FMCSA Safety Program Effectiveness Measurement: Carrier Intervention Effectiveness Model, Version 1.1 Technical Report

U.S. Department of Transportation
Federal Motor Carrier Safety Administration

April 2017

FOREWORD

The Federal Motor Carrier Safety Administration (FMCSA), in cooperation with the John A. Volpe National Transportation Systems Center (Volpe), has developed a quantitative model to measure the effectiveness of motor carrier interventions in terms of estimated crashes prevented, injuries prevented, and lives saved. The model documented in this report is known as the Carrier Intervention Effectiveness Model (CIEM). This model provides FMCSA management with information to address the requirements of the Government Performance and Results Act of 1993 (GPRA), which requires Federal agencies to measure the effectiveness of their programs as part of the budget cycle process. It also provides FMCSA and State safety program managers with a quantitative basis for improving enforcement processes and optimizing the allocation of safety resources in the field. This report documents the technical aspects of the model and presents results for fiscal years (FYs) 2010–12.

NOTICE

QUALITY ASSURANCE STATEMENT

Technical Report Documentation Page

<table>
<tr><td>1. Report No.
FMCSA-RRA-14-012</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
FMCSA Safety Program Effectiveness Measurement: Carrier Intervention Effectiveness Model, Version 1.1
Technical Report</td><td colspan="2">5. Report Date
April 2017</td></tr>
<tr><td colspan="2">6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Author(s)
Segev, Eran; Meltzer, Neil</td><td colspan="2">8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
John A. Volpe National Transportation Systems Center
Safety Management Systems Technical Center
System Measurement and Analysis Division
55 Broadway
Cambridge, MA 02142</td><td colspan="2">10. Work Unit No. (TRAIS)</td></tr>
<tr><td colspan="2">11. Contract or Grant No.</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
U.S. Department of Transportation
Federal Motor Carrier Safety Administration
Office of Analysis, Research, and Technology
1200 New Jersey Ave. SE
Washington, DC 20590</td><td colspan="2">13. Type of Report and Period Covered
Technical Documentation: FY10, FY11 and FY12 Results</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code
FMCSA</td></tr>
<tr><td colspan="4">15. Supplementary Notes
Contracting Officer's Representative: Richard Gruberg</td></tr>
<tr><td colspan="4">16. Abstract
The Federal Motor Carrier Safety Administration (FMCSA), in cooperation with the John A. Volpe National Transportation Systems Center (Volpe), has developed a quantitative model to measure the effectiveness of motor carrier interventions in terms of estimated crashes prevented, injuries prevented, and lives saved. The model documented in this report is known as the Carrier Intervention Effectiveness Model (CIEM). This model provides FMCSA management with information to address the requirements of the Government Performance and Results Act of 1993 (GPRA), which requires Federal agencies to measure the effectiveness of their programs as part of the budget cycle process. It also provides FMCSA and State safety program managers with a quantitative basis for improving enforcement processes and optimizing the allocation of safety resources in the field. This report documents the technical aspects of the model and presents results for fiscal years (FYs) 2010–12.</td></tr>
<tr><td colspan="2">17. Key Words
Motor carrier, compliance review, CSA, intervention, safety, program effectiveness, commercial motor vehicle, offsite investigation, onsite investigation</td><td colspan="2">18. Distribution Statement
No restrictions</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td>20. Security Classif. (of this page)
Unclassified</td><td>21. No. of Pages
51</td><td>22. Price</td></tr>
</table>

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized.

SI* (MODERN METRIC) CONVERSION FACTORS

Approximate Conversions to SI Units				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		Length		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		Area		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yards	0.836	square meters	m^2
ac	Acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		Volume (volumes greater than 1,000L shall be shown in m^3)		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		Mass		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2,000 lb)	0.907	megagrams (or “metric ton”)	Mg (or “t”)
		Temperature (exact degrees)		
°F	Fahrenheit	5(F-32)/9 or (F-32)/1.8	Celsius	°C
		Illumination		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		Force and Pressure or Stress		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

Approximate Conversions from SI Units				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		Length		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		Area		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
Ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		Volume		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		Mass		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or “t”)	megagrams (or “metric ton”)	1.103	short tons (2,000 lb)	T
		Temperature (exact degrees)		
°C	Celsius	1.8c+32	Fahrenheit	°F
		Illumination		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		Force and Pressure or Stress		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003, Section 508-accessible version September 2009.)

TABLE OF CONTENTS

LIST OF ACRONYMS, ABBREVIATIONS, AND SYMBOLS.. viii

EXECUTIVE SUMMARY x

1. INTRODUCTION...1

1.1 BACKGROUND 1

1.2 PROJECT SCOPE 1

2. MODEL METHODOLOGY 3

2.1 MODEL STRUCTURE..3

2.2 CARRIERS INCLUDED IN THE MODEL 5

2.3 CARRIER INTERVENTIONS 6

2.4 USE OF THE FIRST INTERVENTION..6

2.5 DATA REQUIREMENTS...7

2.5.1 Crashes.. 7

2.5.2 PUs.. 7

2.5.3 Interventions 8

2.5.4 Severity Statistics.. 8

2.6 DATA QUALITY CHECKS..8

2.7 STATISTICAL APPROACH..10

2.7.1 Treatment Group.. 11

2.7.2 Comparison Group.. 12

2.7.3 Calculation of Crash Rate Reduction.. 14

2.7.4 Test for Statistical Significance 14

2.7.5 Calculation of Treatment Group Crashes Prevented 15

2.7.6 Extrapolation to Entire Carrier Population 15

2.7.7 Calculation of Direct Safety Benefits 16

3. RESULTS OF IMPLEMENTING THE MODEL..19

3.1 RESULTS CONSIDERING ALL INTERVENTIONS...19

3.1.1 Crash Rate Reduction 21

3.1.2 Safety Benefits.. 22

3.2 LIMITATIONS OF THE MODEL..22

3.2.1 Comparison Group versus Control Group 22

3.2.2 Overall Impact Rather than Impact by Intervention Type 23

3.2.3 Data Quality and Timeliness... 23

3.2.4 Seasonality 24
3.2.5 Excluded Intervention Types 24
3.2.6 Benefit Duration.. 24
3.2.7 Variations of the Model 24

4. SUMMARY AND NEXT STEPS 25

LIST OF APPENDICES

APPENDIX A – DATA FILTERS..27

APPENDIX B – TEST FOR STATISTICAL SIGNIFICANCE 35

LIST OF FIGURES (AND FORMULAS)

Figure 1. Diagram. CIEM process flow. 4
Figure 2. Diagram. Timeframes for pre- and post-intervention periods for a carrier with a first intervention on August 15, 2012. 12
Figure 3. Diagram. Timeframes for pre- and post-intervention periods for FY 2012 comparison group carriers. 13
Figure 4. Formula. Calculation of crash rate reduction. 14
Figure 5. Formula. Calculation of crashes prevented. 15
Figure 6. Formula. Step 1 in performing extrapolation at the carrier size-group level. 15
Figure 7. Formula. Step 2 in performing extrapolation at the carrier size-group level. 16
Figure 8. Formula. Step 3 in performing extrapolation at the carrier size-group level. 16
Figure 9. Formula. Step 4 in performing extrapolation at the carrier size-group level. 16
Figure 10. Formula. Number of fatal crashes prevented. 17
Figure 11. Formula. Number of injury crashes prevented. 17
Figure 12. Formula. Number of lives saved. 17
Figure 13. Formula. Number of injuries prevented. 17
Figure 14. Venn diagram. Illustration showing intersection of carriers passing the CIEM's first two data filter criteria. 28
Figure 15. Graph. CIEM crash rate filter criteria for carriers previously screened for driver-to-carrier ratios and changes in carrier size. 28
Figure 16. Formula. Driver-to-PU ratio filter. 29
Figure 17. Graph. Change-in-carrier-size filter: primary criteria (carrier size groups 1 and 2). 31
Figure 18. Graph. Change-in-carrier-size filter: expanded conditional boundary limits for carrier size groups 1 and 2. 32
Figure 19. Formula. Estimating net crash rate reduction for a carrier size group. 35
Figure 20. Formula. Variance of the net reduction in crash rates. 36
Figure 21. Formula. Standard error of the net crash rate reduction. 36
Figure 22. Formula. 95-percent confidence interval for each estimate of net crash rate reduction. 36

LIST OF TABLES

Table 1. Adjusted crash rate reduction by carrier size group, FY 2010–12. xiv
Table 2. Estimated safety benefits, all interventions, FY 2010–12. xiv
Table 3. CIEM data filters. 9
Table 4. Carrier interventions by type and number of carriers by first intervention. 19
Table 5. Carriers excluded from treatment group by data quality filters and resulting treatment group totals. 20
Table 6. Treatment and comparison group carriers by size group, FY 2010-12. 20
Table 7. Treatment and comparison group crash rate reductions by size group, FY 2010–12. 21
Table 8. Net crash rate reductions by size group, FY 2010–12. 21
Table 9. Estimated safety benefits, treatment group and all carriers receiving interventions, FY 2010-12. 22
Table 10. Filter for change in carrier size—criteria for identifying outliers in carrier pre- to post-intervention PU ratio. 30
Table 11. Crash rate filter criteria 33

LIST OF ACRONYMS, ABBREVIATIONS, AND SYMBOLS

Acronym	Definition
ATET	Average Treatment Effect on the Treated
BASIC	Behavior Analysis and Safety Improvement Category
CIEM	Carrier Intervention Effectiveness Model
CMV	commercial motor vehicle
CR	compliance review
CREM	Compliance Review Effectiveness Model
CSA	Compliance, Safety, Accountability
CSP	cooperative safety plan
DNOC	direct notice of claim
DNOV	direct notice of violation
FMCSA	Federal Motor Carrier Safety Administration
FY	fiscal year
GPRA	Government Performance and Results Act of 1993
HM	hazardous materials
MCMIS	Motor Carrier Management Information System
MCS-150	Motor Carrier Identification Report
MCSAP	Motor Carrier Safety Assistance Program
PRISM	Performance and Registration Information Systems Management
PU	power unit
RIEM	Roadside Intervention Effectiveness Model
STD	standard deviation
USDOT	U.S. Department of Transportation
Volpe	John A. Volpe National Transportation Systems Center

[This page intentionally left blank.]

EXECUTIVE SUMMARY

This report provides a technical description of the methodology used by the Federal Motor Carrier Safety Administration (FMCSA) to estimate safety benefits stemming from the carrier intervention components of its motor carrier enforcement program. These benefits are calculated by the Agency's Carrier Intervention Effectiveness Model (CIEM).[1] The CIEM replaces an earlier Agency program effectiveness model, known as the Compliance Review Effectiveness Model (CREM). Major revisions to this previous model became necessary due to a phased redesign of the Agency's enforcement program beginning in 2010, known as Compliance, Safety, Accountability (CSA).

The CSA compliance and enforcement program includes multiple carrier intervention types (ranging from warning letters to onsite comprehensive investigations) that replace the one-size-fits-all compliance review (CR) intervention type previously used by the Agency. It is expected that the new enforcement model will result in an improved level of safety in commercial motor vehicle (CMV) operations.

The CIEM is designed to be implemented on an annual basis, focusing on carriers receiving interventions in a given fiscal year (FY). Assessing results over a period of years will provide an indication of the effectiveness of FMCSA's compliance and enforcement program in terms of safety benefits.

Both the CIEM and the CREM measure program benefits in terms of crash rate reduction, crashes prevented, lives saved, and injuries prevented. However, results from the two models are not directly comparable because they employ different methodologies to assess different safety programs. Therefore, the new model cannot provide a basis for determining whether interventions performed since the rollout of CSA led to an improvement over pre-CSA interventions.

MODEL APPROACH

The CIEM estimates safety benefits attributed to the set of interventions listed below, which are recorded in the Motor Carrier Management Information System (MCMIS).

- New CSA-related interventions types:[2]
  - Warning letters.

[1] Benefits derived from roadside inspections and traffic enforcements are estimated by a different model, known as the Roadside Intervention Effectiveness Model (RIEM).

[2] This version of the model does not include follow-up verifications, direct notices of violations (DNOVs), direct notices of claims (DNOCs), or cooperative safety plans (CSPs) because the data currently in MCMIS were found to be inconsistent in terms of completeness and accuracy. DNOVs and DNOCs generally pre-empt carrier investigations in the near term; they should not be confused with NOVs and NOCs, which are normally post-intervention enforcement types. CSPs are voluntarily initiated by carriers in collaboration with the FMCSA Division Office during or after a primary intervention. Safety audits are not considered a CSA intervention type. Neither are they included separately in this model because safety audits are performed only on new entrant carriers, which do not have a reliable pre-intervention period.

  - Offsite investigations.
  - Onsite focused investigations.
  - Onsite comprehensive investigations.
- CRs, including:
  - Standard CRs.
  - CRs with cargo tank facility review.
  - CRs with security contact review.
  - Non-ratable safety CRs on interstate carriers, including focused CRs (which do not receive a rating) and hazardous materials (HM) reviews.[3]
- Performance and Registration Information Systems Management (PRISM) warning letters.[4]

The model computes carrier crash rates, defined as crashes per carrier power unit (PU), for carriers receiving any of the above interventions, for the 12-month periods both prior to and following these interventions. With the CSA enforcement program, a motor carrier can receive multiple interventions within a short period of time. Due to this, the CIEM defines the motor carrier's post-intervention period as beginning at the time of its first intervention received during the modeled FY.

The difference between these carriers' pre- and post-intervention crash rates represents the change in their safety performance during this timeframe. To control for potential differences in the impact of interventions on small and large carrier operations, carriers are first placed into one of four size groups when calculating this change. The overall change in crash rate is then calculated for each size group. In addition, to remove the effect of confounding factors from the calculation of the change in safety performance, the changes in crash rates are adjusted by corresponding crash rate changes experienced by a comparison group comprised of each size group's general carrier population. A set of carefully designed filters is used to identify and remove missing and outlier carrier data.

Because it is not feasible to separate out the impact of individual interventions within a FY, the CIEM is not designed to estimate the effectiveness of any specific intervention type on its own. Rather, it estimates the crash rate changes associated with all of the intervention types considered in total. However, particular interventions can be excluded from the model for analytical purposes. For example, given a large increase in CSA warning letters issued in FY 2011, the CIEM was also implemented for that year only for 1) carriers whose first intervention was a warning letter, and for 2) carriers not receiving these warning letters. This was done in an effort

[3] Non-ratable review categories not included are limo, shipper, shipper facility/terminal, drug and alcohol, cargo tank, security contact, and commercial reviews.

[4] Further information on PRISM is provided by FMCSA at: https://www.fmcsa.dot.gov/information-systems/prism/performance-and-registration-information-systems-management-prism.

to determine which safety benefits observed in FY 2011 could be associated with warning letters. The results of such analyses can be found in the CIEM Report for FY 2012 Interventions.[5]

The CIEM incorporates the following additional enhancements in comparison with the CREM:

- Enhanced data filters to ensure the validity and sufficiency of the data used in the model, removing outliers while being careful to allow for legitimate data fluctuations.
- Statistical significance testing to help ensure that the benefits calculated by the model are not merely the result of random fluctuations in crash occurrence.
- Extrapolation of results to the entire population of motor carriers receiving interventions by incorporating benefits for carriers intervened upon that are initially excluded from the model due to insufficient or outlier data.
- Use of separate comparison groups for each carrier size group.

These enhancements are discussed briefly below.

Enhanced Data Filters

In the CIEM, outlier tests aim to screen out suspect PU data. Carrier driver-to-PU and PU-to-driver ratios cannot exceed 7.5, both pre- and post-intervention.[6] This value represents a boundary beyond which the vast majority of carriers do not operate. A second filter tests the pre- to post-intervention changes in PU counts—that is, the change in the size of the carrier. Very large changes in carrier size from the pre- to post-intervention periods are rejected by the model primarily because they have frequently been shown to be based on erroneous data; exceptions are made where the corresponding change in number of drivers is consistent with the change in PUs. The third filter examines the dataset for suspiciously low and suspiciously high crash rates, with thresholds tailored for each carrier size group.

Statistical Significance Testing

The CIEM employs statistical testing procedures to determine if the change in treatment group crash rate from the pre- to post-intervention period in each carrier size group, once adjusted for the comparison group crash rate change, is statistically significant at the 95-percent confidence level. Crash rate change estimates that do not pass this test do not reflect statistically robust model findings and are not used when estimating total crashes prevented due to interventions conducted during the FY.

---

[5] Available at: http://ntl.bts.gov/lib/60000/60500/60503/16-017-CIEM_Report-FY12-FINAL-508C.pdf.

[6] As an exception, carriers operating exclusively as driveaway/towaway carriers, transporting motor vehicles as a commodity with at least one set of the towed vehicle's wheels on the roadway surface, do not have to meet this filter's criteria. Such carriers by nature may have many more drivers than PUs for legitimate business and operations purposes. For more on driveaway/towaway operations, see the Federal Motor Carrier Safety Regulations, Part 390 Subpart A—General Applicability and Definitions.

Extrapolating Model Results to the Entire Motor Carrier Population

To account for interventions performed on carriers that do not pass the model's data requirements and outlier filters, the CIEM augments the initial estimates of crashes prevented by applying the observed average crash rate changes to the excluded carriers, within each carrier size group. The sum of the estimated crashes prevented by treatment group carriers and by the filtered carriers represents the total estimated crashes prevented as a result of carrier interventions.

Comparison Group Calculations

The model comparison group was formed from the population of carriers that did not receive interventions during the modeled FY, the FY preceding it, and the FY following it, and met the same crash and driver/PU filtering criteria as the treatment group. Unlike the CREM, separate comparison groups are created for each of the four size groups. Individual comparison groups based on carrier size allow the comparison group carriers to match the characteristics of the treatment group more closely.

In the CREM, the comparison group used 3years of FY data. The pre-intervention period was based on crash data from the current and previous FY, and the post-intervention period was based on data from the current and following FY. In the CIEM, however, there is no overlap between the comparison group's pre- and post-intervention periods; the pre-period is based on 18 months of crash data prior to April 1 of the FY, and the post-period is based on 18 months of data subsequent to April 1. April 1 is the midpoint of the FY being modeled, and extending by 18 months on either side of that date ensures that the comparison group time periods encompass the entire possible pre- and post-intervention periods.

MODEL FINDINGS

All Carriers Receiving Interventions

The model was implemented for carriers receiving interventions in FY 2010, FY 2011, and FY 2012. Total interventions decreased from 58,199 in FY 2011 to 43,275 in FY 2012. The decrease primarily reflects a decline in CSA warning letters from a peak in FY 2011. Statistically significant crash rate reductions were observed for all 3 years for carriers with up to 20 PUs. For carriers with 21–100 PUs, statistically significant reductions in crash rates were observed in FY 2011 and FY 2012; there was a small increase in crash rate in FY 2010, but it was statistically insignificant. For carriers with 100 PUs or more, the results were not statistically significant for any of the 3 years. Table 1 shows the observed crash rate reductions by carrier size group for FY 2010, 2011, and 2012, adjusted for the crash rate changes exhibited by the comparison group.

Table 1. Adjusted crash rate reduction by carrier size group, FY 2010–12.

Carrier Size Group	FY 2010	FY 2011	FY 2012
1 (1–5 PUs)	32.7%	34.3%	37.5%
2 (6–20 PUs)	16.0%	31.2%	33.4%
3 (21–100 PUs)	-1.2%*	16.7%	17.0%
4 (100+ PUs)	-5.2%*	4.0%*	5.5%*

Note: Negative crash rate reductions indicate increases in crash rates.

*Non-statistically significant adjusted reduction.

As shown in Table 2, these reductions are estimated to have resulted in the following safety benefits:

Table 2. Estimated safety benefits, all interventions, FY 2010–12.

FY	Crashes Prevented	Injuries Prevented	Lives Saved
2010	1,830	1,142	59
2011	6,567	4,033	215
2012	5,283	3,235	173

[This page intentionally left blank.]

1. INTRODUCTION

1.1 BACKGROUND

During the 1980s, Congress passed a series of legislative acts intended to strengthen motor carrier safety regulations. These measures led to the implementation of safety-oriented programs at both the Federal and State levels. The Surface Transportation Assistance Act of 1982 established the Motor Carrier Safety Assistance Program (MCSAP), a grants-in-aid program to States for implementing roadside inspection and traffic enforcement programs aimed at commercial motor vehicles (CMVs). The Motor Carrier Safety Act of 1984 directed the U.S. Department of Transportation (USDOT) to establish safety fitness standards for carriers. The USDOT, through the Federal Motor Carrier Safety Administration (FMCSA) and in conjunction with the States, implemented MCSAP to fund roadside inspection and traffic enforcement programs, the safety fitness determination process, and a commercial motor carrier rating system based on onsite safety audits called compliance reviews (CRs).

The Safety Program Effectiveness Measurement Project was established to develop results-oriented performance measures for FMCSA's functions and operations, as called for in the Government Performance and Results Act of 1993 (GPRA). From 2002 through 2009, the benefits of CR activities were assessed using the Compliance Review Effectiveness Model (CREM).[7] In 2010, following an Operational Model Test in select States, FMCSA began a phased implementation of its Compliance, Safety, Accountability (CSA) program, a redesign of the Agency's enforcement program. The CSA enforcement approach includes multiple carrier intervention types that replace the one-size-fits-all CR intervention type previously used by the Agency. It is expected that this new enforcement program will result in an improved level of safety in CMV operations. This program, however, has necessitated an enhanced approach for measuring the effectiveness of Agency interventions at a national level.

1.2 PROJECT SCOPE

The Carrier Intervention Effectiveness Model (CIEM) provides FMCSA with a tool for measuring the safety benefits of carrier interventions currently used by the Agency under its CSA program. The model incorporates both CRs (previously included in the CREM) and additional interventions, including: warning letters, offsite investigations, onsite focused investigations, and onsite comprehensive investigations. This approach yields national-level measurements of the effectiveness of FMCSA's carrier interventions.[8]

While the new model replaces the CREM, results from the two models are not directly comparable because the methodologies used to assess safety benefits in each are not identical.

---

[7] Reports documenting these results are available at http://ai.fmcsa.dot.gov/pe/home.aspx.

[8] The carrier interventions measured by this model are distinct from roadside interventions, measured directly by FMCSA's Roadside Intervention Effectiveness Model (RIEM).

Therefore, the two models do not provide a basis for comparing the safety benefits derived from the CSA program with those obtained from the Agency's prior enforcement program.

This report presents the CIEM methodology and technical approach, as well as results from implementing the CIEM for carriers receiving interventions in fiscal years (FYs) 2010, 2011, and 2012.

2. MODEL METHODOLOGY

FMCSA employs a data-driven approach to oversee and enforce commercial motor carrier safety. This approach utilizes a variety of data sources to assign safety risks to motor carriers, and the assigned safety risks are used to prioritize carriers for interventions intended to improve motor carrier safety. The CSA enforcement program introduced a new and broader set of carrier interventions, giving FMCSA enhanced flexibility to address motor carrier safety problems. The CIEM measures the combined impact of motor carrier interventions used by the Agency in a given FY, including those introduced by the CSA program and pre-CSA carrier interventions, in terms of crashes and injuries prevented and lives saved.

2.1 MODEL STRUCTURE

The CIEM is a statistical impact evaluation model that uses historical data to compare the safety improvement of carriers receiving FMCSA interventions to their past safety performance, prior to receiving interventions.[9] This comparison is used to establish the extent of safety improvement that can be attributed to interventions. Safety improvement for a carrier is defined as the proportional change in its crash rate; that is, the difference between its pre- and post-intervention crash rates as a proportion of its pre-intervention crash rate.[10] The model is designed to be implemented on an annual basis, focusing on carriers receiving interventions in a given FY.

The model computes aggregate carrier crash rates (defined as crashes per power unit [PU]), for carriers receiving interventions, for the 12-month periods both prior to and following each carrier's first intervention of the FY.[11] The difference between these carriers' pre- and post-intervention crash rates represents the change in their safety performance during the model's timeframe. To control for potential differences in how small and large carriers respond to interventions, carriers are first placed into one of four size groups when calculating this change, and the overall change in crash rate is calculated for each size group. These size groups are defined as follows:

- Size Group 1: 1–5 PUs.
- Size Group 2: 6–20 PUs.
- Size Group 3: 21–100 PUs.
- Size Group 4: more than 100 PUs.

[9] For a more detailed explanation of statistical evaluation design, see the discussion provided by the World Bank on Impact Evaluation Methods and Techniques: http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,contentMDK:20188242~menuPK:415130~pagePK:148956~piPK:216618~theSitePK:384329,00.html.

[10] As shorthand in the remainder of this report, this difference is referred to as the "difference between crash rates" or the "change in crash rates."

[11] PU values are used as a proxy for carrier exposure to crashes. While vehicle miles traveled have the potential to serve as a useful proxy for exposure in the model at a future point in time, FMCSA considers PU information currently in the Motor Carrier Management Information System (MCMIS) to be more reliable.

Carriers in these different size groups have distinct characteristics, and therefore, may respond differently to interventions.

To remove the effect of confounding factors from the calculation of change in safety performance, the difference between pre- and post-intervention crash rates in each size group is adjusted by the change in crash rates experienced by a comparison group (representing those that did not receive interventions) during a similar timeframe. This adjustment removes the potential effect of historical trends and events (such as the national recession that occurred during the timeframe represented by the model results presented in this report).

Figure 1 shows the steps of the model in a process flow diagram.

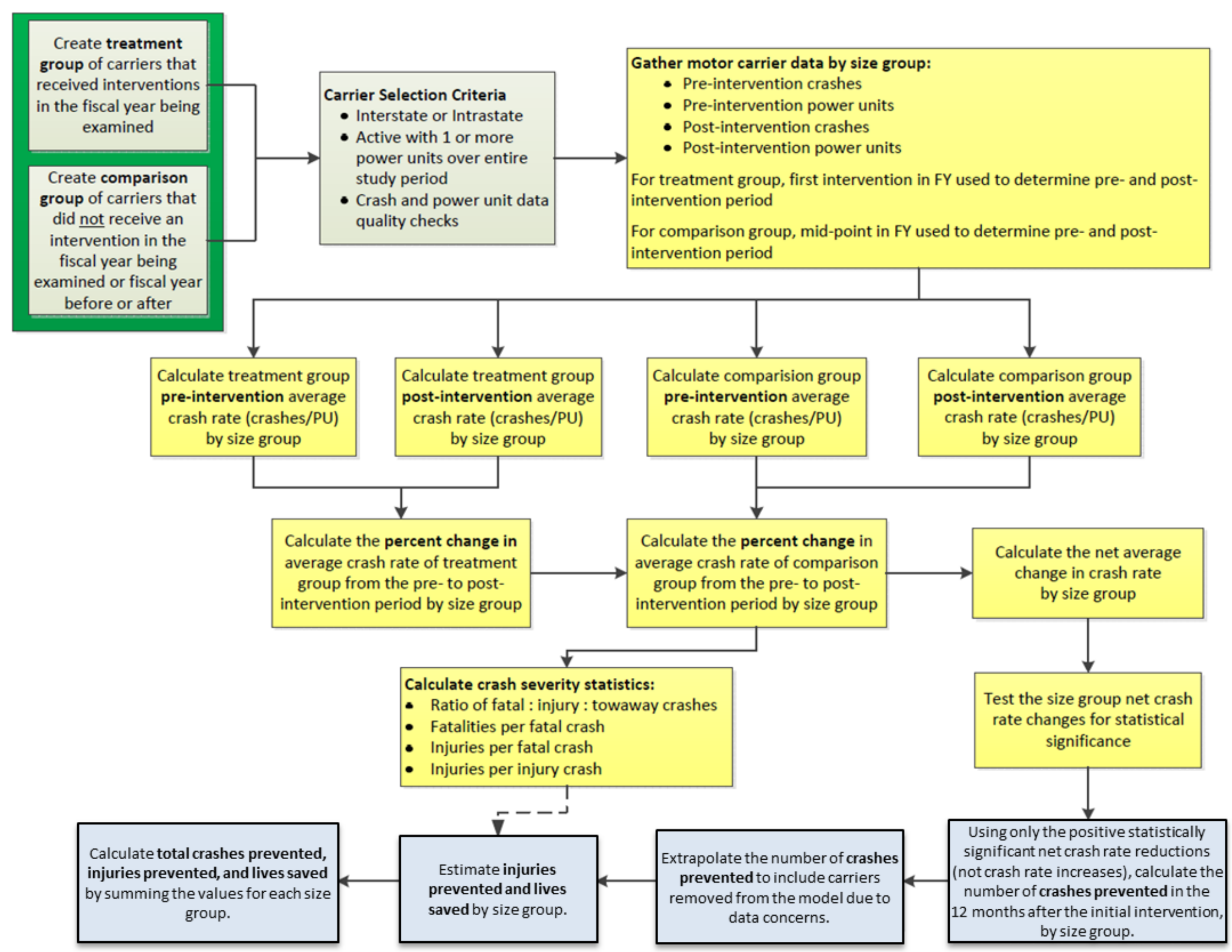

Figure 1. Diagram. CIEM process flow.

As shown in Figure 1, the first step in the model is to create a dataset of active carriers that received interventions during the modeled FY (treatment group) and a dataset of carriers that did not (comparison group). These carriers are then screened for data quality. For carriers passing the data quality filters, crash and PU data are added to each dataset. This information is used to compute crash rates (crashes per PU) by size group for the pre-intervention and post-intervention periods.

The changes in crash rates from the pre-intervention period to the post-intervention period are then calculated for each carrier size group. These treatment group crash rate changes are adjusted by the corresponding comparison group's change in crash rate, within each size group, to account for external factors that may influence carrier safety performance. The net crash rate change is then tested for statistical significance, by carrier size group. Non-statistically significant crash rate changes are excluded from the estimation of safety benefits calculated in the model. In addition, the model does not allow for the possibility that an intervention may increase the likelihood of a motor carrier having a crash, assigning a value of zero benefits to size groups showing net increases in crash rates.

Statistically significant crash rate reductions are then used to estimate the number of crashes prevented as a result of the interventions. The benefits are then extrapolated to carriers initially excluded from the computation as a result of the data quality filters. As a final step, crash severity data from FMCSA's Motor Carrier Management Information System (MCMIS) are used to estimate the total numbers of injuries prevented and lives saved from the estimated crashes prevented in each size group.

2.2 CARRIERS INCLUDED IN THE MODEL

The intent of the model is to provide estimates of safety benefits for carriers that received specific FMCSA interventions. To eliminate potential bias from carriers subject to other programs or jurisdictional rules and regulations, initial filters are applied to screen carriers for inclusion in the model.

Carriers included in the model's treatment or comparison groups are all carriers eligible for FMCSA interventions, including interstate, intrastate non-hazardous materials (HM), and intrastate-HM, as well as foreign-domiciled carriers. To avoid confounding model benefits with those stemming from new entrant safety audits, both treatment and comparison group carriers must not have been classified as "new entrant" carriers during any part of the pre- and post-intervention periods associated with each respective group.

In addition to the requirements described above, the carriers' information in MCMIS is also screened for erroneous data that could introduce bias into the model. Carriers must be active as indicated by the status field in MCMIS throughout the pre- and post-intervention periods and must have nonzero PU data documented in MCMIS.[12] Carriers are also subject to outlier tests to identify suspect crash and PU data; these tests are discussed in more detail in Section 2.6.

The model's treatment group consists of carriers that received at least one FMCSA carrier intervention (as defined in Section 2.3) during the FY. The comparison group consists of carriers that did not receive any of the model's interventions during the pre- and post-intervention periods, which include the modeled FY, the prior FY and the following FY.

[12] Carriers that received an out-of-service order during the pre- or post-intervention periods are removed as well, to account for the fact that such carriers may have been practically inactive during a portion of the evaluation timeframe.

2.3 CARRIER INTERVENTIONS

The CIEM estimates safety benefits attributed to the set of interventions listed below. These interventions are recorded in MCMIS and used to select treatment group carriers:

- CSA interventions.[13]
  - Warning letters.
  - Offsite investigations.
  - Onsite focused investigations.
  - Onsite comprehensive investigations.
- Compliance reviews (CRs), including:
  - Standard CRs.
  - CRs with cargo tank facility reviews.
  - CRs with security contact reviews.
  - Non-ratable safety CRs on interstate carriers, including focused CRs (which do not receive a rating) and HM reviews.[14]
- Performance and Registration Information Systems Management (PRISM) warning letters.[15]

2.4 USE OF THE FIRST INTERVENTION

Because carriers may receive multiple interventions in a FY, the model uses the date of each carrier's first intervention in the FY to delineate its pre- and post-intervention periods.[16] The pre-intervention period is then defined to be the 12-month period immediately prior to this date, and the post-intervention period is defined to be the 12-month period immediately following this date.

Because the first intervention during the FY is used as a demarcation point, the safety impacts of subsequent interventions in the same year may be reflected in the model results. Specifically, those subsequent interventions that occur before the end of the carrier's post-intervention period

---

[13] This version of the model does not include follow-up verifications, direct notices of violation (DNOVs), direct notices of claims (DNOCs), or cooperative safety plans (CSPs) because the data currently in MCMIS were found to be inconsistent in terms of completeness and accuracy. DNOVs and DNOCs generally pre-empt carrier investigations in the near term; they should not be confused with NOVs and NOCs, which are normally post-intervention enforcement types. CSPs are voluntarily initiated by carriers in collaboration with the FMCSA Division Office during or after a primary intervention. Safety audits are not considered a CSA intervention type. Neither are they included separately in this model, because safety audits are performed only on new entrant carriers, which do not have a reliable pre-intervention period.

[14] Non-ratable review categories not included are limo, shipper, shipper facility/terminal, drug and alcohol, cargo tank, security contact, and commercial reviews.

[15] Further information on PRISM is provided by FMCSA at: https://www.fmcsa.dot.gov/information-systems/prism/performance-and-registration-information-systems-management-prism.

[16] If the first intervention for the fiscal year is a warning letter, the letter date is used to demarcate the pre- and post-intervention periods.

may have a sizable impact upon the carrier's safety performance during this same period, as reflected in the post-intervention crash rates calculated by the model.[17] The impact of interventions that take place after the post-intervention period are not accounted for in the model estimates for the current FY, but rather in the next annual implementation of the model. Thus in the next model year, the first of those follow-up interventions would be used to delineate new pre- and post-intervention periods.

Because some carriers receive multiple interventions within the modeled year, the model cannot determine the precise impact of each individual intervention type; rather, it estimates the aggregate benefits of all interventions performed during the modeled year on motor carrier safety.

2.5 DATA REQUIREMENTS

The model requires the data inputs detailed below to produce the estimated safety benefits.

2.5.1 Crashes

Motor carrier crash data are reported to the Agency by the States and housed in MCMIS. A crash report is submitted for any crash in which a motor carrier is involved, and which results in at least one fatality, an injury resulting in immediate medical treatment away from the crash scene, or towing of any vehicle from the crash scene due to disabling damage. These data, originating from State reporting systems, are continuously entered into MCMIS via an automated interface. Consequently, statistics for previous time periods may change, depending on the timeliness and completeness of the initial crash reporting. In the CIEM, MCMIS data snapshots from the most recent month—which include the most current updates for prior months—are used to provide the most complete and accurate crash data available for the model.[18]

2.5.2 PUs

This data element refers to the total number of PUs owned, trip-leased, and term-leased by the carrier.[19] PU data are obtained during interventions, or, in the absence of intervention data, from Motor Carrier Identification Report (Form MCS-150) information submitted by carriers when registering with the Agency. These data are available in MCMIS. Carrier PU counts based on the MCMIS snapshot taken immediately following the interventions are used to classify carriers into size groups and determine pre-intervention carrier crash rates. Carrier PU counts based on the last MCMIS snapshot within the post-intervention periods are used to determine post-intervention carrier crash rates.

---

[17] Such subsequent interventions number approximately 1,500 to 2,000 per year, representing approximately 3–7 percent of each year's total interventions.

[18] The December 2014 MCMIS data snapshot was used for this report.

[19] The model considers the following types of vehicles as PUs: straight trucks, truck tractors, HM cargo tank trucks, motorcoaches, school buses, mini-buses, 9–15 passenger vans, and limousines with 9 or more seats.

2.5.3 Interventions

Interventions are FMCSA actions aimed at improving a carrier's safety performance. FMCSA aims to correct high-risk behaviors by contacting carriers with interventions tailored to their specific safety challenges. Intervention data are input by safety investigators and stored in MCMIS. The CIEM estimates the safety benefits of the carrier-focused interventions listed in Section 2.3.

2.5.4 Severity Statistics

The model uses historical crash severity data from MCMIS to calculate the estimated number of lives saved and injuries prevented as a result of the crashes prevented by performing interventions on motor carriers. This model uses 2-year average statistics for estimating the probability of a crash involving an injury or fatality, along with 2-year average values for estimating the number of injuries and fatalities in such crashes. These values, based on crashes occurring during the modeled year and previous year, are multiplied by the estimated crashes prevented. This step yields estimates of the numbers of injuries prevented and lives saved.

2.6 DATA QUALITY CHECKS

The efficacy of the model results depends on the reliability of the carrier-level data. To ensure the quality of the model inputs, data filters are used to remove seemingly erroneous and outlier data from both treatment and comparison groups. The filters were developed through methodical analysis to minimize inadvertent removal of carriers with accurate data. A summary of the data filters is provided in Table 3, and a more technical discussion of the filters can be found in Appendix A.

Table 3. CIEM data filters.

	Definition	Size Group	Lower Limit	Upper Limit	Exceptions to the Primary Filter Conditions
Primary Filters					
Driver/PU Ratio	Total drivers divided by total PU (pre- and post-intervention)	All	0.1334	7.5	No driver/PU ratio limit for driveaway/towaway carriers
Change in Carrier Size (Pre- to Post-Intervention PU Ratio)	Pre-intervention PU divided by post-intervention PU	1 & 2	0.3334	3.0	If a corresponding change in driver count is observed, then: Lower limit: 0.2 Upper limit: 5.0
		3	0.571	1.75	If a corresponding change in driver count is observed, then: Lower limit: 0.4 Upper limit: 2.5
		4	0.571	1.75	No exception
Secondary Filter					
Carrier Crash Rates	Difference from mean crash rate, by size group	1	-5 STD	+5 STD	Up to 5 crashes
		2	-5 STD	+5 STD	Up to 6 crashes
		3	-5 STD	+5 STD	Up to 6 crashes
		4	-5 STD	+5 STD	Up to 6 crashes; Carriers with 500 or more PUs must have at least 1 crash

The first outlier test aims to screen out bad PU data. Carrier driver-to-PU ratios must be between 7.5 and 1/7.5, in both pre- and post-intervention periods.[20] These values represent thresholds beyond which the vast majority of carriers do not operate.

The second filter tests the pre- to post-intervention and post- to pre-intervention changes in PU counts; that is, the change in the size of the carrier. Very large changes in carrier size from the pre- to post-intervention periods are rejected by the model primarily because frequently they have been shown to be based on erroneous data. In some extreme cases, such changes are legitimate (e.g., in the case of a merger) but represent a substantial confounding factor in terms of measuring crash rate changes between the two periods. For such cases, a conditional filter is employed to prevent removal of carrier data that may be legitimate. This filter uses two data measurements in combination to determine data validity. Carriers that do not pass an initial

[20] An exception to this filter is made for carriers operating exclusively as driveaway/towaway carriers, transporting motor vehicles as a commodity with at least one set of the towed vehicle's wheels on the roadway surface. Such carriers by nature may have many more drivers than PUs for legitimate business and operations purposes. For more on driveaway/towaway operations, see the Federal Motor Carrier Safety Regulations, Part 390 Subpart A—General Applicability and Definitions.

outlier test are judged again, based on additional criteria, before being removed from the sample (see Table 3, last column).

The primary condition for this second filter is the ratio between pre-and post-intervention PU counts. In most circumstances this ratio cannot exceed a factor of 3 or be less than one-third (1 divided by 3) for size groups 1 and 2. For size groups 3 and 4, the ratio cannot exceed a factor of 1.75 or be less than 0.571 (1 divided by 1.75). However, exceptions to this filter are made when the change in driver counts between the pre- and post-intervention periods is consistent with the observed changes in PU counts. When these measures are consistent, the observed larger PU changes are more plausible and less likely to be erroneous. These exceptions have been set as follows: size group 1 and 2 carriers can exhibit a factor up to 5 (or down to 0.2) if there is a corresponding change in the pre- to post-intervention or a post- to pre-intervention change in driver count (between a factor of 1.5 and 10); and size group 3 carriers can exhibit a factor up to 2.5 (or down to 0.4) if there is a corresponding change in driver count (between a factor of 1 and 5).[21] This exception allows more variability for smaller carriers because smaller PU changes result in larger proportional changes for these carriers compared to larger carriers; it relies on the change in driver count as an additional data point on which to accept or reject the PU-change value. This approach results in more inclusive and flexible consideration of outlier values, increasing the confidence that outliers are identified as such and not merely as extreme values.

The third filter examines the dataset for suspiciously low and suspiciously high crash rates. The pre- and post-intervention crash rates must be within five standard deviations of the mean crash rate for the size group to which the carrier belongs, once all other filters have been implemented. Based on analysis of carrier crash incidence, this condition can be overridden by the following exceptions: carriers in size group 1 can have up to 5 crashes; carriers in size groups 2, 3, and 4 can have up to 6 crashes; and carriers with 500 or more PUs must have been involved in at least 1 crash.

All the treatment group filters discussed above use driver, PU, and crash data from MCMIS data snapshots current as of the carrier's first intervention date. The comparison group filters are identical to the treatment group filters, but rely on different data snapshots due to the absence of intervention dates for comparison group carriers. As illustrated in Section 2.7.2, the comparison group pre- and post-intervention periods are defined around the midpoint (April 1st) of the modeled FY.

2.7 STATISTICAL APPROACH

The CIEM estimates the reduction in carrier crash rates due to FMCSA interventions. This statistical method of estimating a causal effect is based on an Average Treatment Effect on the

[21] The limits on the change in driver counts are more lenient than on the change in PU counts to allow for driver count changes somewhat below or somewhat above the changes in PUs, since carrier changes in PUs and driver counts cannot be expected to be precisely the same.

Treated (ATET) approach.[22] ATET models assess the effect of a treatment by comparing the performance of a population that receives a treatment with the performance of the same population if they did not receive the treatment. The latter is not an observable value; therefore, the performance of a comparison group is used to approximate. For the purposes of this model, the crash rate change over time for carriers that received an intervention (i.e., the treatment group) is compared to the crash rate change over time for carriers that did not receive an intervention (i.e., the comparison group).

2.7.1 Treatment Group

The model treatment group comprises carriers that received an intervention during the modeled year and passed all filters and quality checks. The change in treatment group crash rates are computed by carrier size groups (see Section 2.1 for size group definitions). The size group's change in crash rate is computed, based on all crashes occurring before and after each carrier's first intervention of the FY, for all carriers in the size group.

For each carrier, the 12-month period preceding its first intervention is defined as its pre-intervention period, while the 12-month period following this intervention is defined as its post-intervention period. For the treatment group, a carrier's pre-intervention PU value is based on the first monthly MCMIS data snapshot following the first intervention it receives during the FY. This particular snapshot contains the most recent PU information for the carrier at the time of its intervention. The final monthly snapshot for a carrier's post-intervention period is used to define its post-intervention PU value. As an example, Figure 2 illustrates the pre- and post-intervention timeframes delineated by a representative carrier in the treatment group for FY 2012, noting the snapshots used for PU data.

[22] See Abadie, Alberto (2005). *Semiparametric Difference-in-Differences Estimators*, Review of Economic Studies (72, 1-19) for further information on ATET.

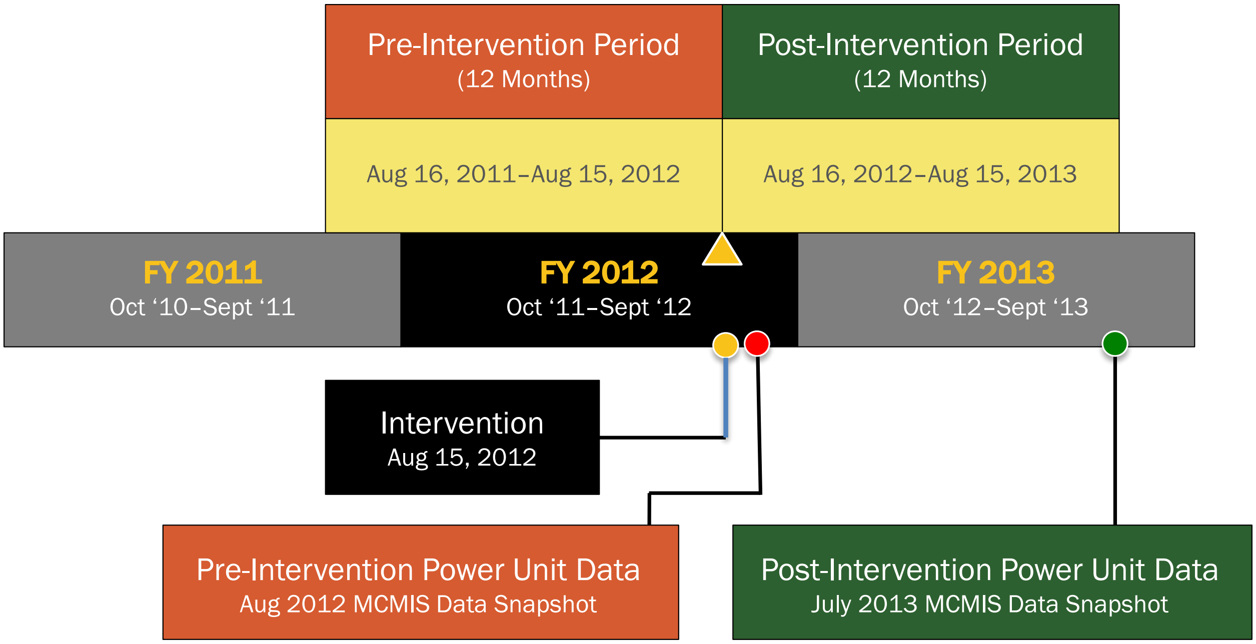

Figure 2. Diagram. Timeframes for pre- and post-intervention periods for a carrier with a first intervention on August 15, 2012.

Pre- and post-intervention crash rates are calculated for each size group as the number of crashes occurring in the size group during these two periods, divided by each period's PU total for the size group. These pre- and post-intervention crash rates are used to calculate the change in the size group's crash rate during the model timeframe.

2.7.2 Comparison Group

The model comparison group was formed from the population of carriers that did not receive interventions during the model timeframe (encompassing the modeled FY and both the prior and following years); carriers used pass a set of filters for missing and outlier data similar to those applied to treatment group carriers.[23] Similar to the treatment group filters, comparison group filters ensure that crash rates are comparable and reliable across carriers and carrier size groups.

Unlike the CREM, which used one comparison group comprising carriers from all size groups, the CIEM utilizes separate comparison groups for each of the four size groups. Matching based on carrier size results in comparison group carriers that resemble the characteristics of the treatment group more closely than a single group. Both the treatment and comparison groups represent a wide range of motor carrier operations, reflecting differences in factors such as weather, congestion, density, and operation types (e.g., short-haul versus long-haul operations).

[23] The comparison group filters are identical to the treatment group filters. However, since the comparison group carriers do not have intervention dates, their power unit data for these calculations are always based on the modeled year's MCMIS April data snapshot for the pre-intervention period and on the subsequent year's September snapshot for post-intervention period.

Having a wide range of carrier operations in both the treatment and comparison groups minimizes the impact of operational factors on the model. Directly stratifying for such factors could achieve similar results. However, such stratification is not feasible, as it would require reducing the size of treatment groups substantially, inhibiting the model's ability to detect statistically significant differences between pre- and post-period crash rates.

For comparison group carriers, periods corresponding to the treatment group's pre- and post-intervention periods are defined as the 18 months preceding and following the midpoint of the FY (March 31st), respectively. Thus, all carriers in this group are assigned the same start and end dates for the pre- and post-intervention periods. Also, the pre-intervention period covers the entire FY prior to the modeled year, while the post-intervention period covers the entire FY following the modeled year. These longer pre- and post-intervention periods for the comparison group (compared to 12-month periods for the treatment group) ensure that the comparison group data cover the entire potential timeframe for the treatment group: a full 36 months.

The April MCMIS snapshot is used as the basis for the comparison group's pre-intervention period PU value, and the September snapshot in the following year is used for that period's PU value. Similar to the treatment group, comparison group carriers' crash rates are calculated as the number of crashes occurring during each period divided by the corresponding PU values. However, to account for the comparison groups' pre- and post-intervention periods being longer than those for the treatment group, (18 versus 12 months) comparison group crash rates are divided by 1.5 to yield equivalent annual crash rates. Figure 3 illustrates the timeframes delineated by these data points for a representative carrier in the comparison group for FY 2012.

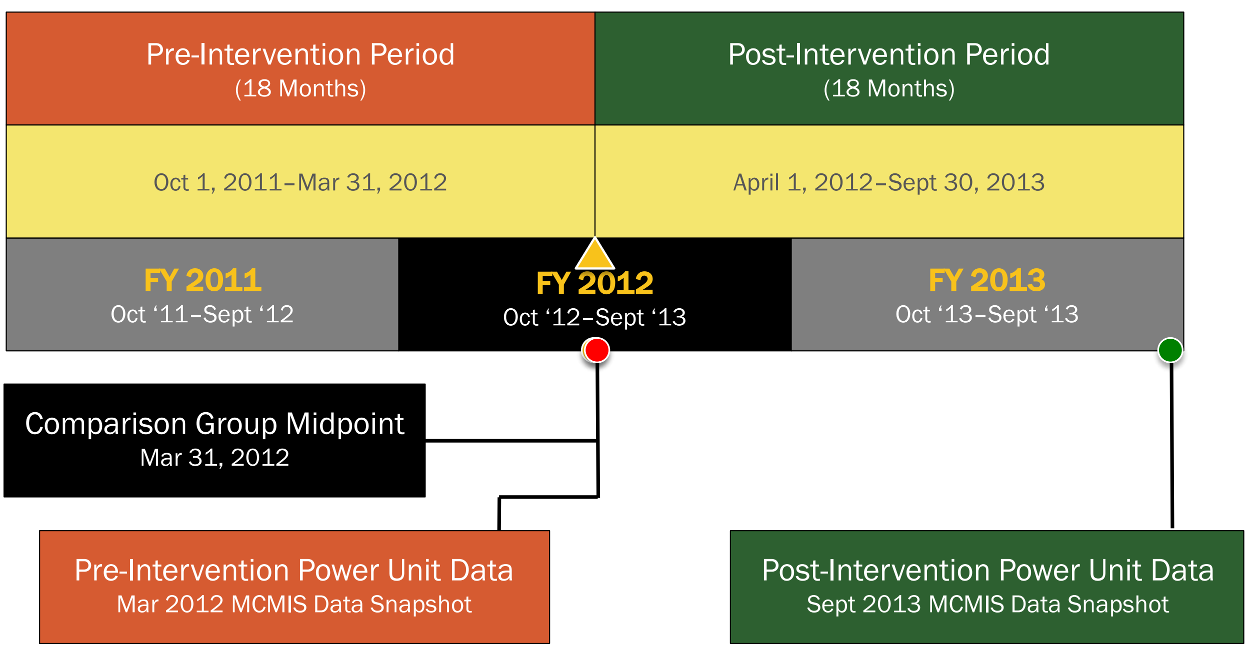

Figure 3. Diagram. Timeframes for pre- and post-intervention periods for FY 2012 comparison group carriers.

2.7.3 Calculation of Crash Rate Reduction

Pre- and post-intervention crash rates are used by the model to determine the change in crash rates, by carrier size group, for the treatment and comparison groups. The absolute difference in crash rates is converted to a percent measure by dividing the difference by the original (pre-intervention) crash rate. This conversion is made to account for the treatment and comparison groups' different starting crash rates; without converting to a percent value, the two groups' measures would be on different scales. The difference between the treatment and comparison groups' crash rate changes (ATET) is the crash rate reduction attributed to interventions. Figure 4 illustrates the steps used to determine this reduction.

$$\begin{array}{l}\textit{Crash Rate Reduction}\\ \textit{Due to Interventions}\end{array} = \boxed{\begin{array}{l}\textbf{Treatment Group}\\ \text{Crash Rate Reduction,}\\ \text{pre- to post-intervention}\\ \qquad \dfrac{[CR_{PRE} - CR_{POST}]}{CR_{PRE}}\end{array}} - \boxed{\begin{array}{l}\textbf{Comparison Group}\\ \text{Crash Rate Reduction,}\\ \text{pre- to post-FY midpoint}\\ \qquad \dfrac{[CR_{PRE} - CR_{POST}]}{CR_{PRE}}\end{array}}$$

Figure 4. Formula. Calculation of crash rate reduction.

where,

CR_{PRE} = Pre-intervention crash rate

CR_{POST} = Post-intervention crash rate

2.7.4 Test for Statistical Significance

To ensure that the statistical estimates of crash rate reduction do not include, to the extent possible, random variations or fluctuations in numbers of crashes from year to year, the results were tested for statistical significance. The goal of the test is to identify which of the initial estimates are statistically significant at a target level of confidence (in this analysis, the 95-percent confidence level). This test determines if treatment group crash rate reductions, adjusted for the comparison group crash rate change, are statistically significantly different from zero in each size group, at the 0.05 statistical significance level (i.e., the 95-percent confidence interval around the estimated effect on crash rates does not include zero).[24] Crash rate change estimates that do not pass this test do not reflect statistically robust findings and thus are not used to estimate crashes prevented. The crash rate reductions found to be statistically significant are used to calculate estimates of crashes prevented that can be attributed to interventions (however, size group crash rate changes that are statistically significant but reflect an increase from the pre- to the post-intervention period are not considered by the model, as noted in Section 2.1). Complete details on the analysis can be found in Appendix B.

[24] See Cochran, William G., *Sampling Techniques*, third edition (1977) for further information.

2.7.5 Calculation of Treatment Group Crashes Prevented

Figure 5 shows how the crash rate reduction due to interventions is converted to a measure of crashes prevented, which also depends on the treatment group's pre-intervention crashes and pre- and post-intervention PU counts. The crash rate reduction is multiplied by the pre-intervention crash rate and post-intervention number of PUs; this yields the estimate of crashes prevented. This reduction is calculated separately for each carrier size group and added across the four size groups, yielding an initial estimate of total FY crashes prevented for the modeled year among treatment group carriers.

$$\text{Crashes Prevented} = \left[\frac{\text{Crash Rate Treatment}_{PRE} - \text{Crash Rate Treatment}_{POST}}{\text{Crash Rate Treatment}_{PRE}} - \frac{\text{Crash Rate Comparison}_{PRE} - \text{Crash Rate Comparison}_{POST}}{\text{Crash Rate Comparison}_{PRE}} \right] \text{X} \; \frac{\text{Crashes Treatment}_{PRE}}{\text{Treatment PU}_{PRE}} \; \text{X} \; \text{Treatment PU}_{POST}$$

Figure 5. Formula. Calculation of crashes prevented.

2.7.6 Extrapolation to Entire Carrier Population

The next step for estimating the intervention safety benefits across the motor carrier population is to estimate the crashes prevented by carriers that received interventions, but were initially excluded from the treatment group due to missing or outlier data. Such carriers, on average, can be assumed to exhibit a response to interventions similar to that of the observed treatment group. Therefore, the results from the observed treatment group crash rate reductions are extrapolated to account for the additional safety benefits resulting from performing interventions on these additional carriers. Without this step, the safety benefits from performing interventions on these carriers would not be included in the model estimates. The sum of the prevented crashes for the treatment group carriers and the prevented crashes for carriers filtered out of the treatment group represents the total estimated crashes prevented as a result of carrier interventions in a given FY.

Although it may be argued that imputing benefits for carriers excluded by the data filters may be unreliable, analysis of large numbers of carriers provides a basis for estimating average carrier behavior. Failing to make a best estimate of the impact of interventions on these carriers would result in an incomplete accounting of the benefits derived from the total set of FMCSA carrier interventions.

The following steps are followed when performing the extrapolation at the carrier size-group level, where *i* is used to denote size groups 1, 2, 3, and 4, respectively:

For carriers that can be assigned to a size group, based on usable power unit data: divide the number of such carriers receiving interventions in the size group by the number of treatment group carriers in the size group, to generate the first adjustment factor (see Figure 6):

$$\text{Adjustment Factor\#1}_{SizeGroup(i)} = \frac{\text{TOTAL CARRIERS with INTERVENTIONS}_{SizeGroup(i)}}{\text{TOTAL TREATMENT GROUP CARRIERS}_{SizeGroup(i)}}$$

Figure 6. Formula. Step 1 in performing extrapolation at the carrier size-group level.

To account for excluded carriers without usable PU data: divide the number of carriers receiving interventions with zero or unknown pre-intervention PU value by the total number of carriers receiving interventions (summed across all size groups), to generate the second adjustment factor (see Figure 7):[25]

$$\text{Adjustment Factor\#2} = \frac{\text{CARRIERS(ZeroUnknown)}}{\text{TotalCARRIERS}}$$

Figure 7. Formula. Step 2 in performing extrapolation at the carrier size-group level.

Multiply the first adjusted factor derived above by one plus the second adjustment factor, to generate the total adjustment factor (see Figure 8):

$$\text{Total Adjustment Factor}_{(SizeGroup(i))} = \text{Adjustment Factor\#1}_{(SizeGroup(i))} \times \left[1 + \text{Adjustment Factor\#2}\right]$$

Figure 8. Formula. Step 3 in performing extrapolation at the carrier size-group level.

Multiply the treatment group's estimate of crashes prevented by the final adjustment factor derived above. For example, total crashes prevented are calculated as follows (see Figure 9):

$$\text{Total Crashes Prevented}_{(SizeGroup(i))} = \text{Treatment Group Crashes Prevented}_{SizeGroup(i)} \times \text{Total Adjustment Factor}_{SizeGroup(i)}$$

Figure 9. Formula. Step 4 in performing extrapolation at the carrier size-group level.

The steps above yield each size group's total estimated safety benefits. Adding across the carrier size groups yields each FY's total extrapolated estimated safety benefits.

2.7.7 Calculation of Direct Safety Benefits

Injuries prevented and lives saved as a result of the crashes prevented can be calculated using historical crash severity data. This model uses 2-year averages calculated from historical MCMIS crash data, to estimate the probability that a crash involves an injury or fatality, and to estimate the number of injuries and fatalities in such crashes. For each model year, the 2-year average is calculated using MCMIS crash data for the current and prior FYs. Figure 10, Figure 11, Figure 12, and Figure 13 present the formulas for these calculations.

[25] Carriers with zero or unknown pre-intervention PU counts represent a small percentage of carriers during the three model years: 0.83 percent in FY 2010; 0.51 percent in FY 2011; 0.63 percent in FY 2012.

$$
\begin{gathered}
\text{Number of } \textbf{fatal crashes} \text{ prevented } = \\
\text{probability of a fatal crash given a crash occurred } \boldsymbol{x} \text{ number of crashes prevented}
\end{gathered}
$$

Figure 10. Formula. Number of fatal crashes prevented.

Note: All averages are for the 2-year period encompassing the modeled fiscal year and the prior year.

$$
\begin{gathered}
\text{Number of } \textbf{injury crashes} \text{ prevented } = \\
\text{probability of an injury crash given a crash occurred } \boldsymbol{x} \text{ number of crashes prevented}
\end{gathered}
$$

Figure 11. Formula. Number of injury crashes prevented.

Note: All averages are for the 2-year period encompassing the modeled fiscal year and the prior year.

$$
\begin{gathered}
\textbf{Lives saved} = \\
\text{number of fatal crashes prevented } \boldsymbol{x} \text{ average number of fatalities per fatal crash}
\end{gathered}
$$

Figure 12. Formula. Number of lives saved.

Note: All averages are for the 2-year period encompassing the modeled fiscal year and the prior year.

$$
\begin{gathered}
\textbf{Injuries prevented} = \\
(\text{average number of injuries per fatal crash } \boldsymbol{x} \text{ number of fatal crashes prevented}) \\
+ \ (\text{average number of injuries per injury crash } \boldsymbol{x} \text{ number of injury crashes prevented})
\end{gathered}
$$

Figure 13. Formula. Number of injuries prevented.

Note: All averages are for the 2-year period encompassing the modeled fiscal year and the prior year.

This calculation of safety benefits is based on an implicit assumption that the distribution of crashes by severity is the same for all carrier types and all treatment categories (i.e., types of intervention received). In the absence of contrary findings, this assumption is reasonable at the national carrier fleet level.

As with crashes prevented, the estimated numbers of injuries prevented and lives saved are calculated by size group for both the treatment group and the carriers filtered out of the treatment group. The sum of each provides the overall estimate of safety benefits due to interventions during the modeled year.

[This page intentionally left blank.]

3. RESULTS OF IMPLEMENTING THE MODEL

3.1 RESULTS CONSIDERING ALL INTERVENTIONS

The model was implemented for carriers receiving the specified intervention types in FY 2010, FY 2011, and FY 2012. Table 4 presents the data for the three FYs. The first three columns show the number of interventions conducted by FMCSA and its State partners and are considered as input into the model, by type, for each of the three FYs. The next three columns report the number of carriers receiving these intervention types as their first intervention in each FY. As explained in the previous section, the model uses the number of carriers with one or more interventions in a given year, and counts the total number of such carriers by only considering its first intervention during that year (that is, carriers with multiple interventions during the fiscal year are not counted multiple times). Thus, the totals in the last three columns represent the total number of carriers considered by the model for each modeled year.

Table 4. Carrier interventions by type and number of carriers by first intervention.

Intervention Type	Number of Interventions FY 2010	Number of Interventions FY 2011	Number of Interventions FY 2012	Number of Carriers Receiving Specified Intervention as a First Intervention FY 2010	Number of Carriers Receiving Specified Intervention as a First Intervention FY 2011	Number of Carriers Receiving Specified Intervention as a First Intervention FY 2012
CSA Warning Letter	5,790	39,004	23,835	5,765	38,918	23,806
Offsite Investigation	687	639	618	620	623	608
Onsite Focused Investigation	1,199	6,246	10,470	1,090	5,427	9,809
Onsite Comprehensive Investigation*	15,393	9,663	7,039	13,952	8,995	6,665
PRISM Warning Letter	7,415	1,764	-	7,390	1,754	-
Non-ratable Review	881	883	1,313	774	754	1,180
Total	**31,365**	**58,199**	**43,275**	**29,591**	**56,471**	**42,068**

*Often called CRs prior to 2012.

Total interventions declined by approximately 25 percent in FY 2012, primarily reflecting a decline in warning letters issued that year. Based on the set of carriers receiving interventions, the treatment group for each year was determined by applying the data quality filters discussed in Section 2.6. Table 5 displays the number of carriers failing each data quality filter, and the resulting number of treatment group carriers for the three model years.

Table 5. Carriers excluded from treatment group by data quality filters and resulting treatment group totals.

Filter Criteria	Carriers Excluded by Filter FY 2010	Carriers Excluded by Filter FY 2011	Carriers Excluded by Filter FY 2012
Inactive during the pre- or post- periods	3,351	3,482	2,839
Zero PUs during the pre- or post- periods	3,172	3,079	2,491
New entrant during the pre- or post- periods	6,012	9,043	9,590
Fails driver-to-PU ratios	140	198	190
Fails change in pre-PU to post-PU or pre-driver to post-driver ratios	532	822	709
Carriers with 500 PUs or more and zero crashes	6	14	6
Fails crash rate thresholds	16	12	16
Having an out-of-service order during the pre-period or post-period	35	45	80
Totals	**FY 2010**	**FY 2011**	**FY 2012**
Total excluded carriers*	8,513	12,161	11,868
Total carriers receiving interventions	29,591	56,471	42,068
Percent excluded	28.8%	21.5%	28.2%
Total carriers in treatment group	21,078	44,310	30,200

* A carrier may be excluded by multiple criteria; therefore, the total excluded carriers do not equal the sum of the carriers excluded by each of the individual filtering criteria listed in the table.

The first three filters in Table 5 account for the majority of the fluctuation in the percentage of total carriers excluded across the three years (from 28.8 percent down to 21.5 percent, and rising to 28.2 percent). The remaining filters impact a much smaller number of carriers, and the combined proportion of total carriers screened out by these filters during each FY is relatively constant.

Table 6 presents the number of treatment and comparison group carriers for each FY, by size group. The number of treatment group carriers in all four size groups increased from FY 2010 to FY 2011, and then decreased in FY 2012, following the general trend observed above for the treatment group as a whole.

Table 6. Treatment and comparison group carriers by size group, FY 2010-12.

Carrier Size Group	Treatment Group Carriers FY 2010	Treatment Group Carriers FY 2011	Treatment Group Carriers FY 2012	Comparison Group Carriers FY 2010	Comparison Group Carriers FY 2011	Comparison Group Carriers FY 2012
1 (1–5 PUs)	10,706	25,179	16,650	704,115	756,996	809,135
2 (6–20 PUs)	6,897	12,485	8,897	65,287	66,247	67,975
3 (21–100 PUs)	2,912	5,291	3,769	12,503	12,169	12,299
4 (≥100 PUs)	563	1,355	884	1,685	1,489	1,446
Total	**21,078**	**44,310**	**30,200**	**783,590**	**836,901**	**890,855**

3.1.1 Crash Rate Reduction

Table 7 presents the initial treatment and comparison group crash rate reductions calculated by the model, by year and carrier size group.

Table 7. Treatment and comparison group crash rate reductions by size group, FY 2010–12.

Carrier Size Group	Treatment Group Crash Rate Reduction FY 2010	Treatment Group Crash Rate Reduction FY 2011	Treatment Group Crash Rate Reduction FY 2012	Comparison Group Crash Rate Reduction FY 2010	Comparison Group Crash Rate Reduction FY 2011	Comparison Group Crash Rate Reduction FY 2012
1 (1–5 PUs)	34.6%	38.1%	35.6%	2.0%	3.8%	-1.9%
2 (6–20 PUs)	19.3%	28.3%	24.9%	3.2%	-2.9%	-8.5%
3 (21–100 PUs)	3.7%	17.1%	14.7%	4.8%	0.4%	-2.3%
4 (100+ PUs)	-2.8%	10.9%	2.6%	2.4%	6.9%	-2.9%

Note: Negative crash rate reductions indicate increases in crash rates.

Note that the crash rate reductions for the comparison group in Table 7 are negative for all size groups in FY 2012. These crash rate increases will amplify the crash rate reductions of the treatment group in the subsequent step of the model, when net crash rate reductions due to interventions are calculated for each size group (see Table 8). The negative comparison group crash rate reductions are consistent with observed crash rate increases nationally in FY 2013, which represents the bulk of the comparison group post-intervention period; that year, total CMV crashes rose about 6.7 percent when compared with FY 2012.[26]

Table 8. Net crash rate reductions by size group, FY 2010–12.

Adjusted Crash Rate Reduction by Carrier Size Group	FY 2010	FY 2011	FY 2012
1 (1–5 PUs)	32.7%	34.3%	37.5%
2 (6–20 PUs)	16.0%	31.2%	33.4%
3 (21–100 PUs)	-1.2%*	16.7%	17.0%
4 (100+ PUs)	-5.2%*	4.0%*	5.5%*

Note: Negative crash rate reductions indicate increases in crash rates.

*Non-statistically significant adjusted reduction.

As noted in Section 2.7, the adjusted crash rates represent the pre- to post-intervention change in treatment group crash rates adjusted for changes in the corresponding comparison group crash rates. This means that, for example, size group 1 carriers with interventions in FY 2012 experienced a 37.5 percent crash rate reduction as a group, after subtracting out the crash rate change for comparison group carriers in the same size group in the same modeled year. The table suggests that smaller carriers generally exhibit greater net crash rate reductions from

[26] MCMIS, as reported on FMCSA Analysis & Information (A&I) website as of July 14, 2016: 138,099 crashes in FY 2013, and 129,427 crashes in FY 2012 (available at: http://ai.fmcsa.dot.gov/CrashStatistics/rptSummary.aspx).

interventions than their larger counterparts. This is consistent with results obtained from the previous effectiveness model, CREM, used to calculate safety benefits for years 2002–09, as well as from the FY 2009 implementation of the CIEM.

3.1.2 Safety Benefits

Table 9 presents safety benefits associated with FMCSA carrier interventions for FY 2010, FY 2011, and FY 2012. The left side of the table presents estimated crashes prevented, injuries prevented, and lives saved among treatment group carriers, for carriers that passed the model's data filters. The right side of the table extrapolates these benefits to all carriers receiving interventions, including those that were screened out of the initial model calculations by the data filters. These benefits declined slightly in FY 2012 compared to FY 2011, reflecting the decline in total interventions, although this decline was in large part offset by larger adjusted crash rate reductions experienced in FY 2012 by carrier size groups 1, 2, and 3.

Table 9. Estimated safety benefits, treatment group and all carriers receiving interventions, FY 2010-12.

Fiscal Year	Treatment Group Carriers (Number)	Treatment Group Carriers Crashes Prevented	Treatment Group Carriers Injuries Prevented	Treatment Group Carriers Lives Saved	All Carriers Receiving Interventions (Number)	All Carriers Receiving Interventions Crashes Prevented	All Carriers Receiving Interventions Injuries Prevented	All Carriers Receiving Interventions Lives Saved
2010	21,078	1,281	800	42	29,591	1,830	1,142	59
2011	44,310	5,394	3,313	176	56,471	6,567	4,033	215
2012	30,200	4,021	2,462	132	42,068	5,283	3,235	173

The safety benefits reported in Table 9 reflect only those associated with adjusted crash rates that were both positive and statistically significant, as reported in Table 8. Carrier size groups not yielding statistically significant positive crash rate reductions during the post-intervention period, after adjusting for crash rate changes in the comparison group, are assumed to experience no safety benefits (neither positive or negative).

3.2 LIMITATIONS OF THE MODEL

The model and data used have limitations that should be considered when interpreting the results. These limitations include the following:

3.2.1 Comparison Group versus Control Group

The model uses a comparison group to adjust for external factors that may affect carrier crash rates, such as economic trends, weather, implementation of new safety regulations and policies by FMCSA and other agencies, and industry technological advances. The comparison group is used as a proxy for a statistical control group; it comprises carriers that did not receive one of the defined interventions during the modeled year or the year prior. However, unlike a precisely matched statistical control group, the comparison group has different characteristics from the treatment group; most important, it is generally a safer group of carriers[27] (the more benign safety profile of the comparison group carriers is the main reason these carriers received no intervention during the period in consideration).

While a true control group with a safety profile matching that of the treatment group would be ideal, it is not feasible to construct one. To begin with, with few exceptions, carriers having a safety profile similar to that of the treatment group (i.e., with similar violation and crash rates), would normally receive interventions.[28] Second, a carrier's safety profile reflects scores across a number of Behavior Analysis and Safety Improvement Categories (BASICs).[29] With the many potential combinations of BASIC rankings, it is highly unlikely that a sufficiently large set of carriers could be found to match the treatment group across these BASICs.

3.2.2 Overall Impact Rather than Impact by Intervention Type

The model does not report the impacts of each individual intervention type; rather, it estimates the combined impact of all interventions performed during the modeled year. This is due to some carriers receiving more than one intervention type, both within each FY and across FYs. This makes it infeasible to construct pre- and post-intervention periods free of confounding effects from the various intervention types. Furthermore, separating out carriers by the intervention types they received would result in small sample sizes, reducing the accuracy and statistical significance of the estimates produced. It is possible, however, to implement the model for groups of carriers receiving different first interventions within a FY. While not fully isolating the benefits associated with specific intervention types, such analysis can provide insight into their likely impact. Results from two such analyses are reported in the CIEM Summary Report for FY 2010–12, for carriers whose first intervention in the modeled year is a warning letter, and for carriers whose first intervention was anything but a warning letter.[30]

3.2.3 Data Quality and Timeliness

The validity of the model results depends on the quality of the crash and carrier information reported to MCMIS. The model results can be influenced by changes in the accuracy and timeliness of crash reporting. For example, improvements in the timeliness of crash reporting could conceivably increase the post-intervention average crash rate and produce a smaller crash rate reduction than actually occurred, while the opposite effect would occur if crash reporting timeliness worsened. Although the use of the comparison group in the model helps to mitigate

---

[27] The FY 2012 pre-intervention crash rates for the treatment and comparison group are shown in the table below.

Size Group	Treatment Group Crash Rate(Crashes/PU)	Comparison Group Crash Rate(Crashes/PU)
1	0.078	0.006
2	0.051	0.006
3	0.039	0.008
4	0.028	0.011

[28] In any given year, there may be a small number of such carriers identified. However, a small number of carriers do not provide a robust control group; moreover, such data may be suspect.

[29] For more information on the BASICs, see: https://csa.fmcsa.dot.gov/about/basics.aspx.

[30] CIEM annual reports are available at: http://ai.fmcsa.dot.gov/pe/home.aspx.

these effects, issues relating to the quality and timeliness of crash reporting may still have an impact upon the model. Similarly, improvements in the accuracy and consistency of PU data would impact estimated crash rate reductions.

3.2.4 Seasonality

The longer (18-month) pre- and post-intervention periods for the comparison group result in uneven seasonality: two fall/winter periods are in the "before" period, while two spring/summer periods are in the "after" period. It is at least conceivable that such differences confound the relative change in crash rates for the comparison group, although there is currently no direct evidence to suggest this is the case.

3.2.5 Excluded Intervention Types

During the initial implementation and national rollout of CSA, some intervention types were recorded inconsistently. As noted in Section 2.3, the following intervention types were excluded from the model due to this inconsistency: follow-up verifications, DNOVs, DNOCs, and CSPs. These intervention types do not make a carrier eligible for the model's treatment group. During FYs 2010–12, these interventions represent a small fraction of the total set of interventions; therefore, the overall impact of their exclusion is likely small. However, improvements in data reporting that would allow for the inclusion of these interventions in the model would allow it to account for any additional potential benefits stemming from these activities.

3.2.6 Benefit Duration

The model does not consider multi-year benefits. In other words, benefits from interventions occurring during the FY are assumed to last no more than 1 year from the time of the first intervention. Thus, if a carrier dramatically improves its safety performance after an intervention and never regresses, subsequent crashes prevented in future FYs are not accounted for in the model. Without further research, the magnitude of this potential limitation is unclear.

3.2.7 Variations of the Model

As noted in Section 3.2.2, the CIEM was not designed to measure benefits resulting from specific intervention types. However, to obtain additional insight into the interventions' effectiveness, the model can select treatment group carriers based on specific first intervention types within the modeled FY. For example, the CIEM has recently been implemented separately for carriers whose first intervention was a warning letter. While some intervention types are not performed frequently enough to yield sample sizes sufficient for statistically significant results, the large number of warning letters issued in recent years makes this variation of the model feasible. Given that the vast number of carriers receiving warning letters did not receive additional interventions in the subsequent 12 months, this additional analysis sheds light on the extent to which safety benefits observed in the modeled year are associated with warning letters as compared with all other interventions. The results of these analyses can be found in the CIEM Summary Report for FYs 2010–12.[31]

---

[31] Available at: http://ntl.bts.gov/lib/60000/60500/60503/16-017-CIEM_Report-FY12-FINAL-508C.pdf.

4. SUMMARY AND NEXT STEPS

CIEM Version 1.1 provides FMCSA with a means for measuring the safety benefits of motor carrier interventions, considering most interventions currently performed by the Agency.[32] As such, it can be used for annual measurement of safety benefits during the phased CSA implementation and beyond.

The model builds on the approach of the CREM, previously used to measure the effectiveness of CRs. However, in contrast to the CREM, the CIEM now incorporates the various intervention types that comprise FMCSA's overall enforcement program, which has expanded with CSA. The model also introduces a component addressing statistical significance and an approach for extrapolating directly measured safety benefits to carriers with missing or outlier crash or PU data.

Overall, the set of FMCSA intervention types specified in the model are shown to have reduced motor carrier crash rates in FY 2012, as well as in prior years. Consistent with CREM results in prior years, crash rate reductions are generally more pronounced for the smaller carrier size groups. Total carrier interventions declined substantially in FY 2012, driven by a reduction in the number of warning letters issued. In contrast, overall percent reductions in crash rates for carriers receiving interventions were higher in FY 2012 than in the previous year. The result of these two opposing trends is a moderate decline in total safety benefits estimated by the model for FY 2012.

Future CIEM implementation will enable FMCSA to continue to measure the impacts of carrier interventions. In addition, potential exploratory analysis may yield model refinements addressing some of the current model limitations. Such analysis and potential enhancements may:

- Help determine whether a distinction can be made between interventions' short- and long-term (beyond 12 months) impacts.
- Address the comparison groups' possible seasonality bias.
- Yield insight into the effectiveness of some individual or combinations of interventions.
- Incorporate some of the intervention types currently not included in the model.

[32] See Section 2.3 for a list of remaining intervention types currently not used in the CIEM.

[This page intentionally left blank.]

APPENDIX A – DATA FILTERS

The CIEM applies both carrier data sufficiency criteria (described in Section 2.2) and outlier data filters (summarized in Section 2.6) to both treatment and comparison group carriers. Together, these tools are intended to: 1) prevent the calculation of erroneous annual crash rates that may result from incomplete carrier data; 2) identify complete but inaccurate data that may bias the model results; and 3) minimize the influence of confounding factors. This is described in more detail below.

1. **Prevent the calculation of erroneous annual crash rates that may result from incomplete carrier data.** For example, a carrier with no PU data in a given MCMIS snapshot may have not operated during the corresponding time period, meaning its safety exposure and potential to experience a crash were zero.
2. **Identify complete but inaccurate data that may bias the model results.** PU data may be out of date or simply incorrect, in some cases by an entire order of magnitude (i.e., an extra digit or a missing digit). Crashes can also be assigned to a carrier incorrectly; this can happen when a motor vehicle operates under a lease arrangement, or it can be due to human error at the time of crash reporting. Both types of data errors would directly impact the model results.
3. **Minimize the influence of confounding factors.** Most carriers remain largely stable over the course of the model's 2- to 3-year timeframe. However, some carriers experience large shifts in size and/or ownership, some of which result from carrier mergers or acquisitions. The operations overhaul experienced by such carriers may influence their safety record, including their crash rates. Attributing their change in crash rates to FMCSA interventions may be misleading.

Following the initial selection of carriers with sufficient data (see Section 2.2), a set of three data filters is employed to remove outlier carriers from the model's treatment and comparison groups. This set of filters represents a substantial enhancement to the outlier filtering employed by the CREM; that model included only one automated filter—based on the change in carrier size from the pre-intervention time period to the post-intervention time period—and used relatively lenient criteria for its outlier limits. The three filters used in the CIEM are based on the following measures:

- Driver-to-PU ratio.
- Change in carrier size.
- Carrier crash rates.

A carrier that fails any one of these filters is deemed an outlier. Together, the three filters provide a basis for removing outlier carriers, yielding a core group of carriers with reliable data for the model's treatment group. This outlier removal process is illustrated in Figure 14 and Figure 15. In Figure 14, the pink circle (left) represents the set of carriers that meet the 'Driver-to-PU' ratio filter, while the light blue circle (right) represents carriers that meet the 'Change in Carrier Size' filter. The intersection of the two circles represents carriers that pass the criteria for both filters, which makes them eligible to be considered by the third and final filter.

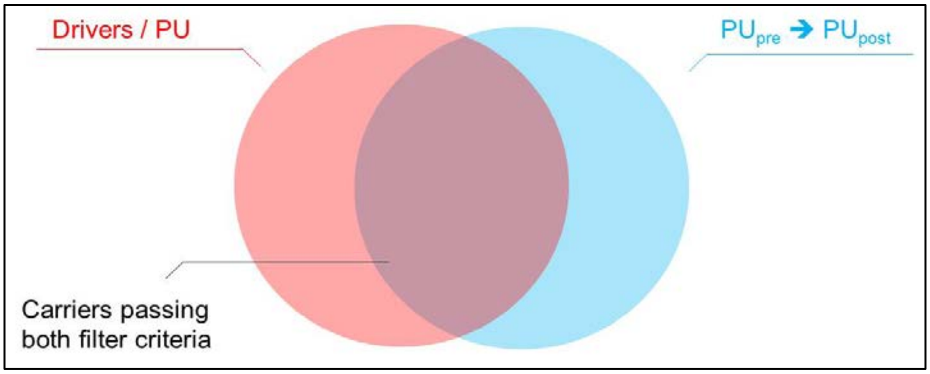

Figure 14. Venn diagram. Illustration showing intersection of carriers passing the CIEM's first two data filter criteria.

Figure 15 illustrates the third filter, where carriers with a crash rate too far below or above their size group's mean crash rate are identified as outliers. The distribution of the carriers considered by this third filter, illustrated in the figure as a curve with most carriers concentrated around the mean crash rate, is dependent on the outcome of employing the first two filters; that is, size groups' average crash rates, likewise, are only calculated after outliers identified by the first two filters are removed. Hence, carriers already deemed outliers by the first two filters based on absolute measures (i.e., without the need to be compared to their peers), do not influence the cutoff points defined by the final filter.

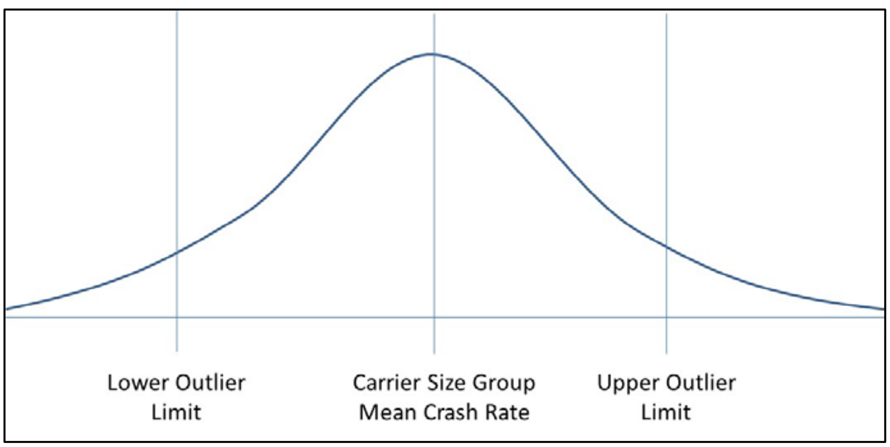

Figure 15. Graph. CIEM crash rate filter criteria for carriers previously screened for driver-to-carrier ratios and changes in carrier size.

Note: The curve is drawn to represent a generic distribution; no implication of normal or other symmetric distribution is implied.

DRIVER-TO-POWER UNIT RATIO

This filter is largely intended to help identify erroneous carrier-level PU data. It relies on each carrier's observed ratio between drivers and PUs, in both the pre- and post-intervention time periods. While driver counts are not used directly in any of the model's calculations of crash rate changes and safety benefits, the driver count values provide one measure of the reasonableness of the carrier's PU information in MCMIS. Aside from leasing operations and driveaway/

towaway carriers (for which this filter makes an exception), most carriers do not have a high or low driver-to-PU ratio. After testing numerous boundary values (to see how many carriers would fail the filter for each boundary value considered), a ratio of 7.5:1 was identified as a reasonable maximum. Since this same ratio of 7.5:1 was also determined to be a reasonable maximum for the PU-to-driver ratio, the reciprocal of 7.5, or .1334, was used for the data filter's driver-to-PU ratio minimum value. Carriers with a ratio outside this range are removed from the model's treatment group. Figure 16 depicts these outlier cutoff boundaries that meet the filter's criteria for inclusion in the CIEM treatment group.

$$\boxed{0.1334 \leq \frac{\text{Drivers}}{\text{Power Units}} \leq 7.5}$$

Figure 16. Formula. Driver-to-PU ratio filter.

Other boundaries considered for this filter include ratios of 2:1, 5:1, and 10:1, along with the respective inverse of each. For each boundary considered, additional data were examined for the excluded carriers, including data from additional MCMIS snapshots both within and outside of the pre- and post-intervention periods. These data were used to evaluate the extent to which the various filter boundaries were identifying outlier carriers with invalid data. While this was a largely manual exercise, much of it relied on examining minimum, maximum, and average values across groups of carriers. In FY 2011, 198 carriers failed the criteria for this filter.

CHANGE IN CARRIER SIZE

Most carriers remain largely stable in terms of overall operation size over the course of a 2- to 3-year timeframe. Small changes in carrier size are nonetheless common, and the impact of such changes on estimates of average crash rates is not a concern. Large changes, however, are a concern for two reasons:

- They are a red flag about the validity of a carrier's data.
- As noted earlier, if the data are in fact correct, they may be an indication of confounding factors, such as acquisitions, mergers, or other changes in ownership.

Because the model uses these PU values directly in the estimation of crash rate changes from the pre- to the post-intervention periods, this filter measures the ratio between each carrier's pre- and post-intervention PU value.

In examining the extent and nature of changes in carrier size over time across the U.S. carrier population, it was observed that small carriers often undergo much larger proportional growth or reduction than larger carriers. This observation is consistent with expectations, since even small absolute changes for small carriers can result in large proportional changes. For example, a carrier with two PUs that purchases two additional trucks experiences 100 percent growth. In contrast, such an acquisition for a carrier with 100 PUs represents just 2 percent growth. Therefore, the boundaries for this filter were set separately for the model's four carrier size groups.

Various sets of outlier limits were considered for this filter across the model's four carrier size groups. The ratio limits tested for pre- to post-intervention period changes in carrier PU counts include 1:1.5, 1:1.75, 1:2, 1:3, 1:4, 1:5, and 1:10, along with the respective inverse of each.[33] In testing these outlier limits, it was observed that a substantial number of carriers identified as outliers by the filter appeared to exhibit consistent changes in driver counts. Therefore, a two-part approach was developed whereby the pre- to post-intervention PU ratios would be used in conjunction with the corresponding change in driver counts. Table 10 lists the ratio limits used by this filter.

Table 10. Filter for change in carrier size—criteria for identifying outliers in carrier pre- to post-intervention PU ratio.

Size Group	Primary Filter Lower Limit	Primary Filter Upper Limit	Exceptions to the Primary Filter Conditions: If Statement	Exceptions to the Primary Filter Conditions: Then Statement
1 & 2	0.334	3.0	If corresponding driver change ratio is: up to 0.667	Then change primary filter limits, as follows: Lower limit: 0.2
1 & 2	0.334	3.0	If corresponding driver change ratio is: 1.5 to 10.0	Then change primary filter limits, as follows: Upper limit: 5.0
3	0.571	1.75	If corresponding driver change ratio is: up to 1.0	Then change primary filter limits, as follows: Lower limit: 0.4
3	0.571	1.75	If corresponding driver change ratio is: 1.0 to 5.0	Then change primary filter limits, as follows: Upper limit: 2.5
4	0.571	1.75	No exception	N/A

The filter limits are more lenient for the smaller carrier size groups than they are for the larger size groups because, as noted earlier, smaller carriers more often experience large proportional changes in PUs as part of normal operations.

It should be noted that the driver ratios corresponding to the conditional component of the filter do not equal the limits used for the PU ratios. Rather, the limits set for the former are more lenient than those on the latter, thus allowing for somewhat larger changes in driver counts from the pre- to the post-intervention periods than in the case of changes in PU counts. This approach is consistent with what is observed in the industry.

The following two figures illustrate how this filter works, using the filter limits for carrier size groups 1 and 2 to illustrate. The filter's primary criteria—that PU change ratios be between one-third and 3.0—is represented by the shaded area in Figure 17 (for simplicity, the figure depicts the filter values for carriers with up to 10 PUs). Carriers whose ratio falls outside of this range

[33] In contrast, the CREM employed a much more lenient 1:100 ratio as the limit for a similar filter across all carrier size groups. Follow-up manual evaluation, which was labor-intensive, was employed to refine the CREM's treatment group composition further.

fall outside of the shaded area and are considered potential outliers, pending the filter's second, conditional component.

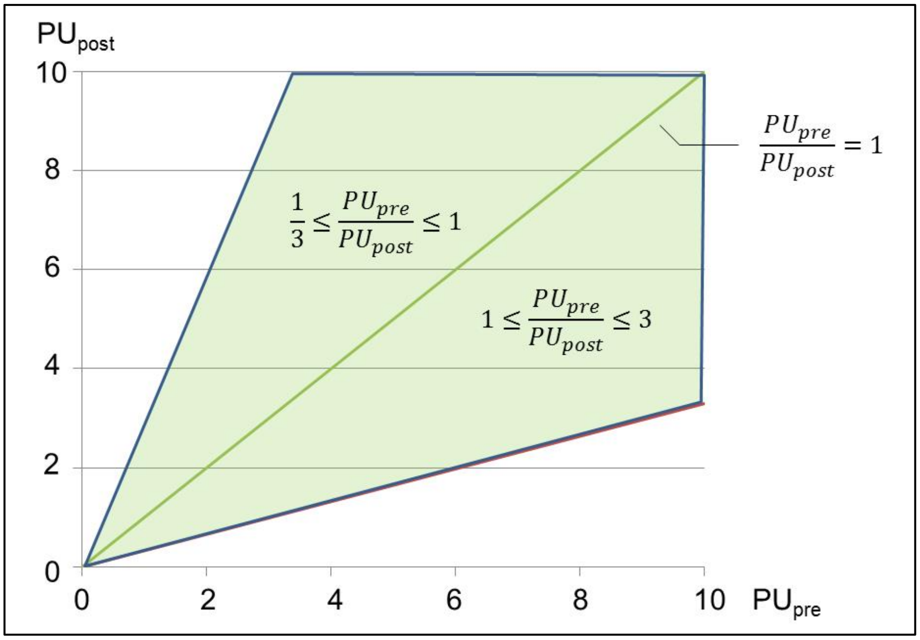

Figure 17. Graph. Change-in-carrier-size filter: primary criteria (carrier size groups 1 and 2).

Figure 18 depicts the filter's secondary, conditional criteria. The shaded areas represent carriers outside the primary criteria but within the expanded conditional limits. Carriers with PU change ratios within the shaded areas are not ruled outliers if their driver change ratios meet the conditions for the second criteria. Carriers represented by the shaded areas whose driver change ratios do not meet the specified conditions are removed as outliers. In FY 2011, 822 carriers failed the criteria for this filter.

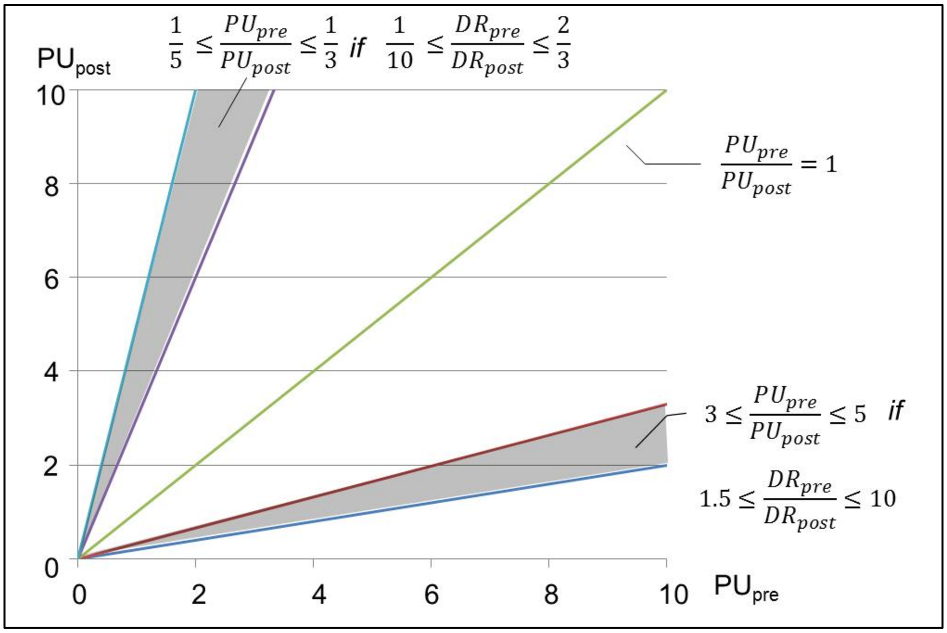

Figure 18. Graph. Change-in-carrier-size filter: expanded conditional boundary limits for carrier size groups 1 and 2.

CARRIER CRASH RATES

The third filter is used to identify outlier carrier crash rates for carriers that have passed all other CIEM data requirements and the two outlier filters described above. This filter relies on the mean and variance of the carrier-level crash rates in each size group. Within each size group, carriers with a crash rate that falls within five standard deviations of the group mean crash rate are generally not considered outliers. However, a number of exceptions are made for carriers that fail to meet these initial criteria. Without these exceptions, numerous carriers with crash profiles that fall within the norm would be removed. Such instances may occur, especially among smaller carriers, where even one or two crashes may generate relatively high crash rates.

Table 11 presents the criteria used by this filter, including the exceptions to the filter's initial rule. Carriers in size group 1 may have up to five crashes without becoming outliers, regardless of their crash rate (crashes per PU). Carriers in size group 2 may have up to six crashes. These crash counts were derived to reflect reasonable conditions for each size group and are consistent with similar types of outlier tests used by FMCSA in other research. In addition, these values provide for a smooth transition from size group 1 to size group 2 in terms of crash rate: at the top end of size group 1 (carriers with five PUs), a carrier can have a crash rate up to 1.0 (i.e., 5/5); at the bottom end of size group 2 (carriers with six PUs), a carrier can also have a crash rate up to 1.0 (i.e., 6/6). This smooth transition avoids inconsistencies in identifying outlier carriers across the size groups.

The final exception made in this filter is for carriers with 500 or more PUs. Such carriers are also ruled outliers if they are not involved in at least one crash. This criterion is based on reasonable expectations for carriers of this size and is consistent with empirical findings. For such carriers, zero recorded crashes suggest either an inaccurate PU count or crashes not properly attributed to the carrier.

Table 11. Crash rate filter criteria

Criteria	Size Group	Lower Limit	Upper Limit	Exceptions to the Primary Filter Conditions
Difference from mean crash rate, by size group	1	-5 STD	+5 STD	Up to 5 crashes
Difference from mean crash rate, by size group	2	-5 STD	+5 STD	Up to 6 crashes
Difference from mean crash rate, by size group	3	-5 STD	+5 STD	Up to 6 crashes
Difference from mean crash rate, by size group	4	-5 STD	+5 STD	Up to 6 crashes; carriers with 500 or more PUs must have at least 1 crash

Note: STD = standard deviation.

[This page intentionally left blank.]

APPENDIX B – TEST FOR STATISTICAL SIGNIFICANCE

Section 2.7.4 summarizes the statistical significance testing conducted on the crash rate reductions computed by the model. The following discussion provides further detail.

The test employed here centers on the estimation of the variance of the carrier-level crash rate reductions for the sample in each group (i.e., treatment and comparison groups for each carrier size group). As explained earlier in Section 2.7.3, the estimated net crash rate reduction for a carrier size group is calculated as follows:

$$\begin{array}{c}\textit{Crash Rate Reduction}\\ \textit{Due to Interventions}\end{array} = \boxed{\begin{array}{l}\textbf{Treatment Group}\\ \text{Crash Rate Reduction,}\\ \text{pre- to post-intervention}\\ \quad \dfrac{[CR_{PRE} - CR_{POST}]}{CR_{PRE}}\end{array}} - \boxed{\begin{array}{l}\textbf{Comparison Group}\\ \text{Crash Rate Reduction,}\\ \text{pre- to post-FY midpoint}\\ \quad \dfrac{[CR_{PRE} - CR_{POST}]}{CR_{PRE}}\end{array}}$$

Figure 19. Formula. Estimating net crash rate reduction for a carrier size group.

where,

CR_{PRE} = Pre-intervention crash rate

CR_{POST} = Post-intervention crash rate

The CIEM estimates treatment group and comparison group crash rates, based on crash and PU data from MCMIS. Although these calculations are based on a virtual census of all in-scope carriers in the treatment and comparison groups, respectively, during the modeled year (except for those that do not pass the CIEM data filters), the treatment group carriers can be thought of as a sample from a super population of carriers (changing over time) that receive interventions. As such, the precision level of the crash rate calculations may be estimated using traditional sampling theory. Estimates of sample means and totals from distributions are assumed to be normal in accordance with the Central Limit Theorem. The crash rate reductions for both the treatment and comparison groups, shown in Table 7, represent the ratio of two separate ratio estimates (i.e., one ratio estimate divided by another).[34] The calculation of the variance of the crash rates in the CIEM first uses a Taylor's series approximation to estimate the variance of the ratio of two random variables, and then employs additional statistical approximations to estimate the variance of two ratios. The variance of the net reduction in crash rates uses a formula

[34] (CR-pre-CR-post)/CR-pre can be expressed as 1- (CR-pre/CR-post). Since the variance of "1" is zero, the total variance stems from the variance of the ratio of CR-pre to CR-post, both of which are, themselves, ratio estimates.

combining the pre- and post-intervention crash rates with their respective variances and the covariance (shown in Figure 20).[35]

$$\textbf{Variance (Net Crash Rate Reduction)} = \textbf{Variance (Treatment Group Crash Rate Reduction)} + \textbf{Variance (Comparison Group Crash Rate Reduction)}$$

Figure 20. Formula. Variance of the net reduction in crash rates.

From the net crash rate reduction's variance, a standard error is computed, as shown in Figure 21:

$$\textbf{Standard Error (Net Crash Rate Reduction)} = \sqrt{\textbf{Variance (Net Crash Rate Reduction)}}$$

Figure 21. Formula. Standard error of the net crash rate reduction.

The 95-percent confidence interval around each estimate of net crash rate reduction is computed, as shown in Figure 22:

$$95\%\ CI = \textit{Net Crash Rate Reduction} \pm 1.96 \times \text{Standard Error of the Net Crash Rate Reduction},$$

where 1.96 is the Z statistic for a 95% confidence interval.

Figure 22. Formula. 95-percent confidence interval for each estimate of net crash rate reduction.

If a confidence interval includes the value zero, then the estimate is not considered statistically significant (because the value of the estimate cannot be ruled out as being zero with 95-percent confidence). If the value is in fact zero, then the net crash rate reduction is equal to zero, meaning there is no difference between the treatment and comparison groups' performance in terms of crash rate reduction over the model's timeframe. When this interval does not include zero, the estimate is considered statistically significant.

The confidence interval calculation is illustrated below, using the CIEM FY 2011 result for size group 3, for which a net crash rate reduction of -16.7 percent, or -0.167, was estimated (see Section 3.1). For this result, the standard error estimate was calculated to be 0.0235, yielding the following 95-percent confidence interval:

---

[35] See Cochran, William G., *Sampling Techniques*, third edition (1977), for further information on these formulas and their derivations.

- Lower Bound = 0.167 – (1.96*0.0235) = 0.121, or 12.1percent.
- Upper Bound = 0.167+ (1.96*0.0235) = 0.213, or 21.3 percent.

Because this interval does not include the value zero, the model estimate for this size group in FY 2011 is statistically significant at the 0.05 significance level. Alternatively, consider the FY 2011 results for size group 4, which exhibited a net crash rate reduction of 0.040 or 4.0 percent. For this result, the standard error was calculated to be 0.0239, yielding the following 95-percent confidence interval:

- Lower Bound = 0.040– (1.96*0.0239) = −0.007, or −0.7 percent.
- Upper Bound = 0.040 + (1.96*0.0239) = 0.087, or 8.7 percent.

Since this confidence interval for the net crash reduction estimate ranges from −10.7 to 8.7 percent (and hence includes zero), the estimated crash rate reduction is not statistically significant at the 0.05 significance level. Accordingly, the FY 2011 change in crash rate initially calculated for this size group is assumed to be zero by the model.