis which was performed on the data collected by States that participated in the Federal Highway Administration (FHWA) Driver Citation/Carrier Data Relationship Project...
>
> ...The pilot States have recommended that continued research into the Driver/Carrier relationship is likely to bring significant results.

The use of citations instead of conviction data was discussed. There was some concern originally that the project data was non-adjudicated. After discussion, it was decided that citation data was more than satisfactory for statistical analysis. Further, it was recognized that, while citation data cannot be used as the basis for sanctions, it is entirely appropriate for use in identifying potential problem carriers for further scrutiny.

Additional topics in this section included the relationship of citations to violations, data collection, data accuracy, and program concerns. A review of the States' input data discussed topics such as: private versus for-hire carriers, hazardous material carriers, types of vehicles and crashes. This report considered carriers with overweight violations, carriers with moving violations and carriers' employment of problem drivers.

The next section was a review of the Multi-State data. This section covered distribution of violations by type, observations and conclusions about drivers with many violations, a review of carriers appearing in multiple States, drivers with many violations for multiple carriers, carriers with many moving violations in multiple States, differing Driver License Number States, and carrier identification issues.

1. Conclusions

The overall conclusions were based on the project's objectives:

- to determine whether there would be value in collecting carrier identification on driver citations, and

- to determine if it was likely that the value of collecting data would be sufficient to warrant the cost of adding carrier identification to citations.

The primary value of collecting carrier data on citations would be in the ability to more accurately determine which carriers were likely to have safety problems.

A critical issue was to determine if there was information in the data which was independent of exposure rates that suggested certain carriers were more likely to have safety problems. The States made several suggestions;

- Ratios of Types of Violations - There was reason to believe that a comparison of the types of violations that a carrier received could be meaningful. However, this was only valid on a State-by-State basis because States were operationally unique. Consequently the ratios of violations assigned varied by State. For example, North Dakota issues very few equipment violations (which are based on a vehicle inspection) during the winter months.

- Log Books - One State focused on log books. A driver who habitually drives over the speed limit will tend to falsify logbooks. This may be an indicator of a problem within the carrier.

- Any Violation - One participating State believed that any violation was a problem because there are carriers that drive many miles without any violations.

- Employment of Drivers with Bad Records - There exists a strong belief that a carrier who is unsafe in any area of its business may be unsafe in additional areas. A carrier who hires unsafe drivers or drivers with deteriorating driving records is engaging in an unsafe practice.

The overall conclusion was that there was strong reason to suspect that carrier identification on driver citations was of significant value, but that the findings did not represent conclusive proof.

2. Recommendations

The States recommended continued analysis of the data. The *Driver/Carrier Summary Analysis Report* recommended further clarification of data be collected, continued compilation of the data be performed, and creation of a structure to target carriers for review be developed. The States also continued and expanded their own use of this data.

B. The *Driver/Carrier Statistical Analysis Report*

The project team recognized that valid statistical results would require the use of violation rates, i.e. violations per some measure of exposure. A number of exposure indicators were reviewed. The project chose to use IRP mileage. Indiana provided a computerized listing of its 1993 IRP mileage records. These mileage records were matched to the carrier, based on name

The *Driver/Carrier Statistical Analysis Report* also explained the statistics used in Phase I of the project in detail.

1. Results

The primary finding of the *Driver/Carrier Statistical Analysis Report* was that the data that was studied supported the conclusion that:

> ***"Drivers for some carriers received violations at a rate significantly higher than the overall average, and the difference was far more than could be explained by random chance."***

This report covered the general statistical approach. Appendix B of this report contains a full explanation of the statistical methods used in Phase I. Hypothesis testing was used extensively in Phase I. The first step was to state a "null" hypothesis, i.e. that the expected condition did not exist. It was then demonstrated that the data did not support the null hypothesis. The hypotheses tested included:

- Distribution of violations among carriers. Stated in the null form, it is:

  "Violations are randomly distributed among the entire population of carriers."

  This turned out not to be the case. **Violations are not randomly distributed among all of the carriers.** Violations per mile are far from randomly distributed among the carriers in the population.

- Distribution of violations among carriers which have violations. Stated in the null form is:

  "Violations are randomly distributed among the carriers which have violations."

  This also turned out not to be the case. **Violations are not randomly distributed among the carriers which had violations.** Among the carriers which had violations, some carriers had much higher violation rates than others.

- Distribution of violations by violation type within groups of interest. Stated in the null form, it is:

  "Carriers with more than 40 violations receive the same proportion of violations by type as the overall population does."

  Again, this turned out not to be the case. **Carriers with more than 40 violations do not receive the same proportion of violations by type over all carriers.** Certain carriers appear to have more of certain types of violations than other carriers.

2. Methodology

This section presents an overview of the methodology that was used in the Phase I statistical analysis. A more detailed presentation of the methodology is included as Appendix B of this report.

General hypothesis testing and related statistical methods were explained. This included a review of the statistical tests;

- comparing actual to expected results,
- a general view of the expected distribution and
- the bell curve.

Basic statistical building blocks, including mean, standard deviation, the Z-statistic, and the Poisson distribution, were explained. These statistical tools were applied to the hypotheses tested. The Work Group developed generalized software to execute these procedures. The software was used to perform consistent analyses on different sets of data.

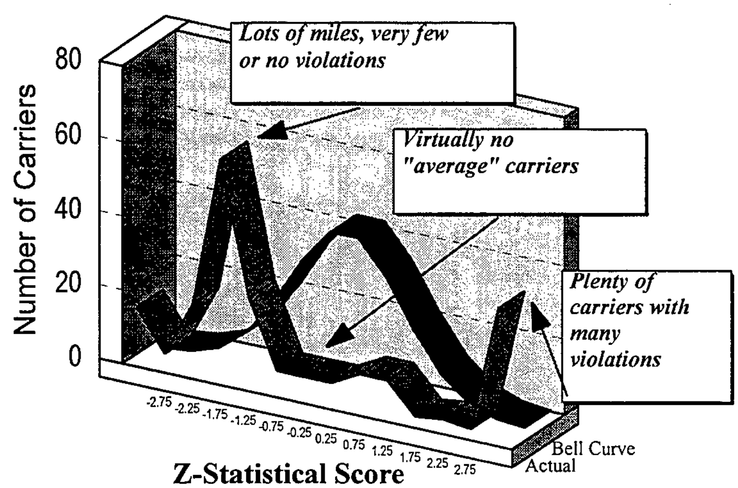
The report results which included, testing on all carriers which had mileage and several subsets, also included the frequency distribution of the Z-Score and Poisson distribution tests. Among the conclusions;

- Carriers that had violations and reported miles were tested. Violations were found not to be distributed randomly among the carriers with mileage data. Violations were also found not to be distributed randomly among the carriers with mileage and violations.

- Carriers that had specific types of violations were also tested. A lack of randomness was discovered within specific types of violations.

- Significant testing was then performed on the distribution of violations by violation type within certain subgroups of interest. Within these subgroups, carriers did not have a random distribution of certain violation types.

Indiana Carriers
With Greater Than 750K IRP Miles



The preceding graph presents the results for Indiana carriers with over 750,000 IRP miles in the year 1993. There are a large number of carriers with many miles and very few violations. Also, there are over 30 carriers with an exceptionally large number of violations.

The bell curve presents a random distribution. If violations were randomly distributed, the actual results would approximate the bell curve. Clearly, violations are not randomly distributed among this population.

3. Conclusions

The key conclusion reached in the *Driver/Carrier Statistical Analysis Report* was that within the data used for the project, violations were not distributed randomly among carriers. Further, within many subgroups of carriers which had violations, the violations were not distributed randomly. The data very strongly suggested that there were significant groups of carriers which may have;

- lower rates of violations than the norm,

- higher rates of violations than the norm, and
- problems with specific types of violations.

Potential data deficiencies were discussed at the end of the *Driver/Carrier Statistical Analysis Report*. Five data concerns were identified and dealt with within the framework of the analysis.

- The data was aged. The data for the mileage was actual data for the year prior to the violations.
- Mileage data was for registrants. The registrant is not always the carrier. The violation data is for the carrier.
- The identification depended on name matching. Manual name matching was used to match IRP registrants to carriers.
- National and regional companies will have only registered a portion of their miles in a specific State. It is possible that the data represented some violations for vehicles that were registered in other States.
- While there is no reason to suspect carrier fraud or misrepresentation, it is always a possibility.

It was the view of the Work Group that none of these factors had enough impact to fully discount the results.

4. Recommendations

The *Driver/Carrier Statistical Analysis Report* recommended that the following next steps be taken.

- Continue to collect data about violations in order to produce a more convincing statistical argument that violations are not distributed randomly among carriers.
- Continue to collect data about violations in order to identify appropriate targets for compliance reviews.
- Based on the collected data, target individual carriers as potentially having problems. Perform Compliance Reviews on these carriers and analyze the results.

Chapter 3 - Validation of Phase I Conclusions

The primary goal of Phase I of the Driver/Carrier project was to determine if it could be demonstrated that some carriers were significantly more inclined than others to get some or all types of violations. The Work Group developed a test plan (which is described in Appendix B of this document) that was designed to demonstrate that violations were not randomly distributed among carriers. The Work Group then executed that plan. The results were included in the *Driver/Carrier Statistical Analysis Report.*

In Phase II of the project, the Work Group analyzed an additional year's data from Indiana as well as a complete data set from Michigan. The Work Group reviewed the statistical tests from Phase I which had led to the conclusion that violations were not randomly distributed among carriers, and ran those same tests with the new data.

The overall conclusion was the same. Among the carriers represented in additional State data, violations were not randomly distributed among carriers.

The remainder of this section is devoted to reporting the results of repeating the Phase I tests with the new data. The reader may refer to Appendix B - Statistical Methods Used in Phase I, for a description of how this methodology was developed.

A. Poisson Distribution Tests

In any group of carriers which travel approximately the same mileage during the year, if violations are randomly distributed, the actual distribution of violations can be expected to be described by a Poisson Distribution. A Poisson Distribution describes the distribution of low probability occurrences in a large number of events.

After the IRP registrants were matched to the carriers in the violation list, the resulting data was analyzed.

In Phase I of this project, the distribution of violations for registrants (carriers) with between 40,000 and 60,000 miles, and for registrants with between 60,000 and 80,000 miles, were calculated. If violations had been randomly distributed among these groups, it would have been expected that the distributions would have been a Poisson Distribution. In Phase II, the same analysis was performed for the data from Michigan, and the data from Indiana for 1994.

1. Michigan Carriers with Between 40,000 and 60,000 IRP Miles

Number of Violations	Number of Carriers
0	462
1	19
2	12
3	10
4	3
5	10
6	1
7	3
8	2
9	1
11	1
13	4
18	1
Total	531

Within the data from Michigan, there were a total of 531 carriers with between 40,000 and 60,000 IRP miles. The distribution of violations is presented in Table 3-1.

In Phase I, the Poisson calculation cut off at a natural breaking point. For carriers with between 40,000 and 60,000 miles, the natural break was that there were no carriers with 5 violations. This was not the case in the new data.

Table 3-2 shows the Poisson Distribution for the 525 carriers with 9 or fewer violation, eliminating the 3 highest violation totals.

Table 3-1

Michigan Carriers with Between 40,000 and 60,000 IRP Miles and Less than Ten Violations

Vio. Count	Poisson Factor[1]	Expected Value	Actual %	Actual Count	Actual as a % of expected	Contribution to Chi Squared
0	67.3%	352.06	88%	462	131%	34.34
1	26.6%	139.34	4%	19	14%	103.93
2	5.3%	27.58	2.29%	12	44%	8.80
3	0.70%	3.64	1.91%	10	275%	11.13
4	0.07%	0.36	0.57%	3	833%	19.36
5	0.005%	.028	1.91%	10	35094%	3,489
6	0.0004%	.0019	0.19%	1	53200%	530
7	0.0000%	.00001	0.57%	3	2822672%	84,674
8	0.0000%	.0000002	0.38%	2	38035621%	760,708
9	0.0000%	0.00	0.19%	1	432448472%	4,324,482
Chi Squared Statistic						5,174,062

Table 3-2

[1] Poisson Factor represents the expected percentage of carriers to have this number of violations, based on the overall number of violations and carriers, if violations were distributed randomly.

These numbers speak for themselves. With 9 degrees of freedom, the total of the Chi Squared Statistic must be less than 19.7[2] for the Work Group to not reject the hypothesis that violations were not randomly distributed. Clearly, the hypothesis that violations are randomly distributed within this group can be rejected.

Over 5% of the carriers in this group, 28 of 531, had the entirely unexpected result of 4 or more violations. This lead to two possible conclusions;

- It is possible that some of these carriers had mileage in the State of Michigan that was not reported through their IRP accounts. For example,
  - They could have been employing owner/operators who carried their own IRP.
  - They could have leased the vehicles, and used the lessor's registration. (However, these were all carriers which did have a significant number of registered miles.)
  - They could have been national carriers with a small local presence (i.e. a small local terminal). In this case, all of the national carrier's violations would have been assigned to the small local base of miles.
- The second possible conclusion is that violations were not randomly distributed, and that some carriers were more likely to have received violations than others.

Because there was such a large population of Michigan carriers with between 40,000 and 60,000 IRP miles and five or more violations, the results of these carriers overwhelmed the test statistic. In order to eliminate this effect, the Poisson Distribution test was re-run with only the carriers which had four or fewer violations.

The results are presented in Table 3-3.

[2] Statisticians prefer to do Chi-squared analysis with expected values in the range of 5 or more when possible. The presence of cells with expected values of less than five brings in additional technical considerations, which can modify the expected Chi Squared. However, the impact of these considerations is minimal, especially as compared to the overwhelming results.

Michigan Carriers with Between 40,000 and 60,000 IRP Miles and Less than Five Violations

Violation Count	Poisson Factor	Expected Value	Actual Percent	Actual Value	Actual as a Percentage of Expected	Contribution to Chi Squared
0	81.2%	411	91.3%	462	112%	6.28
1	16.8%	85	3.7%	19	22%	51.55
2	1.7%	9	2.3%	12	136%	1.12
3	.12%	.61	1.9%	10	1633%	143.92
4	.006	.03	.6	3	9444	277.35
Chi Squared Statistic						480.22

Table 3-3

An inspection of the actual as compared to expected Poisson Distribution of the carriers with four or fewer violations suggested another interesting phenomenon. Even among this group, there were far fewer carriers than expected with only one violation. Of the Michigan carriers with between 40,000 and 60,000 miles;

- 87% had no violations,
- less than 4% had one violation (as compared to the expected value of 16%) and
- over 9% had two or more violations (as compared to the expected value of less than 2%).

The Poisson Distribution suggests that far more carriers will have one violation than will have two or more violations. It would be possible that this could be caused by drivers getting more than one violation in each incident. The actual data was reviewed, and it did not support such an assertion.

> ***"Clearly, at the carrier level, getting the first violation is a strong indicator that the carrier will receive subsequent violations."***

2. Michigan Carriers with Between 60,000 and 80,000 IRP Miles

Number of Violations	Number of Carriers
0	228
1	20
2	9
3	5
4	2
5	4
6	3
7	2
8	0
9	4
10	1
11	1
12	1
14	1
19	1
Total	283

Table 3-4

The Michigan IRP data showed 283 registrants with between 60,000 and 80,000 IRP miles. Table 3-4 presents the distribution of violations for the carriers (which could be identified) among these 283 carriers.

There is no clear statistical breaking point, so the Poisson Distribution analysis was again performed for the carriers with 4 or fewer violations.

Notably, there were again far fewer than expected carriers with one violation.

Considering the results of the prior analysis, the subsequent findings were not unexpected. The results and conclusions were the same. These violations were not randomly distributed among the identified carriers, and the existence of one violation is a strong predictor of a second violation.

Michigan Carriers with Between 60,000 and 80,000 IRP Miles and Less than Five Violations

Violation Count	Poisson Factor	Expected Value	Actual Percent	Actual Value	Actual as a Percentage of Expected	Contribution to Chi Squared
0	58.4%	159.34	84%	228	143%	29.59
1	31.4%	85.80	7%	20	23%	50.46
2	8.5%	23.10	3.30%	9	39%	8.61
3	1.52%	4.15	1.83%	5	121%	0.18
4	0.20%	0.56	0.73%	2	358%	3.73
Chi Squared Statistic						92.57

Table 3-5

3. Indiana Carriers with Between 40,000 and 80,000 IRP Miles

Number of Violations	Number of Carriers, 40,000 to 60,000	Number of Carriers, 60,000 to 80,000
0	533	305
1	42	18
2	7	11
3	8	6
4	7	3
5	4	2
6	1	1
7	1	2
8	0	1
9	2	
12	1	
13	1	
14	1	
31		1
32	1	
Total	609	283

Table 3-6

The number of violations for Indiana carriers, in both mileage ranges, is presented in Table 3-6. The Poisson calculations in Tables 3-7 and 3-8 are for 4 or fewer violations.

The data in these tables supports the same conclusions that had already been developed;

- Some of these registrants must have driven more miles to get the number of citations (such as 31 or 32) that they received, but
- nonetheless, violations were not randomly distributed among these relatively homogeneous groups.

For the overall number of violations and carriers;

- far too few carriers had exactly one violation, and
- far too many carriers had two or more violations.

Indiana Carriers with Between 40,000 and 60,000 IRP Miles and Less than Four Violations

Violation Count	Poisson Factor	Expected Value	Actual Percent	Actual Value	Actual as a Percentage of Expected	Contribution to Chi Squared
0	70.8%	427	88%	533	125%	26.27
1	24.4%	147	7%	42	29%	75.29
2	4.2%	25	1.16%	7	28%	13.34
3	0.48%	3	1.33%	8	274%	8.83
4	0.04%	0.25	1.16%	7	2779%	180
Chi Squared Statistic						303.73

Table 3-7

Indiana Carriers with Between 60,000 and 80,000 IRP Miles and Less than Five Violations

Violation Count	Poisson Factor	Expected Value	Actual Percent	Actual Value	Actual as a Percentage of Expected	Contribution to Chi Squared
0	67.5%	235.69	87%	305	129%	20.38
1	26.5%	92.52	5%	18	19%	60.02
2	5.2%	18.16	3.15%	11	61%	2.82
3	0.68%	2.38	1.72%	6	253%	5.53
4	0.07%	0.23	0.86%	3	1286%	32
Chi Squared Statistic						120.75

Table 3-8

B. Testing of High Mileage Groups

A second set of tests was performed which started by looking at violation rates.

1. Background

The expected variation in rates changes with the size of the sample. For example, the observer should not be surprised if a fair coin comes up heads three times in 10 (30%), but if the coin comes up heads 300 times in 1,000 (30%), it would be suspected that this is not a fair coin.

The Z-Score is a measure (in "standard deviations") of how far an actual observation varies from what was expected. In the examples used in this study, a negative Z-Score indicates that the carrier had fewer violations than expected (average), while a positive Z-Score indicates more violations than average. Z-Score is more fully described in Appendix B.

When looking at the Z-Scores for a group of similar carriers, it would be expected that the distribution of Z-Scores would approximate a bell curve (assuming that there are a sufficient number of expected violations). To the extent that Z-Scores do not approximate a bell curve, it can be postulated that the violations are not randomly distributed among carriers. Appendix B suggests that a Chi Squared statistic of much over 30 will be considered a solid demonstration that violations are not distributed randomly. Due to the anomalies in the data, it would be difficult to reject the null hypothesis if the Chi Squared statistic is even close to 30.

The original report found vivid demonstration that violations were not randomly distributed within at least three significant groupings of Indiana carriers;

- Carriers with over 750,000 miles,
- Carriers with over 750,000 miles and at least one violations and
- Carriers with over 100,000 miles and at least one size or weight violation.

Analyses of these subgroups, using 1994 data, are presented below.

2. Carriers with over 750,000 Miles

Indiana Carriers with Greater than 750,000 IRP Miles (1994 Data)

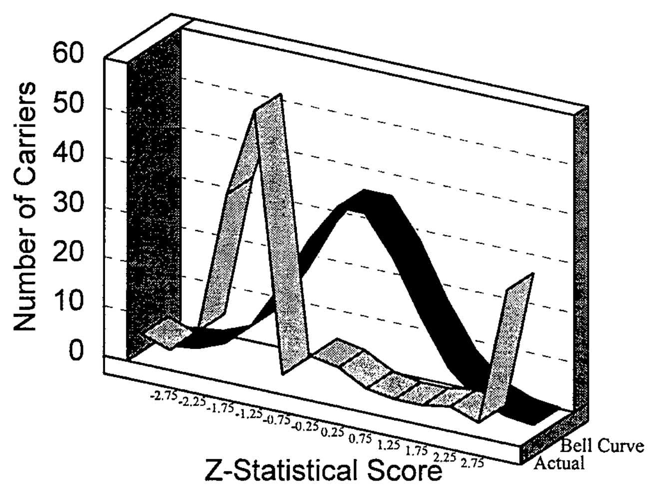


Figure 3-1

This graph presents the results for the later year, for Indiana carriers with over 750,000 miles. The original report showed the same distribution for the previous year's results. It is easy to see visually that the 1994 distribution looks very similar to the 1993 results, presented below.

Total Carriers = 186
Total Violations = 680
Total Miles = 419,406,407
Miles per Violation = 616,744
Chi square = 4,367

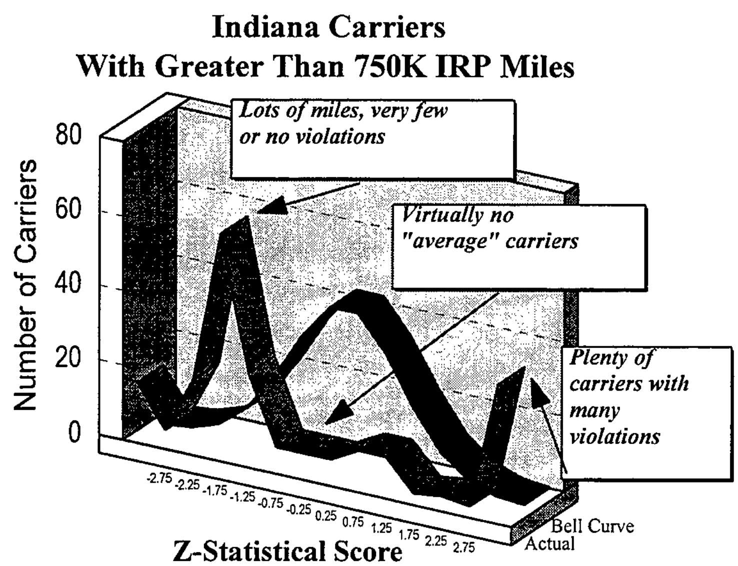


Figure 3-2 (1993 Indiana Data)

Figure 3-2 is repeated from the *Driver/Carrier Statistical Analysis Report.* As noted, it contains 1993 data.

Total Carriers = 217
Total Violations = 1,340
Total Miles = 453,028,713
Miles per Violation = 338,081
Chi square = 6,099

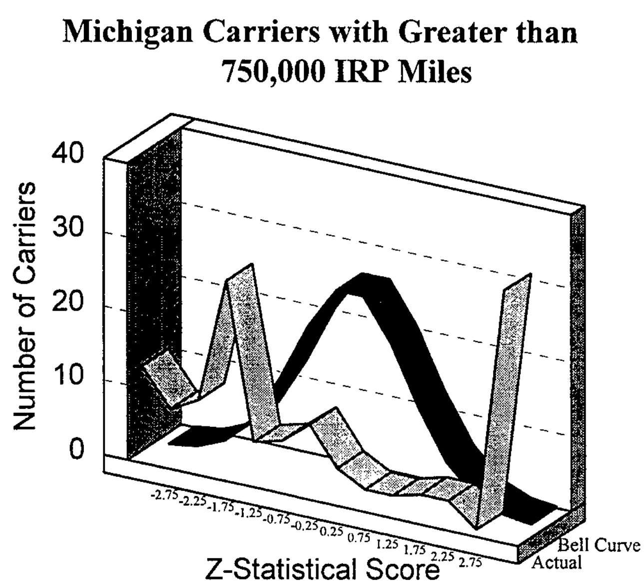


Figure 3-3 Michigan Data

Figure 3-3 shows the same distribution for the Michigan data.

Total Carriers = 135
Total Violations = 1,205
Total Miles = 416,625,525
Miles per Violation = 345,747
Chi square = 7,701

The *Driver/Carrier Statistical Analysis Report* rejected the hypothesis that violations were randomly distributed among carriers with over 750,000 miles. The additional data collected for this report supports the same conclusion.

3. Carriers with over 750,000 Miles and Violations

Carriers with over 750,000 miles and violations were viewed as the group most likely to be homogeneous. Each of these carriers had enough miles that the analysis did not need to be overly concerned with the expectation of too few observations (violations for a carrier). The group should not include many companies which had fewer miles as a registrant than as a carrier due to leasing (although it clearly contains companies that had far more miles as a registrant than as a carrier.) This should be a well established and well understood group of companies.

The hypothesis that was tested was that, within the known group of large carriers which had both IRP mileage greater than 750,000 miles and violations, the violations were randomly distributed. The results were clear.

Violations were clearly not randomly distributed among the carriers with over 750,000 miles and one or more violations.

- There were several carriers which showed very high miles and very few violations. Mostly, these appeared to be leasing companies. In the case of a true leasing company, the company should not have had any violations; these violations may have been improperly assigned. However, at least one of these companies was identified as a company which both ran its own fleet and also leased out vehicles.

- There were also a number of carriers with high miles and high Z-Scores, indicating a high rate of violations over a high number of miles. These appeared to be the carriers of concern.

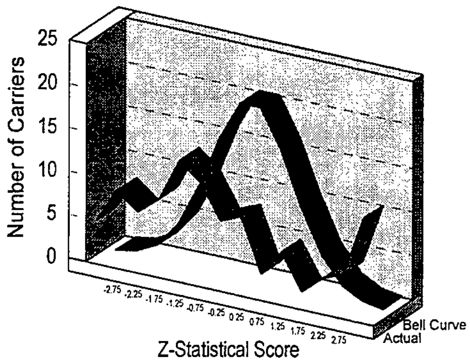
- There was a shortage of "average" carriers. If consideration is given to the fact that there is an overrepresentation of miles in this group, then the "average" carriers should have had Z-Scores somewhat higher than 0. There were valleys in these areas of the graphs, where there should have been peaks.

There is clearly a situation among these carriers where carrier performance, as based on violations, varied significantly.

Figure 3-4 shows distribution for the Indiana Carriers with IRP miles of over 750,000 and Violations from the original *Driver/Carrier Statistical Analysis Report*, along with the more current Indiana and Michigan results. These graphs have been reduced to fit onto one page in order to aide in comparison.

There is a great deal of consistency in this data. Notably, the data from Michigan is exceptionally bipolar.

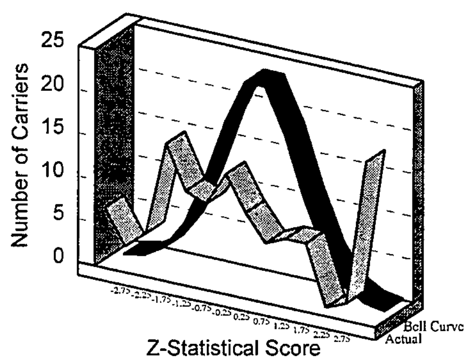
Indiana Carriers, IRP Miles Greater than 750 K and 1 or More Violations



From *Driver/Carrier Statistical Analysis Report* (1993 data).

New (1994) Data

Indiana Carriers ('94), Greater than 750,000 IRP Miles and Violations



Michigan Carriers with Greater than 750,000 IRP Miles and Violations

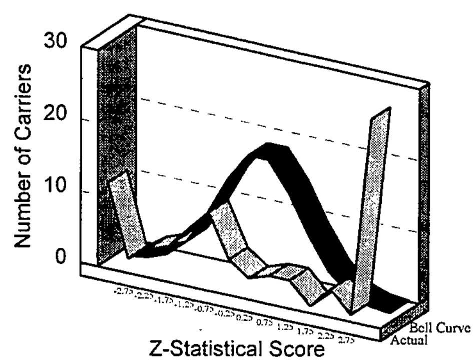


Figure 3-4

Statistics for Figure 3-4

Indiana - 1993

Total Carriers = 108
Total Violations = 1,353
Total Miles = 215,349,877
Miles per Violation = 159,165
Chi square = 1,630.7

Indiana - 1994

Total Carriers = 79
Total Violations = 680
Total Miles = 208,360,617
Miles per Violation = 306,413
Chi square = 1,571

Michigan

Total Carriers = 93
Total Violations = 1,205
Total Miles = 303,991,347
Miles per Violation = 252,275
Chi square = 8,113

4. Carriers with Violations and IRP Miles

Carriers which had IRP miles and (one or more) violations presented the largest group about which sound data was collected. The hypothesis tested was that violations were randomly distributed among these carriers. Because it would be expected that there would have been a number of carriers with low mileage and few violations, a perfect bell curve distribution for these carriers should not have been expected. Rather, the expectation would have been a flattening at the top of the curve, and skewed to the left similar to a Poisson Distribution.

If the actual curves had looked like this, additional statistical treatment would have been necessary to determine the level of impact of low mileage carriers. This was not the case. While the general pattern that had been expected held (especially in the 1993 Indiana data), there was a striking overrepresentation of the carriers with exceptionally high Z-Scores. In each case, over 30% of the carriers showed a Z-Score greater than 3. Some of this effect may have been attributed to:

- carriers which primarily or entirely leased their vehicles. Conventional wisdom was that this represented about 10% of all carriers. Generally, these carriers should not have been in this sampling at all, as they would have had no IRP miles.

- carriers which registered for IRP for the first time, and made exceptionally low estimates of their miles, and

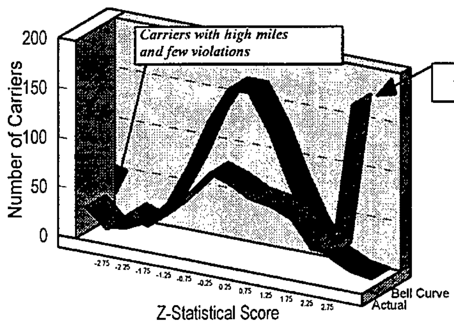
- carriers which radically increased their number of miles traveled.

Nonetheless, there was no reasonable expectation that these three situations affected anything close to a third of the represented carriers. The data is clear,

> ***"Violations were not randomly distributed among carriers which had both mileage and violations."***

It is also interesting to note how consistent these statistics were across the States and years studied. Figure 3-5 presents the graphical representations of these distributions.

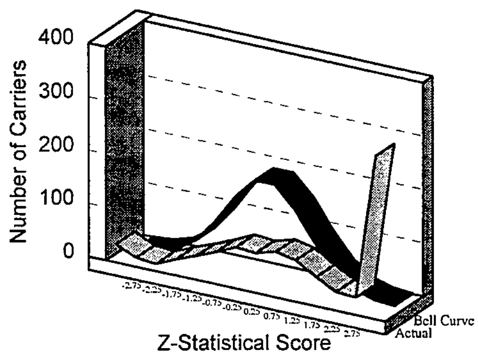
Indiana Carriers With IRP Miles and 1 or More Violations



From *Driver/Carrier Statistical Analysis Report* (1993 data).

New (1994) Data

Indiana'94 Carriers With IRP Miles and 1 or More Violations



Michigan Carriers With IRP Miles and 1 or More Violations

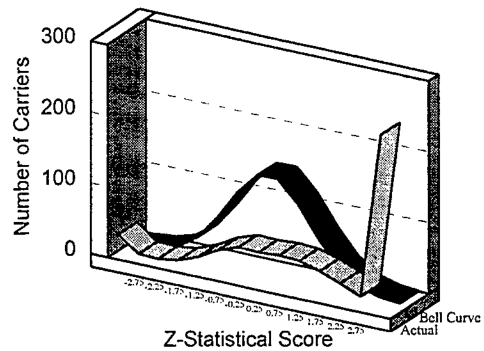


Figure 3-5

Statistics for Figure 3-5

Indiana - 1993

Total Carriers = 865
Total Violations = 3,693
Total Miles = 324,537,146
Miles per Violation = 87,897
Chi square = 35,071

Indiana - 1994

Total Carriers = 936
Total Violations = 3,853
Total Miles = 401,405,593
Miles per Violation = 104,180
Chi square = 77,161

Michigan

Total Carriers = 776
Total Violations = 4,740
Total Miles = 404,440,283
Miles per Violation = 85,325
Chi square = 70,823

C. Carriers with Certain Types of Violations

There were several tests performed on different groups of carriers. The analysis that led to several of the more interesting findings is repeated in this section.

1. Carriers with Over 100,000 Miles and Size and Weight Violations

There was specific interest in size and weight violations, as several members of the group indicated that it was their expectation that size and weight violations are an indication of a certain type of carrier management problem. The *Driver/Carrier Statistical Analysis Report* noted that, in reviewing the size and weight violations, the carriers in the study seemed to fall into three groups;

- Carriers which received no size and weight violations, or virtually no size and weight violations,
- Carriers which received a few size and weight violations, based on the mileage driven, and
- Carriers which received a disproportionately large number of size and weight violations.

The conclusion was that;

> ***"Clearly, size and weight violations were not randomly distributed among all Indiana carriers, and they were not randomly distributed among even those Indiana carriers which had size and weight violations."***

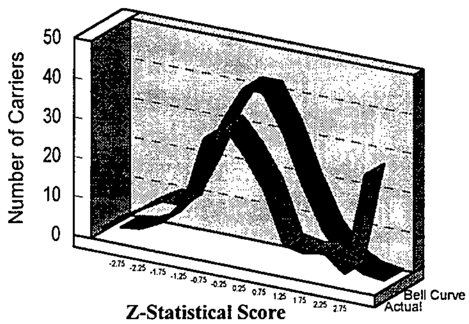
The data from the current study more than supports this conclusion. Even among the carriers with size and weight violations, size and weight violations were far from randomly distributed,

- The 248 carriers in the new Indiana data (which had size and weight violations and over 100,000 miles) had a total of 783 size and weight violations. Of these, 15 carriers had 276 of the violations.
- The 184 carriers in the Michigan data (which had size and weight violations and over 100,000 miles) had a total of 593 size and weight violations. Of these, 16 carriers had 208 of the violations.

The original conclusion is clearly supported by the new data.

The following page presents the graphic presentation from the original Driver/Carrier project along with the same analysis performed on the new data. These statistics were calculated for only size and weight violations, no other violations were considered.

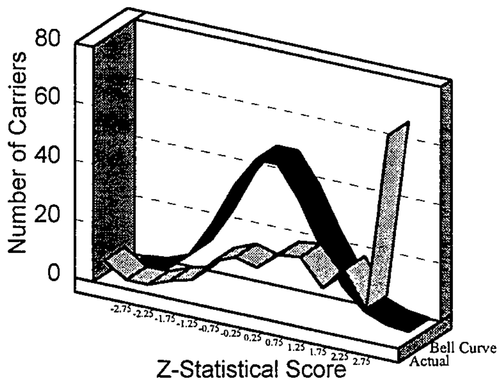
Indiana Carriers, IRP Miles Greater Than 100K & 1 or More Size Violations



From *Driver/Carrier Statistical Analysis Report* (1993 data).

New (1994) Data

Indiana '94 IRP Miles Greater than 100,000 and Size and Weight Vios.



Michigan IRP Miles Greater than 100,000 and Size and Weight Vios.

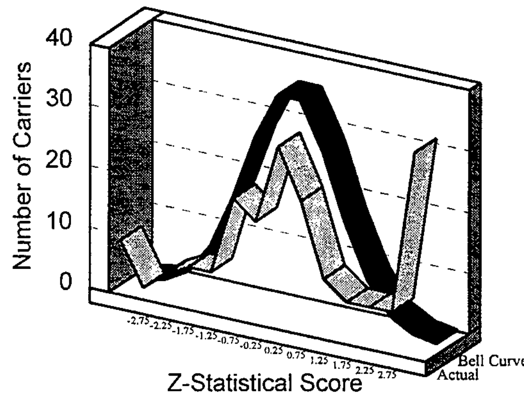


Figure 3-6

Statistics for Figure 3-6

Indiana - 1993

Total Carriers = 220
Total Violations = 929
Total Miles = 211,147,1254
Miles per Violation = 224,854
Chi square = 3,612

Indiana - 1994

Total Carriers = 248
Total Violations = 783
Total Miles = 311,947,684
Miles per Violation = 398,400
Chi square = 16,354

Michigan

Total Carriers = 184
Total Violations = 593
Total Miles = 189,215,495
Miles per Violation = 319,089
Chi square = 5,157

D. Distribution of Violations by Violation Type Within Groups of Interest

The *Driver/Carrier Statistical Analysis Report* reviewed two groups of special interest, carriers with over 40 violations and For-Hire Carriers. These analyses were repeated with the new data.

The results from the 1994 data were entirely consistent with the results of the 1993 data.

1. Carriers with a Large Number of Violations

The hypothesis that carriers with a large number of violations got the same distribution of violations as the average was originally rejected, and it was again rejected. It was clear that,

- in each State, the distribution of violations for the carriers with the most violations was different from the overall distribution, and

- the distribution was different in Indiana and Michigan.

Further, in both States, the carriers with the most violations got more than the average number of size and weight violations. The analyses are presented in the following tables.

Violation Types for Indiana Carriers with 40 or More Violations (1994 Data)

Type of Violation	Expected Result	Actual Result	Difference from Expected	Contribution to Chi Squared
Moving	828.12	653	-21%	37.03
Leaking Load	13.34	12	-10%	0.13
Fuel Tax	93.77	36	-62%	35.59
Size and Weight	1004.78	1461	45%	207.14
Registration	163.45	158	-3%	0.18
Driver License	9.45	6	-37%	1.26
Logbook	279.11	203	-27%	20.75
Equipment	43.78	48	10%	0.41
Other	640.20	499	-22%	31.14
Chi Squared Statistic				333.65

Table 3-9

Violation Types for Michigan Carriers with 40 or More Violations (1994 Data)

Type of Violation	Expected Result	Actual Result	Difference from Expected	Contribution to Chi Squared
Moving	616.70	650	5%	1.80
Size and Weight	479.40	665	39%	71.86
Registration	511.91	478	-7%	2.25
Driver License	521.54	295	-43%	98.40
Logbook	274.41	597	118%	379.24
Equipment	476.63	346	-27%	35.80
Other	358.41	208	-42%	63.12
Chi Squared Statistic				652.47

Table 3-10

This further demonstrates that violations were not randomly distributed among motor carriers. There were other factors, in this case factors that did not appear to be directly related to safety performance, which indicated that there was some relationship between the number of violations that a carrier received and the type of violations that the carrier received.

This data reemphasized the demonstration that the pattern of violations was different in these two States. Further, the difference in the change in pattern was interesting. For example, the carriers with the largest number of violations in Indiana received 27 fewer log book violations than would have been expected. In comparison, the carriers with the most violations in Michigan received over double (118%) the number of log book violations than would have been expected.

2. For-Hire Carriers

It was noted in Phase I that the distribution of violations, by type, was different than would have been expected for For-Hire carriers. Again the distribution of violations for the carriers which were identified as For-Hire carriers (i.e. carriers with an ICC number) was somewhat different than the general distribution.

There was very little difference in Indiana. This could have been because close to 2/3 of the carriers were identified as For-Hire carriers, so the For-Hire carriers made up the major portion of the sample size. Further, although they were less than 2/3 of the number of carriers, the For-Hire carriers had over 3/4 of the violations. This further reduced the opportunity for deviation. This relationship between carrier and violation count supports the generally held belief that For-Hire carriers tend to be the larger carriers, which were focus on trucking as their primary business. In contrast, the private carrier operations tend to be smaller, and are often parts of businesses that was not primarily devoted to trucking.

This thesis was borne out by the Michigan data. In Michigan barely 1/3 of the carriers were identified as For-Hire carriers, but these carriers had well over 1/2 of the total violations. Further, the Michigan For-Hire carriers tended to have;

- fewer registration and driver license violations. These would tend to be violations associated with incomplete or incorrect paperwork, and

- more moving and size and weight violations. These tend to be associated with more aggressive driving and loading practices.

The supporting analysis is presented in the following tables.

Chapter 4 - Violation Rates and Safety Performance

The purpose of this chapter is to respond to the project goal to determine if it can be demonstrated that the lack of randomness in violation rates is due, in part, to a difference in safety performance among carriers.

Phase I of the Driver/Carrier Study demonstrated that the assignment of violations to carriers is not random. It is reasonable to postulate that the drivers for some carriers receive violations at a significantly different rate than drivers for other carriers. Further, this distinction carries across all types of violations, the drivers for some carriers get several different specific types of violations at different rates from the drivers for other carriers. The first phase of this study, as presented in Chapter 3, revalidated these conclusions.

The next issue is whether this difference is due, at least in part, to a difference in safety standards and safety performance.

There is anecdotal evidence that there may be other factors, in addition to safety related factors, which could affect a carrier's violation rate. For example, one carrier's trucks might travel a greater percentage of its miles on roads which have a high enforcement profile, resulting in more violations. It has been suggested that the carriers with the poorest safety performance may have very low violation rates for violations assigned at fixed enforcement locations (scales) because these trucks are actively bypassing the scales.

Further, even if it was true that the only reason that carriers (or drivers) received citations was for engaging in precisely the behavior that leads to crashes, citations would still not be a perfect predictor of crashes, due to the affects of randomness.

Therefore, there is no expectation of a perfect relationship between carrier citation performance and carrier safety performance. Rather, the question is as to whether there is any relationship at all.

This Chapter will be divided into four parts. The first part will include background for understanding the statistical analysis performed. In the second part, the results of the analysis for carriers which had violations will be presented. For the third section, an additional matching was made for carriers which had IRP mileage and carrier census data, but no violations. This section presents the results of the expanded analysis. The final section presents the conclusions of the data analysis.

A. Background

The purpose of this section is to provide background about the data that was used in this stage and the statistical methods that were used for analysis.

1. Coefficient of Correlation

The Coefficient of Correlation is actually a measure of the "goodness of fit" to a straight line developed through linear regression. Given a set of pairs of (x,y) data, it is possible to create a formula which predicts the dependent variable (y), given the independent variable (x). Geometrically, this is represented as the line which most closely fits the data. The Coefficient of Correlation is a measure of how well the line matches the data, in other words how well one variable actually predicts the other variable. A Coefficient of Correlation of 1 shows a perfect match (i.e. given one variable, the other is perfectly predictable), and a Coefficient of Correlation of 0 shows no relationship at all. A coefficient of correlation of -1 also shows perfect predictability, but it indicates that the slope of the line is negative, as one variable increases the other decreases.

Appendix C provides a more complete explanation of the Coefficient of Correlation.

The Coefficient of Correlation does not demonstrate causality. For example, a high correlation between heart disease and high blood pressure does not demonstrate that heart disease causes high blood pressure. In this case, the observer might choose to speculate that the causality is actually the other way around.

In fact, a high Coefficient of Correlation does not actually say anything about causality at all. For example, there might be a high correlation between heart disease and lung cancer. However, there could well be no causality relationship between the two. Both might be caused by cigarette smoking. Nonetheless, if such a correlation existed, a doctor who found heart disease would then have appropriate justification for checking for lung cancer, regardless of the issue of causality.

In this study, we speculate that this is the situation. We speculate that there may not be a causality relationship between most citations and crashes. Rather, we speculate that the same driver practices that lead to citations also lead to crashes, and that at least some of those driver behaviors are related to carrier management practices.

Therefore, the primary objective of this chapter is to determine if there is a statistical relationship, specifically a correlation, between indicators of carrier citation performance and indicators of carrier safety performance. This chapter does not address issues of causation.

2. Available Data

In order to determine carrier citation rates, it has been necessary to gather both;

- measures or counts of citations assigned to a carrier, for use in the numerator of the rate, and
- a measure of exposure, to use in the denominator of the rate.

Three approaches were used to evaluate carrier safety performance;

- the carrier's crash rate, including ranking measures of crashes and exposure,
- the more subjective evaluation of safety performance reflected in the carrier's Safety Rating, and
- the carrier's SCE score, a score used by FHWA to rank carriers for review.

For this study, each of the States that was studied was viewed as an entity unto itself. The analysis was performed on a State by State basis.

a. Citation Data

The citation data for this analysis was provided by the Indiana State Police and by the Michigan State Police. This data represents all of the citations that were issued by the State Police, and for which the motor carrier could be identified, during the time periods in study. For the time period of calendar year 1993, Indiana provided 22,891 citations with a carrier reference. For the calendar year 1994, Indiana provided 23,760 citations with a carrier reference. For a period that included portions of calendar years 1993 and 1994, Michigan provided 33,971 citations with a carrier reference.

In each case, the State Police attempted to accurately identify the motor carrier. While it can be assumed that, in at least some cases, a lessor (rather than the lessee-carrier) was identified, the State Patrol can be considered to be the most highly trained observers in the field, and the organization most likely to accurately identify the carrier. Within the State Patrol, MCSAP trained officers are the most likely to accurately identify the carrier.

The analysis was performed using both overall citation data and subsets based on the type of citation.

b. Mileage Data

For this study, IRP mileage data was used. IRP mileage data has some very specific strengths in this regard. It presents an accurate identification of the actual mileage for the registrant in the State in the prior year. While there are other possible indicators of crash exposure (such as number of power units, or driver hours), mileage is an appropriate measure of exposure.

Unfortunately, the IRP registrant is not always the motor carrier. It has been estimated that about 10% of all trucks on the road are leased. For a leased truck, either the lessor or the lessee can be the registrant. There is no current requirement that an IRP registrant furnish a US DOT number (the primary identification of a motor carrier.) Further, the registrant's vehicles could work for one or several carriers. Commonly, a motor carrier will have more than one IRP account. Speculation as to why a carrier would have multiple IRP accounts is beyond the scope of this study. National carriers may have fleets based in many States. For this study, all of the national carrier's violations were included, but only the mileage of the fleets based in the State.

In order to continue the analysis, it was necessary to match the IRP accounts to motor carriers. This matching was performed manually, using the registrant name on the IRP files and the carrier name on the FHWA files.

For the 1993 tax year, Indiana reported 8,342 registrants. Of these, it was possible to identify 865 distinctly named registrants with carrier violations. For the 1994 tax year, Indiana reported 6,752 registrants. Of these, 189 were exact matches for other registrants in the name and address fields. The mileage on matching records was combined, leaving a total of 6563 distinctly identifiable registrants. (A similar merge process was not performed for 1993 data.) It was possible to identify carrier violations for 936 of these registrants. The Michigan data contained IRP records for 4865 registrants. A similar merge process identified 466 registrant records which matched other records on name and address, reducing the overall count of distinct registrants to 4399. It was possible to identify carrier violations for 729 of these.

SAFETYNET/MCMIS

The SAFETYNET /MCMIS system is FHWA's information system for the Motor Carrier Safety Assistance Program (MCSAP). SAFETYNET is the State level component of the system. MCMIS is the national level Motor Carrier Management Information System component housed at FHWA Headquarters in Washington D.C.

c. Snapshot Data

Snapshot data is data that was generated from the carrier profile in MCMIS. For this project, critical snapshot information was identified as current and immediate prior Safety Rating, SCE rating and MCSIP score. Because so few carriers had MCSIP scores, no further analysis was performed on MCSIP data.

d. Crash Data

MCMIS also contains information about carrier crashes. This information tends to be somewhat inaccurate, as all crash reporting prior to January, 1994 was carrier self-reporting. Further, only crashes that were investigated by the organizations that report through the State police are usually reported to SAFETYNET (and from SAFETYNET to MCMIS.)

Nonetheless, data about crashes present the crux of the issue that this report is examining. The primary purpose of the Motor Carrier Safety Assistance Program (MCSAP) is to prevent crashes, and to minimize the damage of crashes when they occur.

The crash data shows that 3,975 carriers had reportable crashes in either Michigan or Indiana since 1990. Of these 1,723 had crashes in Indiana. 2,793 carriers had crashes in Michigan in the time period. 541 carriers had crashes in both Michigan and Indiana.

3. Usage of Data

Again, the purpose of this chapter is to determine if it can be demonstrated that the lack of randomness in violation rates is due, at least in part, to a difference in *safety* performance among carriers. From a statistical point of view, this would be demonstrated by a correlation between violation rates and indicators of safety performance.

a. Violation Rates

Two indicators were used for the analysis of violation rates;

- Overall violation rate - Which takes into account all types of violations reported, and

- Z-Score - Which focuses on the carrier's performance as compared to other similar carriers. The technical derivation of these Z-Scores is described in Appendix C.

Z-Scores were recalculated based on the particular sub-population that was being studied. In each case, the carrier is only compared to carriers with similar characteristics within the analysis. Therefore, individual carrier's Z-Scores will vary from analysis to analysis.

b. Safety Performance

There are three available indicators of safety performance:

- First, there is the carrier's crash rates. Crash rates are the single most important measurable parameter. It is the goal of safety programs to minimize the number of crashes (as well as to minimize the severity of crashes.).

- Second is the carrier's Safety Rating. This is the result of a review of the carrier. A rated carrier will have one of three ratings;

  - S - Satisfactory,
  - C - Conditional, or
  - U - Unsatisfactory.

  A substantial number of carriers have never received a rating. Newer and smaller carriers tend to be unrated, while most of the larger and older carriers have received at least one rating.

- Third is the SCE ratings. This is a score assigned to carriers to rank them when determining which ones will be scheduled for a review. Carriers with high SCE scores are given a higher priority for review.

Crash rate presents an interesting statistical challenge. The use of only the most recent crashes presents a better timed match to the annual mileage data. However, it was assumed that mileage patterns are fairly stable for most carriers. Usage of additional year's crashes presents a far more robust statistical picture. However, it loses some while precision in rate matching as older data is used. In general, this project accepted fairly old data in order to obtain a richer data history. For crashes, the carrier's total number of crashes from 1990 to the point in time of the data extract has been used. The project did not use data older than 1990.

Safety Rating is a rating assigned as a part of the carrier review process. Safety Ratings have been assigned through two processes;

- a Safety Review, performed by a State or Federal enforcement official, and
- a Compliance Review, performed by a State or Federal enforcement official.

Safety and Compliance Reviews shared a similar evaluation and scoring methodology. Safety Reviews are no longer performed, but many of the existing Safety Ratings are a result of Safety Reviews.

The Safety Rating may be a less accurate indicator of safety performance than crash rate for a number of reasons;

- Safety Ratings are not entirely performance based. Much of the Safety Rating is based on management practices, which are expected to affect safety fitness.
- A portion of the Safety Rating is based on issues related to the carrier's management practices. These practices have not been demonstrably linked to carrier safety.
- Many of the Safety Ratings are dated.

Nonetheless, it is expected that current Safety Rating should provide some insight as to the perceived safety fitness of the carrier. Further, a prior Safety Rating presents a picture of the history of the carrier. For this reason, the project assigned a ratings ranking to each carrier, based on the carrier's current and prior Safety Rating. The goal of the ranking is to provide a method of quantifying the carrier's Safety Ratings. Table 4-1 presents this ratings method.

The SCE scoring methodology had several goals. One was to identify carriers that are unsafe. Another was to identify carriers that pose a greater safety threat to the roads. SCE targeted hazardous waste and passenger carriers. The SCE rating does take into account carriers safety histories. The SCE was not designed simply to predict bad carriers. It ranks the carriers to determine who will be reviewed.

Quantifying Safety Rating		
Ratings Rank	Current Safety Rating	Prior Safety Rating
9	U	U
8	U	C
7	U	S or none
6	C	U
5	C	C
4	C	S or none
3	S	U
2	S	C
1	S	S or none

Table 4-1

c. Use of Ranking

There are several situations where use of ranking data is more appropriate than the use of raw data. The commonly used example of use of ranking methods is for cross-country runners. Because cross-country races are run on different courses, (and because cross-county runners often do not sprint at the end of a race), statistics about cross-country runners are usually based on the runner's position at the end of a race or races, rather than their time.

Further, ranking is relevant to the goals of this project. The objective of the project is not to determine "badness" of a carrier, it is to assist in targeting carriers for review. Thus, the end result of the project will be to assist in ranking. There are at least two conditions in this study where ranking appears to be particularly relevant;

- Z-Score is a statistical abstraction, rather than a raw statistic. Consequently, we chose to consider a ranking as an alternative approach for analysis.

- A review of the data, particularly the data from Indiana, indicates that some carriers, presumably first time filers, make mileage estimates that are very low. While these appear to be adequate for IRP, as the percentage of mileage for fee allocation is acceptable, the resultant crash and violation rates (as well as the Z-Scores) can be far out of proportion. In order to minimize the impact on the statistical analysis, rankings based on actual rates may be used.

4. Expected Results

Overall, it is the goal of this analysis to demonstrate one of two possible results;

- either accept (technically do not reject) the null hypothesis - The difference in carrier violation rates is **not** an indicator of safety performance differences, or

- rejection of the null hypothesis - The difference in carrier violation rates **is** at least to some degree an indicator of safety performance differences.

If the difference in carrier violation rates is an indicator of safety performance differences, then it would be anticipated that there would be some correlation between violation rates and crash rates. A perfect, or even a particularly high, correlation would not be expected for several reasons;

- Crashes and driver citations are still both low probability events with a high degree of randomness. This should always result in a less than perfect correlation. For example, the other driver's behavior affects crashes. While high frequency types of events will mitigate this effect, randomness will have a significant affect on the correlation of low frequency events.

- There may be other factors that affect the carrier's overall violation rate. Examples of this type of extraneous factors are,
  - the carrier being physically located in proximity to a police barracks,
  - general fleet age,
  - the pattern of carrier's operations exposes the carrier to enforcement, or
  - a particularly undeserved reputation with the police.
- Mileage was derived from IRP and manual matching. The mileage rates for analysis are not completely accurate representations of the carrier's exposure.
- Crash history data comes from self-reporting. Further, the time period of the crash history does not match the time period of the violation and mileage histories.

Therefore, it would not be expected that violation rate is a perfect predictor of crashes. Rather, the issue is whether violations are a predictor at all. This will be demonstrated by a correlation between violation rates and indicators of safety fitness, particularly crashes.

However, if the alternative is true, that violation rates are not predictors of crashes, it is expected that the statistical tests will show minimal correlation. It is not always the case that no correlation exists. It is possible that a correlation will exist because both tested factors are correlated to a third factor, or that there is a fault in the design of a particular test. Further, some low level of correlation is expected between totally unrelated events, as a result of randomness. Care must be exercised to avoid conclusions based on spurious relationships.

B. Results

The first set of analyses was performed using only carriers which had been identified as having mileage, violations, and crashes. This analysis is particularly strong in that it uses the best known and identified carriers.

1. Michigan Carriers with Mileage, Violations, and Crashes

Of the 729 Michigan carriers which had identifiable mileage and violations, 334 carriers also had crashes over the period that crash data was maintained. This group was chosen for initial investigation because of the large number of carriers which could be identified as fitting into the group.

Several correlation analyses were performed, using both the actual rates for violations and crashes, as well as the rankings.

Michigan Carriers with Mileage, Violations, and Crashes	
Correlation of Crash Rate and Violation Rate	0.580
Correlation of Z-Score and Crash Rate	0.417
Correlation of Z-Score Rank and Crash Rate Rank	0.600
Correlation of Violation Rate Rank and Crash Rate Rank	0.722
Correlation of Violation Rate Rank and Z-Score Rank	0.964
Correlation of Number of Crashes with Number of Violations	0.384

These results are quite revealing. First, within these Michigan carriers;

"violations are clearly correlated with crashes."

Most revealingly, the violation rate rank and crash rate rank have a Coefficient of Correlation of over .7.

There was no surprise that there is a high degree of correlation between Violation Rate and Z-Score ranking; the Z-Score is based extensively on the violation rate.

In addition to the very high correlation between violation rank and crash rank, there are several other results which bear investigation.

a. Crash and Violation Rates as Compared to Rankings

The correlation between crash rate ranks and violation rate rank (.722) is much stronger than the correlation between crash and violation rates (.580). This is somewhat surprising, especially in light of the findings in the initial Indiana analysis of similar situation. In Indiana, for carriers with miles, violations, and crashes, there was an exceptionally strong correlation between crash rate and violation rate. However, this Indiana result was almost entirely the result of a few carriers with very few miles, and a large number of crashes and violations. This was an IRP/ matching problem. Here, the situation is not so obvious.

The Indiana condition led to an examination of the carriers with large mileage. A cursory review presented a startling observation. Although overall there are less than 42% as many crashes as violations reported, one "carrier" had 83 crashes and 28 violations. This "carrier" is a clearly identifiable company which has a large leasing operational component. One would assume that either these are crashes associated with small household goods moves (an unlikely scenario, since all crash reports in the study were written by the State Police), or that there is some misidentification of lessors as carriers on the roadside.

It appears that poor correlation between rates for the outlying carriers, especially the outlyers at the "good" end of the scale, is a cause of this problem.

Once these data anomalies are explained, there is a more important conclusion. Within this population, the violation rank was a strong indicator of safety performance.

b. Z -Score vs. Violation Rate as a Predictor of Crashes

There are two direct comparisons of violation rate, as compared to Z-Score, as a predictor of crashes. The correlation between violation rate and crash rate is .580, while the correlation between Z-Score and crash rate is only .417. Similarly, the correlation between violation rate rank and crash rate rank is .722, while the correlation between Z-Score rank and crash rate rank is only .600.

Clearly, rank works as a better correlation tool than raw numbers. The poor correlations of the outlyers are at work here. Further, it is worth noting that the correlation for Z-Score rank is much closer to the correlation for violation rank than is the raw data.

A little cogitation explains why violation rate should be a better predictor. Consider, for example, the carrier with a violation rate that is 50% higher than the norm and a relatively small number of miles. This carrier should have a relatively low Z-Score (as compared to a carrier with a large number of miles and a violation rate that is higher than average), as it is likely that, for the individual carrier, this is only a statistical deviation from average performance. But, for a large number of carriers, a violation rate that is 50% higher than average is likely to be a predictor of poor safety performance.

The Z-Statistic assumes that carrier performance is random. But carrier performance is not random. For carriers as a group, a higher violation rate is a strong predictor of a higher crash rate. The Z-Statistic can help assure that small to mid-sized carriers are not singled out for scrutiny because of bad luck. The Z-Statistic is appropriate for use with individual carriers. But at the aggregate level it only serves to obscure the underlying point, carrier violation performance is an indicator of carrier safety performance.

c. Correlation of Number of Crashes with Number of Violations

The Coefficient of Correlation between the number of crashes for a carrier and the number of violations for a carrier was .384. Along with the other results, this strongly validates a number of assumptions that have been made about the data and the underlying business processes.

First, it is no surprise that there is a correlation between the number of crashes and the number of violations. It would be expected that the carriers with more miles would have more violations and more crashes. This would argue for a strong correlation between violation count and crash count, as both would be correlated with mileage.

The "better" motor carriers (that is the carriers with strong safety programs that have resulted in strong safety records) have suggested that violations are, to a larger extent than crashes, the result of some randomness. Therefore, the number of violations that a carrier gets would be, to some degree, a result of exposure. Therefore, they suggest, it would be wrong to simply target carriers for scrutiny based on a count of violations.

This position is clearly supported by the relative weakness of the correlation between the number of crashes and the number of violations for carriers.

The far stronger correlation between crash rate and violation rate, and the stronger yet correlation between violation rate rank and crash rate rank, demonstrate clearly the need to account for exposure.

2. Indiana Carriers with Mileage, Violations and Crashes

For purposes of analysis, only the most recent Indiana data was used. This makes the data comparable to the data from Michigan.

Of the 936 registrants which were identified as carriers, 83 had reported crashes in Indiana. Similar correlation analyses were performed as the analyses for Michigan carriers.

The basic conclusion that this data offers is that

"there is a strong correlation between these carriers violation rates and crashes rates, and an even stronger correlation between the violation rate rank and the crash rate rank. Perhaps the most striking is the strong correlation between the Z-Score ranking and the crash rate ranking."

In addition to this high level conclusion, there are several other results that bear investigating.

a. Very High Overall Correlations

The initial analyses showed incredibly high correlations between crash rate and violation rates, as well as with Z-Scores. A cursory review of the data showed 5 carriers which had clearly understated mileage. Each of the 5 showed 612 miles in Indiana. Among them, there were 78 violations and 16 crashes. Because the violation and crash rates for these carriers were so far out of line, the impact overwhelmed the correlation calculation. Removing these five from the correlation calculations altered the results. Using only the 76 carriers which had over 10,000 miles resulted in a Coefficient of Correlation between crash and violation rates of .650, and a Coefficient of Correlation between Z-Score and crash rate of .399.

Indiana Carriers	
Correlation of Violation Rate and Crash Rate	0.984
Correlation of Z-Scores and Crash Rate	0.984
Correlation of Z-Score and Violation Rate	0.999
Correlation of Violation Rate and Crash Rate, Top 5 Records Excluded	0.904
Correlation of Z-Score and Crash Rate, Top 5 Records Excluded	0.621
Correlation of Z-Scores and Violation Rate, Top 5 Records Excluded	0.790
Correlation between Number of Violations and Number of Crashes	0.620
Correlation of Violation Rate Rank and Crash Rate Rank	0.807
Correlation of Z-Rank and Crash Rate Rank	0.713
Correlation of Z-Rank and Violation Rate Rank	0.947

b. Crash and Violation Rates as Compared to Rankings

Again, for the Indiana data, the ranking appears to be a far stronger statistical tool than the actual rates. This appears to be based on the fact that violation rate is not a good predictor of crash rate for the outlyers.

- Some "good" carriers had far lower crash rates that would be predicted from their violation rates, which may be attributed to the assumption that assignments of violations has a component that is more a random function of exposure, whereas crashes tend to be more of a function of carrier safety performance.

- Some large "carriers" appear to be really leasing companies. The assignment of crashes to carriers appears to be less precise than the assignment of violations. For example, one company has approximately 450,000 IRP miles in Indiana, and one violation; a fairly good accounting. However, this same company has 27 crashes.

3. Indiana Carriers with Violations and Miles

The next group of interest is Indiana carriers with violations and miles. This is 936 carriers, 14% of the 6562 carriers which had mileage.

There are several factors of interest in this data.

a. High Correlation of Crash Rate with Violation Rate and Z-Score

Indiana Carriers	Full Data	Minus 5 Carriers
Correlation of Violation Rate and Crash Rate	0.665	-0.007
Correlation of Z-Score and Crash Rate	0.664	0.005
Correlation of Z-Score and Violation Rate	0.997	0.994
Correlation of Crash count and Violation count	0.258	0.242
Correlation of Z-Rank and Crash Rank	-0.122	-0.156
Correlation of Crash Rank and Z-Rank for Carriers with Crashes	0.700	0.632
Correlation of Crash Rank and Violation Rank for Carries with Crashes	0.825	0.784

There was a surprisingly high correlation between the crash rate and the violation rate, as well as between the crash rate and the Z-Score. This was especially surprising since the crash rate for the overwhelming majority of carriers should have been zero, generating a very low correlation.

Again, there were five carriers which had 612 Indiana miles, with very high crash and violation rates. When these five carriers were removed from the analysis, the correlation between crash rate and Z-Score is negligible. This is not surprising. The overwhelming majority of carriers did not have any crashes, so it would be anticipated that the correlation between crash rate and either Z-Score or violation rate would be negligible. The presence of a few outlying data points skewed the overall analysis.

b. Correlation Between Z-Rank and Crash Rank

Surprisingly, there is a negative correlation between Z-Rank and crash rank. At first appearance, this makes it appear that Z-Score is negatively correlated with safety performance. Clearly, a second look is needed.

In fact, what is occurring is that a low Z-Score and some crashes are both common occurrences for carriers (many of which appear to be leasing companies) which have a very large number of miles. Since over 90% of the identified carriers have no crashes, a carrier with even one crash will receive a very high crash rate ranking. However, these same carriers have very low Z-Scores, as;

- these carriers have a lower than average violation rate, and
- these carriers also have lower than average crash rates.

Again, care is needed to assure that statistics are interpreted correctly, especially in the light of the sparseness of available data.

c. Correlation Between Crash Rate and Violation Rate

There was, for all intents and purposes, no correlation between the violation rate and the crash rate. While initially disconcerting, this should not be a surprising result. Since, in the overwhelming majority of the cases, the crash rate is zero, these cases should contribute strongly to the lack of any correlation. Perhaps, the surprising result is that the correlation is not negative, since the presence of large carriers with very low violation rates and positive crash rates should generate a negative correlation. Apparently, without the contribution of these carriers, the correlation would have been positive.

d. Correlation between Number of Crashes and Number of Violations

There is a surprisingly weak correlation between the number of violations and the number of crashes that a carrier has.

It would have been expected that there would be a fairly strong correlation between the number of crashes and the number of violations.

- Both would be expected to be, to a significant extent, a function of exposure. The larger carriers would be expected to have a higher number of both violations and crashes.

- Smaller carriers with bad safety performance would be expected to have significant numbers of both crashes and violations. Smaller carriers with good safety performance would be expected to have no violations and no crashes.

Both of these factors would be expected to contribute to a higher correlation. Several factors could be contributing to the low correlation between number of violations and number of crashes;

- Most importantly, the very low frequency of crashes (close to 90% of the carriers with violations do not have crashes) makes crashes take on the statistical characteristics of random events, contributing to the low correlation.

- It may be that enforcement officers are targeting certain perceived scofflaw carriers for more intense enforcement. These carriers, even if they have a greater than average crash rate, would have a much greater than average violation rate.

If it is true that officers in the field have identified certain carriers as targets for higher enforcement intensity, then targeting carriers for subsequent scrutiny would be using collected statistics to take advantage of that judgment.

Further, the officers do not have to have made a conscious judgment in identifying targets. It may be that a high violation rate for a carrier is the result of a series of subconscious judgments by many officers at different times and places. Use of violation rates to target carriers is a validation of the value of those judgments.

e. Correlation Between Z-Ranking and Crash Ranking for Carriers with Crashes

As was demonstrated in the previous sections, there is a very strong correlation between Z-Ranking and crash rate ranking, for carriers which have both violations and crashes. There is an even stronger correlation between violation rate ranking and crash rate ranking for carriers which have crashes.

The difference between these statistics is based on the way that Z-Scores are generated. If it is true that violation rate is a strong predictor of crash rate, then this type of result would be anticipated.

This is a very telling statistic; among the carriers which have both violations and crashes, violation rate is a strong predictor of crash rate.

C. Results for Carriers in the Census

An additional attempt was made to identify registrants which did not have violations. The names in the registrant list from IRP were manually matched to the names in the Carrier Census from MCMIS. This provided a significant base of information for additional analysis.

In order to generate the highest possible matching, both the carrier name and the carrier DBA (doing business as) name, as shown on the MCMIS data, were used for matching. The high level of effort necessary for matching, as well as the obvious potential for inconsistency, demonstrates that this is not a feasible approach for a production environment. It was, however, a satisfactory approach for a statistical study.

This task was performed for both Indiana and Michigan carriers.

1. Results for Indiana Carriers

For Indiana, 1202 registrants which had no associated violations were identified as carriers from the census information.

Overall, 2138 of the 6,562, or just less than 1/3 of the registrants were identified. These registrants accounted for 512,238,921, or just over 2/3, of the 762,816,006 registered miles.

a. Overview for Indiana Carriers

A comparison of the 1202 carriers which did not have violations to the 936 which had at least one violation presents some interesting results.

- 127 of the 936 carriers which had violations had less than 1,000 IRP miles in Indiana. Of these, 124 had precisely 612 Indiana miles.

- 259 of the 1202 carriers which had no violations had less than 1,000 IRP miles in Indiana. Of these, 242 had precisely 612 Indiana miles.

The carriers which had less than 1,000 miles are excluded from the following analysis.

Indiana Data from 1994	Carriers Which had Any Violations	Carriers Which had No Violations
Total Number of Carriers with Greater than 1,000 Miles	799	943
Average Mileage for a Carrier	496,078	117,364
Number of Carriers with Crashes	207	60
Average Mileage per Crash	953,272	1,216,203

The results of this analysis confirm two of the conclusions that are critical to this study.

First, there is an element of randomness in the occurrence of crashes, at least in comparison to violations. The stark difference in the mileage per carrier for carriers which have violations as compared to miles per carrier for carriers which do not have violations clearly demonstrates that receipt of violations is, to a substantial degree, a function of exposure. This may be an unsurprising conclusion.

Second, the average mileage per crash is substantially lower for the carriers which have violations (a difference that is statistically significant at the 90%, but not the 95% confidence level). This demonstrates that there is a relationship between violations and crashes using an appropriate test.

Further analysis supports the contention of the larger carriers that, for larger carriers, the mere existence of violations does not indicate a problem. Among the carriers which have violations;

- For carriers with more than 1,000,000 miles, the miles/crash rate is 1,226,525. This is virtually indistinguishable from the overall miles/crash rate of the carriers which did not have violations. (However, there are 16 carriers with between 1,000,000 and 3,000,000 miles which have not violations and a miles/crash rate of 2,277,216. This supports the contention that there are some better than average larger carriers.)

- For carriers with between 1,000 and 1,000,000 miles, the miles/crash rate is 659,826. This is significantly different from the overall rate for all carriers, as well as from the crash rate for carriers without violations with mileage between 1,000 and 1,000,000, which is 1,055,035. Clearly, within this group, violations are an indicator of a propensity toward crashes.

The reader is advised to recall that this analysis uses one year of mileage data and five years of crash data. While these rates are valid for comparison, the actual crash rates are lower than the rate reflected here.

b. Analysis of Indicators for Carriers in the Census

Correlation analysis was performed for all the carriers which had IRP miles and were identified in the census. The results are presented in this table.

i. Crash Rate and Z-Score

There is a high correlation between crash rate and Z-Score. A component of this statistical result may be the impact of a few low mileage carriers with high violation and crash rates. Nonetheless, a substantial portion of this result is associated with the underlying relationship, as is demonstrated by the correlation between the crash and the violation rates.

This presents additional demonstration that the difference in violation rates among carriers is associated with a difference in safety performance.

Coefficient of Correlation of	
Crash Rate and Z-Score	.843
Crash Rate and Violation Rate	.847
Crash Rate and SCE Score	.003
Violation Rate and SCE Score	.003
Crash Rate and Ratings Factor	.023
Violation Rate and Ratings Factor	.007

ii. Crash Rate and SCE Score

The correlation between crash rate and SCE score is virtually non-existent. What this suggests is that SCE score has, in the past, only pointed out carriers which, if they demonstrate poor safety performance, may have a higher level of risk. Among all carriers which have both IRP miles and have been identified in MCMIS, SCE score did not indicate carriers which had poor on-the-road performance.

c. Carriers with Crashes

283 identified Indiana carriers had at least one crash.

In many ways, this is the group of particular interest. This is only a sub-set of all carriers, but it is a particularly critical subset. First, it is the group of carriers for which the most complete information is available. Second, it is a group of particular concern. Especially among the carriers with a history of crashes, it is important to be able to distinguish carriers which get into crashes because of their sheer number of miles traveled, as compared to carriers which get into crashes because of poor safety performance.

As will be demonstrated below, it appears that this group includes at least four sub-groups;

- a number of large carriers which, as a group, have crashes at roughly the average rate,
- a number of mid-sized carriers, which have crash rates significantly higher than average,
- a few small carriers, which have a high crash rate if they have even one crash, and
- a few carriers that have underreported their miles, and have a number of crashes.

The following table presents some of the aggregate data for the carriers with crashes, as compared to the overall population of identified carriers.

	Indiana Mileage	Number of Violations	Logbook Violations	Size and Weight Violations	SCE	Ratings Factor
All Carriers	239,491	1.8	0.13	0.54	9.67	1.54
Carriers with Crashes	1,010,112	7.29	0.5	2.56	14.04	1.6
Factor	4.22	4.05	3.85	4.74		

Two factors in this table appear out of line;

- The difference in the factor of Size and Weight violations for a carrier with crashes points out that carriers with crashes traveled 56% of all of the accounted for miles, but received 62% of the size and weight violations. This may indicate that size and weight violations may have some value as an indicator of poor safety performance in carriers.
- Carriers with crashes have a higher average SCE rating. However, this appears to be a function of the difference in mileage of the carriers which had crashes. The average SCE rating for all carriers with over 250,000 miles (a group with approximately the same average mileage as the carriers with crashes) was 11.40, a bit higher than the average SCE rating for carriers with crashes.

Of the 283 carriers with crashes, 12 had less than 1,000 miles, and at least one violation.

An additional four carriers had less than 1,000 miles and at least one crash, but no violations. Because of the concerns about carriers reporting less than 1,000 miles, a separate analysis was performed eliminating the effect of these carriers.

These results are both enlightening and critical to the conclusions of this report.

Coefficient of Correlation of	All Carrier	Carriers with > 1,000 Miles
Z-Rank and Crash Rank	.623	.578
Violation Rank and Crash Rank (for carriers with violations)	.808	.771
Violation Rate and Crash Rate (for carriers with violations)	.915	.811
Violation Rate and Crash Rate (for all carriers with crashes)	.901	.766
Violation Rank and SCE Score (for Carriers with violations)	-.105	-.025
Violation Rank and Ratings Factor	-.038	-.080
SCE Score and Crash Rank	-.284	-.233
Rating Factor and Crash Rank	-.037	-.007

i. Z-Rank and Crash Rank

Z-Rank and crash rank were correlated for carriers which had crashes, regardless of whether the carriers had violations. The strong correlation between Z-Rank and crash rank (.578) is a clear indicator, within the carriers that had crashes;

"Violation history is a strong indicator of safety performance."

ii. Violation Rank and Crash Rank

This conclusion is strongly supported by the results of the correlation of violation rank and crash rank (.771).

Violation rank was only correlated with crash rank for carriers which have violations and crashes. As was noted previously, there is no significant reason to correlate ranks for zero-valued parameters.

The correlation between violation rank and crash rank was even higher than the correlation between Z-Rank and crash rank. This supports the previous observation that violations are truly critical. Once the hypothesis that violations are randomly distributed is rejected, there may be no reason to continue to use Z-Scores and Z-Ranking for statistical analysis, violation rates and violation rate ranking appear preferable. Nonetheless, so long as it is assumed that there is an element of randomness in assignment of violations, a Z- type of approach may be the fairest approach to carrier regulation.

Similarly, while rate data is appropriate for demonstration of statistical concepts, both overall rate and ranking data can be used for program development. Often, ranking is the easier tool to use.

iii. Violation Rate and Crash Rate

The correlation between violation rate and crash rate for carriers with both crashes and violations (.811) is exceptionally strong. Even when carriers which have crashes and no violations are included into the evaluation, the correlation only drops to .766.

This indicates that,

"The basic relationship between violations and crashes is extremely strong."

The additional factoring of ranking and Z-Scoring, while it does serve a number of purposes, does not contribute any additional strength to the demonstration of the relationship.

iv. Violations Rank, SCE Score and Ratings Factor

The lack of correlation (technically a very small negative correlation) between the SCE score and the violation rank, as well as the lack of correlation between the ratings factor and the violation rank, are presented to demonstrate that the information contributed by the violations data is not redundant with information used in the current (old) SCE score. In fact, SCE can be considered to be independent of violations performance. Further, the carrier ratings were not reflective of violation history.

v. SCE Score and Crash Rank

For carriers which had crashes, the SCE score was slightly, but significantly, negatively correlated (-.284) with crash performance. In other words, for carriers which had crashes, a higher SCE score is (slightly) predictive of a lower crash rate.

This is consistent with the assumption that the goal of the SCE score was to identify carriers whose poor safety performance would result in greater harm, as opposed to identifying carriers with poor safety performance.

vi. Rating Factor and Crash Rank

Within this data, there was virtually no correlation (-.007) between the rating factor and crash rank. In other words, the carrier's rating was not, in any way, predictive of the carrier's safety performance in terms of crash ranking.

Further, the average rating factor of carriers with crashes was virtually the same as the rating factor for carriers which did not have crashes (1.60 v. 1.54). Even among the population of all carriers, there is no indication that the carrier rating is associated with safety performance.

d. Summary of Conclusion based on Results for Indiana Carriers

The goal of this analysis was to ascertain if it can be determined that the lack of randomness among carrier violation rates is due, at least in part, to a difference in safety performance among carriers. In order to make this determination, three factors were used to identify safety performance;

- Carrier crash records,
- Carrier SCE scores and
- the results of carrier reviews, both Safety Reviews and Compliance Reviews.

The results clearly demonstrated that,

"If safety performance is measured by crashes, the lack of randomness in violation rates is associated, in large part, to a difference in safety performance."

There was no clear relationship between carrier violation performance and the results of carrier reviews. Further, there was a small negative relationship between violation performance and SCE scores.

However, these relationships are paralleled by the relationships between crash performance as compared to SCE scores and carrier reviews. There was no discernible relationship between carrier review results and crash performance, and a negative relationship between crashes and SCE scores.

In other words, violation performance appears to be a far stronger indicator of carrier safety performance (in terms of crash history) than either the results of reviews or SCE scores.

2. Results for Michigan Carriers

In addition to the 729 carriers which had violations, an additional 1093 carriers were identified from the profile data.

Overall, 1,822 of the 4399 carriers, or 41% were identified. These registrants accounted for 509,485,805, or 71% of the 720,085,708 registered miles in Michigan.

a. Overview of Michigan Carriers

The overall results for Michigan carriers are similar to the results for Indiana carriers. Again, there is a stark difference in the miles per carrier for carriers with violations as opposed to the miles per carrier for carriers which do not have violations. This continues to support the position that receipt of violations is, at least partially, a result of exposure.

Michigan Data from 1994	Carriers which had Violations	Carriers which had No Violations
Total number of carriers	729	1093
Average mileage for a carrier	555,328	110,075
Number of carriers with crashes	330	184
Average miles per crash	345,716	312,392

There is also a small but insignificant difference in the miles per crash for carriers which have received violations, as compared to the miles per crash for carriers which have not received violations. However, the relationship is the opposite of the one expected, the carriers with violations actually have a lower crash rate (more miles between crashes) than carriers which did not have violations. This effect is due to the fact that there are a number of carriers which have a very large number of miles and some violations, although their violation and crash rates are lower (better) than average. The recognition that this same situation also exists in Indiana makes the Indiana results even stronger.

Perhaps the most striking difference between the Michigan and Indiana data is the overall miles per crash (@ a million IRP miles per crash in Indiana, as compared to a third of a million IRP miles in Michigan.) It is not the case that driving in Michigan is three times more hazardous than driving in Indiana. Rather, Indiana indicates that only about 15% of its crashes are investigated by the State Police, although the number is undoubtedly higher for crashes involving large trucks. Clearly, Michigan's reporting is more complete.

Also, recall that this does not mean that the average large truck is a participant in a reportable crash at either of these rates. The crash totals are for a five year period, while the mileage totals are for a one year period. The actual crash rates are much lower (more miles per crash) than these rates.

b. Analysis of Indicators for Carriers in the Census

The associated information was compiled for the 1822 identified Michigan carriers.

The overall results are consistent with the results from Indiana.

- There is a solid correlation between the violation rate and the crash rate, although the correlation is not quite as strong as was observed in Indiana data.
- As is reflected in similar relationships elsewhere, there is a correlation between the Z-Score and the crash rate, although the correlation is not as strong as the correlation between the violation rate and the crash rate.

Coefficient of Correlation of	
Violation Rate and Crash Rate	0.4250
Z-Score and Crash Rate	0.2755
Violation Rate and Crash Rate for Carriers with Violations	0.4700
Crash Rate and SCE Score	0.0159
Crash Rate and Ratings Factor	-0.0029
Z-Rank and SCE Rank	-0.0323
Z-Score and SCE Score	0.0439
Violation Rate and SCE Score	0.0012

- There is no correlation to speak of between either of the measures of actual carrier performance (violation and crash rates) and the measures assigned by the carrier review process (SCE score and review results.)

c. Carriers With Crashes

514 identified Michigan carriers had at least one crash. Particular attention is paid to these carriers, as these are both the carriers for which the most information is available, and the carriers of most concern. Similar to Indiana, Michigan appears to have at least three distinguishable groups of carriers;

- a number of large carriers which, as a group, have crashes at roughly the average rate,
- a number of mid-sized (and larger) carriers, which have crash rates significantly higher than average, and
- a few small carriers, which have a high crash rate if they have even one crash.

Michigan did not appear to have the small group of carriers reporting very low mileage with a significant number of violations and crashes.

The following table presents some of the averages for Michigan carriers, comparing the statistics for all carriers to the statistics for only carriers which have crashes. The factor indicates the ratios of the two averages.

	Mileage	Total Violations	Z - Score	Logbook Violations	Size and Weight Violations	Rating Factor
Average for All Carriers	279,630	2.61	1.47	0.245	0.48	1.38
Carriers Which had Crashes	711,104	5.48	2.33	1.59	0.95	1.55
Factor	2.54	2.100		6.490	1.979	

Similar to the Indiana results, the carriers which had crashes had, on the average, far more miles than the carriers which did not have crashes. Clearly, exposure plays a large part in determining which carriers have crashes. For the carriers with crashes, the Coefficient of Correlation between mileage and number of crashes was .832.

Further, the factor for violations is lower than the factor for miles, leading to the expectation that the average violation rate for carriers with crashes is actually lower than the overall average violation rate. This is indeed the case; carriers which had crashes received one violation per 130,000 miles, while the overall rate is one violation per 110,000 miles. The single most important factor in predicting which carriers will have crashes (as opposed to their crash rates) is mileage. Further, in the aggregate, the carriers which travel more miles appear to get both fewer violations and fewer crashes. For example, Michigan carriers with over 1,000,000 miles had a crash rate of one crash per 518,000 miles, where carriers with under 1,000,000 miles had a crash rate of one crash per 238,000 miles. (Again, recall that this is an annual mileage figure, whereas the crash is data for a five year period, so the real mileage per crash rate is about 5 times the number presented here. These numbers are, however, valid for comparison.)

The average ratings factor for carriers with crashes is higher than the average rating factor for all carriers. This difference in the average ratings factor cannot be attributed to the difference in the sizes of carriers. An average ratings ranking of about 1.38 is valid across a large number of size groups, and represents the mean ratings factor. In fact, the average ratings factor drops off for higher mileage carriers. The average rating factor of 1.55 for carriers with crashes is demonstrably higher than the mean, and indicates that the carrier's Safety Rating is sensitive to the existence of crashes. [This does not mean that the carrier's Safety Rating is sensitive to the crash rate.]

In Indiana, Size and Weight violations were an indicator of crashes, whereas in Michigan there is no such relationship. Analogously, in Michigan, logbook violations appear to be a strong indicator of poor safety performance, whereas there were no comparable results in Indiana. This underscores the fact that the details of these relationships appear to be valid only on a State by State basis. As was pointed out in the first Driver/Carrier Study, the pattern of violations differs markedly among the States.

The results of the correlation analyses for Michigan carriers with crashes are presented in the table of the right. These results are consistent with the expectations. At the raw data level, there is a strong correlation between crash rate and violation rate. The relationship weakens when comparing Z-Score to crash rate. Surprisingly, the strength of the correlation returns when comparing Z-Rank and crash rank. This is particularly important, as it is often relative ranks which are used to focus attention.

Coefficient of Correlation of	
Violation Rate and Crash Rate	0.5197
Z-Score and Crash Rate	0.3710
Violation Rank and Crash Rank	0.2851
Z-Rank and Crash Rank	0.4839
Crash Rate and SCE Score	-0.0773
Crash Rate and Ratings Factor	-0.0365
Crash Rank and SCE Rank	-0.1806
Crash Rank and Ratings Factor	0.0086
Violation Rate and SCE Score	-0.0686
Violation Rate and Ratings Factor	-0.0571
Violation Rank and SCE Rank	0.0613
Z-Score and SCE Score	-0.0173
Z-Score and Ratings Factor	-0.0612
Z-Rank and SCE Rank	-0.0158

For the carriers with violations and crashes, the relationship between violation rate and crash rate is quite strong, and the relationship is reflected in the ranking.

Again in Michigan, there is no evidence that violations are in any way related to or predictive of either SCE scores or carrier Safety Ratings. However, there is also no evidence that either SCE scores or carrier Safety Ratings are predictive of crash rates.

d. Summary of Results for Michigan Carriers

The results for Michigan are exceptionally consistent with the results from Indiana,

- Violation rates are somewhat predictive of crash rates. If safety performance is measured by crashes, violation performance is related to safety performance.

- There was no relationship between carrier violation performance and either SCE scores or the results of carrier reviews.

- There was also no relationship between carrier crash performance and either SCE scores or the results of carrier reviews.

There was no clear relationship between carrier violation performance and the results of carrier reviews. Further, there was a small negative relationship between violation performance and SCE scores.

These relationships are paralleled by the relationships between crash performance as compared to SCE scores and carrier reviews. There was no discernible relationship between carrier review results and crash performance, and a negative relationship between crashes and SCE scores.

In other words, violation performance appears to be a far stronger indicator of carrier safety performance (in terms of crash history) than either the results of carrier reviews or SCE scores.

D. Conclusions From Data Analysis

The goal of this analysis was to determine whether the difference in violation rates among carriers is due, at least in part, to a difference in safety performances. In other words, the objective of the analysis was to determine if violation rates, or other violation derived statistical measures, are predictive of, or correlated with, safety performance indicators.

Three safety performance indicators were used for this study.

- SCE score, a score that is derived by FHWA to determine which carriers should be subject to review (a closer scrutiny),

- Carrier review results, the score that is assigned to carriers as a result of their reviews, and

Chapter 5 - Alternative Statistical Indicators

A. Introduction

1. Project Objectives

The primary goal of the Driver Citation/Carrier Data Relationship project is to determine the value and uses of identifying the responsible motor carrier on citations to drivers of commercial motor vehicles. Phase I of the Driver Citation/Carrier Data Relationship project addressed the issue of whether the State data indicated that there is a difference in driver performance, based on the employing carrier. This was based on citation information from the State Police and mileage information from the State IRP program. The conclusion was that, based on the data at hand, a difference does exist.

Earlier in Phase II, the same analysis as was performed in Phase I was applied to an additional year's data from Indiana and Michigan. The results of Phase I were revalidated (in Chapter 3).

Additional data related to Safety Ratings, SCE scores and crashes was received from FHWA. Along with the data from Michigan and Indiana, the question was raised as to whether the difference in violation patterns is in some way associated with a difference in safety performance. The conclusion was that there is reason to expect that carriers with higher violation rates also have higher crash rates.

The question then becomes, is this actionable? If the carrier was identified on citations, is there some way that this information could be used to identify potential problem carriers? This is the issue that will be addressed in this chapter.

2. Chapter Objectives

This chapter will follow the path of the Work Group as it reviewed the methodology that was used for research, and show why the Work Group determined that the research methodology could not be used in a production environment. The Work Group then relied on State Police experts to identify an alternative method. The chapter will discuss how the alternative approach was developed, and then provide an explanation of why it is a sound statistical approach. Finally, the benefits that this approach can offer will be reviewed.

This chapter does not require any knowledge of statistics, and is presented in a non-technical manner.

B. Initial Plan -- Similar Methodology

The initial plan was to use a methodology similar to the one that was used in the statistical analyses.

The Work Group originally went through several attempts to determine an appropriate metric for rating or ranking carriers for statistical purposes.

- The first obvious choice was simply the number of violations for the carrier. However, this was clearly not equitable. A large carrier will obtain a number of violations, based on mileage. A rate of some type was clearly required.

- The second suggestion was to use violation rate. This encountered several initial problems. The first problem was that there was no available measure of exposure, i.e. violations per something? For purposes of analysis, IRP mileage was chosen as the best available measure. Carrier violations were manually matched against IRP mileage, to obtain mileage rates.

- There was a second issue with violation rates - they do not take into account differences in size. For example, suppose that the average rate of violations was one per hundred thousand miles. A carrier which had one violation in 15,000 miles would probably not be of great interest, but a carrier which had 10 violations in 150,000 miles would certainly be of interest. Yet both have the same violation rate.

- The Work Group instead used a statistic called the Z-Statistic, which measures how far the result was from the expected result. For round numbers, a Z-Score of greater than 2 indicates it is likely the deviation from average is a result of randomness.. In the above noted case, the Z-Score of the first carrier would be .85, and the assumption that this is an average carrier could not and should not be rejected. In the second case, the carrier's Z-Score is 8.5, and it is clear that this carrier gets violations at a significantly higher rate than average.

1. Strengths

There were several reasons to expect that the Z-Statistic would prove useful in a production environment.

a. Demonstrated Statistical Methodology

The Z-Statistic proved to be instrumental in performing the statistical analyses that demonstrated that some carriers receive violations at a far higher than average rate. Although the explanation is cumbersome and the mathematics complex, the Work Group was able to fully automate the calculation. This approach is more fully explained in Appendix B.

b. Ease of Ranking

A particular benefit of the Z-Score is that it provides a reasonable way to rank carriers, based on deviance. It also presents a fair way of ranking large and small carriers together. The ability to rank is particularly important in targeting for enforcement, as it presents prioritized targets.

2. Weaknesses

Unfortunately, the weaknesses that would be encountered in using this data in a production environment overshadow the strengths.

a. Difficulty in Obtaining Data

Because the data was easy to obtain, citation data was chosen instead of conviction data for the study. It was found by the States that citation data, and in particular, data about citations issued by the State Police, tended to be centralized within each State. Conviction data may be distributed in courts throughout the State. The ease of obtaining the centralized States data made citation data the preferred choice of the States.

However, there was no good mileage data available for use with the original methodology. The project used IRP mileage. Both Michigan and Indiana were able to provide the IRP listings on a floppy disk, for entry into a data base. However, IRP did not capture carrier identification. Consequently, there was a manual matching process.

Clearly, matching the name on IRP to the carrier name was far from a straightforward task.

There are commercial name matching programs available. Research indicated that these also would not be effective. These programs depend on address to help resolve problems. The address on the IRP account is often a terminal. The address on SAFETYNET/ MCMIS is often the address from the MCS-150, which is the central office address, and can be decades out of date.

If this methodology were put into production, it would depend, at least in part, on manual name matching. This would result in a situation where results would not be repeatable. The lack of consistency and repeatability could cause significant problems in the future.

b. Concerns About Quality of Data

The States raised concerns over the quality of exposure data.

IRP mileage is adequate to use for statistical studies. However, it has significant problems with identifying and ranking carriers for review in a high-volume production environment.

- There is a name matching problem. Manual carrier name matching was performed on registrants during data collection. This can be a somewhat arbitrary process. This would not be efficient in a production environment.

- Mileage data may be inaccurate. In many cases, IRP mileage is estimated mileage, which can skew results. Further, The data is over a year old. The operation of the carrier may have changed. If a carrier's operation grew during a year, the mileage that was reported will be lower than the carrier's actual mileage, resulting in a higher violation rate.

- This mileage is for registrants. The registrant may not be the carrier. A leasing company could have registered the vehicle and reported the mileage. The leasing company is not the carrier that was actually responsible for the vehicle when it was transporting the load.

- Companies can have multiple IRP accounts. Some carriers have IRP accounts for each terminal. Some carriers use IRP accounts based on profit centers in their business. An IRP account is supposed to be set up where the carrier's trucks are based.

- Nationwide carriers seem to have inappropriate mileage and violation rates. These carriers may have a terminal in the States. The IRP mileage that is reported is only for that one fleet. But the carrier could have vehicles from many fleets operating in the State. Consequently, their violation rates, as calculated for this study, can be much too high.

The approach that uses IRP mileage as a proxy for exposure has some significant problems for use in a production environment. The Work Group indicated that IRP mileage was not acceptable for use in an actual enforcement environment.

Even if these concerns could be addressed, IRP mileage is operationally unfeasible. The required manual matching of registrants by name is unworkable in a production environment. The age and reporting requirements of the IRP mileage data also render it inappropriate.

c. No Alternative Sources of Exposure Data

The use of other measures of mileage was researched. There were no other available sources of exposure data.

- IFTA mileage was considered, but rejected, based on the laws associated with IFTA reporting and the existing business practices. IFTA mileage is an even less accurate description of carrier mileage than IRP mileage. No other possible measure of mileage was available.

- The Number of Power Units is the number of vehicles that a carrier has in its fleet. The only exposure data currently collected for many carriers is the number of power units in every fleet. This data was collected on the MCS-150 form that the carrier files at the time it applies for a DOT number. For many carriers, this form was filed decades ago. The figure is updated in a Safety or Compliance Review, if the carrier has had a review. Again, this data may be very old and unreliable.

  Note: The Canadian province of Ontario does maintain data on the number of power units. Ontario carriers have a reporting requirement to update the province's data when the size of their fleet changes. Ontario uses number of power units as a measure of exposure. The project did consider doing analysis on Ontario data. However, a strike of the provincial workers during the time of this study made this unfeasible.

Outside of Ontario, the use of power units as an accurate measure of exposure is not feasible.

- The Number of Drivers that a carrier employs was considered. Like the number of power units, This data was collected on the MCS-150 and updated only by reviews. It is aged and unreliable.

3. Recognition of the Unsuitability of the Approach

There appears to be no reasonable source of exposure data for a production environment. Therefore, the use of a Z-Score based rating or ranking had to be rejected.

C. Search for Alternative Approach

The objectives for Phase II of the project required that some statistical indicator had to be found that could be used to identify carriers which are more likely to have less than satisfactory safety performance, based on violation performance. However, there appeared to be no obvious solution.

The State experts were asked to assist in identifying an alternative approach. Various reports were created and supplied to the States. The State experts were asked to review these reports, in hope of identifying some useful pattern. Meetings were held in Michigan and Indiana to review the reports that were produced.

The Work Group hoped that it was moving towards an appropriate approach.

1. Expert Review

a. Produced Reports on Analysis by Carrier

In order to review the problem carriers, the States wanted reports to list possible problem carriers. Reports were produced that contained carrier name and Z-Score based on a variety of criteria including mileage ranges and types of violations. The reports usually displayed carriers ranked by Z-Score (a higher positive Z-Score was considered bad).

b. Produced Reports based on Multiple Violations for Drivers

The data related to drivers with many violations, especially drivers with many violations for more than one carrier, had been an area of high interest during Phase 1. Reports were delivered to the States to reflect this interest.

The first set of two reports was produced. One was called *Drivers With 6 or More Violations Sorted by Driver's License Number (DLN).* Violation records were summarized by driver. The driver records were selected for drivers who had six or more violations. These records were compiled. The data was sorted by Driver DLN. The second report was *Drivers With 6 or More Violations Sorted by Carrier.* Violation records were summarized by driver. The driver records were selected for drivers who had six or more violations. These records were compiled. The data was sorted by carrier name. This report cross referenced the preceding report.

Another report listed driver violation records for carriers with twenty or more violations. For each carrier, it displayed all drivers who had violations. And for each of these drivers, it displayed all of the carriers that the driver had violations for.

Another report included drivers who had received three or more violations for two or more carriers.

c. Visited the States

Meetings were scheduled to review these results with the MCSAP Officers. Meetings were held in Michigan and Indiana to review the project data and the reports. During these meetings the States were continually interested in drivers and carriers which had a high number of violations.

2. Concept

A potentially valuable concept was developed with First Lieutenant Lisa Jacobs and Sergeant Sharon Van Campen of the Michigan State Police. The concept is presented below.

a. Reviewers were Attracted to Violations Per Carrier

A review of the reports pointed out that some carriers had a low number of violations, and others had many violations. The Work Group was looking at total violations per carrier. The officers identified a fairly simple concept.

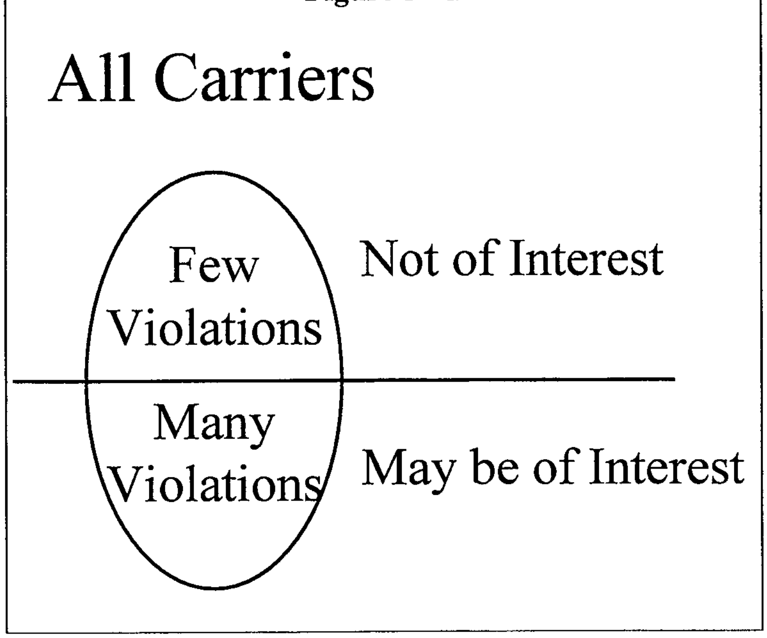
- Carriers with a low number of violations are not of interest.
- Carriers with a high number of violations *may* be of interest.

This is illustrated in Figure 5-1.

The oval represents the entire population of carriers. Some of the carriers have few violations and some of them have many violations.

i. Carriers with Few Violations are Not of Interest

Figure 5 - 1



Carriers with few violations are not of interest. If the carrier is not good, it is too small to tell. It is impossible to tell anything about the carrier's safety performance. These carriers with few violations, are not regarded as crucial to targeting safety performance.

ii. Carriers with Many Violations May be of Interest

While, not all of the carriers with many violations were carriers of interest, the carriers of interest all had many violations. The carriers of concern are in this second group.

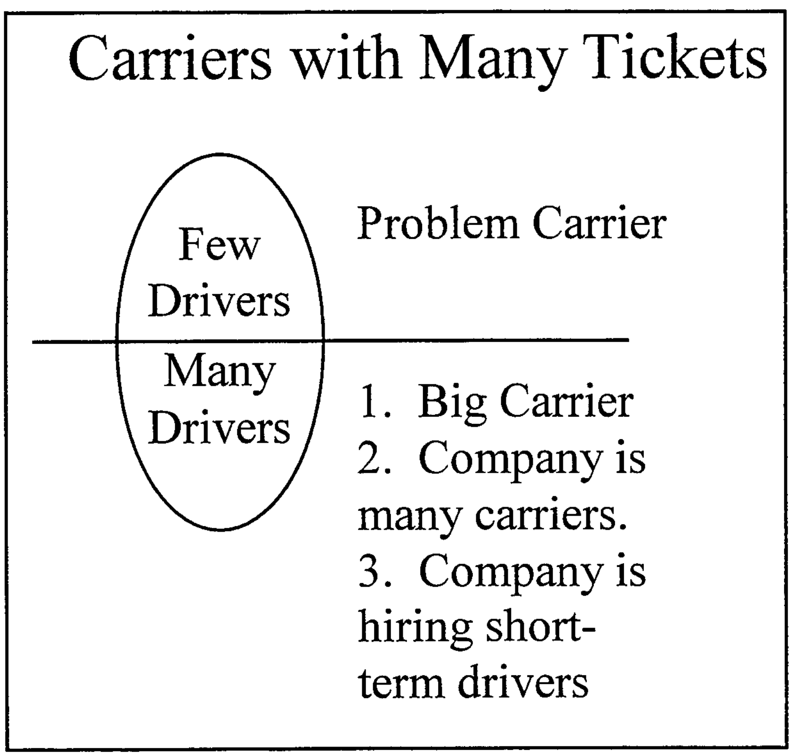
b. Carriers with Many Violations

To determine if carriers with many driver violations are of interest, examination of the carrier's drivers who received the violations was necessary. This group was further divided into,

- carriers which had many violations on a few drivers, and
- carriers which had many violations on many drivers.

i. Many Violations/Few Drivers - High Violations Per Driver

Figure 5 - 1



If the carrier has a large number of violations spread among the same few drivers, there is a high rate of violations per driver. The point of this study was that high driver violation rates point to carriers of interest or concern. Carriers that employ these drivers are therefore of interest. The high violation rate per driver was a sufficient reason for the interest in these carrier's safety performances.

ii. Many Violations/Many Drivers - Low Violations per Driver

If the violations are by many drivers, this means there is a low rate of violations per driver. There may be various reasons to explain a low rate of violations per driver. Some of these carriers may be of interest while others may not. The next objective was to determine what the differences were, and how to distinguish among them.

c. Carriers with Many Violations for Many Drivers

Three general categories of carriers with many violations for many drivers were identified. These classes of carriers are discussed below. Like any general classification, these are guidelines. There is no suggestion that every carrier fits into one of these classifications.

i. Big Carriers

A large carrier may have many violations. It would also have many drivers. In this project this situation obtained the code name "UPS". UPS is considered an example of a company that has an excellent safety record based on a high degree of safety awareness within the company. Nonetheless, UPS drivers will pick up a large number of violations, based on the very large number of miles driven, and the underlying fact that exposure is, to some degree, a contributing factor to receiving violations.

One of the goals of this project was to avoid identifying a UPS-type of carrier as a carrier of interest. Clearly, based solely on number of violations, UPS would be identified as a carrier of interest. It was the goal of any statistic identified or created in this project to avoid identifying the UPS-type carrier (i.e. a large carrier with good safety performance).

The UPS-type of carrier has among the safest drivers on the road. This is evidenced by the fact that UPS has low Z-Scores, even though we can assume that UPS trucks from several out-of-State fleets travel a high number of miles in any given State. UPS has fairly strict rules for its drivers, and assures that its drivers do not drive too many hours. Therefore, not only will UPS' drivers have a low rate of violations, but the drivers who drive for UPS will not get violations while driving for other carriers (since they don't drive for other carriers.)

The key is that these drivers will not have violations for other carriers. Drivers for this type of carrier drive for only this carrier. It is a stable work force. Carriers with a low rate of violations per driver, whose drivers do not have violations for other carriers, are not of interest.

ii. One Company Consists of Many Carriers

This carrier may be a larger company that has and uses more than one USDOT Number. It may be a company that has various subsidiaries. These are various reasons why this carrier has multiple business entities.

One reason for multiple business entities may be to avoid enforcement for unsafe practices. For these companies, violations for their drivers will be spread throughout the business entities. There may also be other business reasons for multiple business entities that are not related to safety performance. There is certainly no legal reason why a company should not use more than one USDOT Number.

From a safety point of view, the real interest is the overall business. In fact, a review of the actual carrier data shows a number of businesses use more than one USDOT Number, which show up as potential problem carriers under a number of the criteria that are discussed in this report, and have drivers who have received violations for more than one of the sibling companies. Existing measures tend to be aimed at the apparently large and bad companies. A company with 60 violations will come to the top of a list, where three companies with 20 violations each tend to be ignored. In many cases, the individual companies have satisfactory Safety Ratings, and are generally ignored. But, when considered together, they add up to a problem company.

The key is that this carrier's drivers are picking up violations for other carriers, generally the sibling company.

iii. Company is Hiring Short-Term Drivers

Not all drivers are employed full time, there are any number of drivers that are employed on a short term basis. In this case, short term can mean anything from a single trip to several months. Usually, these drivers are owner-operators; the carrier hires both the driver and the vehicle. However, the opposite is not true, there are many owner-operators who have long term relationships with carriers.

The hiring practices of these carriers vary. Certainly, there are some carriers which make all of the appropriate checks and assure that they are hiring safe drivers when they hire short-term drivers. However, based on the difficulty in checking on drivers, based on the general attitude of some of these carriers, and based on the fact that there is some indication that drivers with poor driving records are less expensive to hire than drivers with good records, some of these carriers are less careful when selecting drivers for short-term assignments.

The Michigan Officers spoke of a place called "Truck City." Truckers stop there if they need a load. Load brokers are there recruiting drivers for their loads. This practice is also known as Trip Leasing. A driver may get a load fairly easily or the driver may wait, awake, for days looking for a load. The driver may not even know the carrier. The load broker does not always check the safety history of the driver. The driver is on the road, returning home with the new load.

This situation is a potentially dangerous one. Unsafe and tired drivers are able to get loads to haul. Shippers use brokers to move their loads. The carrier may not even be apparent in the paperwork.

In fact, many of these carriers skipped detection in our Phase I analysis. These carriers do not register their vehicles, so they do not have IRP miles. They count on the owner-operator to register the vehicle. A registration based detection or sanctioning scheme will be unsuccessful with these carriers.

Further, the drivers for these carriers will rarely get more than one violation while working for the carrier. The carrier may have a very low number of violations per driver.

The key here is that the carrier's drivers will have violations for many carriers.

d. Conclusion

It is useful to look at the rate that the carrier's drivers are picking up violations for the carrier. But it is not enough. While there is obviously a problem with carriers whose drivers receive a large number of violations, the fact that the carriers' drivers did not receive very many violations, while driving for that carrier, does not necessarily mean that the carrier has good safety practices. It is necessary to look at the carrier's drivers' overall records. The fact that a carrier's drivers have violations for other carriers is a reflection on that carrier.

3. Developing a Statistical Approach

From this understanding, the Work Group was ready to develop a statistical approach. The Work Group reviewed this concept until a statistical approach suggested itself. The Work Group then developed the justification for this approach. This section is presented in a more analytic manner. The section more clearly identifies and specifies the fundamental concepts that were hinted at in the prior section. Next, the basic approach is presented, and finally the actual statistical measure is defined.

a. Fundamental Concepts

There are three fundamental concepts underlying the statistical approach.

- The driver-year is a solid measure of exposure.
- Driver quality really does reflect on the employing motor carrier.
- The measurement of a rate, particularly violations per unit of exposure, meets the primary goal of the project.

Each of these concepts is discussed below.

i. Driver-Year as a Measure of Exposure

Earlier, a driver-year was identified as a desirable measure of exposure. Simply put, in a year, a driver drives a driver-year. In many ways, driver-year is an ideal measure of exposure. A driver who drives in the city will have fewer miles than one who drives on the open road. The higher danger per mile in the city indicates that time, rather than distance, may be the appropriate measure of exposure. Further, a driver-year may be a better measure of exposure than the number of power units. If a power unit is driven in shifts, it will have a higher exposure. It will also have a higher number of driver-years.

Federal regulations, as well as the physical limitations of the human body, limit the number of hours that a driver can work in a day, week or year.

Overall, if it could be accurately measured, driver-year may be the best possible measure of exposure.

There is an implicit measure of a driver-year within the data that was collected for the driver-carrier project. Within the States in the study, the project can identify tens of thousands of drivers who drove during the years in question. For statistical purposes, the project proposes to assume that, if a driver received a violation during a year, the driver drove for a driver-year during the year under study.

It is recognized that there will be some situations where a driver did not drive an entire year. However, these will be relatively few, and the result of overcounting exposure will favor the carrier.

It is also recognized that there were many drivers, the majority of drivers in the States in the study during the data collection period, who did not receive any violations in the States in the study during the study year. Presumably, there is a tendency for the better drivers, working for the better carriers, not to have received violations. Better data would only more effectively emphasize the difference between carriers.

ii. Driver Quality reflects on the Carrier

High violation rates per driver may indicate poor safety performance and bad hiring practices. This is the basic assertion of the Driver/Carrier project. If drivers are receiving a high number of violations, there is not necessarily a safety problem, but the data suggests that it is more likely.

- It has been demonstrated, in Chapter 4 of this report, that there is a relationship between driver violation performance and safety fitness. When a driver receives a citation, it does reflect on the safety fitness of the carrier.

- Carriers who hire bad drivers, i.e. drivers with poor performance records, are also compromising safety fitness. While this report does not address this assertion, it can be supported in two ways. First, it has legislative support. The Commercial Motor Vehicle Safety Act (CMVSA) of 1986 makes specific reference to the requirements for carriers to check the driving record of drivers that they hire, and to not hire drivers with unsound records. Second, it makes intuitive sense; carriers which compromise in one area of safety and compliance are likely to make compromises in other areas.

The States have expressed a particular interest in the drivers with high numbers of violations, or high violation rates (i.e. violations per driver year). These violations not only reflect on the driver, they reflect on the carrier that has hired the driver.

iii. Goal is Violations per Unit of Exposure

In order to statistically analyze or compare a set or sets of data, it is necessary to find a common unit of measure. Often this represents some type of rate. Research was performed to find any usable forms of comparison.

There is no question that violations are the item being measured. However, the question is violation per what exposure unit.

Mileage was chosen to be used as primary exposure for the Driver/ Carrier Study analysis. However, after extensive research, it was found to be unacceptable for use in actual operations. Driver-year appears to be the strongest exposure available for targeting carriers with poor safety fitness. Further, since Z-Score will not work without mileage data, some form of violation rate seems to be the best available measurement tool.

b. Basic Approach - Treat Each Carrier as if Each Driver had Driven the Whole Year for that Carrier

The problem still remained; how to deal with the three general situations (big carriers, one company consisting of many carriers, and a company hiring short-term carriers), and the actual cases these general classifications represent, in order to make sense of the variety of industry practices.

Given the proposition that a driver-year will be used as a measure of exposure, the idea of using violations per driver for the carrier has initial appeal. However, paradoxically, this approach rewards exactly the wrong behavior. The carrier/broker in Truck City will have a rate of close to one violation per driver, since the carrier employs so many drivers for such a short period of time. The company that is disguised as many carriers will also score well, since its drivers' violations will be spread around several companies.

There is a fairly simple way out of this dilemma - treat each carrier as if it employed each of the drivers that worked for it as having worked for it for the whole year. In other words, for ranking and evaluation purposes, for each driver which worked for a carrier for any part of the year, assign to that carrier;

- a driver-year worth of exposure for that driver's year, and
- all of the violations that the driver received during the year, regardless of the carrier that was employing the driver at the time.

This appears to be a fair way of assigning violations to carriers.

- This is eminently fair to the large stable employer; each of its drivers did work for it for the entire year.
- It is also a very fair way to assign violations to the company which is operating as more than one carrier. Essentially, for analysis purposes, this treats the company as one company. (Separate statistics will be generated for other sibling companies.)

- This seems reasonably fair to the companies in Truck City. As was suggested in the CMVSA of 1986, a company is responsible for the practices of its drivers. This is a way to spot the more honest companies in Truck City. (The less honest ones will disguise the identity of the carrier on the shipping paper and cannot be caught without some form of law enforcement act.)

It has been pointed out that there is one situation where a company could be unfairly assigned violations. A company could hire a driver with a clean record. During the year, the driver receives a violation for the company, and later still the driver's employment terminates. If the same driver gets additional violations during the year, it would reflect on the original employer. However, this is expected to be a relatively rare occurrence, and certainly does not invalidate the concept.

c. Statistical Approach - Carrier's Drivers' Violation Rate

The new statistic is called the Carrier's Drivers' Violation Rate (CDVR). The calculation is shown in Figure 5-3. The numerator is the total number of the carrier's drivers' violations. The denominator is the total number of drivers for that carrier.

A carrier's driver is defined as any driver who received a violation for the carrier during the period of evaluation. Since the underlying assumption is that the carrier will receive credit for a driver-year for every driver associated with the carrier, the denominator is, in effect, the carrier's assigned driver-years.

A carrier's drivers' violation is defined as any violation received by the carrier's driver, regardless of what carrier was associated with the violation. (Another way of looking at this statistic is that it is the average number of violations for each of the carrier's drivers.)

Carrier's Drivers' Violation Rate

$$CDVR = \frac{\text{Total Number of Carrier's Drivers' Violations}}{\text{Total Number of Drivers for that Carrier}}$$

Figure 5 - 2

This approach has the two key elements necessary for implementation as a tool for ranking and evaluating carriers.

- There is reason to expect that it will provide the capability to effectively rank carriers in a way that is associated with safety performance.
- It can be calculated from data which can be reasonably obtained.

The following section will discuss key benefits which this statistical approach offers. Chapter 6 presents a review of the approach, to determine if there is evidence to suggest that it really is a valid way to identify potentially problem carriers.

D. Expected Benefits of Carrier's Drivers' Violation Rate

There are a number of expected benefits to be found by using the Carrier's Drivers' Violation Rate to identify the problem carriers.

Among the benefits;

- CDVR matches its methodology to existing business practices,

- the calculation of CDVR is independent of carrier mileage, and
- CDVR can be easily implemented incrementally.

All of these are significant benefits. Each of the preceding benefits has an expected importance to the CDVR. The importance of each will be explained.

1. Matches Methodology to Business Practices

CDVR directly matches the methodology to the underlying business. While there are certainly examples of effective measures that are not related to the underlying business practice, it always appears to be more effective when a statistical measure is clearly related to the item that is being measured. This is the case with CDVR:

CDVR rates a carrier based on its drivers on-the-road performance.

2. Independent of Mileage

CDVR does not depend on mileage. Since use of mileage has a set of concerns associated with it, use of mileage can be avoided. Collecting the mileage data is very time consuming. Matching carrier names to registrants and determining if the data is correct is virtually impossible.

In the highway safety industry, there are many measures of fitness and safety that use mileage in rates. However, in the current business and regulatory environment;

- There is no effective way of determining carrier mileage. This information is simply not collected anywhere, much less stored in a retrievable manner.
- There is no efficient way of gathering the information. The two systems that do gather mileage information (IRP and IFTA) for interstate carriers do not gather information at the carrier level and do not keep information about the carrier.

Under CVIS, there is some consideration being given to have IRP mileage associated with a USDOT number. To even consider using IRP mileage, it would be necessary to gather data from all 48 contiguous States (and the District of Columbia). Since carriers register their vehicles at the fleet (terminal) level, IRP data from every State is necessary to determine the registrant's actual total mileage in a State. This data is currently exchanged on paper. There is a current project which is aiming to automate the electronic exchange of this data.

Even if all 48 contiguous States could electronically exchange data, and even if all 48 States kept track of a USDOT number at the point of registration, IRP mileage for a carrier would still be suspect, unless and until there is some way to track the leasing of vehicles and to assign the responsibility and the mileage to the lessee.

Therefore, for an implementation in the near future, it would be highly desirable to use some measure of exposure other than mileage.

CDVR offers this opportunity.

3. Easily Implemented Incrementally

A key benefit of the CDVR is that it can be implemented incrementally. As will be demonstrated in Chapter 6, a valid CDVR can be calculated for Indiana or Michigan today. It does not depend on SAFETYNET, although the inclusion of SAFETYNET data will certainly add to the results. Any State which collects carrier identification on citations, such as California, Idaho, or North Dakota, could calculate the CDVR based on current data.

Further, CDVR can become valuable incrementally. As data is added from more States, the quality can be expected to improve. Violations of drivers from any State can be assigned to the carrier in the base State.

There may also be value in integrating Commercial Driver License Information System (CDLIS) data. Once a carrier is identified (as being associated with a driver), it should be possible to query CDLIS to obtain added information about the driver's record.

These benefits are significant to the objectives of the project, to determine the value and uses of identifying the responsible motor carrier on citations to drivers of commercial motor vehicles. Targeting of unsafe carriers may now be possible using the CDVR.

This chapter has introduced the CDVR, and presented some possible benefits and uses. The goal of the next chapter is to more rigorously review and evaluate the CDVR. This goal has been achieved by using expert review and statistical analysis. Some parts of the next chapter will present more complex statistical concepts than were presented in this chapter.

Chapter 6 - Validating Carrier's Drivers' Violation Rate

The goal of this chapter is to review and evaluate the newly identified statistic, the Carrier's Drivers' Violation Rate (CDVR). The Work Group performed this evaluation through two methods, expert review and statistical analysis. The results of this evaluation are presented below.

A. Expert Review

The goal of the expert review was to validate the Carrier's Drivers' Violation Rate (CDVR) through expert review of the results of carrier ranking on CDVR.

1. Method

a. Produced Reports

As a result of the meeting with the Michigan State Police Officers, the key review report was developed. The report was titled *Carriers With 20 or More Violations*. The violations were totaled by driver using DLN (as in previous reports.) For each carrier, all of the drivers' violations were totaled and divided by the number of drivers that had violations for that carrier. This produced the Carrier's Drivers' Violation Rate. To get the average violations per driver for each carrier (Carrier's Violation Rate - CVR) the total number of violations that drivers got for that carrier was divided by the number of drivers who received violations for the carrier.

The report listed the carrier and then all of its drivers. The report was sorted in descending order of CDVR. For each carrier, the report listed the carrier's name, driver count and the drivers that the carrier employed. Further, for each driver who had a violation for the carrier, the report listed all of the carriers this driver got violations for and the carrier's street address.

Using this report,, it was possible to see every carrier that the carrier's drivers had worked for, grouped together. This allowed identification of situations where

- a carrier might have changed its name (and USDOT number), or
- one company may have been presented as more than one carrier, with different names and/or addresses, or

- One company may have employed the same driver at more than one of its subsidiary carriers.

A review of the report clearly demonstrated several cases of closely related or interrelated companies.

Any carriers whose drivers had only one driver and one violation were excluded from the report. This meant that the CDVR and the CVR were both equal to 1. This report became a key to the analyis for the remainder of the project.

This report will be referred to as the "list" or "listing" for the remainder of this chapter. It is a list of carriers in descending order of CDVR. The worst carriers were suspected to be at the top.

b. Listing Used in Review in Indianapolis

Originally the listing was presented to the Work Group at the May 1996 meeting. The first and last 20 pages of the report were reviewed. If an officer knew or recognized a carrier, they spoke out. As the States members spoke out, the comments were recorded. The States were so interested in the report that, if no comment was made, they indicated that they would have a subsequent interest in the carrier.. During the comment period, the States also mentioned what business the carrier was in. Their familiarity with these carriers was quite clear.

The States representatives observed that carriers who worked in certain industries seemed to practice similar levels of safety fitness. The predominant industries for unsafe practices were steel and waste hauling for the States under study. Many of the carriers who made the top of the list were in these industries. Several of these companies had the same drivers drive for more than one of these companies.

Starting with one identified bad carrier, and tracing through all the carriers that its drivers drove for starts create a web of bad drivers and bad carriers. This web can be continued through connecting the second carriers drivers to still other carriers.

2. Results

a. Top Twenty Carriers

The results from the top twenty carriers were included to demonstrate the CDVR statistic and its uses. These carriers will be referred to as A, B, C, etc. A table will be used to easily display these results. Comments noted were based on the States' experiences with the carrier.

Carrier	CDVR	CVR	Current Safety Rating	Comments
A	6.667	6.667	none	This carrier was the subject of an investigation as a result of a complaint. The carrier received numerous logbook violations during investigation. The carrier had no other known problems.
B	2.900	2.900	none	The State recognized this carrier as questionable.
C	2.778	2.778	none	The State was not familiar with this carrier.
D	2.625	2.625	none	The State recognized this carrier as being a particularly bad carrier.
E	2.563	2.375	none	The State was familiar with the carrier but did not state whether or not the carrier was good or bad. The State did note that one of the drivers drove for another carrier known to be bad.
F	2.538	2.308	none	The State recognized this carrier as being a bad carrier.
G	2.500	2.500	none	The State commented that this carrier currently may be the subject of a hearing.
H	2.385	1.538	none	The State recognized this carrier as being a particularly bad carrier. The State also noticed that all of the drivers also drove for other subsidiaries. All of the subsidiary carrier names were recognized as bad. This was a company that was spreading its violations among several carriers. This had apparently avoided detection under Federal algorithms, although the carriers were known to State officers.
I	2.300	2.300	none	The State did not recognize this carrier.

Carrier	CDVR	CVR	Current Safety Rating	Comments
J	2.250	2.250	S	The State recognized this carrier as being a bad carrier.
K	2.136	1.591	none	The State recognized this carrier as being a bad carrier. The State also noted that the carrier had undergone a name change.
L	2.129	1.484	none	The State recognized this carrier as being a bad carrier. This carrier also appeared as a possible problem carrier in the discussion in the Phase I reports.
M	2.115	1.577	none	The State recognized this carrier as being a bad carrier, and noted that this carrier is currently under investigation, and has been for a long time. The State also noticed that the drivers for this carrier received violations for numerous other carriers that were considered bad. These carriers were said to all be the same company.
N	2.091	2.000	S	The State recognized this carrier as being a bad carrier and also noted that it thought the carrier was under investigation.
O	2.000	2.000	none	The State recognized the carrier as being an exceptionally bad carrier.
P	2.000	1.944	S	The State noted that this carrier tows for the State police. The State said that some officers may not like this carrier but that there was no indication that this carrier is particularly bad.
Q	2.000	1.933	S	The State recognized this carrier as being a bad carrier.
R	2.000	1.923	none	The State did not recognize this carrier.

Carrier	CDVR	CVR	Current Safety Rating	Comments
S	2.000	1.909	C	The State recognized this carrier as being a particularly bad carrier. This carrier was also mentioned in the Phase I reports.

Table 6-1

Overall it was clear from the review of the top 20 carriers on the list that most of these really were known problem carriers.

- There was a high recognition rate. The States knew these carriers, and knew them to be problem carriers.
- For the few that were not known carriers, the States had a high degree of interest in finding out about these carriers.
- In several cases, it was mentioned that "the Feds" were not on to these carriers, it was the State enforcement people who have targeted the carriers.

The current Safety Rating is listed, an "S" means satisfactory, a "C" means conditional and a "U" means unsatisfactory. If there is an "none" in the table, there was no Safety Rating available in the data collected.

b. For Selected Carriers at the End of the List

The results from the bottom twenty carriers are included to further demonstrate the CDVR statistic and the benefit of not targeting large safe carriers. These carriers will be referred to as AA, BB, CC, etc.

Carrier	CDVR	CVR	Safety Rating	Comments
AA	1.008	1.008	none	The State recognized this carrier as an excellent carrier.
BB	1.024	1.024	none	The State recognized this carrier as a well-known UPS-type carrier and has a good reputation.
CC	1.029	1.029	none	The State recognized this carrier as being a good carrier.
DD	1.034	1.034	none	The State did not recognize this carrier.
EE	1.037	1.037	none	The State did not recognize this carrier.
FF	1.037	1.037	none	The State did not recognize this carrier.
GG	1.038	1.000	none	The State did not recognize this carrier.
HH	1.039	1.020	none	The State recognized this carrier as being a good carrier.
II	1.040	1.020	none	The State did not recognize this carrier.
JJ	1.040	1.020	S	The State recognized this carrier as being a good carrier.
KK	1.040	1.040	none	The State did not recognize this carrier.
LL	1.042	1.000	none	The State did not recognize this carrier.
MM	1.042	1.000	S	This carrier has a big operation near Cincinnati and has had problems in the past. The low CDVR is probably based on the low percentage of miles that each individual driver spends in Indiana.
NN	1.043	1.000	none	The State did not recognize this carrier.
OO	1.043	1.043	none	The State did not recognize this carrier.
PP	1.043	1.043	none	The State did not recognize this carrier.
QQ	1.045	1.000	none	The State did not recognize this carrier.

Carrier	CDVR	CVR	Safety Rating	Comments
RR	1.045	1.023	none	The States recognized this carrier to be a large national carrier with questionable practices. It is likely that this carrier had many drivers come through Indiana and Michigan. These drivers may have had a poor national level record, but they have very few miles, and violations, in Indiana and Michigan.
SS	1.045	1.045	none	The State did not recognize this carrier.

Table 6-2

There were many well known names in the last 20 pages of the report. Many were recognized as national carriers. A key portion of this approach is that it does not inappropriately identify good carriers for further scrutiny.

3. Conclusions of Expert Review

The expert review has been summarized in Tables 6-1 and 6-2. There were two main conclusions as a result of the States' review of the list. The first was that it effectively identified bad carriers and the second is that it was better than other lists. The experts were interested in reviewing the carriers that one of them was familiar with.

- Michigan has been approved to use this list for selecting carriers for further review in the upcoming year in place of SCE score rankings. An expert from Michigan mentioned that these are problem carriers that haven't been identified by the SCE process.

- This list appears to identify bad carriers more effectively than other listings, including the Z-Score listings from this project, Compliance Review results and SCE scores. The State agreed that the list provided in this study was better than any other targeting process developed so far.

B. Statistical Analysis

1. Goals of the Statistical Analysis

The statistical analysis had specific goals.

- The first goal of the statistical analysis is to demonstrate that there is a relationship between the CDVR and crash rates. This goal will reveal the strength of CDVR for targeting carriers for review. This relationship will be analyzed for various groups of carriers.

- The second goal is to determine if the CDVR statistic is a better indicator than others.

Ranking was the primary tool used to compare the crash rate with selected indicators. Ranking of carriers gives a list that may be ordered by the carriers level of safety fitness. Other methods for targeting carriers use rankings. An example is the SCE score. This is a ranking for carriers so they may be selected for review.

The data collected is composed of 1994 citation data from Indiana and Michigan. The analysis was performed on specific States, and also on the combined State data.

a. Determine if Carrier's Drivers' Violation Rate is a Potential Predictor of Safety Performance

The first goal of the statistical analysis is to demonstrate that there is a relationship between the CDVR and crash rates. This relationship will be analyzed for all carriers with known mileage, and for all carriers with mileage that had crashes. The relationship may show that the CDVR can point to unsafe carriers.

Ranking was used to compare crash experience with the CDVR. Ranking of carriers by crash rate gives a list that is ordered by the carriers' apparent level of safety fitness. Other methods for targeting carriers use rankings. CDVR rank and crash rate rank is used in the following statistical analysis results.

b. Determine if Carrier Violation Rate is a Better Predictor

Carrier Violation Rate (the number of a carrier's violations divided by the number of the carrier's drivers is the Carrier Violation Rate) seems to be easier to capture statistic. Therefore, the review also checked on CVR, to determine how its use compared to the

CDVR. The statistical analysis will reveal that the CDVR is a better indicator of crash performance than the Carrier Violation Rate (CVR).

Ranking was also used for analysis of the CVR statistic. No other indicators were found to require further analysis or discussion.

2. Results of Statistical Analysis

Correlation analysis was performed on various subsets of carrier data. The results of the correlation analysis for selected variables in each subset will be presented. The Coefficient of Correlation was calculated for the CDVR rank and the CVR rank against crash rate rank. An examination of selected groups of carriers follows.

a. Indiana Carriers With More than 10 Violations

The results of the analysis for Indiana carriers with more than 10 violations is displayed in Figure 6-3. There are two results displayed, one is for all carriers with known IRP mileage, the other is for carriers with IRP mileage who had crashes.

Indiana Carriers

Number of Violations	Mileage	Crashes	Correlation Coefficient Between CDVR Rank and Crash Rate Rank
More Than 10 Violations	yes	maybe	.29
	yes	yes	.45

Figure 6-3

For carriers with more than 10 violations and IRP mileage, the Coefficient of Correlation between the CDVR rank and the crash rate rank is .29. (For this group of carriers the Coefficient of Correlation for the CVR rank and the crash rate rank is .16. This correlation is below the correlation of the CDVR rank and the crash rate rank, it is not as strong.)

For carriers with more than 10 violations, IRP mileage and crashes, the Coefficient of Correlation is .45. Among the carriers with crashes, the CDVR has an even better correlation to crash rate rank. This tends to confirm the validity of the CDVR as an indicator of safety fitness, or lack thereof, among troubled carriers.

For Indiana carriers, there were 33 (relatively few) carriers with a total of more than 10 violations, and only 19 with more than 15 violations. However, Michigan appears to have several larger carriers. There were 69 carriers that had more than 15 violations, and 49 of these carriers had more than 20 violations.

b. Michigan Carriers With More than 15 Violations

The results of the analysis for Michigan carriers with more than 15 violations is displayed in Figure 6-4. As previously mentioned there are two results displayed, one is for all carriers with mileage, the other is for carriers with known IRP mileage who had crashes.

Michigan Carriers

Number of Violations	Mileage	Crashes	Correlation Coefficient Between CDVR Rank and Crash Rate Rank
More Than 15 Violations	yes	maybe	.17
	yes	yes	.30

Figure 6-4

Again, the correlation between the CDVR rank and the crash rank is stronger among the carriers that had crashes.

For this group of carriers the correlation between the CVR rank and the crash rank is .07. This level of correlation is approaching meaninglessness. Among carriers with more than 15 violations, the effect of associating all of the carrier's drivers' violations with the carrier makes a substantial difference in the outcome.

c. Michigan Carriers With More than 19 Violations

The results of the analysis for Michigan carriers with more than 19 violations is displayed in Figure 6-5. As mentioned previously, there are two results displayed; one is for carriers with mileage, the other is for carriers with mileage who had crashes.

Michigan Carriers

Number of Violations	Mileage	Crashes	Correlation Coefficient Between CDVR Rank and Crash Rate Rank
More Than 19 Violations	yes	maybe	.24
	yes	yes	.37

Figure 6-5

For Michigan carriers with more than 19 violations and IRP mileage, the Coefficient of Correlation between the CDVR rank and the crash rate rank is .24. For this group of carriers, the Coefficient of Correlation for the CVR rank and the crash rate rank is .09. Clearly, the CDVR is a much stronger indicator of crash performance than CVR.

For Michigan carriers with 20 or more violations, mileage and crashes, the Coefficient of Correlation is .37. Again, the CDVR is a strong indicator of safety fitness among carriers who had at least one crash.

d. Overall Carriers With More Than 24 Violations

One advantage of the CDVR is that the concept applies equally well across multiple groupings. Once a driver is identified as being with a carrier, that driver's violations from any number of sources could be gathered together.

The violations from Indiana and Michigan were grouped. Figure 6-6 presents the results for the carriers.

Michigan and Indiana Carriers

Number of Violations	Mileage	Crashes	Correlation Coefficient Between CDVR Rank and Crash Rate Rank
25 or More Violations	yes	maybe	.25
	yes	yes	.25

Figure 6-6

Overall these correlations are comparable to the other results. This helps confirm that the statistical approach is valid. It also helps confirm that data from different States can be merged with some success.

3. Conclusions of Statistical Analysis

The conclusions of the statistical analysis are as follows:

- The conclusions of the expert review are validated by statistical success.
- The CDVR is better correlated with crash rate rank than the CVR rank.
- The results of the study are preliminary, additional study is needed.

Overall, these conclusions may be leading towards the objectives of the project. Each of the preceding bullet points is discussed below.

a. Some Statistical Success

In the several groups that were analyzed, there was a significant correlation between the Carrier's Drivers' Violation Rate ranking and the crash ranking. In other words, it can be expected that the CDVR can be used as an effective tool, or as a portion of a tool, to prioritize carriers for additional scrutiny.

Interestingly, the CDVR rank was more closely correlated with crash rate rank for the carriers that had crashes. One possible conclusion that could be drawn is that within the carriers which have had crashes, the CDVR is particularly good at distinguishing between carriers which have had one or more crashes due to poor safety fitness practices as compared to carriers which have had one or more crashes due to the effects of large exposure.

b. The CDVR is Better Than The CVR

The Carrier's Drivers' Violation Rate (CDVR) is statistically better than the Carrier's Violation Rate. In predicting crash behavior, the more violations that are used in the analysis, the better the CDVR becomes as compared to the CVR. While this result is not particularly surprising, it is worthy of note. The underlying methodology, based on the existing business practices, suggested that the CDVR was an effective way to identify the carriers which were;

- using possibly deceptive business practices, such as multiple identities (i.e. multiple USDOT numbers) for the same business, and

- generally engaging in business practices, such as using owner-operators and short term leasing, which have made the carrier difficult to identify and review in the past.

The use of the CVR does not identify these carriers. The tools that these carriers have used to avoid detection still worked when only the violations assigned to the carrier are taken into effect.

However, once the view is expanded to include all of the citations that were received by the carrier's drivers, a different picture emerges. The real problem carriers are unmasked. The carrier itself may not have very many violations, but the carrier's drivers point straight back to a problem carrier.

c. Results are Preliminary

The results of this study need to be considered preliminary. The expert testimony was clear. The experts were quite impressed with the list of potential problem carriers. One noted that there was a carrier in the CDVR top ten list that in their State was the subject of an on-going investigation. This carrier had had two consecutive Compliance Reviews resulting in two consecutive Satisfactory Safety Ratings, so FHWA does not consider this carrier to be a problem based on its current review criteria.

Further, the experts showed their confidence that the list of carriers ranked by CDVR by requesting and receiving funding for FY '97 to review the top identified carriers.

Nonetheless, the statistical evidence, while solid, is not so strong as to demand acceptance as conclusive. There was an adequate but fairly small number of data points. While a perfect correlation would never be expected, there is clear concern about the use of IRP mileage as the exposure indicator.

These concerns are as follows:

- Manual carrier name matching between IRP registrants and FHWA-registered carriers can be a somewhat arbitrary and capricious.

- IRP mileage may also be estimated mileage and outdated. The operation of the carrier may not reflect the same level of operation when the mileage was reported.

- IRP mileage is for registrants. The registrant may not be the carrier. A leasing company could have registered the vehicle and reported the mileage.

- Nationwide carriers have underreporting of mileage because a large percentage of their miles are reported to other States. This skews violation rates based on mileage.

The CDVR makes the assumption that driver-year is a valid measure of exposure. While this seems to be an obvious approach, it may be appropriate to review this approach with additional experts, to assure that it is valid and explained properly.

It appears clear that the total number of violations that a driver receives in a year represents that driver's violations history. Thus if a driver receives x violations in a year, the driver's violation rate for the year is x violations per year. Given this understanding, the second key concept is that, in order to assure that each carrier's exposure is the same, for the purpose of generating the statistic, each carrier must be assumed to be responsible for the driver for the entire year, and for each of the driver's actions during the year.

This makes a great deal of intuitive sense

- Clearly, the driver's behavior reflects back on the carrier if the driver was working for the carrier at the time.

- Any actions that the driver took before the point of employment reflect on the carrier's hiring practices.

- Actions that the driver took during the period of employment, but not for the employing carrier, may be directly related to a parent company, and certainly relate to the driver's basic behavior patterns.

- Actions taken after employment may still demonstrate something about the hired driver (although they may also demonstrate why the driver was not retained.)

By assigning the driver's entire record to the carrier, the carrier is assigned both a driver-year of exposure and the results of that exposure. While providing a solid basis for comparing carriers, there may be a sense that the carrier will suffer because of a driver's actions while working for another carrier. However, annualizing performance is an appropriate way to normalize the data for carriers which have drivers who worked less than the full year, so this is appropriate.

Because of this, and because the CDVR depends on unadjudicated data, it is clear that the CDVR cannot, by itself, be used as a basis for sanctioning. However, it may still present a strong basis for scrutiny of a carrier.

Further, the overall approach itself needs to be discussed, reviewed, and enhanced. In its simplest form, the CDVR has shown itself to be a potentially valuable tool in identifying potential problem carriers. Far more could be done with this tool.

d. Potential for Additional Refinement/ Implications

The CDVR should be viewed as a concept, not just as a statistic. The underlying concept is that the carrier evaluation based on driver performance is a three step process;

1. First, the driver is associated with the carrier.
2. Second, all of the driver's performance is associated with the carrier.
3. Third, carriers are rated or ranked based on their drivers' performance.

This report will offer several examples of how each of these portions of the method could be refined, enhanced or expanded. It is anticipated that other motor carrier and enforcement experts may offer additional ideas.

i. Associate the Driver with the Carrier

For purposes of this study, the driver was associated with the carrier based on the information on the citation. One obvious problem with this approach is that it muddles any ranking among the better carriers. Drivers with no violations will not be able to be counted.

Clearly, identifying the carrier on the citation is the primary method of associating drivers with carriers, but others may be considered.

The most direct approach would be to have some simple reporting mechanism every time a MCSAP-trained officer performs a business act which results in the association of a driver with a motor carrier. Clearly, roadside inspections fit the bill here. Today, there are millions of traffic enforcement stops that are performed that include all of the enforcement activities defined as a level 3 inspection. However, the paperwork required to report a level 3 inspection is not prepared. If more traffic enforcement stops resulted in documentation of a Level 3 inspection, the data could then be uploaded through SAFETYNET into MCMIS.

The implication is that there is a benefit to retaining the information from every inspection, and from any other contact, in which the driver-carrier relationship was identified (regardless of whether the contact resulted in a violation).

One result would be additional granularity in the resulting statistic. The very best carriers could get a CDVR as low as zero, if they had identified drivers and no violations. (As described in this study, the lowest CDVR is 1, since the only way to identify a driver as being with a carrier is through a violation.) The very worst carriers would get close to the same CDVR, as presumably they have relatively few drivers with no violations. It is possible to envision a situation where this would increase a CDVR.

It is not clear what it would take to implement a program of low probability (e.g. 1 vehicle in 500), random Level 3 inspections at the inspection stations. However, the implementation of such a program (and retention of the data),, along with collection of other data that associates drivers with carriers, would result in a very solid data base connecting drivers with carriers.

ii. Associate the Driver's Performance with the Carrier

This study associated only the driver's in-State performance with the carrier. However, it is possible to envision a situation where other elements of the driver's performance could be associated with the carrier.

1. Inspections

The driver's inspection performance could be associated with the carrier. This may not add a great deal, as the carrier's inspection performance is already associated with the carrier, and the driver's inspection performance for one carrier may not accurately reflect on a different carrier. Nonetheless, it is a potential area for research for two reasons. First, even if the inspection data is currently collected, this may offer a different and better way to process it. Second, it has not been demonstrated that the inspection performance for one carrier does not reflect on a second carrier.

It is not clear that there is a great deal here, but it is also not clear that there is not.

2. *Other Citations*

The area of other citations presents a particularly fertile area for research, and is an area where collection of data associating a driver with a carrier is particularly valuable.

There has been consideration given to asking local (city, county, etc.) police to identify the motor carrier on any violation that they write. Under the NGA guidelines these Officers are required to identify the motor carrier on the crash supplement.

Any number of observers have noted that, in many cases, the local police are not and will not be able to effectively identify the carrier. This assertion, that local police will not be able to accurately identify the carrier, was supported by the data analysis performed on this project. The percentage of crashes that the non-MCSAP trained State police assigned to (apparent) leasing companies was *far* greater than the percentage of violations that the MCSAP trained State police assigned to these same leasing companies. Clearly, the MCSAP trained officers were able to accurately identify the carrier, where other State police officers, given the same situation, identified the lessor as the carrier.

If, however, drivers were associated with carriers through a separate program, the driver action could then be able to be associated with a carrier (or even with more than one carrier) based on the year's history. This would allow violations assigned by local Police to be associated with an appropriate carrier, based on information already in the system about the driver, without the officer having to identify the carrier.

This may not be a perfect approach. Clearly, a violation or crash written up by a local office should be associated with only one motor carrier, while this approach allows an individual driver's actions to be associated with multiple carriers.

This is still far more likely to accurately identify the carrier, and at a far lower cost, than trying to train all of the police officers in the United States to be able to accurately identify a carrier.

Further, it seems that it is a better policy to identify two carriers as being responsible, when only one is, than identifying only one "carrier" as responsible, when the carrier is a leasing company. In the unlikely event that a carrier is reviewed as a result of several of these incorrect assignments, the carrier should have the paperwork to unwind the problem. But, if the responsible carrier is not identified, there is an unidentified potential problem waiting to happen.

The accurate identification of the responsible motor carrier on citations issued by local Police is an issue that does not appear likely to be resolved soon. The downside of having local Police incorrectly identify carriers is clearly perceived, but not fully defined, within the community.

Use of violation data, crash data, and possibly other data, from sources that collect driver identification but do not collect carrier identification, is a significant long-term use of the CDVR concept. A side effect of the CDVR approach would be the gaining of the benefits of matching the driver and carrier on these actions, without requiring the Officer at the site to make the identification.

3. Rate (and Rank) Carriers Based on Drivers' Performance

The CDVR statistic that is used in this report gives equal weight to each violation. Clearly, this is the simplest method of calculation.

However, it is possible to envision the addition of weighting parameters to the method. Within the Driver/Carrier project it was suggested that certain types of violations, particularly logbook and size and weight violations, may be more indicative of carrier related problems. There is also a sense that moving violations, especially serious moving violations, may be stronger indications of problems.

It was beyond the scope of the current project to explore these relationships. Further, there was not enough data collected to be able to extract a valid refinement of these relationships. However, there is clearly a potential for refinement in this area.

Chapter 7 - Phase II Conclusions

Phase II of the Driver Carrier Project had three objectives,

- to validate, if appropriate, the conclusion from Phase I of the project, that violation rates differ among carriers,
- to determine if the difference in violation rates relates to a difference in safety performance and
- to determine if the inclusion of carrier identification on a citation could be used in identifying carriers that are more likely, than average, to have safety performance problems.

The project was successful in meeting all three of these objectives.

A. Violation Rates do Differ by Carrier

The revalidation of the Phase I project goals using new data was successful. In this phase, the Work Group received an additional year's data from Indiana as well as a complete data set from Michigan. The Work Group reviewed the statistical tests that had led to the initial conclusion. The same tests were run on the new data and the same conclusions were reached.

The revalidation of the Phase I project goals used IRP mileage in the statistical analysis.

The overall conclusion is the same as reported in Phase I. Among the carriers represented in the Michigan data, and in Indiana data from 1994, **violations are not randomly distributed among carriers.**

B. Violation Rate is an Indicator of Safety Performance

The second objective was to determine if the difference in violation performance can be related to a difference in safety performance. The method chosen was to determine if a relationship exists between violation rates and safety indicators - crash rates, SCE score and Safety Rating.

To derive the crash rates, the number of crashes in the last five years, as reported through MCMIS, was divided by the number of reported IRP miles in the immediate prior year. SCE score and Safety Rating were provided by FHWA from MCMIS. A matching was also performed to obtain IRP miles for carriers which had crashes, SCE scores and/or Safety Ratings, but which did not have violations.

The results of the correlation analyses between violation rate and safety indicators are as follows:

- There was a solid correlation between the violation rates and the crash rates in both Michigan and Indiana.

- No relationship was demonstrable between violation rate and SCE or Safety Rating.

Initially, this was a distressing result. The charge of the project was to use a correlation between violations (or Z-Score) and the known indicators of carrier performance, SCE score and Safety Rating. However, a more surprising result was found - there was no correlation between crash rates and either SCE score of Safety Rating.

- Further analysis suggested that this is not a terribly surprising result for SCE scores. SCE scores are designed to suggest carriers for review. Since hazardous materials and passenger carriers can cause the most damage if they are unsafe (and therefore should be targeted for review), and since it is desirable that these should be the safest carriers, hopefully a piece of the SCE score correlates negatively with crash results (i.e. haz-mat and passenger carriers are both safer than average and higher priority targets for review.) In fact, SCE scores did consistently correlate more poorly with crash rate than did Safety Rating.

- Nonetheless, there was no significant correlation between Safety Rating and crash results. In other words, this project demonstrated that, for the data that the project received, Safety Rating was entirely ineffective in predicting carriers which had bad crash results.

The project had no choice but to discount SCE score and Safety Rating as predictors of safety fitness. The strong correlations between violation rates and crash rates was considered the demonstration of the relationship.

The conclusion is that **the difference in violation rates is related, at least in part, to a difference in safety performance.**

C. Carrier Violation Rate can be Used to Identify Problem Carriers

The next project goal was to determine if there would be any effective and efficient way to use carrier identification from driver citations.

1. Violation Rates based on Mileage is Impossible

For statistical purposes, this project used IRP mileage. IRP mileage was satisfactory for the statistical analyses. However, for a number of reasons, IRP mileage is not feasible to use in a production environment. Notably, without electronic data from all 49 contiguous jurisdictions, it would be impossible to obtain a fair mileage profile for a carrier in any one State.

The only other system that gathers mileage on a nationwide basis is IFTA. For both technical and business function reasons, attempting to use IFTA mileage would be less effective than the use of IRP mileage.

The project did demonstrate that, at least to some extent, information about driver violations and the associated carrier can be captured. But, in order to perform any analysis on rates, some measure of exposure other than mileage is required.

2. CDVR - Based on Driver Year - is a Reasonable Alternative

Fortunately, the project developed an alternative, the Carrier's Drivers' Violation Rate (CDVR). For this project, the CDVR was calculated as the total number of violations that a carrier's drivers' received divided by the number of drivers that were associated with the carrier. The numerator, total number of violations that the carrier's drivers' received includes all violations that the driver received, not just those violations that were specifically associated with that carrier.

Many of the observers of the motor carrier enforcement area have suggested that a move to a more performance based method of reviewing carriers is appropriate. The CDVR is an appropriate performance based measure.

The CDVR is based on the underlying assumption that the performance of a carrier's drivers presents a valid view of the carrier's performance. This is the conclusion that was reached in this project, and demonstrated in Chapter 4 of this report.

The CDVR is also based on the assumption that a driver-year is a reasonable, and in fact an entirely appropriate measure of exposure. Since there is no way to determine how much of the year the driver worked for any particular carrier, and since the driver's behavior was considered to reflect on the carrier, the statistic is calculated based on the driver's performance for the entire year, regardless of who the driver was working for at the time.

Within the data that was analyzed, ranking carriers based on the CDVR was vividly demonstrated to be a successful way of identifying many serious problem carriers. The top several dozen carriers, ranked by the CDVR, were divided by the State police experts into three roughly equal groups;

- carriers that were currently identified by both FHWA and the States as likely problem carriers,
- carriers which had been identified by the States as problem carriers, but which had not drawn Federal interest and
- carriers which had not been previously identified as problems.

The State police representatives were quite interested in looking into the carriers that they were not familiar with. Based on the other carriers on these lists, these looked to be good candidates for further scrutiny.

The CDVR should be viewed as a concept, not just as a statistic. The underlying concept is that carrier evaluation based on driver performance is a three step process;

(1) First, the driver is associated with the carrier.

(2) Second, all of the driver's performance is associated with the carrier.

(3) Third, carriers are rated or ranked based on their drivers' performance.

In addition to the validation of the CDVR presented in the active review process, the association between CDVR was validated statistically.

3. Potential for Expansion

As was demonstrated in this project, the CDVR presents a useful methodology for evaluating carriers' safety performance using just the data that was inherent in the violation statistics. However, the CDVR could be expanded for even more sophisticated analysis.

a. Expanded Matching of Drivers with Carriers

Once the concept of CDVR is accepted, it can be improved substantially.

The first step would be to increase the sources of data. Among other sources, this could include violations from additional States as well as violations from additional sources, such as SAFETYNET.

The next step would be to add to the process by associating drivers with carriers. By using this approach, at every step of the enforcement process, the driver should be associated with carrier **if** the officer is competent and trained to make the association. For example, the driver would be reported with every vehicle inspection, including inspections which did not result in violations. This would allow the capture of information about drivers who did not have violations, and increase the effectiveness of the CDVR statistic.

b. Use Data Matching Drivers and Carriers

When the ability to match drivers to a carrier has been expanded, it would then make sense to consider sources of information about drivers which do not have an associated carrier. For example, there has been consideration given to asking local police to identify the carrier on a violation. However, experience shows that the local police will not be able to effectively identify the carrier.

If, however, drivers were associated with carriers through a separate program, the driver action could then be able to be associated with a carrier (or even with more than one carrier) based on the year's history. This would allow violations assigned by local police to be associated with an appropriate carrier, based on information already in the system about the driver.

This may not be a perfect approach. Clearly, a violation or crash written up by a local Officer should be associated with only one motor carrier, while this approach allows an individual driver's actions to be associated with multiple carriers. However, it seems that it would be much better to associate a citation or crash with two carriers, only one of which was the correct carrier, than to associate the citation or crash with only the leasing company, and never be able to identify the carrier.

The information garnered in this way could be used to select carriers for review or scrutiny, and further investigation would be performed before sanctioning or enforcement action would be taken based on this type of assignment.

4. CDVR Has Tremendous Potential

One of the key goals of the Driver/Carrier Project, Phase II was to determine if there is any value in retaining the carrier's USDOT Number in citation records, in order to more effectively target potential problem carriers.

Use of the Carrier's Drivers' Violation Rate offers the potential for a performance based measure that will be able to use the relationship between violations and crashes to identify carriers which may have a higher than average risk of poor safety performance.

D. Summation

Phase II of the Driver/ Carrier Project had two major new goals:

- Determine if the difference in carrier violation performance could be related to a difference in carrier safety performance, and, if so
- Determine if there is a statistical treatment which, using carrier identification on citations and other currently collected or reasonably easy to collect data, can be used to identify potentially problem carriers.

The answer to both of these questions is Yes.

> **The difference in carrier violation performance is a predictor of a difference in carrier safety performance.**
>
> **There is a statistical treatment, specifically the Carrier's Drivers' Violation Rate, which does suggest a difference in carrier safety performance, and which can be calculated using carrier identification on citations and other currently collected or reasonably easy to collect data.**

Appendix A - Glossary of Terms and Acronyms

A. Terms and Definitions

Term/Acronym	Definition
AAMVA	American Association of Motor Vehicle Administrators.
AAMVAnet, Inc.	American Association of Motor Vehicle Administrators Network, Incorporated, a subsidiary of AAMVA.
Carrier	see Motor Carrier.
CDLIS	Commercial Driver License Information System.
CDVR	**Carrier's Drivers' Violation Rate**
CVR	**Carrier's Violation Rate**
Citation	Issued by a law enforcement agency. The assignment of a violation to a person or organization cites a violation. May consist of multiply violations. It is usually called a ticket.
CMV	Commercial Motor Vehicle.
CMVSA	Commercial Motor Vehicle Safety Act of 1986.
FHWA	Federal Highway Administration.
IFTA	International Fuel Tax Agreement.
IRP	International Registration Plan. The program for registering heavy trucks that are used in interstate commerce.
Lessee	The person or company who leases the vehicle from the lessor.

Lessor The person or company who leases a vehicle to a lessee.

Mean The arithmetic average of a set of values. The sum of the values of the target parameter divided by the number of occurrences. Usually represented by the Greek letter *m*.

MCMIS Motor Carrier Management Information Systems.

MCSAP Motor Carrier Safety Assistance Program.

MCSIP Motor Carrier Safety Improvement Program.

Motor Carrier A person or company who is transporting goods by use of a commercial motor vehicle. The motor carrier responsible for the safety fitness of a vehicle and its load.

Owner The person or company that owns the vehicle.

OMC Office of Motor Carriers.

Person or Company Any business entity, including an individual, partnership, corporation, trust or fiduciary.

Poisson Distribution A statistical method used to measure the probability of a low frequency event to occur.

Registrant The person or company that has registered the vehicle. This may not be the owner, or the driver, or the motor carrier.

SAFETYNET The computerized system for collecting State safety data.

SAFETYNET 2000 The computerized system currently in the design and development stage. SAFETYNET 2000 will replace version 8.0 of the SAFETYNET system.

SCE A derived statistic that is used to rank carriers to determine which will be selected for review.

Standard Deviation
A measure of the spread of the values of a parameter. The square root of the variance. Usually represented by the symbol s.

USDOT
United States Department of Transportation.

Variance
A measure of the spread of the values of a parameter of interest. The square of the standard deviation. Usually represented by the symbol s^2.

Violation
A cited occurrence of an infraction of the law. Or the actually law or regulation which is identified (or cited) on a citation.

Z-Score
A measure of the value of a parameter as compared to the mean based on the standard deviation. $(y-m)/s$.

Appendix B - Statistical Methods Used in Phase I

A. Summary

The *Driver/Carrier Statistical Analysis Report* from Phase I included two sections which addressed the statistical methods.

- Chapter 3 addressed the General Statistical Approach. It described both the methods that were employed and how those methods were used.

- Appendix C to the document provided a more detailed and complete explanation of the statistical approach. It was prepared for the more mathematically and statistically oriented reviewer.

This information is specifically useful in reviewing Chapter 3 of this report, which applies the statistical methods used in the *Driver/Carrier Statistical Analysis Report* to the newer data that was collected.

This Appendix contains reprints of the two relevant sections from the original Driver/Carrier report.

B. General Statistical Approach

The purpose of this section is to present a general description of the statistical approach that was used in this analysis. Appendix C, which follows, will present a more detailed explanation of the actual statistical tools that were used. It should be recognized that statistics never "prove" anything, in the most rigorous sense. Rather, statistics will provide us with evidence to support or demonstrate a position.

Seven Steps in Hypothesis Testing

1	Establish the hypothesis.
2	Establish the confidence level.
3	Determine the statistical test.
4	Establish the critical values.
5	Gather the data.
6	Analyze the data.
7	Reach a conclusion.

The general approach to the statistics used in this study was hypothesis testing. Each of the steps will be described below. Hypothesis testing is not particularly effective in asserting that a particular hypothesis *is* supported by the data, but it offers a powerful capability to demonstrate that a hypothesis *is not* supported by the data.

Therefore, the statistician generally tests the **null hypothesis** (the hypothesis that nothing significant has occurred) and demonstrates that the null hypothesis cannot be supported by the data.

For example, a starting assumption would be that all carriers have the same likelihood of receiving violations, i.e. the same underlying violation rate. If this is true, then the actual differences in violation rates would be attributable to randomness. The statistician would then look at the differences in rates, and determine whether the difference might be attributed to randomness. Based on the results of the statistical test, the hypothesis that all carriers might have the same underlying violation rate would be either rejected or not rejected. If the hypothesis that the difference could be due to randomness is rejected, then it is assumed that there is some factor (or factors) other than randomness which accounts for the differences in violation rates.

1. Steps in Hypothesis Testing

The Work Group addressed each of the steps in the hypothesis testing methodology.

a. Establish the Hypothesis

This is the step to determine what is being tested. Generally, the project tested the null hypothesis. The null hypothesis would be (generally) that violations are distributed randomly among carriers.

b. Establish the Confidence Level

In general, the project planned to use a 98% confidence level. In other words, the null hypothesis will not be rejected unless there is at least a 98% chance that the differences cannot be explained by random chance. This 98% confidence level is more conservative than the 95% level used in, for example, political polls. In general, a more stringent confidence level, such as a 99% or 99.5% level, would only be used in instances where:

- the results of an incorrect rejection of the null hypothesis could be catastrophic, and

- the result of an incorrect acceptance of the null hypothesis is not costly.

However, the confidence level never became a factor in this analysis. In each actual case tested, either there was an underlying problem (generally the problem was not enough data), or the statistical measures showed overwhelming results.

c. Determine the Appropriate Statistical Test

A number of statistical tools were used in this project. These are described in general in Section B-3 below, and more completely in Appendix C.

d. Establish the Critical Values

This is the application of the desired confidence level to the actual statistical test, to determine the data point for the critical values.

e. Gather the Data

Data gathering for these analyses was done in three steps.

(1) Violation data was gathered by the pilot States.

(2) Exposure data was also provided by the pilot States.

(3) Exposure data was matched to the violation data by the Work Group.

f. Analyze the Data

In this step, the appropriate statistical test is performed on the data.

g. Reach a Conclusion

In this step, the previously determined critical value is compared to the actual results.

Section B-2 below addresses the various hypotheses that were tested. Section B-3 discusses the various statistical tests that were used. Section B-4 describes how the statistical tests were applied to the hypotheses.

The reader should recognize that an understanding of the statistical methodology may enhance an understanding of the results. However, a complete understanding of the methodology is not critical to an understanding of the results.

2. Hypotheses Tested

There were three general areas that were tested;

(1) Distribution of Violations Among Carriers,

(2) Distribution of Violations Among Carriers which Have Violations, and

(3) Distribution of Violations by Violation Type Within Groups of Interest.

There were at least two separate levels of interest in this study, the State level, and the National level. While the ultimate concern of the program was at the national level, the Driver/Carrier project accumulated State level data. Therefore, the statistical analysis, as described, was performed several times, on data from each State.

Further, the *Driver/Carrier Summary Analysis Report* demonstrated that the patterns and rates of violation assignment differed significantly among the pilot States. Therefore, an analysis across State data was not appropriate.

a. Distribution of Violations Among Carriers

The first hypothesis to be tested was whether violations are randomly distributed among the entire population of carriers. Stated in the null form, the hypothesis was:

Violations are randomly distributed among the entire population of carriers.

These tests depended on the use of data related to the entire population of carriers, not just the carriers which received violations. This was the test which most precisely tested the underlying issues.

The tests associated with this hypothesis, as well as the tests associated with the second hypothesis, were performed against both overall violation data and appropriate subsets of the carriers.

b. Distribution of Violations Among Carriers Which have Violations

The second hypothesis tested was whether violations were randomly distributed among the carriers which had at least one violation. As compared to the first test, which used all carriers which reported mileage, this test only applied to carriers which also had reported violations. The null hypothesis was:

Violations are randomly distributed among the carriers which have violations.

This hypothesis was not as precisely on-point as the first, as it only tested a subset of all carriers. However, it did offer a key benefit; it took advantage of a type of data that was collected by the pilot States.

There was another benefit in this test. Theoretically, it would be possible for violations to be randomly distributed among carriers which had violations, but not randomly distributed among all carriers. In this situation, it would be postulated that there were two classes of carriers; carriers which had violations and carriers which didn't. Thus, even if the null hypothesis that violations were randomly distributed among carriers with violations could not be rejected (i.e. it appeared that violations were randomly distributed among carriers with violations), it was still possible that violations were not randomly distributed among all carriers.

The converse was not the case. If this null hypothesis, that violations were randomly distributed among carriers with violations, was rejected; the higher level null hypothesis, that violations were randomly distributed among all carriers, must also have been rejected. Therefore, testing for randomness among carriers which had violations may have directly affected the conclusions about all carriers.

c. Distribution of Violations by Violation Type within Groups of Interest

Several States suggested that there was a difference in type of violation received by some classes of carriers. This study looked at two cases of this situation, carriers with more than 40 violations and for-hire carriers. The expectation was that carriers with more than 40 violations may have received different types of violations. The null hypothesis tested was:

Carriers with more than 40 violations receive the same proportion of violations by type as the overall population does.

A similar test was also run to compare for-hire carriers to the population. The hypothesis for this test was that:

For-hire carriers receive the same proportion of violations by type as the overall population does.

The purpose of this test was to determine if there was a demonstrable difference in the behavior of carriers based on the type of carrier.

This testing did not depend on exposure data. It is included in the statistical report (as opposed to the Summary report) because it used the same statistical tools as described in Section C, below.

3. Statistical Tests

This section presents an overview of three statistical concepts that were critical to the statistical analysis that was performed. These tests were

- comparing actual to expected cell contents (Chi Squared),
- a general view of expected distribution (Bell Curve), and
- distribution of infrequent events (Poisson Distribution).

An overview of each of these statistical concepts is presented below. The topics are more completely reviewed in Appendix C.

a. Comparing Actual to Expected Cell Contents (Chi Squared)

In any number of situations, there is a need to look at expected results, compare them to actual results, and determine if the difference between the expected results and actual results can be ascribed to randomness.

In general, the situation can be viewed as creating a table such as the following.

	Range 1	**Range 2**	**Range 3**	**Range 4**
Expected Results	Expected Results for Range 1	Expected Results for Range 2	Expected Results for Range 3	Expected Results for Range 4
Actual Results	Actual Results for Range 1	Actual Results for Range 2	Actual Results for Range 3	Actual Results for Range 4

Table B-1

A statistical test called Chi Squared presents a statistic which evaluates the difference between the expected and actual results. A small Chi Squared statistic supports the hypothesis that the actual results could have been taken from the same population as the expected results, and that the differences are attributable to randomness. If the Chi Squared statistic is too large, then the likelihood is that the actual results did not come from the population and the distribution that was used to define the expected results. (Interestingly, if Chi Squared is too small it suggests that there is not enough randomness, and that the actual results match the expected results too closely.)

The actual test value for Chi Squared depends on the number of ranges. The following table, Table B-2, presents the Chi Squared values at the 98% confidence limit;

Number of Ranges	**Chi-Square for Rejection**
5	13.388
6	15.033
7	16.622
8	18.168
9	19.697
10	21.161
11	22.618
12	24.054
13	25.472
14	26.893

Table B-2

In other words, a Chi Squared statistic is calculated to determine whether the differences between the actual results and the expected results can be attributed to randomness. If the Chi Squared value is below the value in the table, the hypothesis that the difference is attributable to randomness cannot be rejected. However, the actual cause of the difference in not addressed.

For example, if Chi Squared had been calculated for a table with 14 rows and the resulting Chi Squared value was 24 (below that of 26.893) then the stated hypothesis that the difference is attributable to randomness, cannot be rejected. The statistician can say that the difference between the expected and actual results *is not* attributable to randomness. There may well have been other factors at work. The statistician simply cannot reject the hypothesis that the difference was attributable to randomness.

However, if the Chi Squared value is above the value shown in Table B-2, then the hypothesis that the difference between the expected values and the actual values is attributable to randomness, can be rejected.

Of course, it is still up to the statistician to determine the correct expected value (for each cell) for the item being tested. The next two sections describe the two methods that were used in this project to determine expected values of the distribution of violations.

b. General View of Expected Distribution (Bell Curve)

When there are a large number of occurrences of events and the expected results can be described as a mean (or average), then the actual results will have a frequency distribution that takes the shape of a bell curve. This is called a normal distribution.

For example, if 100 people flip a coin 100 times, we can reasonably expect that many of the 100 people will not get exactly 50 heads (and 50 tails.) In fact, if we graph the frequency of the number of heads, we expect it to take the shape of a bell curve.

Figure B-1 presents a graphic representation of a bell curve. The bell curve was used extensively in this report, as it represented expected frequency distribution.

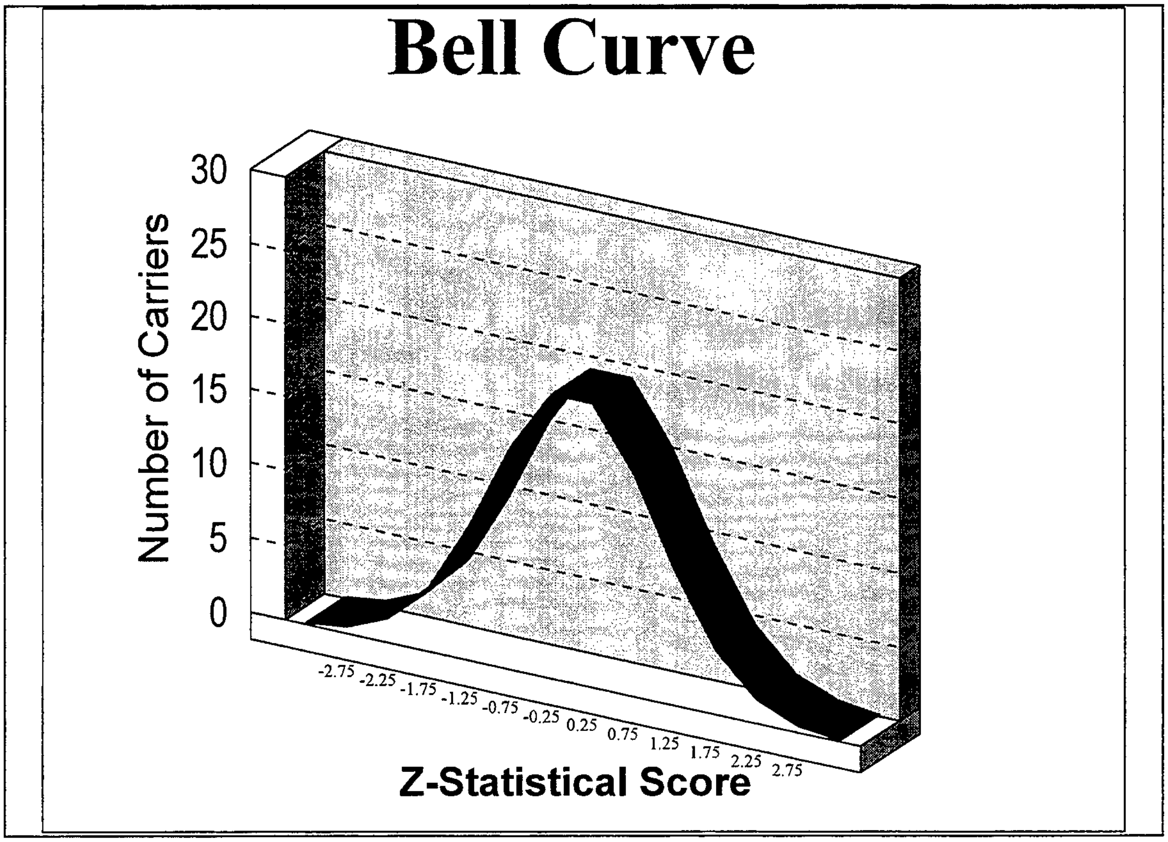


Figure B-1

The terms "mean" and "standard deviation" are used in describing any statistical distribution. These terms are particularly useful in describing a normal distribution, or bell curve. The term "Z-Statistic" is only used in describing a normal distribution. The following sections will define these terms, and describe how they are specifically used in describing a bell curve.

i. Mean

Mean is the statistical term for what is usually called "average". In the above example, the mean of heads is expected to be 50.

If a frequency distribution is a bell curve, the mean is the highest point in the bell curve.

In many of the statistical tests in this study, the mean was the average number of miles traveled per violation.

ii. Standard Deviation

Standard Deviation is a statistical measure of how far an actual occurrence (or event) is from the mean. Standard deviation provides a mathematical description of the expected spread.

For example, if we flip a coin 10 times, and it comes up heads 8 times, we would not be at all surprised. However, if we flip the coin 100 times, and it comes up heads 80 times, we would find the result so shocking that we would be quite certain that we were not dealing with a fair test. In the first case, (for 10 coin flips the standard deviation is 1.58) the event of 8 heads is less than 2 standard deviations from the mean. In the second case, (the standard deviation for 100 coin flips is 5) the event of 80 heads is 6 standard deviations from the mean.

The following, Table B-3, describes the frequency distribution found in a normal distribution, or bell curve.

Frequency Distribution of Bell Curve

Number of Standard Deviations (n) from the mean	Occurrences within a standard deviation of n below the mean	Occurrences within a standard deviation of n above the mean	Total occurrences within n standard deviations below and above the mean
0.5	0.19	0.19	0.38
1	0.34	0.34	0.68
1.5	0.435	0.435	0.87
2	0.48	0.48	0.96
2.5	0.495	0.495	0.99
3	0.498	0.498	0.996

Table B-3

Standard deviation is an applicable statistic for any sampling. It does not depend on a normal distribution for its validity. The normal distribution is just one example of the relationship between a standard deviation and a frequency distribution.

iii. Z-Statistic

A Z-Statistic is, quite simply, a tool to describe the results for a single event in terms of mean and standard deviation. Given that the actual result, mean and standard deviation are known, the Z-Statistic is calculated as follows;

Z-Score

$$Z = \frac{\text{Actual Result - Mean}}{\text{Standard Deviation}}$$

Figure B-2

The Z-Statistic is a powerful tool for converting the results of a test into a statistically meaningful number. For the purposes of this report, it will allow us to compare the number of violations among carriers, even if the carriers have different numbers of miles. Within the context of any particular test, a low Z-Score represents a carrier with good performance, and a high Z-Score presents a carrier with a bad performance.

In the cases above, the Z-Statistic for 8 heads is 1.9[(8-5)/1.58]. The Z-Statistic for 80 heads is 6 [80-50)/5].

The frequency distribution shown in Table B-3 is used extensively throughout this project. The number of standard deviations from the mean translates directly to a Z-Score. Note that the bell curve graphs are labeled with the centers of the frequency ranges. For example, the range of 1.50 to 1.99 is labeled 1.75.

The concept of a frequency distribution for a bell curve, as defined by a Z-Statistic, along with the statistic to measure whether an actual set of results matches an expected set of results (Chi Squared) provides an effective set of tools to determine whether an actual frequency distribution matches the results that would be expected from a random process. (Standard Deviation was calculated on a carrier by carrier basis, as described in Appendix C.)

Based on the central limit theorem, it can be assumed that the Z-Statistics for a group of carriers would approximate a normal distribution when there are sufficient violations, if violations were randomly distributed among carriers. An exceptionally high Z-Score can be taken as evidence that the carrier has an exceptionally high violation rate. A low Z-Score for a carrier demonstrates that either the carrier has a low violation rate, or that the carrier has not presented a sufficient number of miles to obtain an accurate evaluation.

c. Distribution of Infrequent Events (Poisson Distribution)

A bell curve describes the frequency distribution of events when there are enough events that the expected number of "successes" (measurable events, e.g. heads) is fairly high. Any time that an average of at least 5 "successes" can be expected, a bell curve is a good descriptor of the expected frequency. A bell curve is an adequate description of the frequency distribution when the mean is even two "successes".

However, a bell curve does not describe the frequency distribution when there are a large number of low probability events, and the expected mean is particularly small.

For example, if there are 1000 carriers, each of which travels 50,000 miles, and the average for the group is 1 violation every 250,000 miles, then the average carrier will get 1/5 of a violation (the violation rate is .2 violations/carrier). Obviously, there will not be any carriers which have exactly 1/5 of a violation. An observation of the situation indicates that the distribution is not a bell curve. It could be expected that, if violations are randomly distributed, just a few carriers will have 2 or more violations. Therefore close to 200 carriers will have 1 violation, a little over 800 of the carriers will have no violations.

There is a statistical distribution which describes this situation, the Poisson Distribution. Given a large number of low probability events, the Poisson Distribution presents an expectation of the actual distribution of events. In the example cited above, the Poisson Distribution would suggest that 819 carriers will have no violations, 164 carriers will have one violation, 16 carriers will have two violations and 1 carrier will have three violations.

Given a number of carriers with similar mileage amounts and a group violation rate, we can use a Poisson Distribution to predict the number of carriers which will have each count of violations. The Chi Squared function can then be used to determine if the actual results match the prediction.

A key drawback to the Poisson Distribution is that it depends on each carrier having the same number of miles driven. The actual statistical analysis was performed using mileage ranges. The violation rate was calculated using the average mileage within the selected group. As long as reasonably tight mileage ranges were used, this situation very closely approximates the requirements for a Poisson Distribution.

Poisson Distribution statistics for standard ranges can be found in most statistics textbooks. However, for this project, the Quattro Pro function of @Poisson was used to precisely calculate the Poisson Distribution value on a case by case basis.

4. Application of the Tools to the Hypotheses to be Tested

This section will describe how the tools, described in Section 3 above, were used to test the hypotheses, described in Section 2 above.

a. Distribution of Violations Among Carriers

The first hypothesis tested was whether violations were randomly distributed among the entire population of carriers. Stated in the null form, the hypothesis was:

Violations are randomly distributed among the entire population of carriers.

Using North Dakota as an example, the following procedure was followed:

i. Calculate the overall violation rate.

(1) Match the contents of the file of North Dakota violations to the carriers which have North Dakota IRP miles (North Dakota carriers). Data for carriers which had violations in North Dakota but which did not have North Dakota IRP miles were not used in this test.

(2) Using the North Dakota IRP file, determine the total IRP miles for North Dakota carriers.

(3) Determine the total number of violations for IRP carriers.

(4) Divide the total miles [determined in (2) above] by the total violations [determined in (3) above] to obtain the sample violation rate [*overall violation rate*.] Note that this violation rate only applies to the data from this sample.

ii. For each carrier;

(1) Calculate the *carrier violation rate* (*carrier miles/carrier violations*)

(2) Calculate the *carrier's standard deviation* using the binomial distribution as described in Appendix C.

(3) Calculate the *Z-Score* [(*carrier violation rate-overall violation rate*)/*carrier's standard deviation*]

iii. Accumulate the Z-Scores into the frequency distribution shown in Table B-4.

This compares the actual results to a bell curve.

Comparison Chart

Z-Statistic	Expected Count	Actual Count
< -3.00	.00139*n	
-2.50 to -2.99	.00500*n	
-2.00 to -2.49	.01691*n	
-1.50 to -1.99	.04481*n	
-1.00 to -1.49	.09298*n	
-0.50 to -0.99	.15098*n	
00 to -0.49	.18793*n	
00 to 0.49	.18793*n	
0.50 to 0.99	.15098*n	
1.00 to 1.49	.09298*n	
1.50 to 1.99	.04481*n	
2.00 to 2.49	.01691*n	
2.50 to 2.99	.00500*n	
> 3.00	.00139	

Table B-4.

iv. Determine the Chi Squared value for Cell Probability (as described in Appendix C.)

v. In this case, there were actually 13 degrees of freedom. Therefore the test value for Chi Squared is 25.472.

If the Chi Squared statistic is greater than 25.472, the null hypothesis will be rejected, and it may be concluded that violations are not randomly distributed among all carriers in North Dakota.

This same test was repeated for Indiana. It could not be performed for Idaho, as the Work Group did not have an automated file of all Idaho carriers.

The Work Group developed generalized software to execute this procedure. The software was used for both this test and the next ones, in which the same procedure was performed on different data.

A review of the actual results indicated several concerns with the tests;

- There were relatively few actual results, especially for North Dakota.
- There were many carriers which had very few miles. Most of these carriers did not, and should not have been expected to, have violations. Therefore, the results could not be expected to form a bell curve.

Two methods were used to resolve these concerns;

- Only carriers with a large number of miles, generally about enough miles to expect two violations, were used in a subsequent test. The results of this test should have very closely matched a bell curve.
- Carriers were stratified, and Poisson Distributions were used to determine expected values for the stratified groups.

There was also interest in the community to have the results reviewed with respect to moving violations only. Again, the same set of procedures was executed using only moving violations. Further, reflecting other interests and suggestions, the analysis was performed using only size and weight violations, and using only log book violations.

b. Distribution of Violations Among Carriers Which have Violations

The second hypothesis tested was whether violations were randomly distributed among the carriers which had at least one violation. The null hypothesis was:

Violations are randomly distributed among the carriers which have violations.

The same set of tests and testing procedures was used as described in Section 1 above. However, only data for carriers which had violations was used. Again, a wide variety of different conditions were tested.

c. Distribution of Violations by Violation Type Within Groups of Interest

In the first case in this group, the null hypothesis tested was:

Carriers with more than 40 violations receive the same proportion of violations by type as the overall population does.

This test was run for Michigan and Indiana data, as they were the only States to report both a number of carriers with more than 40 violations and totals for the entire population. The following procedure was used;

(1) For each type of violation (within each State), its proportion of the total violations was calculated.

(2) The percentage of the total violations that were received by carriers with more than 40 violations was calculated.

(3) For each type of violation, the proportion of total violations by type was multiplied by the percentage of violations received by carriers with more than 40 violations to get the expected number of violations for carriers over 40 by type.

(4) The expected number of violations was compared to the actual number of violations, by type. A Chi Squared analysis was performed to determine if the violations for carriers with over 40 violations could have been drawn randomly from the total of all carriers.

This is a classic application of the Chi Squared test.

The same procedure was used for determining if there was a significant difference in the violations received by For-Hire carriers.

Appendix C - Statistical Methods Used in Phase II

A. Purpose

The purpose of this Appendix is to provide a more complete presentation of the statistical approach that was used in this phase of the study. This Appendix is appropriate for the more mathematically and statistically oriented reviewer. An understanding of the statistics presented in this Appendix is not critical to the understanding of the information presented in this report.

The introductory material for this Appendix is presented in Chapter 4 of this report. This Appendix should be viewed as a supplement to Chapters 4 through 6.

This Appendix will present more complete views of;

- Statistical Terms Used in this Document,
- Coefficient of Correlation,
- Computation of Standard Deviation using a Binomial Probability Distribution,
- Confidence Level, and
- Chi Squared (C^2) for Cell Probability.

Some of the material is similar to the material in Appendix B, particularly for statistical tools that were used in both phases of the project.

B. Statistical Terms Used in this Document

The following table provides definitions of statistical terms that are used within this document. Appendix A contains a complete glossary for the report.

Statistical Terminology

Term	Definition
Occurrence of a parameter	A real occurrence of value. Represented by the symbol y.
Mean	The arithmetic average of a set of values. The sum of the values of the observations divided by the number of occurrences. Represented by the letter *m*.
Variance	A measure of the spread of the values of a parameter of interest. The square of the standard deviation. Generally, easier to calculate directly than the standard deviation. Represented by the symbol s^2.
Standard Deviation	A far more useful measure of the spread of the values of a parameter than the variance. The square root of the variance. Represented by the symbol *s*.
Z-Score	A measure of the value of a parameter as compared to the mean based on the standard deviation. $(y-m)/s$.
Large Sample	A sample large enough to prove statistically significant. As a rule of thumb, a sample with 30 or more cases. For Chi-squared analysis, each cell should have an expected value greater than 5.

Table C-1

C. Coefficient of Correlation

This is an indicator that tells the strength of the linear relationship between two variables. The two variables will be independent of their scales of measurement. It measures the linear correlation between two variables.

The Coefficient of Correlation is actually a measure of goodness of fit to a straight line developed through linear regression. Given a set of variables of (x,y) data, it is possible to create a line which closely fits the data. The Coefficient of Correlation is a measure of how well the line matches the data. A Coefficient of Correlation of 1 shows a perfect match, and a Coefficient of Correlation of 0 shows no match at all. A coefficient of -1 also shows a perfect match, but it indicates that the slope of the line is negative, as x increases, y decreases.

The Coefficient of Correlation does not demonstrate causality. For example, a high correlation between heart disease and high blood pressure does not demonstrate that heart disease causes high blood pressure. In this case, the observer might choose to speculate that the causality is actually the other way around.

In fact, a high Coefficient of Correlation does not say anything about causality at all. For example, there might be a high correlation between heart disease and lung cancer. However, there could well be no causality relationship between the two. Both of these might be caused by cigarette smoking. Nonetheless, if such a correlation existed, a doctor who found heart disease would then have appropriate justification for checking for lung cancer, regardless of the issue of causality.

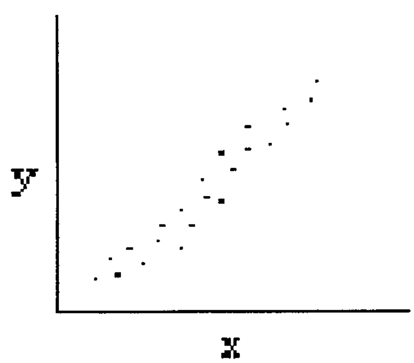


Figure C-1

The linear relationship of two variables is often displayed with a scatter diagram. A scatter diagram has an x and y axis and a set of points plotted on it. Correlation can be examined by looking at the following diagrams.

The first diagram, Figure C-1, exhibits a strong positive linear correlation. This shows that for increasing values of y there are increasing values of x. The points on the diagram begin in the bottom left and move towards the top right.

A strong correlation has a coefficient of near 1.00.

Figure C- 2

The second diagram, Figure C-2, exhibits a strong negative linear correlation. This shows that for increasing values of y there are decreasing values of x. They move in opposing directions somewhat. The points on the diagram begin in the top left and move towards the bottom right.

A strong negative correlation has a coefficient of near -1.00.

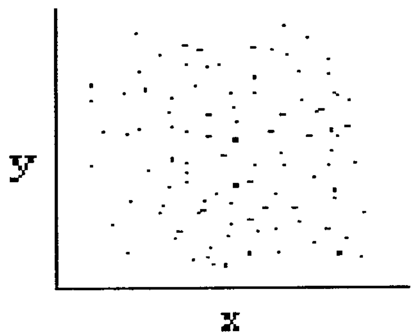


Figure C- 3

The third diagram, Figure C-3, exhibits no apparent linear correlation. This shows that the values of y and the values of x are in no order. The points on the diagram seem to be distributed all over the place.

No apparent correlation has a coefficient of near 0.00.

The Coefficient of Correlation can be any number from -1.00 through 1.00. This number represents a percentage (1.00 being 100 %). If the Coefficient of Correlation is .90 we say that there is a 90% correlation between the variables. This would be a very excellent correlation.

In the Phase II statistical analysis, Quattro Pro and Microsoft Excel were used to compute the Coefficients of Correlation during the research. Both of these application software packages have built in functions that calculate a Coefficient of Correlation.

D. Computation of Standard Deviation using a Binomial Probability Distribution

For many of the tests of randomness in this report, it was necessary to know both the mean of a sample (or population), and how far each of the occurrences varied from the mean. For example, one question that the study addressed was whether violations were randomly distributed among carriers for those carriers which had violation and mileage figures in Indiana.

Once miles to carriers were matched, we could look at the total number of violations for Indiana carriers who had miles and violations [3,693], and the total number of miles for the group [312,230,738]. Consequently, the (mean) number of miles per violation within this group [84,547] could be calculated.

It was clear which carriers had more violations per mile, and which had less, but this raw statistic was not particularly useful. For example, a carrier which went 100,000 miles and had two violations does not appear to be much worse than the average, where a carrier which went 1,000,000 miles and had 20 violations does seem much worse, even though the percentage over the mean is the same.

The reason for this was that the two carriers were a different number of standard deviations from the norm. It became clear that there was a need to calculate the standard deviation for each carrier in each test.

The appropriate statistical method is to use a binomial probability distribution. The binomial distribution describes the probability distribution for multiple occurrence of a case which can come out in two different ways, when the probability of the event in question (a "success") is known. The "fair" coin flip is a binomial event with the probability of a head being equal to .5 (50%).

For any sample group whose distribution can be described as binomial, the variance is defined as follows:

$$s^2 = np(1-p)$$

Where n is the number of samples and p is the probability of the event.

For example, consider a basketball shooter is a 70% foul shooter (i.e. $p = .7$). If she takes 10 shots in a practice session, her expected value is 7 goals (np), the variance is 2.1 and the standard deviation (square root of the variance) is 1.45. If she shoots 100 free throws, the expected value is 70, the variance is 21, and the standard deviation is 4.58. If she shoots 1,000 free throws, the expected value is 700, the variance is 210 and the standard deviation is 14.5.

Consequently, we would not be surprised if she shot 5 of 10 free throws ($z=-1.4$). However, if she shot 50 of 100 free throws ($z=-4.37$), we would reject the hypothesis that she is a 70% foul shooter.

The same logic was applied to motor carriers receiving violations. The project viewed the probability of receiving a violation for each mile driven. In the Indiana example described above, it was assumed, for this particular test, that the probability of a carrier (within the test population) receiving a violation in a mile driven was 1 in 84,547. Given the number of miles driven, the binomial probability distribution was used to calculate the variance and standard deviation. A Z-Score was assigned to each carrier for the test.

Again using the first example, the Z-Score for the first carrier was 0.75 and the Z-Score for the second carrier was 2.38.

A separate calculation for the standard deviation was made for each carrier for each test. Consequently, the Z-Score for a carrier was specific to the test in question. This explains why the reader may observe that a carrier will have a different Z-Score on each different test.

E. Confidence Level

The concept of confidence level appears frequently in statistics. While confidence level always means the same thing, it is interpreted differently in different situations.

In an example of the most common use, a pollster says that the results of the poll show that 30% ± 4.5% of the people support a particular politician, with a confidence level of 95%. What this actually means is that, if the same results came up a large number of times, the actual support for the politician within the total population would be between 25.5% and 34.5% in 95% of the cases. In this case, the pollsters went out of their way to assure that they had a random sample of the population.

In this project, confidence level is used technically in the same way. However, in different situations, it appears to be used in several slightly different ways.

1. Was a Sample Randomly Drawn from a Larger Population

One issue that was faced was whether a particular sample or sub-group was randomly drawn from the larger population. In this type of situation, the null hypothesis is that the sample was randomly drawn from the larger population. We reject the null hypothesis only when the data cannot support it.

In this case, we evaluate the sample group that came from a larger population. Evaluating the target parameters, we use ratios to calculate values of parameters within the sample group that would be expected if the sample group was drawn randomly from the population. In this way, we can accurately predict the expected value.

We know that the actual results will not be exactly what we calculated as the expected values (we would be suspicious if the match was too exact.) However, we can use the Chi Squared statistic to obtain a metric of how far the actual result is from the expected result.

We then used a look-up table, and find out what percentage of the time the Chi Squared can be expected to be less than its actual value, given the assumption that the sample was taken randomly from the population. If the answer is that the overwhelming proportion of the time (e.g. 98% or more of the time) that a sample had been drawn randomly from the population, the resulting Chi Squared would have been less than the actual Chi Squared, we reject the hypothesis that the sample was drawn randomly from the population.

2. Does the Actual Distribution Match the Expected Distribution

This is not a case of using a sample and a population. Rather it is a test of a population. If the value of a parameter is distributed randomly within a population, the distribution of that parameter will form an expected distribution.

- For low probability events, and particularly low probability events with respect to the number of trials, the probability distribution of resulting events will take on a Poisson Distribution. For example, if an event occurs once every 84,000 miles, the event will be considered a low probability event for a carrier which travels 100,000 miles, but not for a carrier which travels 1,000,000 miles.

- For fairly high probability events, the distribution of resulting events will be a normal distribution (bell curve).

- As the expected number of events increases, the Poisson Distribution approaches a bell curve.

For this project, given a particular population, we used the appropriate statistical distribution to project the anticipated distribution. We then used a Chi Squared test to determine if the actual distribution matched the expected distribution.

As noted in Chapter 3, there is a significant advantage to using a normal distribution. Thanks to the use of the binomial probability distribution to calculate the standard deviation, an analysis which uses a normal distribution as a target can use data from carriers with a wide range of actual mileages. A Poisson analysis is limited to comparing only carriers which have (approximately) the same number of miles.

Again, the null hypothesis was that the test parameter (generally violations) were randomly distributed among the population. We rejected the null hypothesis when the actual distribution of the test parameter in the population did not match the expected distribution.

In this case, the proper application of confidence limits was not quite so straightforward. We could have used the statistical standards for Chi Squared, as shown in Table C-2, which is repeated below, in cases where we were certain that we were using the correct distribution. However, in cases where a normal distribution was used for low probability events, the proper anticipated result did not exactly match the bell curve. We had to be careful in rejecting the null hypothesis. The Work Group chose to review the situation on a case by case basis, and consider how much the anticipated results should have varied from the results calculated using the normal distribution, and how much the actual results varied from the calculated anticipated results.

F. Chi Squared (C^2) Statistic

The Chi Squared statistic is used to compare the goodness of a fit between an expected distribution and an actual one. The Chi Squared Statistic (C^2) for Cell Probability is calculated as follows:

$$X^2=\sum_{i=1}^{n} q$$

where $q = \frac{[n_i - E(n_i)]^2}{E(n_i)}$; n_i is the observed cell count, and $E(n_i)$ is the expected cell count.

Note that X^2 is created through a summation process. The critical values of Chi Squared vary depending on how many cells (degrees of freedom) are involved.

The Chi Squared statistic describes the probability that the actual Chi-Squared value will be less than the given value, presuming that the sample is drawn randomly from the population. For example, assume that a number of occurrences are randomly drawn from a population and placed into 6 groups. If the appropriate frequency distribution can be predicted, 98% of the time the Chi Squared value for the actual, as compared to the predicted, will be 13.388 or less.

Table B-2 from Appendix B is repeated below with more technically correct column headings.

Chi Squared Values

Degrees of Freedom	98% Value for X^2
5	13.388
6	15.033
7	16.622
8	18.168
9	19.697
10	21.161
11	22.618
12	24.054
13	25.472
14	26.893

Table C-2

More complete treatments of Chi Squared, and the other topics addressed in this study, are available in most college and graduate level introduction to statistics textbooks.

G. NHSTA Review of Statistics

During the course of this project the Work Group met with various statisticians from the National Highway Safety and Transportation Administration. One meeting was held during Phase I of the project and another during Phase II. During these meetings, the current and future statistical analyses were discussed. The statisticians discussed methodologies, voiced concerns and made suggestions.

The Phase II meeting was attended by statisticians John Winnicki and Ellen Hertz of the National Highway Traffic Safety Administration (NHTSA) and included Paul Alexander of FHWA as the Driver/Carrier Project Manager. In this meeting, the overall project goals were discussed. Also discussed was how the decision was made to use a bell curve and normal distribution in the analysis. These statistical tools were selected for their ability to show that violations and carrier safety performance were not random.

The Indiana crash data was discussed. It was noted that it uses the same data as the Safety system. Also mentioned was a new approach to finding carriers with questionable safety practices.

The discussion also covered the statistical approaches that were used in Chapter 3 and 4 (There were no new techniques introduced in Chapters 5 and 6). The new analysis approach was also discussed. The statisticians expressed a great deal of interest in the analysis. They suggested that good exposure data was necessary to properly conduct the study. They had no objections to the statistical approaches used in this study and both were confident in the statistical methodology.

Appendix D - The States' Comments

This Appendix addresses the comments that were received from the States during Phase II.

A. Data

The States were concerned with properly collecting project data.

1. Data Collection

The States were concerned with the training of officers to properly collect the carrier data. Idaho produced a video that shows officers how to recognize the carrier and record carrier data properly. Indiana (officers) mentioned that certain Officers can only issue certain types of citations. For example, local authorities can issue moving violations but can not perform truck inspections.

Indiana had to correct some of its data before releasing it into the statistical analysis portion of the study. Indiana noted that it was their opinion that the safety rating was not useful. Whether a carrier is Conditional or Unsatisfactory may not mean anything. Indiana also stated that training for Officers varied and that Compliance Reviews were very manpower intensive. Indiana also reported that a carrier is allowed to report 612 IRP miles for its first two years of business. Indiana therefore wanted to get away from using mileage for exposure. They also mentioned that they had a hard time tying the citation to the carrier. "It was much easier to tie the citation to the driver," stated an officer. Indiana suggested that the USDOT number should be required on all trucks. This would make data collection easier.

Many of the States had a number of different projects underway, and had trouble finding the personnel to review the carriers.

2. Data Integrity

The integrity of the data is as good as the people who collected it. Adequate training, or any training in locating and recording carrier information, greatly improves the integrity of the data.

Manual matching of carrier names with the registrant names to locate IRP data had a set of problems associated with it.

3. Information Systems Utilized

The data that was collected was from various systems.

4. Inconsistent Citation Issuance by Officers

Officers issued citations differently. This variation in issuing citations was influenced by many factors. In North Dakota, Officers issue far fewer citations in the winter months than during other times of the year. The time of the day may also have an influence on citations issued.

5. Trends in Issuing Citations

Certain States had different ratios of violation types. There could be many reasons for this. States may participate in different projects which focus on a specific violation type.

In Michigan it was noted that size and weight violations had higher rates of driver license violations associated with them. Overlength violations did not have driver license violations associated with them.

6. Training

As noted, Idaho prepared a video to train officers to locate and record carrier information properly. North Dakota liked this video idea and requested a copy of Idaho's video for its own use.

B. Information Used That was Provided by the Project

The information used in the project was obtained through meetings and reports.

1. Meetings

Phase II project meetings were held in January and May of 1996. These meetings were very beneficial. They bought the States representatives together to discuss their ideas, experiences and issues. States were interested in what other States were doing and what they had done.

2. Reports

Indiana suggested adding address on the "*Carriers with 20 or More Violations*" report. This allowed for an easy comparison of a drivers set of carriers. For each driver there was a list of all the carriers they received violations with. Carriers which had change names may have still had the same address.

C. Carrier Review Process

The section covers the carrier review process and issues, accomplishments and improvements.

1. Issues

a. Lack of Confidence in Compliance Review

Sometimes it was hard to tell which carriers were really problem carriers since the State representatives had little confidence in the Compliance Review process. One State representative noted that a carrier, which was in the top 10 in the CDVR list and was the subject of an on-going investigation had two prior ratings of Satisfactory. Another noted that a company which had two consecutive Unsatisfactory ratings was known to be a fine carrier.

2. Accomplishments

The States reviewed many carriers. The reviews were done based on the reports that were supplied based on the citation data.

3. Improvements

The States agreed that there is a need to improve the current rating system. They think that an SCE score in no way reflects the safety fitness of a carrier in terms of targeting carriers for review.

D. Benefits of Using Carrier Information

Every participant indicate that there was benefit in using carrier information from citations. Each of the States said that they would continue to use carrier information even when the project is over. North Dakota wanted the database software that was used in the overall study so that it could set up its own system for researching citation data.

Acknowledgments

Listed below are the project participants for this Study.

- Paul Alexander, Project Manager, USDOT, FHWA, OMC.
- Patti Coman Drissel, Application Manager, Driver Systems, AAMVAnet, Inc.
- Duncan Macpherson, Chief, Office of Technology (Project Chairperson), New York DMV.
- Duane Sammons, Lieutenant, Idaho State Police.
- Guy Boruff, Lieutenant, Indiana State Police.
- Robert Corkwell, Sergeant, Indiana State Police.
- Lisa Jacobs, First Lieutenant, Michigan State Police.
- Sharon Van Campen, Sergeant, Michigan State Police.
- Kurt Schmidt, Data Processing Coordinator, Highway Patrol, North Dakota Highway Patrol.
- Walter Spiegel, Statistician, Keane Federal Systems, on behalf of AAMVAnet, Inc.

The Project Team is recognized as providing exceptionally high levels of knowledge, cooperation, and teamwork to this project.