Syracuse University
SURFACE

Economics Faculty Scholarship | Maxwell School of Citizenship and Public Affairs

11-1-2004

Ranking and Selection of Motor Carrier Safety Performance by Commodity

William C. Horrace
Syracuse University, whorrace@maxwell.syr.edu

Thomas P. Keane
Federal Motor Carrier Safety Administration

Follow this and additional works at: http://surface.syr.edu/ecn

Part of the Economics Commons

Recommended Citation
Horrace, William C. and Keane, Thomas P., "Ranking and Selection of Motor Carrier Safety Performance by Commodity" (2004). *Economics Faculty Scholarship.* Paper 21.
http://surface.syr.edu/ecn/21

This Article is brought to you for free and open access by the Maxwell School of Citizenship and Public Affairs at SURFACE. It has been accepted for inclusion in Economics Faculty Scholarship by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

This is an author-produced, peer-reviewed version of this article. The published version of this document can be found online in the *Accident Analysis & Prevention* (doi: 10.1016/j.aap.2003.10.004) published by *Elsevier*.

Ranking and selection of motor carrier safety performance by commodity

William C. Horrace, Syracuse University
Thomas P. Keane, Federal Motor Carrier Safety Administration
Keywords: Transportation; Trucks; Accidents; Segment comparisons

Abstract

We use recent safety performance data to rank US motor carrier commodity segments (e.g., Tank segment or Produce segment) in terms of several driver-related, vehicle-related, and crash-related safety measures. Ranking and selection inference techniques are used to determine the best and worst performing commodity segments at the 95% confidence level. The results are mixed, however the Passenger segment is generally best, while the Produce, Intermodal, and Refrigerated segments tend to be worst.

1. Introduction

The primary mission of the Federal Motor Carrier Safety Administration (FMCSA) is to prevent commercial motor vehicle fatalities and injuries. In an effort to better understand the diverse nature of the commercial motor carrier industry and explore safety and operational differences among its major commodity segments, the FMCSA undertook the Motor Carrier Industry Profile Study (MCIPS). The study uses the Motor Carrier Information Management System (MCMIS) and the Motor Carrier Safety Status Measurement System (SafeStat) as its sources for safety and operations data on commercial motor carriers. The MCMIS and Safe-Stat are maintained by the FMCSA and are populated with carrier-level data from roadside inspections, FMCSA and State compliance reviews, crashes, and enforcement cases against motor carriers. The data are used to track carrier-level safety performance, based on several different measures of driver and vehicle performance and compliance.

Using the MCIPS data, this paper explores the following important questions: (1) which commodity segments of the commercial motor carrier industry perform best and worst in terms of safety, and (2) which safety performance measures are the most and least reliable for making this assessment? The policy implications are important: if there exist significant safety performance differentials across segments of the industry, then motor carrier safety regulations and enforcement could be tailored specifically to address the poorest performing segments. We develop safety performance estimates across commodity segments, rank them by segment, and then apply the theory of ranking and selection to determine (in a probabilistic sense) which of the segments are the most and least safe.[1] Generally, we find that while there are significant safety performance differences across some commodity segments, most of the perceived differences may be statistical noise. We also find that some safety measures are less reliable than other in terms of their statistical precision.

2. Data

The Motor Carrier Management Information System (MCMIS) was the primary source of data used in this study. The MCMIS is used by FMCSA to maintain a comprehensive safety record of property and passenger carriers subject to the Federal Motor Carrier Safety Regulations (FMCSR). The MCMIS Census File contains records from over 500,000 entries (e.g., motor carriers, hazardous materials shippers, and registrants), and contains information on each company's identity (name, address), operations classification (type of business), cargo classification (segment), and numbers of trucks and drivers within the company.

The MCMIS data set also contains an Inspection File, which contains the results of roadside inspections (submitted by States), enforcement actions (taken by Federal personnel against a motor carrier), and compliance reviews (conducted by FMCSA and State safety investigators). Enforcement actions may include civil penalties or out-of-service (OOS) orders placed against a carrier. (The majority of driver OOS violations pertain to non-compliance with the hours-of-service rules.) Compliance reviews are on-site reviews of a motor carrier's operations,

[1] Motor carrier safety estimation has a fairly sparse literature. Some recent research in this area can be found in McCartt et al. (2000), Moses and Savage (1994, 1996, 1997), and Scopatz (2001).

conducted by FMCSA and State personnel to determine the level of compliance with the FMCSR. The MCMIS data set also includes a Crash File, which contains data from State police crash reports electronically transmitted to FMCSA.[2] The census, inspection, and crash data are reviewed and updated as new information is collected by FMCSA on a motor carrier, whether through inspections, compliance reviews, enforcement action, or reportable crashes.

This study also uses data produced by the Motor Carrier Safety Status Measurement System (SafeStat), a powerful analytical tool developed by FMCSA and the Volpe National Transportation Systems Center to accurately identify and monitor high-risk motor carriers within the overall motor carrier population. SafeStat became operational in 1995 and uses MCMIS data as input to evaluating a motor carrier's relative safety fitness. (Our data goes up to and includes the year 2000.) SafeStat incorporates current on-road safety performance data with on-site compliance review data collected by FMCSA and State safety investigators to provide a comprehensive evaluation of a motor carrier's safety performance using four types of data: accident, driver, vehicle, and safety management. Safety event data are normalized to account for a carrier's size or amount of exposure using carrier-descriptive data such as number of power units or the number of roadside inspections. Data are also time-weighted, with the most recent events having greater weight than older ones. Crash data are also severity-weighted, based on the number and type of injuries sustained by victims involved in the crashes.

All the SafeStat data serve to measure a carrier's relative safety fitness and assess its risk of having future crashes. It should be noted that not all motor carriers contained in the MCMIS data set are analyzed by SafeStat. To be assessed by SafeStat, a motor carrier must have experienced at least three inspections within the past 30 months. As such, the number of carriers with SafeStat data (e.g., those analyzed in this study) is significantly smaller than the total population of carriers contained in MCMIS. More importantly, the results of this study are only relevant for the subpopulation of motor carriers that have had at least three inspections within the past 30 months; they are not valid for the entire population.

The final data used in this study are at the carrier level. Each observation consists of an individual carrier identification number, carrier name and contact information, several SafeStat carrier-level safety measures along different metrics, and the commodity segments in which the carrier operates. A single carrier can operate in several different (or multiple) segments, even though the safety measures in the data are for the entire fleet. The data also differentiate carriers into two distinct sectors: the For-Hire sector and the Private sector. Generally speaking, the For-Hire trucking sector is comprised of carriers whose primary business it is to haul property or passengers for another person or company in exchange for direct compensation. The Private trucking sector is generally comprised of carriers whose primary business is not trucking, but who choose to haul their own property or passengers.

We examine the recent safety performance of eleven For-Hire and ten Private commodity segments of the motor carrier industry. The ten Private segments examined are: Building Materials, Bulk Freight, Refrigerated (nonproduce), General Freight Truckload, Household Goods, Intermodal, Large Machinery, Passenger, Produce, and Tank. The 11 For-Hire segments examined include the same ten commodity segments, plus the "General Freight Less-Than-Truckload" segment. While most of the commodity segments are self-explanatory, two are less obvious. Bulk freight is an aggregation of the grain and feed, coal and coke, and dry bulk segments. "General Freight Less-Than-Truckload" carriers collect small shipments from local pickup points, move them between terminals as truckload shipments, and break them up at the destination terminal, where the shipments are then made to individual destination points.

Our analysis uses nine safety measures. Safety performance is evaluated with two driver-related measures: driver safety evaluation area (SEA), and driver out-of-service (OOS) rate; two vehicle-related measures: vehicle SEA, and vehicle OOS rate; three crash-related measures: accident SEA, fatal crash rate, and total crash rate; and two safety management performance measures: Safety Management Review Measure (SMRM), and Enforcement Severity Measure (ESM). A complete description of safety measures is in the data Appendix A.

3. Safety estimation

Let Y_i^k represent an observation of SafeStat safety measure k ($k = 1, \ldots, K$) for motor-carrier j ($j = 1, \ldots, J$) in no particular motor-carrier segment. K equals the number of safety measures under consideration (we consider nine such measures). J is the total number of carriers with at least three inspections within the last 30 months (J is approximately 94,000 in the For-Hire sector, and approximately 55,000 in the Private sector). Let $i = 1, \ldots, M$ index commodity segments (this study is limited to $M = 10$ segments in the Private sector and $M = 11$ segments in the For-

[2] The prevailing wisdom is that state police departments transmit only about two-thirds of reportable crashes. This underreporting may draw into question the validity of the crash results of this study.

Hire sector). Notice that Y_j^k is not indexed by commodity segment, so SafeStat compiles safety measure *k* at the carrier-level but not at the segment-level within carriers. Since this study assesses safety performance at the *segment* level, our ideal safety measures would be complied *by segment and by carrier*. Unfortunately, safety measures are available at the carrier-level only (without regard to the segments in which they operate). Consequently, let d_j^i be an indicator variable equal to 1 if the *j*th carrier operates in segment *i* (*i* = 1, . . . ,*M*), equal to "0" otherwise. Then, for a given carrier *j*, the number of segments in which it operates is non-negative integer:

$$\sum_{i=1}^{M} d_j^i \in [1, M].$$

The indicator variable d_j^i makes the multiple-segment nature of carriers explicit. Based on this notation, the number of observations of safety measure *k* in segment *i* is

$$N_i^k = \sum_{j=1}^{J} d_j^i < J.$$

Then a measure of average safety performance of segment *i* for safety measure *k* is

$$\bar{Y}_i^k = \frac{1}{N_i^k} \sum_{j=1}^{J} d_j^i Y_j^k; \quad k = 1, \ldots, K, \; i = 1, \ldots, M.$$

The average implies the sample variance measure:

$$s_{ik}^2 = \frac{1}{N_i^k - 1} \sum_{j=1}^{J} d_j^i (Y_j^k - \bar{Y}_i^k)^2;$$

$$k = 1, \ldots, K, \; i = 1, \ldots, M.$$

The segment safety measures $\bar{Y}_i^k \; (i = 1, \ldots, M)$, imply rank statistics for each safety measure *k*:

$$\bar{Y}_{(1)}^k \le \bar{Y}_{(2)}^k \le \cdots \le \bar{Y}_{(M)}^k; \quad k = 1, \ldots, K.$$

Unfortunately, the measure $\bar{Y}_i^k$ is potentially biased, because it does not account for the effects of multiple-segment carriers in the data. To see this one need only recognize that it is based on the carrier-wide measure Y_i^k which ignores segment. If a carrier performs well in one segment (A) and poorly in another (B), then this difference is masked in the data, and its performance in segment A is understated while its performance in B is overstated. This is less of a problem if a carrier's true safety performance is not segment specific, but more of a carrier-level phenomenon. The data set did contain quite a number of multiple-segment carriers. The For-Hire sector had about 60,000 (64%) operating in a single commodity segment, while the Private sector had about 42,000 (about 78%). The rest operated in multiple segments. For example, the For-Hire sector had about 16,000 carriers (17%) operating in two segments while 9000 (9%) operate in three. These numbers imply that multiple-segment carriers may be a significant percentage of the data set and should be addressed in the assessment of safety across segments.[3] (Note: There were also a number of carriers with no segment designation; these carriers were ignored in the analysis.)

4. Ranking and selection procedures

Ranking and Selection Theory is attributed to Gupta (1965), and it is a subset of a larger body of statistical inference procedures called "Multiple Comparison Procedures." The technical Appendix B provides some background information on what follows, however the interested reader is referred to this literature for a complete understanding of the theory and proofs.

Let λ be a value on the unit interval (0, 1), then λ can be thought of as an inferential error rate and (1−λ)×100% a confidence level. For example, for this study we select λ = 0.05, so that our inferential confidence level is 95%. Define the set of all segment indices {1, . . . ,M}k for safety measure k. Ranking and Selection Theory defines two subsets from the set {1, . . . , $M\}_k$, denoted S_+^k and S_-^k, each containing some subset of segment indices from {1, . . . , $M\}_k$. Call S_+^k the "subset of the best," and S_-^k the "subset of the worst."
Then,

with probability at least 1 − λ, the set S_+^k *contains the best motor-carrier segments for safety measure* k, *and with probability at least 1 − λ, the set* S_-^k *contains the worst motor-carrier segments for safety measure* k.

[3] To control for multiple-segment carriers in the data, a carrier-level regression was estimated to capture the marginal effects of segment membership on safety performance. Estimated marginal effects were used to predict safety scores for a typical firm in each segment. Results are not reported but were similar to the average measure results, suggesting that multi-segment carriers are, perhaps, not problematic.

Therefore, motor carrier segments that fall into the subset of the best can be deemed "best," and those in the subset of the worst can be deemed "worst" at the $(1 - \lambda) \times 100\%$ confidence level for each safety measure.

These probability statements are extremely powerful. They allow us to better understand the significance of the segment-by-segment rankings for each safety measure. It should be noted that these subsets will generally contain more than one segment (but not always). It should also be noted that as λ decreases, the probability associated with each statement increases. For this study we select the standard $\lambda = 0.05$ implying that the probability statements are made at the 95% confidence level. At larger values of λ, the confidence level decreases and the inference procedure can produce a single segment in the each of the subsets. Therefore, confidence level experiments could be conducted to single out one best and one worst segment for each safety measure, *k*.

Table 1
For-Hire trucking segment-by-segment results: $\bar{Y}_i^k$ and 95% subsets of the best and worst[a,b,c]

Safety measure[d]	Building Materials	Bulk Freight	Refrigerated	General Freight Less Than Truckload	General Freight Truckload	Household Goods	Intermodal	Large machine	Passenger	Produce	Tank
Driver-related measure											
Driver SEA	40.41	37.08	51.03 W	36.52	43.67	44.89	39.04	39.06	21.36 B	48.57	36.49
Driver OOS rate	9.36	8.86	12.15	4.76 B	11.18	13.78 W	8.08	9.64	6.93	12.06	7.32
Vehicle-related measure											
Vehicle SEA	50.23	47.30	46.02	41.33	47.53	44.07	57.23 W	51.33	30.87 B	47.49	44.51
Vehicle OOS rate	27.04	25.53	23.82	20.12 B	25.25	23.28	28.61 W	28.95 W	18.50 B	25.35	23.33
Crash-related measure											
Accident SEA	11.67	9.96	12.85	30.80 W	11.23	7.86 B	14.28	9.96	7.67 B	10.86	15.82
Fatal crash rate	0.015	0.022 W	0.017 W	0.003 B	0.015	0.013 W	0.014 W	0.014	0.007 B	0.020 W	0.009
Total crash rate	0.342	0.411 W	0.363	0.091 B	0.352	0.302	0.241	0.329	0.219	0.418 W	0.205
Management											
SMRM	20.45	22.88	21.26	9.44 B	22.17	29.32 W	20.85	21.53	17.09	23.31 W	12.66 B
ESM	2.50	2.29	3.12 W	2.50 B, W	2.60	2.51	2.62 W	2.43	1.57 B	3.14 W	2.51

[a] B denotes "segment is in the subset of best with probability at least 95% for this safety measure."
[b] W denotes "segment is in the subset of the worst with probability at least 95% for this safety measure."
[c] B, W occurred due to small sample size, N_i^k, making inference less sharp.
[d] Definitions of safety measures are contained in the data Appendix A.

Table 2
Private trucking segment-by-segment results: $\bar{Y}_i^k$ and 95% subsets of the best and worst[a,b,c]

Safety measure[d]	Building Materials	Bulk Freight	Refrigerated	General Freight Truck Load	Household Goods	Intermodal	Large machine	Passenger	Produce	Tank
Driver-related measure										
Driver SEA	25.89	25.63	28.53	29.40	30.05 W	28.69 W	31.17 W	34.60 W	31.58 W	20.70 B
Driver OOS rate	9.98	8.33	9.11	12.24 W	13.84 W	8.75 B	11.61	11.75 W	10.54	6.43 B
Vehicle-related measure										
Vehicle SEA	47.32	46.53	31.58 B	38.16	28.67 B	50.50 W	56.03 W	51.05 W	37.16	42.72
Vehicle OOS rate	27.90	28.56	16.41 B	21.45	15.18 B	28.25	33.66 W	28.81 W	19.96	24.17
Crash-related measure										
Accident SEA	4.81	5.32	7.28 W	3.81	2.74 B	4.82 B	4.13	2.74 B	5.83	5.57
Fatal crash rate	0.013 W	0.016 W	0.011 W	0.018 W	0.003 B	0.011 B, W	0.014 W	0.000[e]	0.010 B	0.010 W
Total crash rate	0.277	0.359 W	0.224	0.372 W	0.310 W	0.199	0.286	0.080 B	0.286	0.207
Management										
SMRM	31.53 W	32.73 W	35.05 W	33.89 W	28.22 B, W	39.34 B, W	33.84 W	26.67 B, W	33.43 W	20.52 B
ESM	1.48 W	1.30	2.40 W	0.98	1.34 B, W	1.08 B, W	1.45	0.14 B, W	2.19 W	1.50 W

[a] B denotes "segment is in the subset of best with probability at least 95% for this safety measure."
[b] W denotes "segment is in the subset of the worst with probability at least 95% for this safety measure."
[c] B, W occurred due to small sample size, N_i^k, making inference less sharp.
[d] Definitions of safety measures are contained in the data Appendix A.
[e] Inference could not be performed due to a complete lack of variability in the data (all observations the same).

5. Results

The results for the best and worst subsets for $\bar{Y}_i^k$ for both the For-Hire and Private sectors are in Tables 1 and 2, respectively. In the tables the columns contain the segments (*i*) and the rows contain the safety measures (*k*). For each safety measure the cells are populated with "B" if the segment was contained in the subset of the best with 95% confidence, with "W" if the segment was contained in the subset of the worst with 95% confidence, and with "B, W" if it was contained in both. Blank designations imply that the segment was neither "best" nor "worst" for that measure.

In Table 1 (For-Hire), for the driver SEA measure (first row) the single best segment was the Passenger segment and the single worst was the Refrigerated segment. This is a strong inference statement because it identifies single segments as the safety extrema at the 95% confidence level. This is also the case for the driver OOS rate and the vehicle SEA measures. In Table 1, the only occurrence of a "B, W" classification is for the Enforcement

Severity Measure (ESM) in the "General Freight Less Than Truckload" segment. This occurred because the measure was only based on N_i^k = 19 observations, so the precision of the inference for that measure is suspect; there is not enough data for the ESM measure to be statistically meaningful (in a ranking sense). In Table 1, the results are generally good because the sample sizes tended to be large. Overall, the Passenger and the "General Freight Less Than Truckload" segments performed the best, being in the best subset six times each. The worst segments were the Intermodal and Produce segments, being in the worst subset four times each. The least reliable measure for ranking the segments was the fatal crash rate and the ESM, which tended to have multiple segments in the best and worst subsets.

Turning to the Private sector results of Table 2, the inference is less sharp, due to generally smaller sample sizes than in the For-Hire analysis. For example, in the driver SEA category, the best segment was the Tank segment, but the worst subset consisted of five segments (Household, Intermodal, Large Machine, Passenger and Produce), but again this may be due to small sample sizes (e.g., Intermodal only had N_i^k = 166 observations). The driver OOS rate is an interesting case here. Notice that the segments with the lowest (best) safety scores were Tank (6.43), Bulk Freight (8.33), and Intermodal (8.75). However, the inference only selected Tank and Intermodal to be best at the 95% level. Why? Tank was selected because it had by far the lowest score. The Bulk Freight and Intermodal segments had similar scores, but only Intermodal was selected because it was measured with less precision than Bulk Freight (the Bulk Freight measure was based on N_i^k = 10, 061 observations, while the Intermodal measure was based on N_i^k = 261), implying that the inference could not reject the hypothesis that the Intermodal driver OOS rate was the smallest. This makes clear the dangers of drawing conclusions on rank statistics without statistically valid inference procedures.

Again the fatal crash rate and the ESM tended to be the least reliable safety measures in terms of discerning a best and worst segment. However, the Safety Management Review Measure (SMRM) was also fairly unreliable. All three measures had a multiplicity of segments that were both in the best and in the worst subsets, and often suffered from small sample sizes within certain segments. For example, in the ESM category the Passenger, Intermodal and Household segments only had N_i^k = 7, 12, and 29 observations, respectively. The sharpest overall inference was for the driver OOS rate, vehicle SEA, vehicle OOS rate and the accident SEA. The only safety measures that produced a single best or worst segment were the driver SEA, which determined that the Tank segment was the single best, the accident SEA, which determined that the Refrigerated segment was the single worst, and the total crash rate, which determined that the Passenger segment was the single best. The Household Goods segment was best six times but worst five times. The Passenger segment was best four times but worst six times. The Intermodal segment was best five times, but worst five times.

6. Conclusions and future research

This inferential study has highlighted the limitations and the strengths of the MCIPS safety measures in various motor carrier segments, and, perhaps more importantly, it has highlighted the strengths and weaknesses of the data itself. Without rehashing the specific results of Section 5, some strong conclusions have surfaced. First, the subsample of MCIPS data used in the analysis only contains carriers who have had at least three inspection in the last 30 months, therefore the results are only interpretable for this subpopulation of carriers. Second, rank statistics by themselves need to be interpreted with caution. This is not to say that the segment-by-segment ranks are wrong, but it is to say that inferential procedures are necessary to get a true sense of which segments are best and worst. If identifying the best and the worst performing segments is important, then it is even more important to recognize that there may be more then one best or worst segment along any safety metric, even though the point estimates suggest otherwise.

It may also be useful to reexamine the crash data results (fatal and total crash rates), since crashes are clearly important safety measures for policy and enforcement. These data tend to have relatively smaller sample sizes than other measures. While increasing the sample sizes of crash data is clearly *not* feasible, perhaps the information provided in the SafeStat data set for crashes should be expanded to improve the accuracy and sharpen the inference for these measures. One obvious improvement would be to provide information on what trucks are carrying when crashes occur; this would allow disaggregating of the data along segments, mitigate the multiple-segment carrier problem, and potentially improve the accuracy of the results.

Concerns over the effects of multiple-segment carriers must be meaningfully addressed. For instance, the study could be repeated for single-segment carriers only. However, the smaller sample sizes might make the inference less sharp. Alternatively, methods that quantify and control for "multi-segment carrier effects" in the data

themselves could be developed. Currently, the data do not admit this quantification, so perhaps the data collection process could be altered to make this feasible.

Acknowledgements

The support and assistance of Cyrus Baghelia (*Econometrica, Inc.*, Bethesda, MD) is gratefully acknowledged. Comments of government officials from the Federal Motor Carrier Safety Administration (FMCSA) and Department of Transportation are also acknowledged.

Appendix A. Data

Nine measures were selected from a larger list of 23 safety performance measures for which data were collected in this study. Many of the 23 measures are closely correlated, and several serve as direct inputs to others. In an effort to reduce the potential effect of this correlation on the results, researchers selected nine "core" measures in this analysis. These nine measures are fairly representative of the four safety areas of interest in this study: driver-, vehicle-, crash-, and safety management-related areas. The specific safety performance measures examined include:

Driver safety evaluation area (*SEA*) is a SafeStat composite value calculated from driver inspection indicator (DII), driver review indicator (DRI), and moving violation indicator (MVI). The DII is based on driver roadside out-of-service (OOS) inspection violations, the DRI is based on violations of driver-related acute and critical regulations discovered during a compliance review, and the MVI is based on serious moving violations recorded in conjunction with roadside inspections. Each inspection is weighted by its age and the number of driver OOS violations found, and then normalized by the number of driver inspections within the last 30 months.

Total driver OOS rate is derived from MCMIS data. The total number of driver OOS violations, divided by the total number of driver inspections experienced by the motor carrier.

Vehicle safety evaluation area (*SEA*) is a SafeStat composite value calculated from the vehicle inspections indicator (VII) and the vehicle review indicator (VRI). The VII is based on the number of vehicle roadside OOS inspection violations and the VRI is based on violations of vehicle-related acute and critical regulations discovered during compliance reviews. Each inspection is weighted by its age and the number of vehicle OOS violations found, and then normalized by the number of vehicle inspections within the last 30 months.

Total vehicle OOS rate is derived from MCMIS data. The total number of vehicle OOS violations divided by the total number of vehicle inspections experienced by a motor carrier.

Accident safety evaluation area (*SEA*) is a SafeStat composite value calculated based on accident involvement indicator (AII) and the recordable accident indicator (RAI). The AII uses measures derived from state-reported crash data normalized by the number of power units owned/leased by the motor carrier from MCMIS. The RAI uses measures based on recordable crashes and annual vehicle miles traveled (VMT) data gathered at the most recent compliance review.

Fatal crash rate is derived from MCMIS data. The number of fatal crashes experienced by a carrier divided by the number of power units owned or leased by that carrier.

Total crash rate is derived from MCMIS data. The number of total recordable crashes experienced by a carrier divided by the number of power units owned or leased by that carrier.

Safety Management Review Measure (*SMRM*) is a compliance review measure that uses the rate of safety management-related acute and critical violations of FMCSA regulations discovered during a compliance review.

Enforcement Severity Measure (*ESM*) is derived from MCMIS data and is weighted number of past (closed) enforcement cases brought against a carrier by FMCSA. This is a general indicator of the commitment to safety by the motor carrier over time (within the last 6 years).

Appendix B. Technical

The following can be found in Horrace (1998). Assume that the data for safety measure *k* can be represented by $Z_k' = [Z_1, \ldots, Z_M]$, an *M*-variate normal distribution with *M* finite means, $\mu_k' = [\mu_1, \ldots, \mu_M]$, and finite, covariance matrix $\Omega^k\ (M \times M)$. Define L_s, *s* = 1, 2, . . . ,*M* as an (*M* − 1) ×M matrix such that

$$L_s Z_k = [Z_s - Z_1, \ldots, Z_s - Z_{s-1}, Z_s - Z_{s+1}, \ldots, Z_s - Z_M]$$

Then, the joint density of $L_s Z_k$ is

$$
\begin{aligned}
&f(z_1, \ldots, z_M, \Omega^k, L_s) \\
&\quad = (2\pi)^{-M/2}(\det L_s \Omega^k L_s')^{-(1/2)} \\
&\qquad \times \exp\left\{-\frac{1}{2}(z-\mu_k)' L_s'(L_s \Omega^k L_s')^{-1} L_s(z-\mu_k)\right\}
\end{aligned}
$$

Standardize $L_s Z_k$, then $L_s \Omega^k L_s'$ is a correlation matrix. Define the critical value z_Ω^λ as the solution in z of the equation:

$$
P\{\max_i |Z_i| \le z\} = \int_{-z}^{z} \cdots \int_{-z}^{z} f(z_1, \ldots, z_M, \Omega^k, L_s) \times dz_1 \cdots dz_M = 1 - \lambda.
$$

Let $\bar{Y}_i^k$ $i = 1, \ldots, M$ be any independent, unbiased estimates of the commodity segment means for safety measure k. Let the elements of Ω^k be ω_{nt}^k, $n, t = 1, 2, \ldots, M$. Define the "Gupta subsets," S_-^k and S_+^k:

$$
S_-^k = \{s : \bar{Y}_s^k - \bar{Y}_i^k + z_\Omega^\lambda(\omega_{ss} + \omega_{ii})^{1/2} \ge 0; \quad \forall i \ne s\}
$$

and

$$
S_+^k = \{s : \bar{Y}_i^k - \bar{Y}_s^k + z_\Omega^\lambda(\omega_{ss} + \omega_{ii})^{1/2} \ge 0; \quad \forall i \ne s\}.
$$

For this study, the degrees of freedom are large enough to allow $\omega_{ii}^k = s_{ik}^2 / N_i^k$ in the above formulae. The z_Ω^λ was calculated for each k, using an iterative search at three decimal places and the Simpson numerical integration feature in the Gauss programming language.

References

Gupta, S.S., 1965. On some multiple decision (selection and ranking) rules. Technometrics 7 (2), 225–245.

Horrace, W.C., 1998. Tables of percentage points of the k-variate normal distribution for large values of *k*. Commun. Stat. Simul. Comput. 27 (3), 823–831.

McCartt, A.T., Rohrbaugh, J.W., Hammer, M.C., Fuller, S.Z., 2000. Factors associated with falling asleep at the wheel among long-distance truck drivers. Accid. Anal. Prev. 32 (4), 493–504.

Moses, L.N., Savage, I., 1997. A cost–benefit analysis of US motor carrier safety programmes. J. Transport. Econ. Policy 31 (1), 51–67.

Moses, L.N., Savage, I., 1996. Identifying dangerous trucking carriers. Risk Anal. 16 (3), 359–366.

Moses, L.N., Savage, I., 1994. The effect of firm characteristics on truck accidents. Accid. Anal. Prev. 26 (2), 173–179.

Scopatz, R.A., 2001. Crashes involving long combination vehicles—data quality problems and recommendations for improvement. Transport. Res. Rec. 1779, 162–172.