MOTOR CARRIER SAFETY

A Statistical Approach Will Better Identify Commercial Carriers That Pose High Crash Risks Than Does the Current Federal Approach

G A O
Accountability·Integrity·Reliability

Highlights

Highlights of GAO-07-585, a report to congressional requesters

Why GAO Did This Study

The Federal Motor Carrier Safety Administration (FMCSA) has the primary federal responsibility for reducing crashes involving large trucks and buses that operate in interstate commerce. FMCSA decides which motor carriers to review for compliance with its safety regulations primarily by using an automated, data-driven analysis model called SafeStat. SafeStat uses data on crashes and other data to assign carriers priorities for compliance reviews.

GAO assessed (1) the extent to which changes to the SafeStat model could improve its ability to identify carriers that pose high crash risks and (2) how the quality of the data used affects SafeStat's performance. To carry out its work, GAO analyzed how SafeStat identified high-risk carriers in 2004 and compared these results with crash data through 2005.

What GAO Recommends

GAO is recommending that FMCSA use a negative binomial regression model to identify carriers that pose high crash risks.

In commenting on a draft of this report, the Department of Transportation agreed that the use of a negative binomial regression model looked promising for selecting carriers for compliance reviews, but expressed some reservation about the greater sensitivity of this approach to problems with reported crash data.

www.gao.gov/cgi-bin/getrpt?GAO-07-585.

To view the full product, including the scope and methodology, click on the link above. For more information, contact Sidney H. Schwartz at (202) 512-7387 or schwartzsh@gao.gov, or Susan A. Fleming at (202) 512-2834 or flemings@gao.gov.

What GAO Found

While SafeStat does a better job of identifying motor carriers that pose high crash risks than does a random selection, regression models GAO applied do an even better job. SafeStat works about twice as well as (about 83 percent better than) selecting carriers randomly. SafeStat is built on a number of expert judgments rather than using statistical approaches, such as a regression model. For example, its designers decided to weight more recent motor carrier crashes twice as much as less recent ones on the premise that more recent crashes were stronger indicators of future crashes. GAO estimates that if FMCSA used a negative binomial regression model, FMCSA could increase its ability to identify high-risk carriers by about 9 percent over SafeStat. Carriers identified by the negative binomial regression model as posing a high crash risk experienced 9,500 more crashes than those identified by the SafeStat model over an 18 month follow-up period. The primary use of SafeStat is to identify and prioritize carriers for FMCSA and state compliance reviews. FMCSA measures the ability of SafeStat to perform this role by comparing the crash rate of carriers identified as posing a high crash risk with the crash rate of other carriers. Using a negative binomial regression model would further FMCSA's mission of reducing crashes through the more effective targeting of compliance reviews to the set of carriers that pose the greatest crash risk.

Late-reported, incomplete, and inaccurate data reported to FMCSA by states have been a long-standing problem. However, GAO found that late reported data had a small effect on SafeStat's ability to identify carriers that pose high crash risks in 2004. If states had reported all crash data within 90 days after occurrence, as required by FMCSA, a net increase of 299 carriers (or 6 percent) would have been identified as posing high crash risks of the 4,989 that FMCSA identified. Reporting timeliness has improved, from 32 percent of crashes reported on time in fiscal year 2000, to 89 percent in fiscal year 2006. Regarding completeness, GAO found that data for about 21 percent of the crashes (about 39,000 of 184,000) exhibited problems that hampered linking crashes to motor carriers. Having complete information on crashes is important because SafeStat treats crashes as the most important factor for assessing motor carrier crash risk, and crash information is also the crucial factor in the statistical approaches that we employed. Regarding accuracy, a series of studies by the University of Michigan Transportation Research Institute covering 14 states found incorrect reporting of crash data is widespread. GAO was not able to quantify the effect of the incomplete or inaccurate data on SafeStat's ability to identify carriers that pose high crash risks because it would have required gathering crash records at the state level—an effort that was impractical for GAO. FMCSA has acted to improve crash data quality by completing a comprehensive plan for data quality improvement, implementing an approach to correct inaccurate data, and providing grants to states for improving data quality, among other things.