United States Government Accountability Office

Report to Congressional Requesters

July 2015

MOTOR CARRIER SAFETY

Additional Research Standards and Truck Drivers' Schedule Data Could Allow More Accurate Assessments of the Hours of Service Rule

GAO-15-641

GAO Highlights

Highlights of GAO-15-641, a report to congressional requesters

July 2015

MOTOR CARRIER SAFETY

Additional Research Standards and Truck Drivers' Schedule Data Could Allow More Accurate Assessments of the Hours of Service Rule

Why GAO Did This Study

FMCSA—within the Department of Transportation (DOT)—issues rules to address safety concerns of the motor carrier industry, including on truck drivers' HOS. In July 2013, FMCSA began to enforce three new provisions of its HOS rule. GAO was asked to review a 2014 FMCSA study on the rule, as well as the rule's assumptions and effects. This report (1) compares the study to generally accepted research standards, and (2) identifies the assumptions used to estimate the rule's costs and benefits and the rule's driver-operation, economic, safety, and health effects.

GAO identified research standards that professional associations, academics, and GAO's prior work have used. GAO evaluated the 2014 FMCSA study against these standards. GAO also compared FMCSA's assumptions about how drivers would be affected by the HOS rule against actual drivers' schedule data from 16 for-hire carriers that cover the years 2012 through 2014. These data include information on over 15,000 drivers per year, but are not generalizable to the motor carrier industry as a whole.

What GAO Recommends

GAO recommends that FMCSA adopt guidance outlining agency research standards. FMCSA agreed with GAO's recommendation. GAO also suggests that Congress consider directing DOT to study and report on how electronically collected driver schedule data can be extracted, stored, and analyzed in a way that addresses cost and privacy concerns.

View GAO-15-641. For more information, contact Susan Fleming, (202) 512-2834, or flemings@gao.gov

What GAO Found

GAO found that the January 2014 study issued by the Federal Motor Carrier Safety Administration (FMCSA) to examine the efficacy of its hours of service (HOS) rule—a regulation that governs how many hours truck drivers transporting freight can work—followed most generally accepted research standards. However, FMCSA did not completely meet certain research standards such as reporting limitations and linking the conclusions to the results. For example, by not adhering to these standards, FMCSA's conclusion in the study about the extent to which crash risk is reduced by the HOS rule may be overstated. GAO found that FMCSA has not adopted guidance on the most appropriate methods for designing, analyzing, and reporting the results of scientific research. Without such guidance, FMCSA may be at risk for excluding critical elements in research it undertakes to evaluate the safety of its rules, leaving itself open to criticism.

FMCSA made several assumptions and anticipated certain effects of the HOS rule in the regulatory impact analysis. Specifically, to estimate the economic costs of the rule, FMCSA assumed that some drivers would lose a certain amount of driving and on-duty time and then estimated the amount and cost of the work time lost. Further, FMCSA assumed that reduced work time could increase a driver's opportunity to sleep, leading to safety and health benefits. Assessing the effectiveness of the HOS rule is difficult because of the limited availability of representative driver schedule data (i.e., records of drivers' work hours). Nevertheless, GAO's analysis of a limited sample of available data provides some insight into the rule's effects and the extent to which they aligned with FMCSA's assumptions and estimates. For example, according to GAO's analysis, some drivers at a sample of 16 for-hire carriers who worked the longest hours (over 65 hours per work week) reduced their work hours after the rule went into effect, a finding consistent with FMCSA's assumptions that drivers working over 65 hours were more likely to be affected. However, GAO's analysis found that drivers who worked less than 65 hours per work week also changed their schedules after the rule went into effect, a result not anticipated by FMCSA.

The ability of FMCSA and others to assess the effects of rules, such as the 2011 HOS rule, is impacted by the limited availability of representative driver schedule data. No organization collects or maintains a centralized database with such data that can be generalized to the motor carrier industry as a whole. Collecting schedule data has historically been difficult, but a recent statutory change that requires carriers to electronically record and store these data provides a potential data source for the future. However, before these data can be used for research purposes several challenges would have to be addressed. First, there are statutory limits on the use of these data for purposes other than enforcing motor carrier safety regulations. Additionally, privacy and cost concerns must be resolved before these data could be made available for analysis. According to FMCSA officials, they do not plan to study how to use these data in a way that will address privacy and cost concerns, in part, because of the statutory limits. Given the potential value of these data to future regulatory analysis, it may be important to provide Congress with information on how these data can be extracted, stored, and analyzed while addressing any privacy and cost concerns.

Contents

Tables

Figures

Abbreviations

ANOVA	analysis of variance
ATRI	American Transportation Research Institute
BEA	Bureau of Economic Analysis
DOT	Department of Transportation
FAID	Fatigue Audit InterDyne™
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
HOS	hours of service
KSS	Karolinska Sleepiness Scale
MAP-21	Moving Ahead for Progress in the 21st Century Act
MCMIS	Motor Carrier Management Information System
OMB	Office of Management and Budget
PVT	psychomotor vigilance test
RIA	regulatory impact analysis
WIM	Weigh-in-Motion

U.S. GOVERNMENT ACCOUNTABILITY OFFICE
441 G St. N.W.
Washington, DC 20548

July 29, 2015

The Honorable Bill Shuster
Chairman
Committee on Transportation and Infrastructure
House of Representatives

The Honorable Deb Fischer
United States Senate

The United States economy depends on commercial motor vehicles (e.g., large trucks) to move goods across the country. Commercial motor vehicles were responsible for shipping approximately 13-billion tons of freight in the United States in 2012—a figure that is estimated to increase over forty percent by 2040. However, the large number of commercial motor vehicles on the road poses safety risks. In 2013, there were over 2,500 fatal crashes that involved large trucks.[1]

The primary mission of the United States Department of Transportation's (DOT) Federal Motor Carrier Safety Administration (FMCSA) is to reduce crashes, injuries, and fatalities involving large trucks and buses. To achieve this mission, FMCSA issues regulations to address safety concerns, including rules governing the number of hours that property-carrying commercial motor vehicles can work and drive per day and week.[2] These hours of service (HOS) rules are similar to those found in other modes of transportation, such as aviation and rail, and are designed to ensure that those operating commercial vehicles do not become fatigued and, as a result, crash.

[1]A large truck is defined as a truck with a gross vehicle-weight rating greater than 10,000 pounds. The Federal Motor Carrier Safety Administration has authority to regulate motor carriers that travel between states or "interstate" motor carriers as well as all motor carriers that transport hazardous material. A motor carrier is defined as a person providing motor vehicle transportation for compensation. 49 U.S.C. § 13102 (14). This report focuses on commercial interstate and intrastate hazardous material trucks that transport freight.

[2]Federal law gives the Secretary of Transportation the authority to issue hours of service regulations "when needed to promote safety of operation." 49 U.S.C. § 31502. Similar, concurrent authority is provided by 49 U.S.C. § 31136(a).

In 2003, FMCSA published several changes to the HOS rule for drivers of property-carrying commercial motor vehicles. For example, FMCSA increased the daily driving limit from 10 to 11 hours, but also extended the amount of time drivers must spend off-duty following time spent driving or working from 8 to 10 hours. The 2003 final rule led to a decade of legal challenges from organizations representing motor carriers, which believed the HOS rule would negatively impact economic productivity, and from safety advocacy organizations, which believed that the rule did not sufficiently protect safety and health. In an effort to address these concerns, FMCSA revised the HOS rule three more times between 2005 and 2011, publishing a final rule on December 27, 2011.[3] FMCSA began to enforce the new provisions on July 1, 2013 and in August 2013, the United States Court of Appeals for the D.C. Circuit upheld all but one relatively minor provision of the rule.[4]

The 2011 final rule made several key changes to the 2003 rule, including altering the "restart" provision that allows commercial drivers to reset or begin at zero the on-duty hours counted against maximum work week requirements (up to 70 hours over 8 days) by taking 34 hours off-duty.[5] Specifically, the new rule added the following provisions:

- **The two-night provision:** Drivers choosing to use the restart provision must include two-nighttime rest periods from 1 a.m. to 5 a.m.

- **The 168-hour limit:** Drivers can only use the restart provision once per week (every 168 hours).

[3] 76 Fed. Reg. 81134 (Dec. 27, 2011).

[4] The court vacated the provision that would have subjected short-haul drivers—those who drive in a local area—to the requirement to take a 30-minute rest break. The court concluded that DOT had a rational basis for the rule and found there was no serious flaw in the rule's cost/benefit analysis. *American Trucking Associations v. Federal Motor Carrier Safety Administration,* 724 F.3d 243 (D.C. Cir. 2013). Due to the court's ruling on this matter, we are not reevaulating the conclusions of the rulemaking; instead this report focuses on the effects of the implemented rule.

[5] 76 Fed. Reg. 81134 (Dec. 27, 2011). Drivers are not required to take this 34-hour break, but may choose to do so to increase their scheduling flexibility, among other reasons. The background section of this report provides more detail on the restart provision.

- **30-minute rest break:** Drivers are required to take a 30-minute, off-duty break at some point during the first eight hours of work, i.e., a work shift, to continue to operate a commercial vehicle.

FMCSA has been required to conduct two studies on the 2011 HOS rule. Specifically, in July 2012 FMCSA was required by law to complete a field study examining the impact of the two-night provision of the HOS rule on commercial drivers in a naturalistic environment.[6] In December 2014, enforcement of two of the new provisions was suspended by law—the 168-hour limit and the two-night provision—and FMCSA was required to complete an additional naturalistic study examining driver work schedules and fatigue levels before and after the 2011 final rule went into effect.[7,8]

Given the high level of interest in the 2011 HOS rule and its potential impacts, you asked us to examine research that evaluated the rule and assess the rule's effects. This report examines (1) the strengths and limitations of FMCSA's field study that examined whether the two-night provision reduces commercial driver fatigue and (2) the key assumptions FMCSA used, at the time the 2011 HOS rule was promulgated, in estimating its effects and what is known about the economic, safety, and health effects of the rule.

To assess the strengths and limitations of the field study, we identified generally accepted research standards and compared them to information included in the study, statements made by FMCSA officials and the authors of the study, and our analysis of study data, which were collected between January and July 2013. We also assessed the reliability and accuracy of data and results reported in the field study by reviewing the data collected for each of the study participants.[9] We

[6]Moving Ahead for Progress in the 21st Century Act, Pub. L. No. 112-141, § 32301(a), 126 Stat. 405,786 (2012).

[7]A naturalistic study observes the behavior of participants in their natural environment without interference by researchers.

[8]FMCSA was also required to assess safety critical events and driver health outcomes, among other things. See the Consolidated and Further Continuing Appropriations Act, 2015, Pub. L. No. 113-235, div. K, tit. I, § 133(a) 128 Stat. 2130, 2711 (2014).

[9]Specifically, our review included a reliability assessment of the computer programs used to process the data, replication of the data reported in the field study, and a sensitivity analysis in which we altered some of the methodologies used in the study to identify how choices made by FMCSA might have impacted the results.

determined that these data were reliable for the purposes of assessing the field study, including the methods used. We also conducted a literature review of scholarly peer reviewed materials, government reports, and conference papers that discussed the validity and reliability of the fatigue measures used in the field study, as well as assessed the relationships between these measures and on-the-road safety outcomes (e.g., crashes).

To identify key assumptions FMCSA used to estimate the effects of the 2011 HOS rule, we first identified those assumptions that likely had an important effect on estimated costs and benefits. We then categorized these assumptions into five areas: which drivers were affected, how drivers were affected, economic costs, safety benefits, and health benefits. FMCSA officials confirmed that our characterization of key assumptions was accurate. To evaluate what is known about the economic, safety, and health effects of the 2011 HOS rule, we conducted a comprehensive search for existing databases with information on motor carrier operations and asked officials from FMCSA and other stakeholders for recommendations of databases we could use to analyze the effects of the rule. Through this process we identified three relevant data sources suitable for analysis: (1) driver logbook data from the American Transportation Research Institute that covers the years 2012, 2013, and 2014 and includes information on over 15,000 drivers from 16 for-hire carriers; (2) vehicle count data from the Federal Highway Administration that covers the years 2012 and 2013 and includes data on non-commodity-carrying (including passenger) vehicles, and commodity-carrying (i.e., commercial) vehicles from select locations in 14 states; and (3) FMCSA data on all reported crashes involving commercial motor vehicles in the United States that covers the years 2008 through 2014. We reviewed the reliability of these data by examining documentation, talking with knowledgeable officials, and checking the data for completeness and reasonableness and determined that the data were sufficiently reliable for the specific purposes of our data analysis. We analyzed data from each of these sources to identify any possible changes in motor carrier behavior and safety outcomes as a result of the HOS rule. While these were the best data available, the results of the analysis of driver logbook data cannot be generalized to the entire motor carrier industry. In addition, we used a biomathematical fatigue model that estimates human alertness based on working schedules to determine possible fatigue levels for drivers before and after the 2011 HOS rule went into effect in July 2013. We also reviewed provisions in the Moving Ahead for Progress in the 21st Century Act (MAP-21) requiring motor carriers to electronically record drivers' schedules and compared

FMCSA's lack of such schedule data to data collection practices of other DOT agencies that regulate hours of service.

We also spoke with FMCSA officials, the authors of the field study, and other selected experts and stakeholders, including fatigue experts, organizations representing motor carriers, safety advocacy organizations, motor carrier companies, commercial vehicle drivers, and customers of the motor carrier industry (e.g., companies that ship freight). We selected fatigue experts based on academic articles they published related to the fatigue measures used in the field study, as well as our prior work on fatigue science.[10] We selected organizations representing motor carriers, safety advocacy organizations, and shipping companies to interview based on our prior work on commercial motor carrier safety and recommendations from various stakeholders. We selected motor carrier companies to speak with based on several criteria, including how they operated (i.e., the distance they traveled), their fleet size, and the cargo they transported. Motor carrier companies we spoke with also put us in contact with drivers whom we interviewed to understand how the 2011 rule had impacted them. The views and information collected from interviews of these stakeholders are not generalizable to other motor carrier stakeholders. We also conducted several literature reviews of academic peer reviewed materials, government reports, and conference papers that studied the potential impacts of the HOS rule. Appendix I contains a more detailed explanation of our scope and methodology. Appendix II contains details about our work to test the sensitivity of the results of FMCSA's field study to various alternative methodological choices. Appendix III contains more detail on the process we used to identify FMCSA's key assumptions. Appendix IV contains details on our analysis of driver schedule data. Appendix V contains details on our use of a biomathematical fatigue model. Appendix VI contains details on our analysis of vehicle count data. Appendix VII contains details on our analysis of crash data.

We conducted this performance audit from June 2014 to July 2015 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our

[10]GAO, *Freight Railroad Safety: Hours of Service Changes Have Increased Rest Time, but More Can Be Done to Address Fatigue Risks,* GAO-11-853, (Washington, D.C.: September 2011).

findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on the audit objectives.

Background

The commercial motor-carrier industry is a vital part of the U.S. economy. As of December 2013, FMCSA estimates that there were more than 539,000 active motor carriers and 5.6 million commercial drivers operating in the United States that transport freight and move people,[11] numbers that fluctuate over time due to the tens of thousands of carriers that enter and leave the market annually. Truck carriers are responsible for transporting approximately 13 billion tons of cargo annually, representing 67 percent of all cargo weight that was shipped in 2012.

The United States' commercial motor-carrier industry represents a range of businesses, including private and for-hire freight transportation,[12] passenger carriers, and specialized transporters of hazardous materials. These carriers can be small and involve a single vehicle that is owned and operated by a single individual, or large corporations that own tens of thousands of vehicles. Similarly, commercial motor-vehicle operations differ greatly, including whether a carrier operates locally, regionally, or across the country (i.e., "over-the-road"), as well as the time of day when a driver operates. In addition, some carriers move goods as part of a single shipment to one destination (called "truckload" operations), while other carriers make numerous deliveries to many different destinations (called "less-than-truckload" operations).

Because commercial motor vehicles travel hundreds of billions of miles every year on our nation's highways, alongside citizens in passenger vehicles, safety is an important concern. In general, the number of commercial motor-vehicle crashes has declined over the past decade, but increased more recently. Commercial motor-vehicle crashes can be caused by a variety of factors. For example, a crash can be the result of vehicle defects (e.g., inadequate tire pressure or a faulty brake system) or driver error by either the commercial driver or the driver of a passenger

[11]FMCSA, *2014 Pocket Guide to Large Truck and Bus Statistics* (October 2014).

[12]Private carriers run an internal trucking operation to support a primary business in another industry, such as a retail store chain, while for-hire carriers sell their trucking services on the open market.

vehicle. FMCSA has reported that the most commonly cited driver-related factor in a fatal truck crash in 2012 was speeding.[13] That same year, driver impairment, including by fatigue, was the fourth most commonly cited factor.

Research indicates fatigue can lead to a state of diminished capacity, which can have ramifications for humans such as having more difficulty maintaining attention, becoming less communicative, and having reduced situational awareness. People are then at greater risk of committing errors in their work, which can ultimately lead to more crashes. Managing fatigue is particularly critical for tasks that require constant attention, such as driving a commercial vehicle.

FMCSA is responsible for overseeing the large and diverse commercial motor-vehicle industry. To do so, FMCSA establishes safety standards for interstate motor carriers as well as intrastate hazardous-material carriers operating in the United States.[14] To enforce compliance with these standards, FMCSA partners with state agencies to perform roadside inspections of vehicles and investigations of carriers.[15] In fiscal year 2014, FMCSA had a budget of approximately $572 million and almost 1,100 FMCSA staff members located at headquarters, four regional service centers, and 52 division offices.

Drivers of commercial motor vehicles in the United States have been subject to driving and working hour restrictions for almost 80 years.[16] In 1937, the entity that previously carried out certain of FMCSA's current functions—the Interstate Commerce Commission—adopted the first HOS rule. While the specific requirements in several of the provisions of that

[13]FMCSA, *Large Truck and Bus Crash Facts 2012*, June 2014.

[14]49 U.S.C. §§ 31136 (Interstate), 5103 (Hazardous materials).

[15]State agencies include state highway patrols, departments of transportation, and public utility commissions. FMCSA employs full-time vehicle inspectors on the southern border of the United States. In addition, all FMCSA safety investigators, safety auditors, and inspectors must conduct a minimum number and certain types of inspections annually to maintain certification.

[16]The HOS regulations are found at 49 C.F.R. Part 395. Appendix B to 49 C.F.R. Part 386 provides details on the penalties associated with violating the HOS regulations, including a maximum civil penalty of $11,000 for knowing falsification of HOS records, and a maximum penalty of $16,000 per violation for certain non-recordkeeping violations, among others.

rule have changed over time, there are three general requirements that are still in place today:

- **Daily off-duty period:** Drivers must be off-duty—not working or driving—a minimum number of hours per day.[17]

- **Daily driving limit:** Drivers can only drive a maximum number of hours per day.

- **60/70 hour on-duty limit:** Drivers are restricted from driving when they reach a total of 60 or 70 hours of on-duty time[18] over a rolling 7- or 8-day period, respectively.[19]

Because the 60/70 hour on-duty limit is a rolling period, drivers must calculate these hours by looking back over the prior 7- or 8-day period. For example, according to the schedule below, the driver has accumulated 68 hours in the 8-day rolling period from Thursday to Thursday, and 72 hours of on-duty time from Friday to Friday (see fig. 1). In this example, the driver worked under 70 hours (68 hours total) and is in compliance with the 60/70 hour on-duty limit from Thursday to Thursday, but not from Friday to Friday when he worked more than 70 hours (72 hours total) over an 8 day period.

[17]The current specific hourly requirements for the daily off-duty period and daily driving limit provisions, which were put into place in 2003, are discussed later in this background section.

[18]On-duty time is the total working hours, including driving. For example, on-duty time may include time spent at a plant, terminal, or other facility of a motor carrier or shipper waiting to be dispatched and time loading, unloading, supervising, or attending the truck or handling receipts for shipments.

[19]Whether a driver falls under the 7- or 8-day period is determined by the carrier for which the driver operates. If a driver's carrier (employer) does not operate commercial motor vehicles every day of the week, the driver may not drive after being on-duty 60 hours during any 7 consecutive days. If a driver's carrier does operate commercial motor vehicles every day of the week, a 70-hour/8-day schedule is permitted. The driver may not drive after being on duty 70 hours in any 8 consecutive days. 49 CFR 395.3(b).

Figure 1: Example of Calculating On-Duty Hours for a Driver over an 8-Day Rolling Period

Day	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri
On-duty hours	8	12	0	0	12	11	13	12	12

8 + 12 + 0 + 0 + 12 + 11 + 13 + 12 = **68** on-duty hours in 8 days (Thu-Thu)

12 + 0 + 0 + 12 + 11 + 13 + 12 + 12 = **72** on-duty hours in 8 days (Fri-Fri)

Source: GAO. | GAO-15-641

In 2003, FMCSA altered the HOS rule for drivers transporting freight in several significant ways:

- Created a 14-hour “driving window” that restricts driving beyond the 14th consecutive hour after a driver comes on-duty.
- Increased the daily driving limit (from 10 to 11 hours).
- Increased the off-duty period (from 8 to 10 hours).

Collectively these provisions allow drivers to operate a commercial motor vehicle for 11 hours a day and be on-duty or work up to 14 hours per day, but also require drivers not to work or drive for at least 10 hours per day.

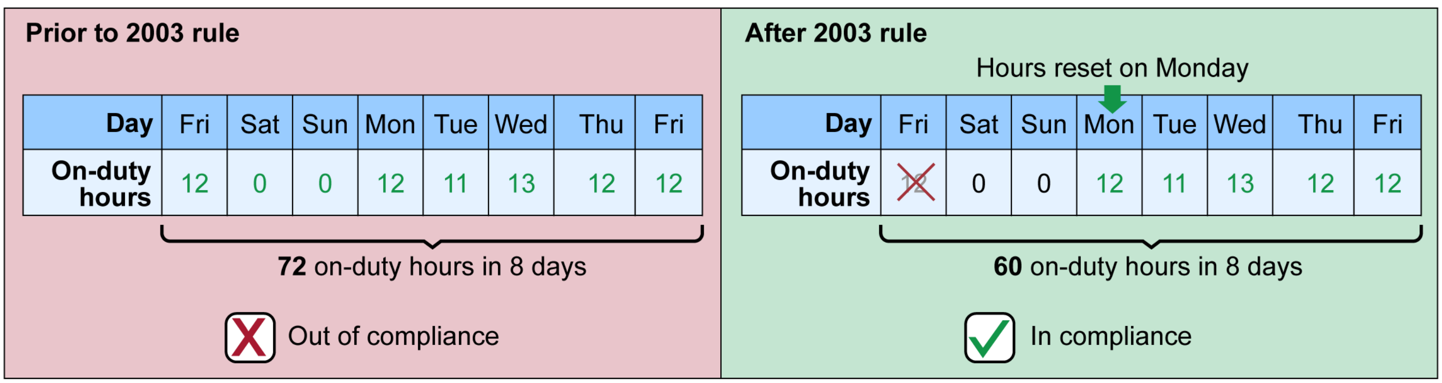
FMCSA retained the 60/70 hour on duty limit, but allowed drivers transporting freight to “restart,” or begin at zero, their on-duty hours following an off-duty period of 34 hours or longer. Drivers are not required to take this 34-hour break, but may choose to do so to increase their scheduling flexibility, among other reasons. The provision, according to FMCSA, was adopted because it would provide drivers experiencing sleep loss or significant sleep debt with an opportunity to recover and, therefore, prevent a significant number of fatigue-related crashes.[20]

To understand the restart provision, consider the example depicted below in which a driver does not have any on-duty hours for Saturday and

[20] 68 Fed. Reg. 22456, 22479 (Apr. 28, 2003).

Sunday (48 hours). Prior to the restart provision, this driver would have reported working 72 hours from Friday through Friday and would be out of compliance with the HOS rule. Under the restart provision, the driver may now reset his/her total accumulated hours to zero on Monday because the driver was not driving or working over the weekend, a period greater than 34 hours. Therefore, while the driver was not in compliance with the 60/70 hour on-duty limit under the prior rule, he/she would be in compliance using the new restart provision, having accumulated only 60 hours for the 8 day period from Friday to Friday. (See figure 2).

Figure 2: Example of Calculating 60/70 On-Duty Hour Limit before and after 2003 Hours of Service Rule



Source: GAO. | GAO-15-641

As previously discussed, in 2011 FMCSA made several significant changes to the HOS rule, including the restart provision. Specifically, the rule added the following provisions:

- **Two-night provision:** For the 34-hour off-duty period to count as a restart, it must include two nighttime periods from 1 a.m. to 5 a.m.[21]
- **168-hour limit:** Drivers can only take a 34-hour restart every 168 hours (once per week).
- **30-minute rest break:** Drivers must take a 30-minute, off-duty break during the first 8 hours of their work day in order to continue to drive.[22]

[21]Nighttime periods are based on the driver's home terminal time zone. For example, a driver who lives in the Pacific Time zone, but is working in the Eastern Time zone would have nighttime periods running from 4 a.m. to 8 a.m. Eastern Time.

According to FMCSA, the 2011 changes were made for several reasons. Specifically, the two-night requirement was intended to allow drivers more nighttime rest, which is more restorative than daytime rest because it better accords with the body's circadian rhythms. The 168-hour limit was put into place to reduce fatigue by limiting the ability of drivers to work close to the maximum number of on-duty hours permitted on a continuing basis. For example, prior to the 2011 rule changes drivers who used a restart could work more than 80 hours in an average week. By limiting use of the restart to once every 168 hours, drivers could work no more than 70 hours in an average week. Lastly, the rest break requirement was designed to alleviate fatigue and fatigue-related performance degradation.

FMCSA asserted that, while the 2011 rule would result in some economic costs, the changes would affect only a small percentage of drivers who regularly work long hours and that the safety and health benefits would outweigh any loss in productivity.[23] Several organizations that represent motor carriers, however, have stated that the rule negatively impacted a larger portion of the trucking industry than was anticipated by FMCSA and that the rule had several unintended consequences, such as forcing a large amount of commercial vehicle traffic onto the roads during congested periods of the day.

In July 2012, FMCSA was required by federal law to conduct a study examining the efficacy of the restart rule, which was to be completed by March of 2013.[24] The study was to include statistically valid analysis and data on drivers in real world conditions affected by the maximum driving-time requirements. FMCSA was required to develop a methodology consistent with prior research examining the two-night provision that was conducted in a laboratory. Based on these requirements, FMCSA completed a field study (i.e., looking at drivers in real world conditions) that focused on the two-night provision, but did not examine the other two new HOS provisions: the 168-hour limit and the 30-minute rest break. To complete the field study, FMCSA contracted with Washington State University and Pulsar Informatics—a technology development and

[22]76 Fed. Reg. 81134 (Dec. 27, 2011).

[23]76 Fed. Reg. 81134 (Dec. 27, 2011).

[24]While FMCSA was to complete the field study by March 2013, it was not required to submit a report until September 2013. Pub. L No. 112-141, § 32301, 126 Stat. 405, 786.

research company.[25] Researchers from these institutions collected large amounts of data from 106 commercial vehicle drivers during two duty cycles that included two different restarts of at least 34 hours or longer. Data collected included information on sleeping and activity patterns, caffeine use, and three fatigue measures:

- the psychomotor vigilance test (PVT), which was the primary outcome measure of the study and measured a driver's ability to react to visual stimuli, e.g., flashing lights, on a smart phone within a specific amount of time;
- the Karolinska Sleepiness Scale (KSS) that includes a subjective self-assessment of how tired a driver was at a specific point in time; and
- lane deviation, which measured lateral vehicle movements within a driving lane.

Researchers used these data to compare fatigue levels for drivers taking a restart that included one-nighttime period versus drivers taking a restart that included two-or-more nighttime periods. Based on this evidence, the study concluded that a restart with two-nighttime breaks helps to mitigate fatigue and provides support of the efficacy of the new restart rule. The final report was submitted to FMCSA in September 2013 and issued publicly in January 2014.[26]

In December 2014—approximately 18 months after the HOS rule went into effect—congressional action resulted in the suspension of two provisions: the 168-hour limit and the two-night provision.[27] These provisions are suspended until September 30, 2015, or FMCSA completes a new, naturalistic study—currently underway—examining the operational, safety, health, and fatigue impacts of the restart provisions in the HOS rule, whichever comes later. While similar in many respects to the prior field study, the new study is required to include more robust data

[25]The Virginia Tech Transportation Institute was awarded the primary contract for this work, but subcontracted the research to Washington State University and Pulsar Informatics.

[26]Van Dongen, Hans P.A., Mollicone, Daniel J. *Field Study on the Efficacy of the New Restart Provision for Hours of Service* (2014).

[27]See the Consolidated and Further Continuing Appropriations Act, 2015. The provision provides that no appropriated funds shall be used to enforce the two-night provision (49 C.F.R. § 395.3(c)) or the 168-hour provision (49 C.F.R. § 395.3(d)). Pub. L. No. 113-235, div. K, tit. I, § 133(a), 128 Stat. 2130, 2711.

and examine several issues not discussed as part of the prior field study. For example, both studies collected or will collect PVT data to assess driver fatigue. However, the new study will collect data over a significantly longer time period (approximately two weeks of data for each driver for the prior field study versus 5 months of data for the new study) and will examine crash and driver health data that were not considered as part of the field study. FMCSA is required to provide a copy of the new study to Congress by December 2015.

The Field Study Followed Several Accepted Research Standards, but Did Not Report Its Limitations or Make Conclusions That Were Fully Linked to the Results

Based on our prior work evaluating research programs, our internal expertise in research design, and established guidelines and reports for conducting research and program evaluations, we identified six generally accepted research standards that are critical for designing, analyzing, and reporting the results of scientific research.[28] FMCSA's 2014 field study followed three of these standards, including using measures that are valid and reliable, employing quality controls, and reporting results that are supported by data (see table 1). The field study partially met the standards of using a methodology that supports the study's objectives and reporting conclusions that are linked to the results. The field study did not meet the standard of reporting methodological limitations and their impacts.

[28]To develop these standards we used *Guiding Principles for Evaluators (July 2004)*; and Thomas D. Cook and Donald T. Campbell, *Quasi-Experimentation: Design & Analysis Issues for Field Settings (1979)*, among other guidelines and reports. For more information on our sources and methodology for developing research standards, see appendix I.

Table 1: Extent to which the Federal Motor Carrier Safety Administration (FMCSA) Followed Several Accepted Standards for Designing, Analyzing, and Reporting the Results of Scientific Research in its 2014 Field Study

Standard	Field study followed?
The measures chosen for evaluation should be valid and reliable.	Yes
There should be quality controls in place to identify data inconsistencies or errors.	Yes
The results and analysis reported in the study should be supported by the data.	Yes
The methodology, including statistical techniques, should support accomplishing the objectives of the study.	Partially
The reported conclusions of a study should be linked to the results.	Partially
The study should report all methodological limitations including their potential impact on the results.	No

Source: Analysis of GAO identified generally accepted research standards and FMCSA's Field Study on the Efficacy of the New Restart Provision for Hours of Service (Jan. 2014). | GAO-15-641

The primary fatigue measure used in the field study is valid and reliable and supported by two secondary measures: According to our literature review and interviews with fatigue experts, the primary fatigue measure used in the field study—the PVT—is a reliable and valid measure of fatigue and is widely used in fatigue science research.[29] While we identified limitations on the reliability and validity of the secondary measures used in the field study—the KSS and lane deviation—we also found they are widely used in fatigue science.[30] We also discussed their use with FMCSA officials and determined that the KSS and lane deviation measures were used to support the primary findings from the PVT. Specifically, FMCSA officials stated, and we agree, that using multiple measures with different strengths and weaknesses can help to reinforce findings. In the case of the field study, each of these three measures independently supported the finding that drivers taking a one-night restart were more fatigued than those taking a two-or-more night restart.

The field study included controls to identify data inconsistencies and errors: Specifically, the field study involved a visual data check that

[29]The most commonly used method for applying the PVT is through a 10-minute test on a computer. To accommodate the commercial-trucking environment, the field study used a modified version of this test that lasted for 3 minutes on a smart phone. Our literature review found no evidence that these modifications impact the reliability or validity of the PVT.

[30]One limitation of using the KSS and lane deviation was the possibility of introducing unexplained variability into the data. For example, we identified several articles which noted that the KSS is a self-reported measure of fatigue that relies on individuals' opinions of their current fatigue. As such, individuals with objectively similar fatigue may report different scores on the KSS.

compared data for each driver across nine variables, including activity—as recorded by a wrist monitor, test results on the PVT and KSS, and GPS location.[31] Researchers used this data check to identify any obvious errors and possible contradictions in the data, such as instances when the data might indicate a driver was coded as driving and sleeping simultaneously.[32] Researchers also debriefed the study participants at the end of the study to discuss inconsistencies, mistakes, or missed entries as well as any outliers on the fatigue measures. This approach allowed researchers to account for incomplete or obviously biased data. Based on our review of the data and statistical coding, we believe that these quality controls were sufficient to ensure the data used in the field study are reliable.

The results and analysis reported in the study are supported by the data: To assess whether the results and analysis were supported by the data, we replicated several aspects of the study. First, we acquired the raw data and the computer programs used in the field study to process the data and run statistical analyses. As a check on the reliability of the computer programs, we applied the programs to the data for one driver and verified that the results were consistent with numbers reported in the field study. Because the same computer program was applied to all drivers by the authors of the field study, this check provided assurance that the analysis for all drivers was reliable. Second, using data files provided by the authors, we replicated demographic information reported in the study, including average age, years of experience, and sex of the driver, among others. Similarly, using data files provided by the authors, we replicated the main study results. For example, our replication showed that the study accurately reported the number of reaction time errors or "lapses" on a PVT test for drivers following a one-night restart and a two-or-more night restart.[33]

[31]The visual data check was based on a graphical representation of these nine variables that used lines and other symbols to identify data points at specific times of each day.

[32]FMCSA contracted with the Virginia Tech Transportation Institute, Washington State University, and Pulsar Informatics—a technology development and research company—to complete the field study. When we refer to the researchers or authors of the study, we mean employees of these organizations.

[33]A PVT lapse occurred when drivers did not react to visual stimuli, e.g., flashing lights, on a smart phone within 355 milliseconds.

The methodology used in the field study partially supports the study's objectives: The field study examined drivers in a naturalistic environment, driving their normal routes, and collected fatigue related data in a way that was consistent with the study's objectives. The study collected data on sleep patterns, fatigue, and driving behavior using wrist monitors, smartphones, and sensors attached to truck equipment. These methods did not require drivers to interact repeatedly with researchers or change their driving behaviors in response to study requirements, which otherwise might have biased measurements or prevented the results from applying to real world conditions. In addition, we determined the statistical method used to evaluate the study's findings—mixed-effects regression modeling—was appropriate.[34] Mixed effects modeling is an accepted method of analyzing longitudinal data, which, in this study, involved the same drivers taking the PVT at various times of day over several weeks. In addition, the models reasonably allowed for the possibility that the rule's effects could vary according to the time of day and period of observation.

However, we did find one area in which the methodology used did not support the study's objectives. Specifically, the purpose of the study was to examine the efficacy of the 2011 HOS rule, which required drivers choosing to take a restart include two-nighttime periods (from 1 a.m. to 5 a.m.). This rule was expected to reduce fatigue for drivers that previously took restarts with one-nighttime period as part of their restart. Therefore, the study should have been designed to compare the ability of drivers to perform fatigue tests, e.g., the PVT, after a one-night or a two-night restart. However, the field study actually compared drivers with one-night restart breaks to those with two or more nights of rest. One concern with including drivers taking more than two nights rest is the possibility of biasing the results. For example, drivers with more rest could be less fatigued and have fewer lapses on a PVT test. Thus, including such drivers does not directly support the objective of the study to determine the impact of the changes to the restart provision from a one-night to a two-night restart. According to FMCSA officials, the decision to include drivers taking more-than-two night restarts was a judgment call made by researchers, and was likely due to the timeframe mandated by law for completing the study.

[34]The authors of the field study used a different term, mixed-effects analysis of variance (ANOVA), for this statistical method.

Following the study's issuance FMCSA officials stated that they conducted additional analysis and determined that the inclusion of drivers with a two-or-more night restart did not influence the results. We conducted a similar analysis and also found that the inclusion of drivers with two-or-more night restarts did not bias the results. Specifically, we found that the number of on-duty, PVT lapses for drivers who took a two-night restart was virtually the same as those taking a two-or-more night restart.[35] While we did not identify any bias, the decision to include drivers with more than two-night restarts was not appropriate given the objectives of the study.

<u>The results of the field study support the report's conclusions about fatigue, but conclusions about crash risk may be overstated:</u> The study found a statistically significant difference across several measures between drivers with a one-night restart and those with a two-or-more night restart. Specifically, drivers with a one-night restart had more lapses on the PVT, rated themselves as more sleepy on the KSS, and had greater lane deviation than those with a two-or more night restart.[36] FMCSA used these results to support two points in the conclusions section of the study: (1) the two-night provision mitigates fatigue, and (2) the two-night provision supports the efficacy of the restart rule and has implications for real-world driving performance, road safety, and crash risk. As discussed above, our review of relevant research and discussions with fatigue experts shows that the primary measure used in the field study—the PVT—is a reliable and valid measure of fatigue and is supported by the KSS and lane deviation, which are widely used in fatigue science. Therefore, the evidence supports FMCSA's conclusion in the study that the two-night provision mitigates fatigue.

Regarding the implications of the study for real-world driving performance, road, safety, and crash risk, FMCSA reported that on-duty drivers taking a one-night restart had 2.0 lapses per PVT test, while on-duty drivers taking a two-or-more night restart had 1.7 lapses per PVT

[35]The study found that drivers with a one-night restart had .33 more lapses per PVT test during duty cycles than drivers with a two-or-more night restart, while our analysis found that drivers taking only a two-night restart had .32 lapses more lapses per PVT bout. The results from the study and our analysis are both statistically significant. For more information on our analysis, see appendix II.

[36]Lane deviation, as defined in the field study, could mean lateral movement within a lane and did not imply crossing a line into another lane.

test—an average difference of 0.3 lapses. In addition, FMCSA found a larger difference in the number of lapses between drivers taking a one-night versus a two-or-more night restart when operating at night—an average difference of 0.8 lapses per PVT test.[37] FMCSA used these results to support its conclusions that this number of additional lapses, i.e., 0.8 more lapses, has "implications for real-world driving performance, road safety threat detection, evasive maneuvering and braking response speed and crash risk." Further, FMCSA also provided a hypothetical example of the distance a truck could travel during a single lapse, noted that such inattention may mean the driver is not seeing critical information—such as road signs, traffic signals or other vehicles—and concluded that "an increase in lapses of attention increases crash risk."

The conclusion of the report—which implies a direct relationship between the study's results and crashes—may overstate the findings and erroneously imply that this study provides evidence of reduced crash risk. As we discuss later in this report, while there is evidence that fatigue and crashes are generally related, we found little research that attempted to quantify this relationship using any of the fatigue measures in the field study. In particular, through our review of literature and conversations with fatigue experts, it was uncertain the extent to which, if any, a difference of 0.3 or 0.8 lapses on the PVT might be related to on-road safety outcomes.[38] According to FMCSA officials, its study does not directly link results to crash risk due to challenges associated with

[37]As we discussed earlier, a PVT lapse occurred when drivers did not react to visual stimuli within 355 milliseconds.

[38]More specifically, the safety implications of a difference between roughly 1 lapse per PVT and roughly 2 lapses—which was the difference observed in the field study between drivers who had taken a 2-or-more-night restart and those who had taken 1 night—was uncertain. We did find one study that used performance on a driving simulator to quantify the relationship between the KSS and on-road safety. Results of this study suggest that the KSS difference observed in the field study (about 3.1 after a two-or-more-night restart versus about 3.3 after a one-night restart) would be associated with minimal difference in crash risk. Ingre, Michael, Torbjőrn Ǻkerstedt, Bjőrn Peters, Anna Anund, Gőran Kecklund, and Andrew Pickles. Subjective sleepiness and accident risk avoiding the ecological fallacy," *Journal of Sleep Research,* 15 (February 2006). We also reviewed one study that attempted to directly relate performance on the PVT to performance on a driving simulator. This study concluded that the PVT has limitations as a predictor of driving performance and that further research is needed to develop more complex measures. See Baulk, S. D., S.N. Biggs, K.J. Reid, C.J. van den Heuvel, and D. Dawson (2008), Chasing the silver bullet: Measuring driver fatigue using simple and complex tasks. Accident Analysis and Prevention, 40, 396-402.

performing such a study. Instead, FMCSA used lane deviation as a proxy measure for crashes and driving performance.[39] In addition, FMCSA officials stated that, since fatigue is connected to crash risk, increases in fatigue, such as those shown in the field study, are associated with crash risk. While we agree that evidence generally supports that fatigue and crash risk are related, we are uncertain how fatigue differences of the size reported in the field study would be associated with crash risk. Thus, the safety implications and policy importance of the study's estimated effects on fatigue may be overstated.

The field study did not report all of its methodological limitations, including how these limitations might have affected the results: As with any research effort, FMCSA made methodological choices to achieve the goals of the study under the requirements outlined in MAP-21. For example, according to FMCSA, many of the decisions on recruiting drivers to participate were dictated by the short time frame under which FMCSA was required to complete the study—approximately 9 months—and the availability of resources. After considering several of these methodological choices that we reviewed, such as the sample size used, they appear to be reasonable given these constraints. However, by not reporting how certain decisions may have limited the study and impacted the results, FMCSA did not follow accepted research standards, as discussed below:

- *The field study did not describe its driver recruitment process.* According to FMCSA officials, participants in the study were recruited from three carriers that performed several different types of operations (i.e., local, regional, and over-the-road). One of these motor carriers was selected because of a prior relationship it had with FMCSA and the other two carriers were chosen after being identified as relatively safety conscious. A third-party data provider used logbook data from these carriers to identify drivers that met several criteria—including working near the maximum allowable weekly hours—which FMCSA used to identify specific operating locations that were likely to employ drivers meeting the study's requirements. By not including information on the recruiting process in the field study, FMCSA's study prevents

[39]The field study found that on-duty drivers with a one-night restart had 0.1 centimeter greater lane deviation than those taking a two-or-more night restart. As noted earlier, lane deviation meant any lateral movement, including movement that was entirely within a lane; it did not imply crossing a line into another lane.

readers from assessing potential bias in the collection of data and invites skepticism of the recruitment process. For example, FMCSA did not disclose that it recruited a substantial proportion of its sample of drivers with a one-night restart from a single, large U.S. city, according to one of the carriers selected to participate in the field study. Because these drivers likely operate in a similar environment, such a sample raises questions about whether these drivers' measured fatigue levels after one-night restarts might have been related to other factors with that location. Such other factors might include characteristics of the local traffic or trucking operations. According to FMCSA officials, they recruited drivers in this manner because they needed to complete the work in a relatively short time frame—approximately 9 months—and had difficultly identifying drivers likely to take a one-night restart. While this method of recruiting drivers does not necessarily bias the study's results, it has the potential to do so and should have been reported by FMCSA.

- *The field study did not report the implications of its sample size.* The field study collected data on a sample of drivers that produced statistically significant results. However, the sample size was not sufficient to address questions that have been raised about how the restart provision affects the diverse commercial-trucking industry. Several stakeholders we spoke with said that the restart provision likely has different effects across industry segments, including local, regional, and over-the-road drivers. These stakeholders said that the restart provision may more successfully reduce fatigue for over-the-road drivers who are away from home for weeks at a time than for drivers who operate locally and sleep in their own beds every night.

  We analyzed the study's data to assess the implications of its sample size. We found that, while the field study included a diverse population of drivers from industry segments, its sample did not include enough drivers from each of these segments to estimate differences in PVT at all or to identify statistically significant differences.[40] In addition, we assessed the sample size that might have been required to estimate significant differences across industry segments. We found that the field study's sample size was insufficient to estimate significant differences in the primary fatigue measure—the PVT—for each of

[40]We assessed the feasibility of making separate estimates for three operation subgroups: local, regional, and over-the-road.

several segments of the industry.[41] According to FMCSA officials, decisions on the sample size for the field study were made to accommodate the deadline for when it must be completed. Given the importance of understanding how the results of the field study might have varied by segment, FMCSA should have stated how the study's sample size limited the generality of its findings and conclusions.

- *The study did not report how a critical analytical decision might impact the results.* As was discussed above, the field study compared drivers with one- versus two-or-more night restart breaks. According to accepted research standards, FMCSA should have reported why drivers taking a more than two-night restart were included in the study, even though the rule should have been designed to compare ability of drivers to perform fatigue tests, e.g., the PVT, after a one-night or a two-night restart. Furthermore, FMCSA should have reported the potential for this decision to bias the results and, in particular, to exaggerate the difference in fatigue between the two groups of drivers.

According to FMCSA officials, researchers followed several standards when developing and conducting the field study, including Office of Management and Budget's (OMB) guidelines on the use of independent peer review panels and the DOT Scientific Integrity Policy. The OMB guidelines FMCSA used require agencies to use a peer review process for important scientific information to enhance its quality and credibility. FMCSA officials reported using the guidelines to evaluate the research design prior to implementation as well as to assess the report findings. In addition, DOT's Scientific Integrity Policy lays out nine elements that primarily detail the importance of communication, transparency, and integrity when using scientific information for decision making. While these standards address several important considerations, including enhancing credibility and transparency, they do not provide specific guidance on how to determine appropriate methods for designing, analyzing, and reporting the results of scientific research. Without such guidance, FMCSA research may not include critical elements, as occurred in the reporting of the field study. Without these elements, FMCSA leaves itself vulnerable to criticism over the integrity of its

[41]For more information on this "power analysis," including our methodology and results, see appendix II.

research—an important consideration, given the heightened profile of HOS regulations.

FMCSA Anticipated Several Effects of the 2011 Hours of Service Rule, and While Available Data Provide Some Insight, Data Limitations Hinder the Ability to Fully Assess the Rule's Effects

FMCSA used several key assumptions to estimate the costs and benefits of the 2011 HOS rule.[42] For example, FMCSA assumed that the rule would result in schedule adjustments for approximately 15 percent of drivers who work more than 65 hours per week on average.[43] These assumed schedule adjustments drove FMCSA's estimates of the economic costs and the health and safety benefits of the 2011 HOS rule. Our analysis of available data provides some insights into the potential effects of the 2011 HOS rule. For example, we found that about 12 percent of drivers in our dataset worked an average of over 65 hours per 8 day work week before the rule went into effect.[44] The lack of nationally representative driver schedule data limits the ability of researchers to fully evaluate the rule's impact. Although soon all carriers will be required to collect schedule data electronically and these data could be useful for evaluating the effects of this rule and in formulating future rules, current legal restrictions limit how these data can be used.[45]

[42]Executive Order 12866, "Regulatory Planning and Review" directs federal agencies to assess the potential costs and benefits of their significant regulatory actions when certain criteria are met. 58 Fed. Reg. 51735 (Oct. 4, 1993). OMB Circular A-4 provides guidance to federal agencies on the development of these regulatory analyses, which states that in the absence of adequate data, agencies will need to make assumptions and provides guidance for their use.

[43]DOT, FMCSA, *2010-2011 Hours of Service Rule – Regulatory Impact Analysis*, RIN 2126-AB26 (December 2011).

[44]Throughout this report, we describe hours worked per week. As explained in the background section, drivers may work according to the 60-hour/7-day limit or the 70-hour/8-day limit depending on the weekly operations of their employer. In our analysis of driver schedule data, we used an 8-day work week because the majority of drivers in our dataset likely operate on an 8-day schedule. For comparison purposes, we also analyzed the dataset using a 7-day work week. Using a 7-day work week, less than one percent of the drivers in our dataset worked more than 65 hours per week and no driver worked more than 75 hours per week. Other analyses using similar driver schedule data conducted by the American Transportation Research Institute used a 7-day work week and found similar results, specifically less than one percent of drivers working more than 65 hours per week. For more information on our analysis of driver schedule data, see appendix IV.

[45]FMCSA expects to publish the final rule on Electronic Logging Devices and HOS supporting documents on September 30, 2015.

FMCSA Used Several Key Assumptions to Estimate the Potential Effects of the Rule

As part of its rule-making process, FMCSA developed assumptions about the motor carrier industry and its operations to estimate the economic costs and the safety and health benefits of the 2011 HOS rule.[46] For example, FMCSA assumed that drivers working more than 65 hours per week would incur the majority of schedule changes as a result of the rule.[47] FMCSA estimated that such drivers made up approximately 15 percent of the driver workforce. FMCSA then made assumptions about how drivers would likely be affected by the rule to estimate a weekly per driver work time reduction of 0.05 hours (3 minutes) for drivers working an average of 60 hours per week; .38 hours (23 minutes) for drivers working an average of 70 hours per week; and 8.7 hours for drivers working an average of 80 hours per week. To estimate the economic costs of the rule, FMCSA calculated the cost of this lost work time and estimated $430 million in annual costs.[48] Further, FMCSA assumed that reduced work time would increase a driver's opportunity to sleep leading to safety and health benefits. FMCSA estimated the rule would save between 10 and 26 lives per year and result in $150 to $390 million in annual safety benefits. FMCSA also estimated between $70 million and $850 million in health benefits.[49] (See figure 3).

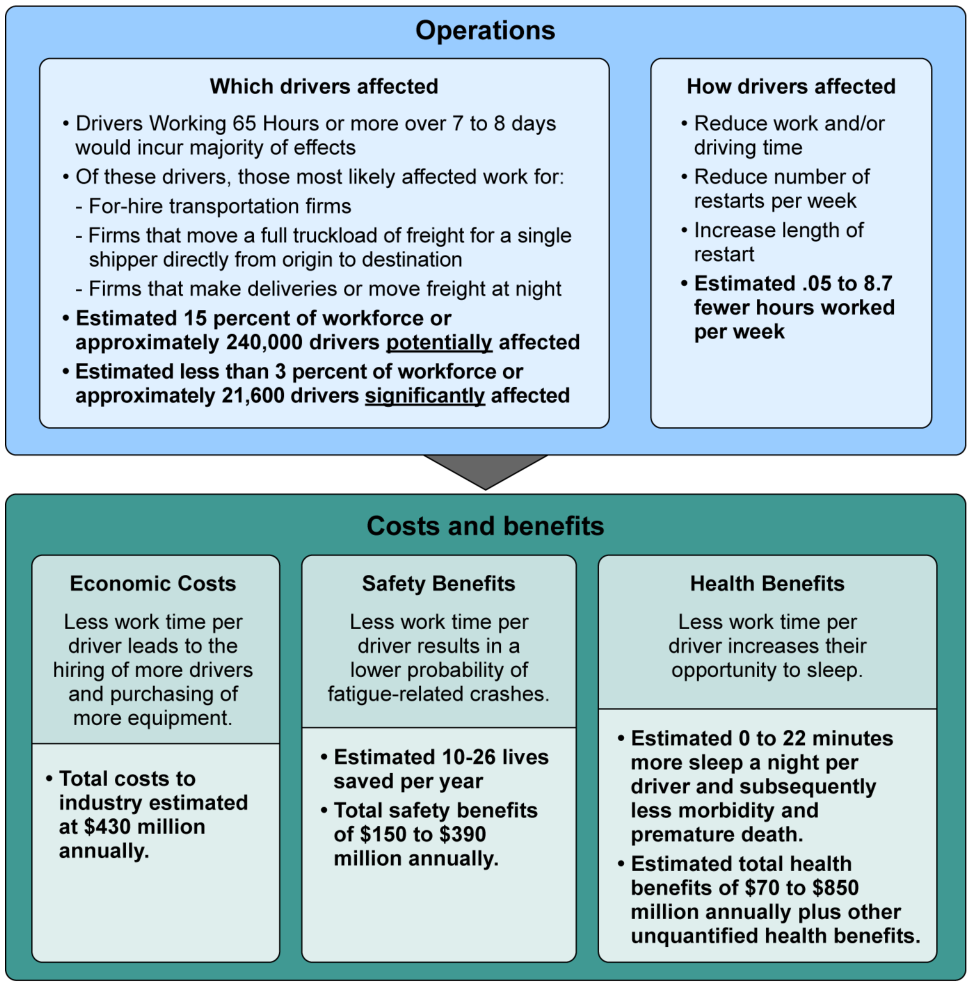
[46]The assumptions made by FMCSA are outlined in its regulatory impact analysis (RIA) for the 2011 final HOS rule. DOT, FMCSA, *2010-2011 Hours of Service Rule – Regulatory Impact Analysis*, RIN 2126-AB26 (December 2011)/ We identified key assumptions in this analysis and confirmed them with FMCSA officials. For the full list of assumptions and estimates we identified and confirmed by FMCSA officials as important to the cost and benefit analysis, see appendix III.

[47]As part of its RIA, FMCSA put drivers into four driver intensity groups based on their weekly on-duty hours: *moderate* with an average of 45 hours including drivers with weekly averages from 20 to 55 hours; *high* with an average of 60 hours including drivers with weekly averages from more than 55 to 65 hours; *very high* with an average of 70 hours including drivers with weekly averages from more than 65 to 75 hours; and *extreme* with an average of 80 hours including drivers with weekly averages greater than 75 hours.

[48]While the primary economic costs are those associated with less driver work time, FMCSA also assumed economic costs due to training and reprogramming as drivers and carriers adjusted to the rule. These costs include training drivers on the new rule, updating software programs, and other transition costs.

[49]FMCSA estimated health benefits assuming different baseline levels of sleep for drivers: low, medium, and high. Assuming drivers have low baseline levels of sleep results in higher estimated health benefits from the rule. Combined with FMCSA's use of two discount rates, 3% and 7%, to estimate the future health benefits, these differing assumptions result in a range in potential health benefits. All estimated costs and benefits amounts are in 2008 dollars.

Figure 3: Key Assumptions and Estimates Used by the Federal Motor Carrier Safety Administration (FMCSA) to Determine the Costs and Benefits of the 2011 Hours of Service Rule



Source: GAO identification of key assumptions in FMCSA's 2010-2011 Hours of Service Rule Regulatory Impact Analysis | GAO-15-641

Analysis of Available Data Provides Some Insight into the Effects of the Rule

Analyzing available data can help provide some insight into the rule's potential effects and the extent to which they aligned with FMCSA's assumptions and estimates.[50] Specifically, we conducted a comprehensive search for available data that might provide insight into the rule's effects in the five main areas in which FMCSA made assumptions: (1) which drivers were affected, (2) how drivers were affected, (3) economic costs, (4) safety benefits, and (5) health benefits. The results from our analysis of schedule data from 16 for-hire carriers is consistent with FMCSA's assumption that drivers working more than 65 hours per work week would be more likely to reduce their work hours. We also found support for FMCSA's assumption that certain schedule changes could reduce driver fatigue and while total crashes, and crashes with injuries did not appear to change, crashes involving fatalities may have declined after the rule went into effect. However, we were unable to assess, due to a lack of data, FMCSA's assumptions on economic and health effects. All of the data and analyses we used had limitations, including the extent to which the data are representative of the motor carrier industry and the extent to which we can attribute observed differences to the rule change. In addition, we could not fully analyze the possible effects of the 2011 HOS rule due to the relatively short time frame during which the rule was in place—approximately 18 months. (See table 2).

[50]As discussed later in this report, other data that are currently unavailable would be required to better analyze the effect of the rule.

Table 2: Summary of GAO's Findings on the Effects of the Hours of Service Rule (July 2013 to December 2014)

Effect area	Summary of results	Limitations
Which drivers were affected	• Drivers working the longest hours, over 65 hours per 8-day work week, likely affected • Drivers working less intense schedules, 35 hours up to 65 hours per 8-day work week, also likely affected • Motor carriers reported a wide variety of their operations were potentially affected	• Findings are not representative of the motor carrier industry and are not generalizable • Observed changes in behavior cannot be attributed to the rule change
How drivers were affected	• Drivers reduced their weekly hours • Restart use declined • Traffic volume shifted from night time hours to other parts of the day • Motor carriers reported a range of operational effects or changes attributable to the rule, including decrease in driver hours, less frequent restart use, and more driving during period of congestion, among others	• Findings are not representative of the motor carrier industry and are not generalizable • Shifts in traffic volume based on limited data and cannot be directly attributed to the rule change • Observed changes in behavior cannot be attributed to the rule change
Economic costs	• There are no industry-wide data to assess the economic effects of the rule • Motor carriers, drivers, and other industry stakeholders reported negative economic impacts, including productivity loss, decrease in delivery timeliness, less profitability, and increased driver turnover	• Results are based on interviews with selected motor carriers and drivers and are not generalizable
Safety benefits	• Fatigue can have a negative impact on driver performance and can lead to safety critical events, including crashes. • Drivers following the rule could have lower risk of fatigue • We found no change in the number of total crashes, or crashes with injuries, however the number of crashes involving fatalities may have declined after the rule went into effect • There was no statistically significant change in the likelihood of truck crashes occurring between 5 a.m. and 9 a.m. after the rule went into effect	• Fatigue analysis is based on simulated schedules, is not representative of the motor carrier industry, and is not generalizable • Crash analysis is based on only 15 months of data • Observed changes in behavior cannot be attributed to the rule change
Health benefits	• There are no data available to assess the health effects of the rule • Literature suggests that insufficient amounts of sleep have negative health impacts, including obesity, diabetes, hypertension, and cardiovascular disease • Motor carriers and drivers reported no noticeable positive health effects from the rule • Safety groups believed the rule could improve driver health, but those impacts would be difficult to attribute to the rule change	• Results are based on literature and interviews with selected motor carriers and drivers and are not generalizable

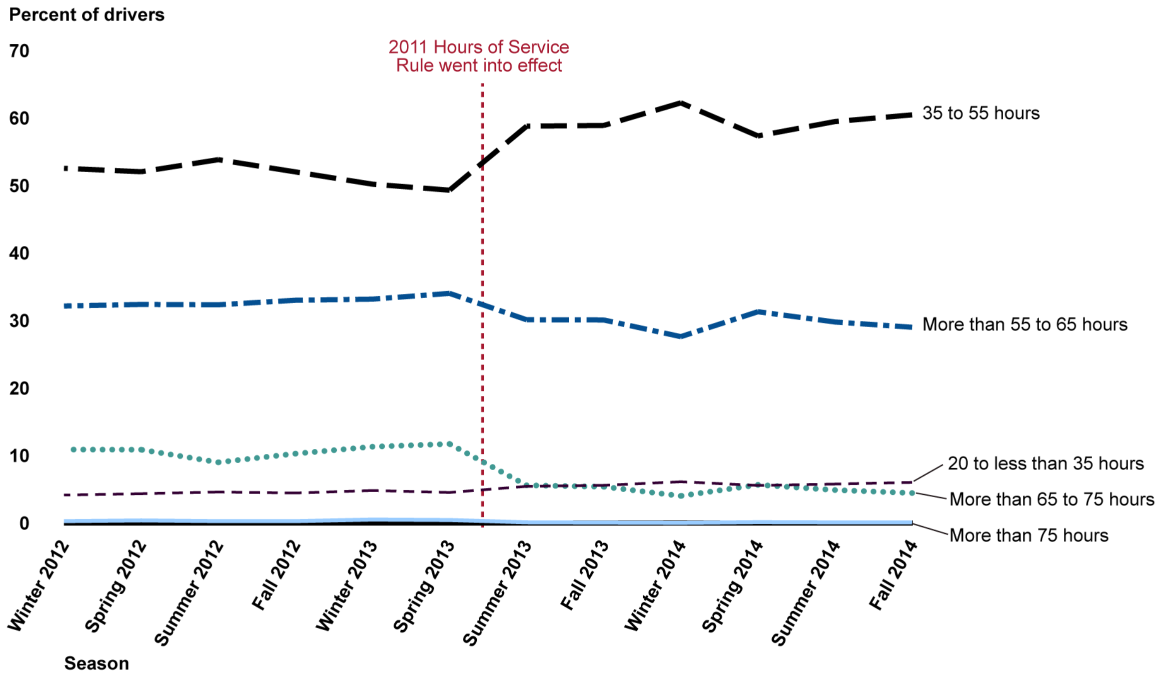
Source: GAO. \| GAO-15-641

Which drivers were affected: Our analysis of 2012 to 2014 driver schedule data from 16 for-hire motor carriers is consistent with FMCSA's expectation that the percentage of drivers working the longest hours—over 65 hours per 8-day work week—decreased after the rule went into effect on July 1, 2013.[51] Specifically, we found that the percentage of drivers in our dataset working over 65 hours per 8-day work week decreased from 12 percent immediately before the rule went into effect to 6 percent afterwards.[52] Our analysis also found that the percentage of drivers working between 55 and 65 hours per 8-day work week decreased from 33 percent to 29 percent. In contrast, the percent of drivers working 35 to 55 hours per 8-day work week increased from 48 percent to 57 percent after the rule went into effect. (See figure 4). However, the driver schedule data used in our analysis are not representative of the entire motor carrier industry and do not include any information on drivers who work in some segments of the trucking industry, such as private motor carriers. Therefore, our findings cannot be generalized to the motor carrier industry as a whole.

[51]We purchased these data from the American Transportation Research Institute. The data include observations on more than 15,000 drivers working for 16 for-hire motor carriers. These data are not publicly available. The data measure a driver's total recorded on-duty hours for each day in our dataset, but not the time of day when the driver began or ended his or her on-duty hours. As such, we could not analyze whether a recorded restart included one or two nights. However, we did analyze the length of restart periods, which provides limited insight into how drivers used restarts in our dataset before and after the 2011 HOS rule went into effect. For more information on our analysis of schedule data and additional results, see appendix IV.

[52]As discussed earlier in this report, a driver's work week may span a 7- or 8-day period. While we analyzed the driver schedule data using both 7- and 8-day weeks, the results reported in this section use 8-day work weeks. In addition, our analysis uses drivers' average work week to calculate the percent of drivers' working certain average weekly hours. We also analyzed work time using the sum of each work week and categorized each 8-day work week into a weekly on-duty hours category. Using an 8-day work week, we found that before the 2011 HOS rule went into effect about 24 percent of the weeks worked in our dataset had a sum of greater than 65 hours per week and around 5 percent of the weeks worked had a total on-duty hour sum of greater than 75 hours. This finding demonstrates that while a driver may on average work less than 65 hours per week, he/she may have work weeks of greater intensity (e.g., more than 65 hours per week).

Figure 4: Percentage of Drivers by Average Weekly On-Duty Hours before and after the 2011 Hours of Service Rule for 16 For-Hire Motor Carriers (2012–2014)



Source: GAO analysis of driver schedule data from 16 For-Hire Motor Carriers. | GAO-15-641

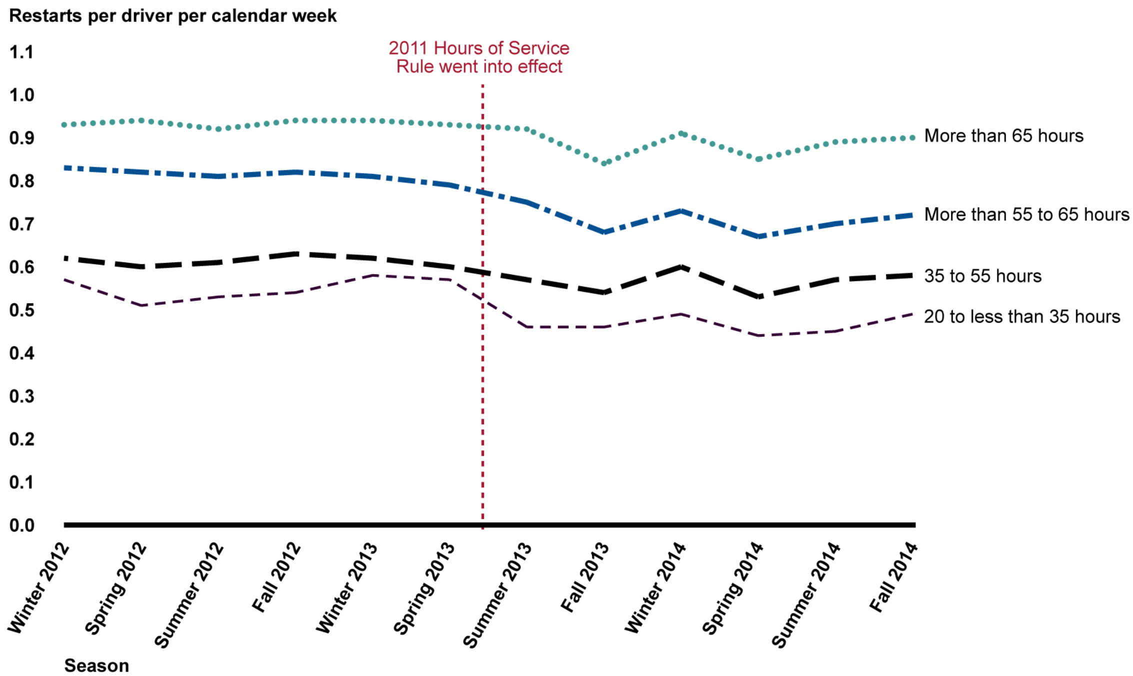
Note: For the purposes of this analysis, a work week was defined as 8 consecutive days because the majority of drivers in our dataset likely operate on an 8 day schedule.

In addition, our interviews with representatives from 20 motor carriers suggest that a wide variety of motor carrier operations were potentially affected by the HOS rule. Specifically, 13 motor carriers we spoke with, including 10 for-hire carriers and three private carriers, stated they were affected by the rule. These carriers described themselves as truckload carriers, less-than-truckload carriers or a combination of both and employed drivers working both more than and less than 65 hours per week. In addition, seven motor carriers reported they were not affected by the HOS rule, including three for-hire carriers, three private carriers, and one carrier that used a truck on a part-time basis and did not fall into either category.

How drivers were affected: Our analysis of schedule data from 16 for-hire carriers supports the assumption that some drivers may be working fewer hours and using restarts less often. Specifically, as shown above in figure 4, these schedule data show that fewer drivers were working more than 55 hours per 8-day work week after the rule went into effect in July 2013.

We also estimated that drivers, on average, worked approximately 1 to 2 fewer hours per work week after the rule went into effect.[53] In addition, the number of restarts drivers took per calendar week (168 hours) generally declined after the rule went into effect, including for drivers working more than 65 hours per 8-day work week (see fig. 5). We also found that drivers took approximately 6 percent fewer restarts per week, on average, after the rule went into effect.

Figure 5: Restarts per Driver per Calendar Week before and after the 2011 Hours of Service Rule For 16 For-Hire Motor Carriers (2012–2014)



Source: GAO analysis of driver schedule data from 16 For-Hire Motor Carriers. | GAO-15-641

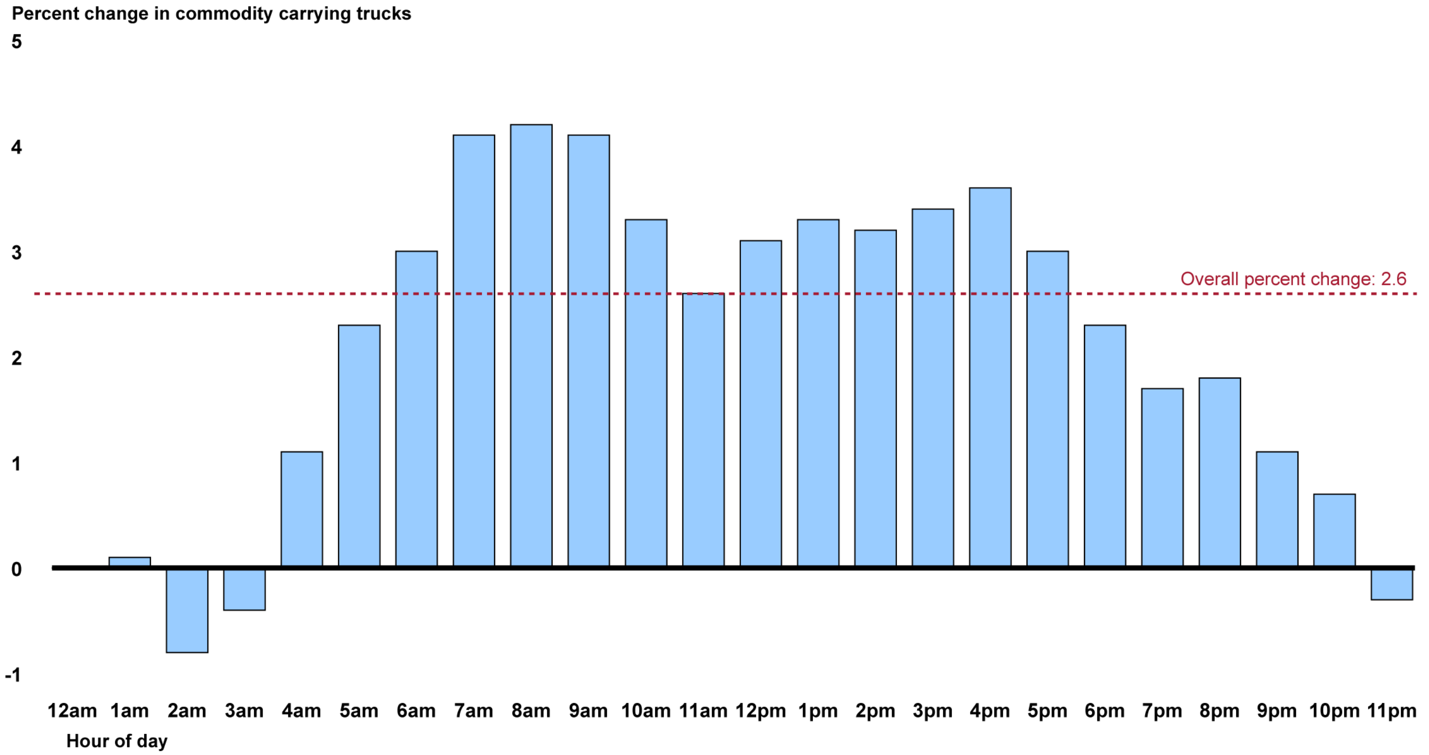
Note: For the purposes of this analysis, a work week was defined as 8 consecutive days because the majority of drivers in our dataset likely operate on an 8-day schedule. Restarts greater than 72.99 hours were excluded.

[53]Our analysis of driver's average hours worked per work week was based on a subset of the over 15,000 driver records per year included in our dataset. Specifically, we analyzed schedule data for the approximately 7,000 drivers who were included across all 3 years of our data set to identify how individual drivers had been impacted by the changes to the 2011 HOS rule. Using statistical methods, we found that the differences in on-duty hours (1 to 3 fewer hours per week) and restart use (6 to 7 percent fewer restarts per week) in these drivers' schedules before and after the rule went into effect were statistically significant.

We also found evidence supporting a shift in traffic volume from night time hours to other parts of the day following the effective date of the HOS rule. These shifts cannot be directly attributed to the rule change and the data we analyzed from the Federal Highway Administration were limited to 14 states. Our analysis showed that commodity carrying, i.e., commercial, motor vehicle traffic volume increased by 2.6 percent between November and December 2012 and November and December 2013.[54] However, as figure 6 shows, this increase was not consistent across the time of day. Specifically, commodity carrying motor vehicle traffic mostly increased during daytime hours (6 a.m. to 6 p.m.) and decreased during evening and nighttime hours (6 p.m. to 6 a.m.), relative to the increase in overall commodity carrying vehicle traffic. The largest differences from the growth in traffic volume occurred in the middle of the night (11 p.m. to 4 a.m.), when volume declined or remained flat despite an increase in commodity carrying vehicle traffic overall and in the mid-morning hours (7 a.m. to 10 a.m.) when volume grew faster than the overall growth rate.

[54]We selected these time periods for several reasons, including controlling for the possible effect of seasonality. For more information on our analysis of vehicle count data and additional results, see appendix VI.

Figure 6: Percentage Change in Commodity Carrying Vehicle Volume after the 2011 Hours of Service Rule, by Hour of the Day (November and December, 2012–2013)



Source: GAO analysis of Federal Highway Administration data. | GAO-15-641

Note: Our analysis was restricted to 324 stations from 14 states out of 987 total stations across the country. We cannot attribute changes in traffic volume to the 2011 hours of service rule.

Consistent with our data analysis, representatives from 13 of the 20 motor carriers we interviewed reported a range of operational effects or changes they attributed partially or entirely to the 2011 HOS rule. The other seven motor carrier representatives did not report any operational effects due to the rule. For example, motor carriers reported a decrease in driving hours per driver (10 carriers), less frequent restart use (10 carriers), more driving during periods of traffic congestion (8 carriers), hiring more drivers (6 carriers), and an overall decrease in operational flexibility (10 carriers). Several of these operational changes, such as a decrease in driving hours per driver and subsequent hiring of more drivers to make up for reduced driver productivity, could lead to more commercial vehicles on the road.

Economic Costs: We interviewed motor carriers and industry stakeholders about the economic impacts of the 2011 HOS rule but did not find a data source that would allow us to reliably analyze these impacts for the industry. Representatives of 11 out of the 20 motor carriers we interviewed mentioned negative economic impacts due to the 2011 HOS rule. Ten motor carriers reported a productivity loss, six reported a decrease in delivery timeliness, six reported less profitability, and four reported an increase in driver turnover.[55] The other nine motor carriers we interviewed reported no negative economic effects. In addition, four of the eight drivers we spoke with reported a decrease in their income due to the 2011 HOS rule. These drivers reported that reduced work time—fewer or shorter work shifts—was the reason for their loss in income. The four drivers that reported no decrease in pay either did not have a schedule change or were working more hours per week to compensate for schedule changes due to the rule. Other industry stakeholders we interviewed, including 12 out of 13 associations representing segments of the motor carrier industry, shippers, and warehouse operators, also reported negative economic impacts on their respective industries due to the 2011 HOS rule, such as reduced productivity. For example, some industry stakeholders told us their members must hire more drivers to do the same amount of work and keep more inventory in stock to mitigate delays and delivery uncertainty.[56]

Safety Benefits: According to academic literature we reviewed, fatigue can negatively impact the ability to perform tasks. Fatigue, as a result of being awake for long periods of time or insufficient sleep, can have a negative impact on driver performance and can lead to safety critical events, including crashes. However, the literature is less clear on which HOS interventions (e.g., rest breaks and limiting driving and on-duty time) best minimize fatigue and reduce crash risk. For example, research has found that crash risk is higher when drivers start their shifts or after taking

[55]Productivity is generally measured in terms of output obtained for a given level of input. For the trucking industry, productivity can be measured in terms of cost per mile or per ton-mile, labor hours required to deliver a cargo shipment, revenue per tractor, revenue per employee, and other measures.

[56]Industry stakeholders said their members adjusted, as needed, to the rule when it went into effect.

an extended break; but research also shows that drivers who do not take breaks have a higher crash risk than drivers who do take breaks.[57]

Our analysis suggests that drivers' operating after the HOS rule went into effect on July 1, 2013, could have a lower risk of fatigue.[58] To assess the extent to which the 2011 HOS rule might affect driver fatigue, we conducted two separate analyses using a biomathematical fatigue model.[59] First, we used the model to compare hypothetical driver schedules that complied with two of the 2011 HOS rule provisions—the 168-hour limit and the two-night provision—to similar schedules that did not comply with the rule but were in compliance with the previous HOS rule.[60] We found that some schedule changes that would be required after the rule went into effect would result in lower fatigue scores and,

[57]Richard J Hanowski, Jeffrey S. Hickman, Rebecca L. Olson, and Joseph Bocanegra, "Evaluating the 2003 revised hours-of-service regulations for truck drivers: The impact of time-on-task on critical incident risk," *Accident Analysis and Prevention* 41, no. 2 (2009): 268-275.; K. Wu and Paul Jovanis. "Effect of Driving Breaks and 34-Hour Recovery Period on Motor Carrier Crash Odds," *Driving Assessment* (2011): 606-613.; Chen Chen and Yuanchang Xie, "Modeling the safety impacts of driving hours and rest breaks on truck drivers considering time-dependent covariates," *Journal of Safety Research* 51 (2014): 57-63.

[58]Without representative driver schedule data we could not assess how drivers' changed their schedules in response to the HOS rule and whether these changes resulted in a lower risk of driver fatigue.

[59]The model we used, the Fatigue Audit InterDyne™ (FAID) model, calculates scores based on hours worked. These scores indicate the likelihood of impairment due to fatigue—the higher the score, the higher the likelihood an individual working a modelled schedule could be fatigued. FAID accounts for the duration of work time and breaks, when during the day work and breaks occur, work history over the past 7 days to measure accumulated fatigue, and the biological limits on recovery sleep (e.g. humans sleep longer and have sleep of higher quality at night). FAID and other biomathematical models are used by government regulators, industry, and researchers to assess possible fatigue risks associated with work schedules. In addition, the FAID model has been independently validated by the Department of Transportation. For more information on the model we used and the results of our analysis, see appendix V.

[60]Creating hypothetical schedules allowed us to test the protective effects of the rule for schedules that may appear extreme. While we do not know how many drivers regularly work the maximum-allowed hours on a day or night schedule, these schedules are theoretically possible.

therefore, a lower risk of driver fatigue.[61] Specifically, the hypothetical schedules we modelled that had to change after July 1, 2013, in order to comply with the rule were:

- *Maximum Allowed-Hours Day and Night schedules: 14-hour shifts over 5 consecutive days or nights.* Drivers following a daytime schedule of 14-hour shifts would have to modify their schedules to comply with the 168-hour limit. Specifically, they could no longer take a restart with less than 168 hours (7 days) between when they began their last restart period and the beginning of the next restart period. To comply with the 168-hour limit, drivers following this daytime schedule must take 2 days off instead of one day off between work cycles. For night drivers working 14-hour shifts, working 5 consecutive days with one day off for a restart would not comply with either the 168-hour limit or the two-night provision. These drivers must also take 2 days off instead of one day off between work cycles. The longer restart period (2 days versus one day) taken by these hypothetical day and nighttime drivers working the maximum hours allowed each day results in lower fatigue scores and, therefore, lower risk of driver fatigue over the five day work cycle.

- *Sixty- and 70-Hour Night Schedules: 10- to 12-hour shifts over six consecutive nights.* Drivers working through the night but working less than the maximum allowed 14 hours each day would also have to change their schedules after July 1, 2013. Specifically, these drivers would have to take 2 days off between work cycles in order to comply with the two-night provision. Again, the longer restart period (2 days versus one day) results in the hypothetical driver having lower fatigue scores and, therefore, a lower risk of fatigue over the 6-day work cycle.

Our second analysis using the biomathematical fatigue model estimated possible fatigue effects of the 2011 HOS rule based on information provided by motor carriers, industry stakeholders, and drivers we interviewed. We found that some of the real world schedule changes

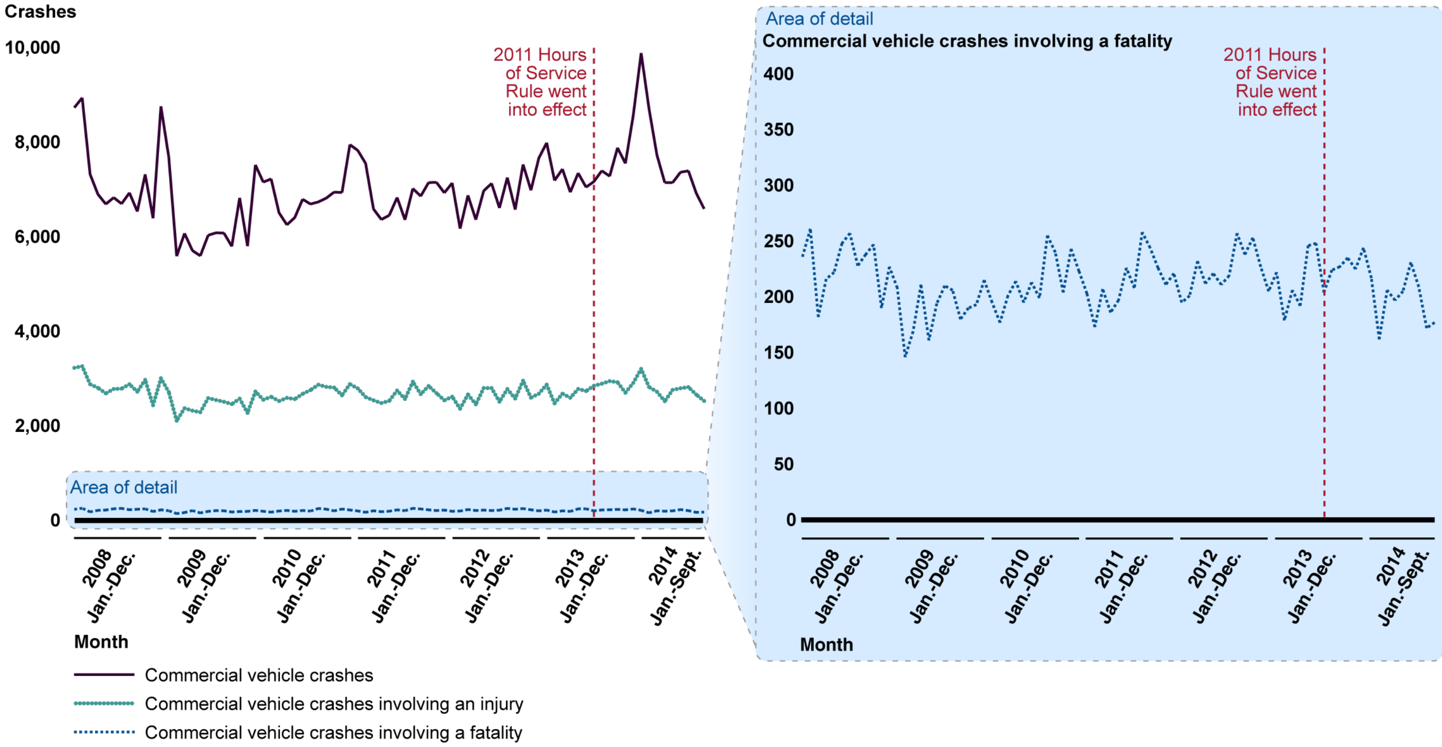
[61]Several schedules did not have to change in order to comply with the 2011 HOS rule. For example, as we modelled them drivers working regular daytime schedules with an average of 70, 60, or 50 hours per week would not need to change their schedules to continue using the restart provision. Also, we modelled the schedules of drivers working night schedules with an average of 50 hours per week, which would not need to change their schedules to comply with the 2011 HOS rule.

described by motor carriers resulted in a lower risk of driver fatigue, while other schedule changes showed a higher risk of driver fatigue. For example, a motor carrier told us that, to comply with the two-night provision, several of its nighttime drivers switched from a 6-day schedule to a 5-day schedule. Our analysis of this schedule change shows that drivers under this schedule have a lower risk of fatigue after the rule went into effect. In contrast, a different motor carrier told us that, to comply with the 168-hour limit, several of its nighttime drivers switched from a 4-day schedule with 3 days off to an alternating schedule of 5 days working and 2 days off and then 5 days working and 3 days off. Our analysis of this schedule change shows that these drivers have a higher risk of fatigue on certain days in their schedule after the rule went into effect. This may be due to more work time and less off-duty time for drivers following the schedule that complies with the 2011 HOS rule.

We also collected data on crashes involving commercial trucks from 2008 to 2014. As shown in figure 7, the monthly incidence of truck crashes and crashes involving fatalities varies considerably over time. To assess whether the rule had any discernable effect on crashes, we conducted two types of analyses. We examined (1) whether the rule affected the number of monthly crashes and (2) whether the rule resulted in any shift in the time of day in which crashes occurred.[62]

[62]For further detail regarding our analysis of crash data see appendix VII.

Figure 7: Commercial Motor Carrier Crashes from January 2008 to September 2014



Source: GAO analysis of Federal Motor Carrier Safety Administration data. | GAO-15-641

Trend in crashes over time: Many factors influence the number of crashes over time. For example, as the economy grows, typically there will be more trucks on the highways moving freight and this would be expected to be correlated with a higher incidence of crashes, all other things being equal. We used a statistical model to examine how the rule's implementation may have affected the number of monthly crashes, crashes with injuries, and fatal crashes, while controlling for other factors that affect the incidence of crashes, such as a proxy for trucking volume and seasonal variation. Our analysis showed a strong correlation between our measure of trucking volume and crashes—crashes occurred with more frequency when the industry was moving more freight—which is consistent with other previous work.[63] In addition, our initial analysis found a positive statistical association between the implementation of the rule and the number of all crashes, suggesting that crashes rose—holding other factors constant—after the rule when into effect. However, our discussions with industry participants indicated that the winter of 2013 to 2014 was unusually harsh and made operations very difficult for the industry. These weather conditions may have been a contributing factor, irrespective of the rule change, to a short-term rise in crashes in December 2013 through February 2014. When we used statistical methods to control for these months in a second iteration of our analysis, we no longer found a statistically significant increase in the number of crashes after implementation of the rule. That is, the incidence of crashes would appear to not have changed due to the rule in this scenario. Our analysis of crashes with injuries and crashes with fatalities was similarly structured—we included variables to account for truck volume, seasonal variation, and in a second iteration of the model, the unusual 3 months of weather during the winter of 2013 to 2014. We found that the implementation of the rule was not associated with any change in the incidence of crashes with injuries, whether or not we statistically controlled for the winter of 2013 to 2014. However, our analysis does indicate that crashes involving fatalities may have become less frequent after the implementation of the HOS rule.

Time of day shifts in crashes: Some stakeholders told us that the two-night provision of the HOS rule would potentially result in an increase in crashes involving commercial trucks during morning rush hours (i.e., 5 a.m. to 9 a.m.). They believed that some truckers who might have previously driven during the middle-of-the-night hours would be shifted to the early and mid-morning timeframe due to provisions of the new HOS rule. To test this possibility, we developed statistical models to examine whether, conditional on the number of crashes that occurred, there was a change in the likelihood of truck crashes occurring between 5 a.m. and 9 a.m. after the HOS rule went into effect.[64] Our analysis found no

[63]See Tom Brijs, Dimitris Karlis, and Geert Wets, "Studying the Effect of Weather Conditions on Daily Crash Counts Using a Discrete Time-Series Model*," Accident Analysis and Prevention* 40 (3) 1180-1190, Jan. 2008; Paul P. Jovanis and Hsin-Li Chang, "Modeling the Relationship of Accidents to Miles Traveled," *Transportation Research Record*, No. 1068, (1986): 42-51; and Thomas F. Golob and Wilfred W. Recker, "Analysis of Truck-Involved Freeway Accidents Using Log-Linear Modeling," *Journal of Safety Research*, Vol. 18,(1987): 121-136.

[64]The analysis actually coded the time frame as starting at 5:00 a.m. until 8:59 a.m.

statistically significant change in this likelihood associated with the implementation of the rule.[65]

An important caveat to our statistical analyses of crash data is that there are only 15 months of available data following the implementation of the rule. These limited data, which have non-trivial variation month to month, may not suffice to discern trends in the incidence of crashes. For example, as we explained above, our post-rule time frame covers only one winter—a time when crashes tend to rise—and we were told that this was an unusually difficult winter for the industry. These data limitations affect the ability to draw empirical inferences on the effects on crashes of the rule's implementation.

The 20 motor carriers we interviewed said they believed the rule had no noticeable effect or a negative impact on safety. Of the eight drivers we interviewed, one said the rule had no effect on safety, and seven said they had seen a negative impact on safety. For example, drivers mentioned that the time pressures they stated were associated with the 2011 HOS rule—particularly the 30 minute rest break—had resulted in more speeding. Drivers said that speeding occurs as drivers try to make deliveries or mileage goals before taking a 30-minute rest break. Safety groups we interviewed generally supported the 2011 HOS rule and thought it could improve safety for certain segments of the industry, but had concerns about ensuring compliance with the rule.

Health Benefits: Academic literature we reviewed indicates that insufficient amounts of sleep have negative health impacts. These effects include obesity, diabetes, hypertension, and cardiovascular disease. However, we did not find a data source that would allow us to assess health impacts of the rule. According to FMCSA officials and other stakeholders we spoke with, it could take years to identify health impacts following an intervention, such as the HOS rule. In addition, representatives from 16 of the 20 motor carriers we interviewed said they believed the 2011 HOS rule had no noticeable health effects. The other four motor carriers said they either noticed negative health effects— due to increased stress or fatigue as drivers tried to make up for lost income, among other reasons—or had no way to measure health impacts. In

[65]Models developed for this analysis controlled for road conditions at the time of the crash (i.e. daylight, nighttime, dry, wet, snow, and so forth) and whether the crash occurred on a weekday or a weekend day. See appendix VII for more detail.

addition, of the eight drivers we interviewed, six said their health was negatively impacted by the 2011 HOS rule, and two said there has been no impact. Those reporting a negative health impact cited, for example, being more stressed as they adjust to new schedules or try to make up for lost income. Safety groups we spoke with generally thought the 2011 HOS rule could improve driver health, but that health impacts would be difficult to attribute to the rule change.

Lack of Representative Driver Schedule Data Limits Analysis of the Effects of the HOS Rule

The ability of FMCSA and others to assess the effects of the 2011 HOS rule is impacted by the limited availability and representativeness of driver schedule data (i.e., records of drivers' work hours). No organization, including FMCSA, collects or maintains a centralized database with representative driver schedule data that can be generalized to the entire motor-carrier industry. For example, while we obtained and analyzed drivers' schedule data from a private research organization that includes over 15,000 usable driver records from 16 for-hire trucking companies, findings from that dataset cannot be generalized to the estimated 539,000 active motor carriers and 5.6 million drivers who operate in the United States. As a result it is not possible to determine the extent to which different types of commercial motor carriers, such as private carriers or drivers who own and operate their own vehicle, were affected by the rule because they are not included in the dataset. Similarly, FMCSA does not collect or use representative driver schedule data. Instead it uses roadside inspection, compliance review, and safety audit data as well as special studies to analyze motor carrier and drivers' safety behaviors.

Collecting representative schedule data is well recognized as a valuable analytical resource. Other DOT agencies that regulate the hours of service for transportation operators have demonstrated the importance of using representative schedule data. For example, the Federal Railroad Administration has collected logbook data from a representative sample of rail employees to examine fatigue and safety concerns since 2001. The Federal Aviation Administration also has access to electronic flight data on aircraft operations that is used to identify and analyze national trends, target agency resources, and identify and reduce or eliminate safety risks. In addition, FMCSA itself has cited the need to evaluate the actual impacts of its regulations, including the 2011 HOS rule, in order to meet executive requirements on regulatory review.[66] We have also found that sufficient and appropriate data are critical for rulemaking, decision making, and assessing the effects of regulations and programs.[67] As we have previously found, using data is especially critical for FMCSA, which has limited resources to oversee a large and diverse industry.[68]

Collecting these types of data has historically been difficult, but a recent statutory change to how driver schedule data will be collected provides a potential opportunity to do so. Many carriers use paper logbooks, a practice that makes collecting and analyzing schedule data administratively unmanageable. MAP-21 required FMCSA to develop a rule that will require motor carriers to stop using paper logbooks and instead, install electronic devices on individual vehicles to record and store drivers' schedule data.[69] This shift to electronic data will provide a potential source of representative data that could more easily be accessed and analyzed. However, FMCSA officials said that they do not currently plan to collect and use such data for research purposes because MAP-21 limits their use to the enforcement of laws.[70]

[66]Executive Order 13563, "Improving Regulation and Regulatory Review" directs federal agencies to measure, and seek to improve, the actual results of regulatory requirements. 76 Fed. Reg. 3821 (Jan. 21, 2011).

[67]GAO, *Managing for Results: Data-Driven Performance Reviews Show Promise But Agencies Should Explore How to Involve Other Relevant Agencies*, GAO-13-228 (Washington, D.C.: February 2013); and GAO, *Managing for Results: Executive Branch Should More Fully Implement the GPRA Modernization Act to Address Pressing Governance Challenges*, GAO-13-518 (Washington, D.C.: June 2013).

[68]GAO, *Motor Carrier Safety: New Applicant Reviews Should Expand to Identify Freight Carriers Evading Detection*, GAO-12-364 (Washington, D.C.: March 2012); and GAO, *Federal Motor Carrier Safety: Modifying the Compliance, Safety, Accountability Program Would Improve the Ability to Identify High Risk Carriers*, GAO-14-114, (Washington, D.C.: Feb. 2014).

[69]MAP-21 required the Department of Transportation to issue regulations requiring the use of an approved electronic logging device to accurately record commercial driver hours of service and record drivers' location. The rule is to require commercial vehicles to comply 2 years after the final rule is published. 49 U.S.C. § 31137(b)(1)(C). FMCSA expects to publish the final rule on electronic logging devices and HOS supporting documents on September 30, 2015.

[70]The Secretary may utilize information contained in an electronic logging device only to enforce the Secretary's motor carrier safety and related regulations, including record-of-duty status regulations. 49 U.S.C. § 31137(e)(1).

As we have shown, representative data on the motor carrier industry—specifically the number of drivers nationwide, the amount they drive, and how they use the restart provision—would provide critical information to FMCSA about the operations of the industry it oversees. Access to this type of data would also support data-driven rulemaking efforts and regulations. According to FMCSA officials, access to driver schedule data would enhance program evaluation, rulemaking, and the analysis of existing rules, including the HOS rule. We recognize, however, that there are concerns to collecting, storing, and analyzing these data, for example:

- **Privacy:** MAP-21 put protections in place to ensure driver privacy, including provisions to limit the harassment of drivers, protect personally identifiable information, and preserving the confidentiality of any personal data. Additionally, as we have previously found, collecting motor vehicle data raises concerns about sharing data with third parties, tracking drivers' movements, and vehicles' speed and location.[71]

- **Cost:** FMCSA officials expressed concerns about the costs to the agency and industry to collect, store, and analyze representative schedule data. For example, FMCSA officials believe that obtaining, standardizing, and depersonalizing data for millions of drivers would pose substantial costs on the agency, motor carriers, and drivers.

There may be ways of mitigating these privacy and cost concerns, but FMCSA has not examined the costs and benefits of collecting electronic driver-schedule data on a large-scale and currently does not plan to do so. For example, to address privacy concerns, information can be "de-identified" for the purposes of analyzing data. Specifically, the Federal Aviation Administration uses de-identified electronic flight data to identify and reduce or eliminate safety risks. In addition, our analysis of the effects of the HOS rule used schedule data from several thousand drivers, which we analyzed without any personally identifiable driver or motor carrier information. Cost concerns could potentially be addressed by collecting a nationally representative sample of the data—rather than from the entire universe of drivers. Sampling is a commonly used social science practice that allows researchers to apply conclusions drawn from

[71]GAO, *Intelligent Transportation Systems: Vehicle-to-Vehicle Technologies Expected to Offer Safety Benefits, but a Variety of Deployment Challenges Exist*, GAO-14-13 (Washington, D.C.: November 2013).

a subset of a population to the entire population. Although FMCSA officials told us that they have no plans to examine the possibility of collecting, storing, and analyzing electronic driver schedule data because of the MAP-21 provision, they thought privacy concerns, for example, could be addressed through stripping raw data of drivers' personally identifiable information before being shared with the agency or others. Until FMCSA evaluates the potential for collecting, storing, and analyzing data; the potential benefits of using the data; and privacy and cost concerns as well as the potential for FMCSA to mitigate them, the agency's ability to use electronic schedule data in the future will remain unknown.

Conclusions

The HOS rule has now been debated for over a decade with supporters and critics arguing over the extent to which it improves safety and health outcomes and negatively impacts the economy. FMCSA has attempted to quantify these effects through research efforts, but many continue to question the results. This is why it is critical for FMCSA to evaluate potential effects of rules, such as the HOS rule that has a heightened public profile, using established research standards and representative schedule data. FMCSA's field study demonstrated that the agency did a reasonable job in designing the study and analyzing data. However, because the agency did not use guidance outlining specific standards for conducting and reporting the results of scientific research, the field study fell short in reporting several limitations and did not fully link the results to its overall conclusions. These shortcomings leave the agency open to criticism over the integrity of the study and invite skepticism about the results.

Similarly, by not having access to and analyzing representative driver schedule data, FMCSA may continue to face challenges to the credibility of its rules. The lack of representative schedule data required FMCSA to make a number of assumptions about how the motor carrier industry would be affected by the HOS rule. Due to current data limitations, it is not possible to fully evaluate the rule's impact. For example, our analysis suggests that the rule may have affected a larger population of drivers than FMCSA anticipated, but without representative data there is no way to be certain this is universally the case. Understanding the population of drivers affected by the HOS rule is critical for determining the associated economic costs and safety and health benefits. While electronically collected, representative schedule data that will soon be available to FMCSA would allow it to assess the impact of its rules, MAP-21 places limitations on the use of this data for purposes other than enforcing laws.

Further, collecting, storing, and analyzing the data would require mitigating privacy and cost concerns. Given the potential value of such data for evaluating the impact of future regulatory rules, Congress may benefit from information on how the electronic data to be collected in response to the MAP-21 requirements could be extracted, stored, and analyzed and how privacy and cost concerns associated with the use of these data could be addressed. Such a study could help Congress determine whether limitations on the use of electronic schedule data for purposes other than enforcement could be amended in the future.

Matter for Congressional Consideration

Congress may wish to consider directing DOT to study and provide a report to Congress identifying approaches for extracting, storing, and analyzing electronically collected motor carrier drivers' schedule data, including the potential benefits, privacy, and cost concerns, and options for how such concerns could be mitigated.

Recommendations for Executive Action

To help ensure that FMCSA's future studies follow generally accepted research standards, the Secretary of Transportation should direct the FMCSA administrator to adopt guidance outlining research standards for designing, analyzing, and reporting the results of scientific research.

Agency Comments and Our Evaluation

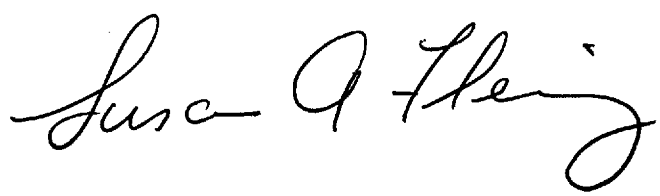
We provided a draft of this report to DOT for review and comment. DOT provided written comments, which are reprinted in appendix VIII. In its written comments, DOT stated that FMCSA agreed with our recommendation, although DOT also stated that FMCSA adhered to standard principles and practices of scientific research in conducting its January 2014 HOS study. As stated in this report, we identified generally accepted research standards that DOT followed; however, we also identified standards that it did not fully follow, namely, using a methodology that supports the study's objectives, reporting conclusions that are linked to the study's results, and reporting methodological limitations. We believe that FMCSA would strengthen its future research and enhance the integrity of future studies by adopting guidance outlining research standards for agency employees to follow.

DOT also stated in its written comments that our report recognized several achievements associated with its 2011 HOS rule, specifically a decrease in the frequency of drivers using long work schedules, a lower risk of driver fatigue, and a reduction in the number of commercial vehicle crashes involving fatalities. As we state in this report, however, these findings cannot be directly attributed to the rule.

In addition to the written comments, DOT also provided technical comments, which we incorporated as appropriate.

We are sending copies of this report to the Administrator of FMCSA, the Secretary of the Department of Transportation, and interested congressional requesters. In addition, the report will be available at no charge on the GAO website at http://www.gao.gov.

If you or your staff has any questions about this report, please contact me at (202) 512-2834 or flemings@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made major contributions to this report are listed in appendix IX.



Susan Fleming
Director, Physical Infrastructure Issues

Appendix I: Objectives, Scope, and Methodology

Our objectives were to examine (1) the strengths and limitations of the Federal Motor Carrier Safety Administration's (FMCSA) field study that examined whether the two-night provision reduces commercial driver fatigue, and (2) the key assumptions FMCSA used, at the time the 2011 hours of service (HOS) rule was promulgated, in estimating its effects and what is known about the economic, safety, and health effects of the rule.

Methodologies for Evaluating the Field Study

To assess the strengths and limitations of the field study, we drew on established guidelines and reports for assessing research and analysis, our reports on evaluating research programs, and our internal expertise in research design. We identified key standards used by professional associations, academics, and our prior work, with an emphasis on standards applicable to the design and methods used in the field study. (See table 3). Specifically, we identified six standards relating to (1) choosing evaluation measures that are valid and reliable; (2) using quality controls to identify data inconsistencies or errors; (3) reporting results and analysis that are supported by the data; (4) using a methodology, including statistical techniques, that support the objectives of the study; (5) reporting conclusions that are linked to the results; and (6) reporting all methodological limitations, including their potential impact on the results. We compared these standards to information included in the field study, statements made by FMCSA officials and researchers contracted by FMCSA to complete the study, and our analysis of study data.

Table 3: Guidance and Reports Used to Identify Generally Accepted Research Standards

Source	Guidance or report
Professional association	RAND Corporation, Standards for High-Quality Research and Analysis (2015)
	American Evaluation Association, Guiding Principles for Evaluators (July 2004)
	American Psychological Association, American Psychologist, Reporting Standards for Research in Psychology (December 2008)
Academia	Cook, Thomas D. and Campbell, Donald T., Quasi-Experimentation: Design & Analysis Issues for Field Settings (1979)
	Nunnally, Jum C., Psychometric Theory (1967)
GAO	*Designing Evaluation: 2012 Revision*, GAO-12-208G (Washington, D.C. Jan. 2012)
	Assessing the Reliability of Computer-Processed Data, GAO-09-680G (Washington, D.C. July 2009)
	Defense Transportation: Study Limitations Raise Questions about the Adequacy and Completeness of the Mobility Capabilities Study and Report, GAO-06-938 (Washington, D.C. Sept. 20, 2006)

Source: GAO. | GAO-15-641

To assess the reliability and accuracy of data and results reported in the field study, we collected and reviewed data provided by the primary researchers.[1] These data, collected between January and July 2013, included between 50 and 60 files for each of the 106 study participants, which the authors of the study used to build summary-level datasets for each driver. We also acquired the computer programs used by the field study's authors to process the data and run statistical analyses. As a check on the reliability of both the computer programs and the data we received, we applied the programs to the data for one driver and verified that the results were consistent with numbers reported in the field study.[2] In addition, we replicated results for 15 driver demographic variables, including age ranges, gender, experience, and type of operation, among others. For example, we confirmed that 100 men and 6 women participated, as the study reported. We also talked with the primary authors of the study about the steps they took to ensure the completeness and accuracy of the data, and checked the data for completeness and reasonableness. We determined that the data were sufficiently reliable for the purpose of our data analysis; in particular, we determined that we could, with sufficient reliability, conduct additional analysis of the data and align our results with those reported in the field study.

We assessed the sensitivity of the study's results to various alternative methodological choices. For example, we tested whether the results varied depending on whether the study included drivers having more than two nights of rest, as suggested by several stakeholders we interviewed.[3] Specifically, we removed drivers having more than two nights of rest from the analysis and generated results for the psychomotor vigilance test (PVT) and Karolinska Sleepiness Scale (KSS) that could be compared to

[1]FMCSA worked with a contractor and two subcontractors to complete the study, including researchers from the Virginia Polytechnic Institute and State University, Washington State University, and Pulsar Informatics. The primary researchers for this study were Hans Van Dongen from Washington State University and Daniel Mollicone from Pulsar Informatics.

[2]We did not run this check for all 106 drivers because of computing time and resources required and because of our assessment that the risk of an error in the programs or the data was low.

[3]More discussion on this theory is included in the body of our report.

those reported in the field study.[4] For a detailed technical discussion on our sensitivity testing and results, see appendix II.

To evaluate the validity and reliability of the fatigue measures used in the field study as well as assess the relationship between these measures and on-the-road safety outcomes (e.g. crashes), we conducted a literature review and spoke with experts in fatigue science. Our literature review used searches conducted by a GAO librarian across bibliographic databases[5] containing scholarly and peer reviewed materials, government reports, and conference papers to identify literature on the PVT, KSS, and lane deviation.[6] These searches returned 144 articles. We read the abstracts of these articles and then identified and reviewed 15 articles that discussed the validity and reliability of the PVT, KSS, and/or lane deviation. In addition, we conducted word searches in the abstracts of the 144 articles to identify and review 81 articles that discussed on-the-road safety.[7] Articles containing information on the reliability and validity of the PVT and the KSS were not evaluated for methodological quality by GAO technical experts. This approach was for two reasons. First, we used these articles as just one of multiple sources of information on this topic; other sources included interviews with experts in fatigue science who had experience with these particular measures of fatigue. Second, a review of the articles revealed a strong consensus on the reliability and validity of the PVT, which was the primary measure of fatigue used in the FMCSA field study.

We also selected three fatigue-science experts to interview on (1) the validity and reliability of the measures used in the field study, (2) the link between on-road-safety outcomes and these measures, and (3) how sleep science relates to changes to the new HOS rule (see table 4). We

[4]We did not conduct sensitivity analyses using the lane deviation data, because we did not have the time or resources to analyze the large volume of data collected.

[5]Databases consulted for these searches include Ei Compendex®, Embase®, EMCare®, MEDLINE®, NTIS: National Technical Information Service, PsycINFO, and Transport Research International Documentation, among others.

[6]This effort included three separate searches. First, we searched for the terms *psychomotor vigilance test* or *PVT,* and variants of *reliable, valid,* and *psychometric.* Second, we searched for the terms *Karolinska Sleepiness Scale* or *KSS* and variants of *reliable, valid* and *psychometric.* Third, we searched for the terms *lane changing* or *lane deviation,* and variants of *sleep, fatigue* and *drowsy*.

[7]Search terms included safety, accident, risk, driving, impairment, and crash.

selected these experts based on their numbers of published articles related to the fatigue measures used in the field study,[8] as well as our prior work on rail's HOS regulations, work that included opinions from several fatigue-science experts.[9]

Table 4: Fatigue-Science Experts Interviewed

Expert	Title	Department/Institution
David Dinges	Professor and Chief of the Division of Sleep & Chronobiology	Perelman School of Medicine, University of Pennsylvania,
Roger Rosa	Deputy Associate Director For Science	The National Institute for Occupational Safety & Health
Torbjörn Åkerstedt	Professor and Division Manager for Biological Psychology and Treatment Research	Stress Research Institute, Stockholm University

Source: GAO. | GAO-15-641

To identify the data collection and analytical methods used in the field study we spoke with officials from FMCSA and individuals contracted to complete the study, including the primary researchers of the study. These officials provided information on the statistical analysis used in the field study, driver recruitment, and data reliability assessments. To better understand the recruitment process and data collection methods, we also interviewed officials from all three motor carriers that participated in the study. Finally, we spoke with numerous industry and safety stakeholders to get their views on the strengths and limitations of the field study. For a list of these stakeholders, see table 6 below.

Methodologies for Evaluating the Anticipated and Actual Effects of the Hours of Service Rule

To identify key assumptions FMCSA used to estimate the effects of the 2011 HOS rule, a GAO economist, not involved in other aspects of our work and with substantial experience in cost-benefit analyses, reviewed the Regulatory Impact Analysis for the 2011 HOS rule. This expert identified an initial list of assumptions. We further categorized these assumptions into five areas: which drivers were affected, how drivers

[8]We excluded two highly cited authors from consideration because one of them was an author of the field study and the other worked at the same academic institution as an author of the field study.

[9]GAO, *Freight Railroad Safety: Hours of Service Changes Have Increased Rest Time, but More Can Be Done to Address Fatigue Risks,* GAO-11-853 (Washington, D.C.: Sept. 2011).

were affected, economic costs, safety benefits, and health benefits. FMCSA officials confirmed that our characterization of key assumptions was accurate. For more information on this process and the full list of FMCSA assumptions, see appendix III.

To identify what is known about the possible economic, safety, and health effects of the 2011 HOS rule, we conducted a comprehensive search for existing databases with information on motor carrier operations. We conducted web searches of academic journals, as well as government and other stakeholder websites. We also used search engines and asked officials from FMCSA and other stakeholders for recommendations on databases we could use to analyze the effects of the rule. We categorized the databases we found by the type of data collected, source, whether that data was publicly or commercially available, and potential costs. Based on this information we identified three different, relevant data sources, which were used for a variety of purposes (see table 5). To assess the reliability of these databases, we reviewed documentation on data collection efforts and quality assurance processes, talked with knowledgeable officials about these data, and checked the data for completeness and reasonableness. We determined that the data were sufficiently reliable for the purpose of our data analysis, though we also identified limitations in each dataset. More information on our analysis of American Transportation Research Institute (ATRI) logbook data, Federal Highway Administration (FHWA) vehicle count data, and FMCSA crash data can be found in appendixes IV, VI, and VII, respectively.

Table 5: Data Sources Used to Assess Possible Effects of the 2011 Hours of Service Rule

Source	Data type	GAO's purpose
American Transportation Research Institute	Driver logbook	Assessed a large sample of driver schedules from 2012 to 2014 to determine how long drivers were on duty and how they used the restart provision before and after the 2011 hours of service rule was implemented.
Federal Highway Administration	Vehicle counts	Examined the extent to which the overall number of commodity-carrying motor- vehicle traffic counts changed before and after the rule went into effect by comparing data from November and December of 2012 and 2013. Special attention was paid to whether any changes in traffic counts occurred during certain hours of the day.
Federal Motor Carrier Safety Administration	Crash	Examine the extent to which the overall number of commercial motor-vehicle crashes (and associated injuries and fatalities) changed before and after the rule went into effect using data from 2008 to 2014. As part of the analysis, we examined whether any changes in crashes occurred during specific hours of the day.

Source: GAO. | GAO-15-641

We spoke with multiple stakeholders, including organizations representing motor carriers, safety advocacy organizations, motor carrier companies, commercial vehicle drivers, and customers of the motor carrier industry (e.g., companies that ship freight) to understand how the 2011 HOS rule affected them and their members (see table 6). We selected stakeholders and shipping organizations to interview based on our prior work on commercial motor-carrier safety and recommendations from these stakeholders. We used FMCSA's Motor Carrier Information Management System to select active, U.S.-based motor carrier companies subject to the hours of service rule. Our selection was based on three primary criteria: (1) operation type (for-hire versus private), (2) fleet size, and (3) cargo type (interstate versus intrastate hazardous material). Using these criteria we created six selection groups to identify target populations that included for-hire and private carriers and had a range of fleet sizes.[10] We also created a seventh group that included only hazardous material carriers. We randomly selected 150 carrier companies from each of these groups for a total of 1,050 carrier companies across all seven groups. We then selected 3 to 5 carrier companies from each group using four additional selection criteria: (1) driver radius in miles (to identify carriers that drove a wide range of distances); (2) FMCSA region (for the purposes of ensuring geographic diversity); (3) HOS violations in the past 2 years (to identify carriers that might have experience with these violations); and (4) years in operation

[10]Fleet sizes were categorized as owner/operator (1 power unit), small (2 to 5 power units), medium (6 to 500 power units), and large (more than 500 power units).

(to identify carriers with a range of experience in the industry). Motor carrier companies we spoke with also put us in contact with eight drivers whom we interviewed to understand how the 2011 HOS rule had impacted them. The information collected from these interviews with representatives of motor carriers and drivers is not generalizable to all motor carrier companies or drivers.

Table 6: Industry and Safety Stakeholders, Shippers, and Motor Carriers Interviewed

Industry and safety stakeholders	Alliance for Driver Safety and Security
	American Trucking Associations
	Advocates for Highway and Auto Safety
	American Transportation Research Institute
	Commercial Vehicle Safety Alliance
	Harbor Trucking Association
	International Brotherhood of Teamsters
	National Association of Truck Stop Operators
	National Private Truck Council
	National Transportation Safety Board
	Owner-Operator Independent Drivers Association
	Truck Safety Coalition
Associations representing shipping, warehouse, and third party logistics companies	International Warehouse Logistics Association
	National Association of Manufacturers
	National Grocers Association
	National Retail Federation
	Transportation Intermediaries Association
	U.S. Chamber of Commerce
	U.S. Poultry and Egg Association
Motor carrier companies	Con-Way Truckload
	Crete Carrier Corporation
	JB Hunt
	JKM Transportation
	JMN Transportation
	K&J Trucking
	Kroger Dedicated Logistics
	LSG Sky Chefs
	Meadow Lark Transport
	Mesa Trucking
	Mountain Transport Inc.

New England Motor Freight
North Coast Truck Line
Process Machinery
Schneider National
Sentinel Transportation
Sides Contracting Company
UPS
USAV Group
Vern's Truck & Trailer Service

Source: GAO. | GAO-15-641

To summarize information from our interviews with motor carriers, we systematically analyzed information they provided on how the 2011 HOS rule had impacted them. To do so, one analyst coded interview responses. A second analyst randomly selected three interviews to code. The results from the two analysts were compared to ensure that the coding method was consistently applied, which confirmed that the coding done by the first analyst was reliable.

To identify possible fatigue effects on driver safety and health as well as the effectiveness of HOS rules in mitigating fatigue we conducted a literature review. Our literature review was also used to provide additional information on the biomathematical fatigue model we selected to analyze simulated and example driver schedules (see below). We identified specific articles for our literature review through searches conducted by a GAO librarian across bibliographic databases[11] containing peer-reviewed materials, government reports, and conference papers.[12] We honed the list of results by reviewing abstracts and focusing on articles published since 2008. This process identified 54 articles, most of which we used for background purposes or to provide additional context. We also identified several articles related to the effectiveness of HOS rules that were used

[11]Databases consulted for these searches include CINAHL, Ei Compendex®, Embase®, EMCare®, MEDLINE®, NTIS: National Technical Information Service, PsycINFO, Transport Research International Documentation, and WorldCat among others.

[12]This effort included three sets of searches. The first two sets of searches included, terms related to fatigue with terms related to driver performance, health and safety. The third set of searches focused specifically on, the FAID model and terms related to drivers, pilots, and railroad employees. Our first two searches examined articles that were from 2005 or later, while our third search included articles from 2010 or later.

to support our findings and required additional review. To select these articles we had two analysts code 10 of the 54 articles as either relating to effectiveness of HOS rules, relationship between fatigue and driver performance and safety, and/or relationship between fatigue and driver health. We compared the coding results and determined that the remaining coding could be reliably completed by a single analyst. Using this method we identified 15 articles related to the effectiveness of HOS rules. The methodologies and findings of these articles were reviewed by GAO's technical experts (research methodologists and an economist) to ensure that the findings were well supported.

To examine the potential effects of the 2011 HOS rule on driver fatigue, we used a biomathematical fatigue model of human alertness response to work and rest patterns. We obtained the Fatigue Audit InterDyne™ (FAID) model from InterDynamics Pty Ltd as the result of a competitive procurement. Our analysis using the FAID model involved two parts:

1. We developed schedules with different work periods (day and night) and different shift lengths (e.g. 14-hour shifts to simulate the maximum allowed daily on-duty hours).We then created different scenarios based on whether a schedule would need to change to comply with the two restart provisions in the 2011 HOS rule. For example, we used the FAID model to compare driver fatigue levels for a driver working 14-hour night shifts 5 days a week with one-night off (not in compliance with the two restart provisions in the 2011 HOS rule) to a driver working the same 14-hour night shifts 5 days a week with two-nights off (in compliance with the two restart provisions in the 2011 HOS rule).
2. We analyzed real world schedules described to us by motor carrier operators and drivers we interviewed regarding how they had changed schedules due to the new rule. For example, we used the FAID model to analyze fatigue levels of drivers for a motor carrier that told us their drivers had changed from six shifts per week (before the 2011 HOS rule went into effect) to five shifts per week (after the 2011 HOS rule went into effect) in order to be compliant with aspects of the rule.

For a technical discussion of the scope, methodology, and additional results using the FAID model, see appendix V.

Appendix II: Analysis of Data Analysis and Collection in FMCSA's Field Study

As part of our review of the strengths and limitations of FMCSA's field study that examined the two-nighttime provision of the 2011 hours of service (HOS) rule,[1] we conducted two statistical analyses, as described below. First, we conducted sensitivity tests to assess how certain methodological decisions in the field study potentially affected its results. Second, we conducted a power analysis to assess the adequacy of the study's sample size for estimating how the rule affected fatigue within each of several industry segments.

Sensitivity Tests

During the course of our review, we asked stakeholders we interviewed to describe their general impressions of the data and statistical methods used to evaluate the rule. Stakeholders specifically questioned two methodological decisions:

1. The study analyzed data for drivers who had taken more than a two-night restart. Some stakeholders expressed concern that this decision did not conform with the purpose of the study to analyze the impact of changing the restart provision from requiring one-night of rest to exactly two nights.
2. The study collected data from what stakeholders characterized as a small sample of drivers, a sample that included local and regional drivers. Some stakeholders said that the study's sample size of 106 drivers was not large enough to provide enough data to draw sound conclusions or be representative of the trucking industry. In addition, several stakeholders we spoke with said that the restart provision likely has different effects across industry segments, including local, regional, and long-haul or over-the-road drivers.

To test the validity of these concerns, we acquired the source data and computer code used in the field study from the authors.[2] We conducted sensitivity tests of these data that assessed whether the study's results might have been different if the researchers had made different methodological choices. Our tests examined whether the results of the psychomotor vigilance tests (PVT) and the Karolinska Sleepiness Scale (KSS) would have been different had the study: 1) excluded drivers taking

[1]Hans P.A. Van Dongen and Daniel J.Mollicone, *Field Study on the Efficacy of the New Restart Provision for Hours of Service* (2014).

[2]We were also provided with a list of all the input and output files used and met with the authors to discuss how the data and code were organized.

a two-or-more night restart and 2) allowed the outcomes to vary by the drivers' industry segments, i.e., local, regional, or over-the-road.

Our first sensitivity test removed drivers who took more than a two night restart from the "treatment" group hypothesized to be affected by the regulation. The study used the following linear mixed effects statistical model to analyze how average PVT and KSS scores varied according to whether drivers took one-night restart versus a two-or-more night restart:[3]

$$Y_{ij} \sim \mathrm{N}(\mathbf{x_{ij}}\boldsymbol{\beta} + \mu_j, \sigma_Y^2)$$
$$\mu_j \sim \mathrm{N}(\mu, \sigma_\mu^2)$$

$\boldsymbol{x_{ij}}$	=	{I($Condition_j$ x $Timing_{ij}$ x $Period_{ij}$)}
$Condition_i$	=	{1 night restart, 2+ night restart}
$Timing_{ij}$	=	{12:00 – 4:00 AM, 4:00 – 8:00 AM, … , 8:00 – 11:59 PM}
$Period_{ij}$	=	{1st duty period, restart period, 2nd duty period}

Y_{ij} denoted the number of PVT lapses or the KSS score for driver j = 1, 2, … , J on measurement occasion i = 1, 2, … , N_j, with the total sample size given by $N = \sum_j N_j$. $\boldsymbol{x_{ij}}$ was a vector of indicator variables for each level of the Cartesian product of $Condition_j$, $Timing_{ij}$, and $Period_{ij}$, similar to a design vector in an experimental study. $Condition_j$ measured the number of nights in the restart period between the two duty periods in the study. $Timing_{ij}$ and $Period_{ij}$ measured the time of day and stage of the study, respectively, when the PVT or KSS outcomes were measured.[4] $\boldsymbol{\beta}$ was a column vector of fixed coefficients. Since $\boldsymbol{x_{ij}}$ did not include an intercept and all covariates were categorical, $\boldsymbol{\beta}$ denotes the cell means for each subgroup in $\boldsymbol{x_{ij}}$.

We increased the granularity of $Condition_j$ in order to estimate the effect of interest, the contrast between mean fatigue after one-night versus two-night night restarts, by adding a separate category for three-or-more night restarts. The study estimated this overall contrast by taking equally weighted linear combinations of the estimated effects across subpopulations defined by $Timing_{ij}$ x $Period_{ij}$, using the "least-squares

[3]The study reports the results of repeated measures analysis of variance (ANOVA) F-tests, but the authors implemented these tests using linear mixed effects models.

[4]We used the study's time periods, which were defined as six disjoint intervals of four hours each (e.g., 12:00 AM – 3:59 AM).

means" method implemented in SAS.[5] We used the same approach, in order to minimize methodological changes other than those we evaluated.

Our test found that including drivers who took more than a two-night restart in the treatment group did not affect the results, as shown by table 7. On-duty drivers taking only a two-night restart had 0.32 lapses more lapses per PVT bout, on average, than drivers taking a one-night restart. This estimate is statistically indistinguishable from the 0.33 lapses reported by the field study, which included drivers taking more than a two-night restart in the effect estimate. The results from the study and our tests are both statistically distinguishable from zero at the 0.10 level of confidence. Similarly, the KSS results reported in the field study for on-duty drivers are statistically indistinguishable to those we found when removing drivers taking more than two-night restarts from the treatment group.

Table 7: Estimated Effects of Restart Nights during Duty Hours, by Treatment Group Construction

Fatigue measure	Source	Effect	Estimated difference in means (90 percent confidence interval)
Psychomotor vigilance test	Field study	One-night restart versus a two-or-more night restart	0.33 [0.09, 0.57][a]
	GAO	One-night restart versus a two-night restart	0.32 [0.08, 0.57][a]
Karolinska Sleepiness Scale	Field study	One-night restart versus a two-or-more night restart	0.21 [0.05, 0.37][a]
	GAO	One-night restart versus a two-night restart	0.19 [0.03, 0.35][a]

Source: GAO analysis of Federal Motor Carrier Safety Administration's Field Study on the Efficacy of the New Restart Provision for Hours of Service (Jan. 2014). | GAO-15-641

[a]Significant at 0.10 level.

Our second sensitivity test estimated the difference in mean PVT and KSS outcomes between drivers with one-night restarts and two-or-more night restarts, separately by industry segments defined by local versus regional and over-the road operations.[6] We attempted to estimate separate effects for the regional segment alone, but the study data did not include any regional drivers who took one-night restarts. As a result, we combined regional and over-the-road drivers, who may behave in similar

[5]"LSMEANS Statement," *SAS/STAT 9.3 User's Guide* (Cary, NC: SAS Institute, Inc., 2011:, 467-482.

[6]We included drivers classified by the study as "van truckload" in the local group. We included drivers classified as "dedicated" or "independent contractor" in the regional and over-the-road group.

ways such as driving longer distances and sleeping away from home with some regularity. Finally, we omitted effects for Periodij to simplify the model, since the estimated treatment effects did not vary when this variable was included or excluded. Thus, we estimated two alternatives:

$\boldsymbol{x}_{ij}$ = {I($Condition_j$ x $Timing_{ij}$ x $Operation_j$)}
$Condition_j$ = {1 night restart, 2 night restart, 3+ nights restart}
$Timing_{ij}$ = {12:00 – 4:00 AM, 4:00 – 8:00 AM, … , 8:00 – 11:59 PM }
$Operation_j$ = {Local, Regional or Over-the-Road}

$\boldsymbol{x}_{ij}$ = {I($Condition_j$ x $Timing_{ij}$ x $Operation_j$)}
$Condition_j$ = {1 night restart, 2+ nights restart}
$Timing_{ij}$ = {12:00 – 4:00 AM, 4:00 – 8:00 AM, … , 8:00 – 11:59 PM }
$Operation_j$ = {Local, Regional or Over-the-Road}

Our test found that the field study did not sample enough drivers from each of these segments to estimate effects for each subgroup that were statistically distinguishable from zero (significant). We found that regional and over-the-road drivers had a significant difference of 0.42 PVT lapses and an insignificant difference of 0.13 KSS scores, on average, between drivers with a one-night and a two-or-more night restart (see table 8). In contrast, local drivers had insignificant differences of -0.15 PVT lapses and 0.26 KSS scores, on average. We found similar effects when comparing drivers with one-night versus two-night restarts. In sum, the variation in the results suggests that the 2011 HOS rule may have had variable effects on different segments of the industry. However, the relatively wide confidence intervals of these estimates, in part due to limited sample sizes within each industry segment, makes the exact degree of variation uncertain.

Table 8: Estimated Effects of Restart Nights during Duty Hours, by Treatment Group Construction and Industry Segment

Fatigue measure	Source	Effect	Industry segment[a]	Estimated difference in means (90 percent confidence interval)
Psychomotor vigilance test	Field study	One-night restart versus a two-or-more night restart	Local, regional, and over-the-road	0.33 [0.09, 0.57][b]
	GAO	One-night restart versus a two-or-more night restart	Local	-.15 [-0.59, 0.28]
			Regional and over-the-road	0.42 [0.16, 0.67][b]
		One-night restart versus a two-night restart	Local	-.08 [-.52, 0.36]
			Regional and over-the-road	0.30 [0.03, 0.57][b]
Karolinska Sleepiness Scale	Field study		Local, regional, and over-the-road	0.21 [0.05, 0.37][b]
	GAO	One-night restart versus a two-or-more night restart	Local	0.26 [-.02, 0.54]
			Regional and over-the-road	0.13 [-0.04, 0.30]
		One-night restart versus a two-night restart	Local	0.21 [-.07, 0.49]
			Regional and over-the-road	0.10 [-.08, 0.28]

Source: GAO analysis of Federal Motor Carrier Safety Administration's Field Study on the Efficacy of the New Restart Provision for Hours of Service (Jan. 2014). | GAO-15-641

[a]Separate estimates for regional drivers were not feasible because the study data did not include drivers in this industry segment who took one-night restarts. We included drivers classified by the study as "van truckload" in the local group. We included drivers classified as "dedicated" or "independent contractor" in the regional and over-the-road group.

[b]Statistically significant at the 0.10 level.

Power Analysis

To assess stakeholder concerns about the adequacy of the field study's sample, we conducted a statistical power analysis to assess how many drivers and PVT measurements might be needed to estimate the HOS rule's effects for each of several segments of the industry. In general, a power analysis estimates the probability that a study will conclude that an effect exists when, in fact, the effect does exist in the population of interest. In our application, we estimated the probability of identifying differences in mean PVT results for drivers taking a one-night versus two-night restart, separately for subpopulations of drivers operating in the local and regional or over-the-road industry segments and at various times of day.

We found that the field study's sample size was insufficient to estimate statistically significant differences in the primary fatigue measure—the PVT—for each of these industry segments and times. By collecting data from approximately 100 drivers, each taking 25 PVTs, the study would have had an approximately 5 percent or lower chance of identifying statistically significant differences as small as 0.3 PVT lapses for each subpopulation. This effect size is similar to what the study found across all time periods, using a sample of a similar size. In contrast, collecting data from approximately 4,000 to 6,000 drivers, each taking 200 PVTs, would have increased the chance of identifying significant differences for each subpopulation separately to approximately 40 to 60 percent. For an effect of 1 PVT lapse—slightly larger than what the study found for overnight observation periods—sample sizes of 800 drivers and 200 PVTs would have achieved approximately 60 to 80 percent power. These results illustrate the range of sample sizes that FMCSA would need to consider, subject to available resources for driver recruitment and study administration, in order to confidently detect effects of this size for each of several industry subpopulations.

Below, we discuss the methods we used to conduct the power analysis and its results in more detail.

HOS Study Analytical Methods and Data

The HOS study evaluated the effects of the rule on three measures of driver fatigue: PVT, KSS, and lane deviation. For the purpose of our power analysis, we focused on scores on the PVT due to the extensive scientific research that has validated this measure of fatigue. PVT scores consist of the number of lapses in attention a study participant experiences across a battery of tests. The study administered the PVT test at various times during the participation period, which spanned one-duty period, a restart period, and a second-duty period. Our power analysis focuses on the duty periods during which drivers would need to be alert in order to safely operate their commercial motor vehicle.

Tables 9, 10, and 11 provide descriptive statistics about the study's data collected during the duty periods from the 106 participating drivers. The number of PVTs varied across drivers from 8 to 53, with a median of 24.5 and a mean of 25.0. The study collected a total of 2,653 non-missing observations during the duty periods at the level of driver-tests, after excluding observations during the restart period. Across all observations, the mean number of PVT lapses was 1.8. PVT lapses had a skewed distribution, with a 90th quantile of 5, a 95th quantile of 8, and a 99th quantile of 18. (See table 9.) The study administered several hundred PVTs to drivers taking one, two, or three-or-more night restarts across various types of operations (e.g., local, regional, and over-the-road). (See table 10.) The study made fewer observations across subpopulations defined by restart nights and time period of observation, however, ranging from 44 to 274 tests per subpopulation (see table 11). This suggests that the three-way cross-classification of restart nights, operation type, and time period may produce limited sample sizes for each subpopulation.

Table 9: Federal Motor Carrier Safety Administration's (FMCSA) Field Study Descriptive Statistics on Psychomotor Vigilance Test's (PVT) Outcomes in Duty Periods

	Number of PVTs per driver	Number of PVT lapses
Mean	25	1.8
Standard deviation	6.3	3.5
Min	8	0
Max	53	35
10th quantile	18	0
90th quantile	32	5
95th quantile	33	8
99th quantile	47.5	18

Source: GAO analysis of FMCSA's Field Study on the Efficacy of the New Restart Provision for Hours of Service (Jan. 2014). | GAO-15-641

Note: Statistics on the number of PVTs per driver are calculated at the driver level, whereas statistics on the number of PVT lapses are calculated at the driver-test level.

Table 10: Federal Motor Carrier Safety Administration's (FMCSA) Field Study Psychomotor Vigilance Tests (PVT) Taken during Duty Periods by Restart Nights and Operation Type

Restart nights	Local, regional, or van truckload	OTR, dedicated, or independent contractor	Total
1 night	255	659	**914**
2 nights	603	692	**1295**
3+ nights	248	196	**444**
Total	**1106**	**1547**	**2653**

Source: GAO analysis of FMCSA's Field Study on the Efficacy of the New Restart Provision for Hours of Service (Jan. 2014). | GAO-15-641

Table 11: Federal Motor Carrier Safety Administration's (FMCSA) Field Study Psychomotor Vigilance Tests (PVT) Taken during Duty Periods by Restart Nights and Observation Time

Restart Nights	12:00 - 4:00 AM	4:00 - 8:00 AM	8:00 AM - 12:00 PM	12:00 PM - 4:00 PM	4:00 PM - 8:00 PM	8:00 – 11:59 PM	Total
1 night restart	213	156	142	122	180	101	**914**
2 nights restart	244	237	152	212	274	176	**1295**
3+ nights restart	44	88	74	88	86	64	**444**
Total	**501**	**481**	**368**	**422**	**540**	**341**	**2653**

Source: GAO analysis of FMCSA's Field Study on the Efficacy of the New Restart Provision for Hours of Service (Jan. 2014). | GAO-15-641

As we discuss above, the authors of the field study analyzed these data using a hierarchical linear model, with normally distributed random effects at the level of drivers but fixed effects otherwise. To analyze power, we used models similar to those we used above to estimate separate effects across industry segments.

We used two separate methods to estimate power. These methods estimated power to detect effects for each of several subpopulations of drivers and for the overall population, respectively. Using the first method, we made distributional assumptions about the data generation process, informed by the field study's results, and used Monte Carlo simulation methods and the actual model of the data to be estimated (from above). Using the second method, we made similar distributional assumptions, but we applied power equations for simplified versions of this model available in the statistical literature.

Power Analysis for Multiple Subpopulations

Our analysis estimates power for estimating separate effects for each of 12 subpopulations defined by two industry segments (local versus regional/over-the-road) and six equally spaced observation periods (e.g., 12:00 AM to 4:00 AM).

Let β_{ck} denote the row of $\boldsymbol{\beta}$ such that c = $Condition_j$ = {1, 2, 3} and k = {$Timing_{ij}$ x $Operation_j$} = {1, 2, … , K}, with $\boldsymbol{\beta}^T = [\beta_{11}, \ldots, \beta_{1K}, \beta_{21}, \ldots, \beta_{2K}, \beta_{31}, \ldots, \beta_{3K}]$. In this application, K = 12. We seek to estimate $E(Y_{ij} \mid Condition_j = 2) - E(Y_{ij} \mid Condition_j = 1)$ for the kth subpopulation. Under the model, these quantities can be expressed as $\boldsymbol{\delta}^T = [\beta_{21} - \beta_{11}, \ldots, \beta_{2K} - \beta_{1K}]$, which implies that $\boldsymbol{\beta}^T = [\beta_{11}, \ldots, \beta_{1K}, \beta_{11} + \delta_1, \ldots, \beta_{1K} + \delta_K, \beta_{31}, \ldots, \beta_{3K}]$. We focus on the comparison between drivers with 1 versus 2 night restarts, because the rule was expected to reduce fatigue among drivers taking 1 night restarts prior to the rule's implementation.

Our approach uses Monte Carlo methods to simulate artificial data from the distributions implied by the hierarchical model and the previous study's data.[7] For simulation purposes, we assume that $\delta_k = \delta > 0$ for all k and set all model parameters and the distribution of X_{ij} to estimates from the HOS study. We draw $m = \{1, 2, \ldots, M\}$ sets of simulated data, each of which includes $i = \{1, 2, \ldots, N\}$ PVT tests administered to $j = \{1, 2, \ldots, J\}$ drivers:

$$\tilde{\mathbf{x}}_{\mathbf{ij(m)}} \sim Multinomial(\widehat{\boldsymbol{\theta}})$$
$$\tilde{\mu}_{j(m)} \sim N(\hat{\mu}, \hat{\sigma}_{\mu}^2)$$
$$\tilde{Y}_{ij(m)} \sim N(\tilde{\mathbf{x}}_{\mathbf{ij(m)}}\widehat{\boldsymbol{\beta}} + \tilde{\mu}_{j(m)}, \hat{\sigma}_Y^2)$$

For the mth simulated dataset, the estimates of interest are $\boldsymbol{\delta}_{(m)}^{\mathrm{T}} = [\hat{\beta}_{21(m)} - \hat{\beta}_{11(m)}, \ldots, \hat{\beta}_{2K(m)} - \hat{\beta}_{1K(m)}]$, which, under the model, have entries equal to the difference in mean outcomes between drivers in subpopulation k having two versus one-night restarts. We use the set of M simulated values of $\boldsymbol{\delta}$ to test composite hypotheses about effects for all subgroups, as well as multiple simple hypotheses about the effects for each subgroup.

Composite Hypotheses

We test the composite linear hypothesis that the effects for all K subpopulations are simultaneously equal to 0:

$$H_0: \quad \mathbf{R}^{\mathbf{T}}\boldsymbol{\beta} = \mathbf{0}$$
$$H_1: \quad \mathbf{R}^{\mathbf{T}}\boldsymbol{\beta} \neq \mathbf{0}$$

$\mathbf{R}^{\mathbf{T}}$ is a K x $3K$ contrast matrix, containing three submatrices of constants associated with each level of $Condition_j$:

$$\mathbf{R}^{\mathbf{T}} = \begin{bmatrix} r_{11}^1, \ldots, r_{1K}^1 & r_{21}^1, \ldots, r_{2K}^1 & r_{31}^1, \ldots, r_{3K}^1 \\ \vdots & \vdots & \vdots \\ r_{11}^K, \ldots, r_{1K}^K & r_{21}^K, \ldots, r_{2K}^K & r_{31}^K, \ldots, r_{3K}^K \end{bmatrix} = \begin{bmatrix} -1, \ldots, 0 & 1, \ldots, 0, & 0, \ldots, 0 \\ \vdots & \vdots & \vdots \\ 0, \ldots, -1 & 0, \ldots, 1 & 0, \ldots, 0 \end{bmatrix}$$

This implies that $\mathbf{R}^{\mathbf{T}}\boldsymbol{\beta} = \boldsymbol{\delta} = \mathbf{0}$ under H_0.

[7] Andrew Gelman and Jennifer Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models* (New York: Cambridge University Press, 2007): 447-455.

Let $\mathbf{V_R}$ = Var($\mathbf{R^T\beta}$) = $\mathbf{R^T V_\beta R}$, where $\mathbf{V}_\beta$ is the covariance matrix of $\boldsymbol{\beta}$. For each simulated dataset, we test these hypotheses using the multivariate Wald test statistic, $= \boldsymbol{\beta}^{\mathrm{T}}\mathbf{R}\mathbf{V}_{\mathbf{R}}^{-1}\mathbf{R}^{\mathrm{T}}\boldsymbol{\beta} \sim \chi^2_K$, with each simulation's value equal to $H_{(m)} = \widehat{\boldsymbol{\beta}}_{(m)}^{\mathrm{T}}\mathbf{R}\widehat{\mathbf{V}}_{\mathbf{R}(m)}^{-1}\mathbf{R}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_{(m)}$. We reject H_0 for simulation *m* if $H_{(m)} > \chi^2_{K,.95}$, the critical value of the chi-squared distribution with *K* degrees of freedom, assuming α = .05.[8]

Since we assume that H_0 is false under the simulated model, power is equal to (1-β) = Pr(Reject H_0) = E[I(Reject H_0)], where the last term is the expected value of binary random variable indicating whether we reject H_0 in simulation *m*. We estimate power as $\left(1-\hat{\beta}\right) = \frac{1}{M}\sum_m \mathrm{I}(H_{(m)} > \chi^2_{K,.95})$, where I() is an indicator function ranging on {0,1}.

Multiple Simple Hypotheses

We also test multiple simple hypotheses about the effect for each subpopulation *k*:

$$H_{0k}:\ \beta_{2k} - \beta_{1k} = 0$$
$$H_{1k}:\ \beta_{2k} - \beta_{1k} \neq 0$$

H_{0k} is a linear function of the parameters, so we use the same Wald test statistic as we did for the composite hypothesis test:

$$\widehat{H}_{k(m)} = \widehat{\boldsymbol{\beta}}_{(m)}^{\mathrm{T}}\mathbf{R_k}\widehat{\mathbf{V}}_{(m)}^{-1}\mathbf{R_k^T}\widehat{\boldsymbol{\beta}}_{(m)} \sim \chi^2_1 \ ,$$

where $\mathbf{R_k^T}$ is the *k*th row of $\mathbf{R^T}$, with all other rows set to zero. As in the composite test, we reject H_{0k} for simulation *m* if $H_{k(m)} > \chi^2_{K,.95}$.[9] For subpopulation *k,* power is equal to (1-β_k) = Pr(Reject H_{0k}) = E[I(Reject H_{0k})], defined analogously to the composite test. We estimate power for the *K* set of simple tests as $\left(1-\hat{\beta}\right) = \frac{1}{M}\sum_m \left(\prod_k \mathrm{I}(\ H_{k(m)} > \chi^2_{K,.95}\right)_m$, noting that the inner product is the joint probability of rejecting all hypotheses.

[8]Stephen W. Raudenbush and Anthony S. Bryk, *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2d ed (Thousand Oaks, CA: Sage Publications, 2002): 57-61.

[9]This test could be conducted using *z*-statistics, consistent with results on the variance and distribution of linear combinations of Normal random variables. We use the Wald statistic to link the two types of tests we conducted.

This is follows from the contrast weights in $\mathbf{R}^{\mathbf{T}}$, whose orthogonality implies independent comparisons.[10]

Sample and effect sizes similar to those in the field study provided low amounts of power to detect effects of -0.3 or -1 PVT lapses, as shown in table 12. The probability of detecting effects of this size for each of the 12 driver subpopulations was approximately 5 percent or less when sampling 100 drivers and administering 25 PVTs, using either a joint hypothesis test or multiple simple tests across subpopulations. Power increased to 0.11 when detecting an effect equal to -1 lapse using a joint test, but it remained approximately zero using multiple simple tests. These results confirm that the field study could not reliably estimate separate effects by type of driver and time of day.

Table 12: Power Analysis of Federal Motor Carrier Safety Administration's (FMCSA) Field Study for Subpopulations

Assumptions			Estimated power					
			Composite test			Multiple simple tests		
Number of psychomotor vigilance tests (PVT)	Number of drivers	Effect size	Estimate	Lower 95% CI	Upper 95% CI	Estimate	Lowe 95% CI	Upper 95% CI
25	100	-0.3	0.05	0.03	0.07	0.00	0.00	0.00
200	100	-0.3	0.06	0.03	0.09	0.00	0.00	0.00
25	250	-0.3	0.06	0.03	0.09	0.00	0.00	0.00
200	250	-0.3	0.07	0.04	0.09	0.00	0.00	0.00
25	800	-0.3	0.11	0.07	0.14	0.00	0.00	0.01
200	800	-0.3	0.07	0.04	0.09	0.01	0.00	0.02
25	4,000	-0.3	0.41	0.36	0.47	0.09	0.06	0.13
200	4,000	-0.3	0.40	0.35	0.46	0.25	0.20	0.30
25	6,000	-0.3	0.58	0.53	0.64	0.23	0.18	0.28
200	6,000	-0.3	0.60	0.54	0.65	0.37	0.31	0.42
25	8,000	-0.3	0.72	0.67	0.77	0.36	0.31	0.42
200	8,000	-0.3	0.74	0.69	0.79	0.53	0.48	0.59
25	10,000	-0.3	0.85	0.81	0.89	0.52	0.46	0.58
200	10,000	-0.3	0.84	0.80	0.88	0.68	0.62	0.73

[10]William L. Hays, *Statistics*, 5th ed (New York: Harcourt Brace College Publishers, 1994): 433-438.

Assumptions			Estimated power					
			Composite test			Multiple simple tests		
Number of psychomotor vigilance tests (PVT)	Number of drivers	Effect size	Estimate	Lower 95% CI	Upper 95% CI	Estimate	Lowe 95% CI	Upper 95% CI
25	100	-1	0.11	0.07	0.15	0.01	0.00	0.02
200	100	-1	0.13	0.09	0.17	0.02	0.00	0.04
25	250	-1	0.27	0.22	0.32	0.06	0.03	0.09
200	250	-1	0.28	0.23	0.33	0.10	0.07	0.13
25	800	-1	0.78	0.73	0.82	0.46	0.41	0.52
200	800	-1	0.78	0.73	0.82	0.62	0.56	0.67
25	1,000	-1	0.88	0.85	0.92	0.61	0.55	0.67
200	1,000	-1	0.88	0.84	0.92	0.79	0.74	0.83
25	1,500	-1	0.99	0.98	1.00	0.84	0.80	0.88
200	1,500	-1	0.99	0.97	1.00	0.94	0.91	0.96

Source: GAO analysis of FMCSA's Field Study on the Efficacy of the New Restart Provision for Hours of Service (Jan. 2014). | GAO-15-641

Note: Effect size is the mean difference in PVT lapses between drivers taking two-night versus one-night restarts. The power estimate is the probability of identifying significant effects for each of 12 driver subpopulations, defined by local versus regional/over-the-road operator types and 6 time periods of 4 hours each, using samples of the sizes assumed. Confidence intervals represent expected error due to the use of Monte Carlo simulation methods, with the number of simulations equal to 300.

Our analysis found that increasing the field study's sample size by several hundred additional drivers and PVTs would be necessary to achieve conventional levels of power, depending on the true effect size to be detected. Using a joint hypothesis test or multiple simple tests when the effect equaled -0.3 lapses, we found that about 6,000 drivers and 200 PVTs would be required to achieve power of 0.6 using a joint hypothesis test and 0.37 using multiple simple tests. To detect an effect of -1 lapse with power of approximately 0.6 to 0.8, samples of more than 800 drivers and 200 PVTs would be necessary. Since the field study reports effects in this range, the appropriate sample size would appear to fall between these bounds, subject to available resources for driver recruitment and study administration.

Power generally increased more strongly with increases in the number of drivers, rather than the number of PVT tests. This result is consistent with the statistical literature on hierarchical linear models in which the marginal effects of interest vary only across clusters.[11] The moderately high intraclass correlation of 0.68 in the study's PVT data within drivers increases the required driver sample size to achieve a given level of power, all else being equal. However, the number of PVTs increases power more strongly when using multiple simple tests, because more PVTs increases the number of observations across time periods and therefore increases power for each test.

Power Analysis for the Overall Population

To provide an alternative power estimate for the overall population of drivers, using less complex methods, we analyzed power for a simplified version of the model using results from the literature on hierarchical linear models.[12]

For this analysis, we calculated power for estimating one treatment effect δ as defined above, without assuming that the effect varied across subpopulations. This reduced the model to xij = {I(1 night restart), I(2+ night restart)}. We further assumed that the sample had the same (balanced) number of observations for each of the two treatment conditions. This simplified the model so that variance estimation, hypothesis testing, and power calculation became tractable using the closed-form methods developed by the literature on hierarchical linear models (largely by avoiding many covariates). The more realistic Monte Carlo analysis summarized above compensated for these simplifications, since using two different methods checked the robustness of results from either one.

We tested the following simple hypothesis:

$$H_0: \beta_2 - \beta_1 = 0$$
$$H_1: \beta_2 - \beta_1 \neq 0$$

Following the literature, we use F-tests formed using the ratio of mean-squares for the treatment contrast and clusters within treatment groups:[13]

[11]Stephen W. Raudenbush, "Statistical Analysis and Optimal Design for Cluster Randomized Trials," *Psychological Methods* 2, no. 2 (1997): 173-185.

[12]Raudenbush, "Statistical Analysis;" and Jessaca Spybrook et al., "Optimal Design for Longitudinal and Multilevel Research: Documentation for the 'Optimal Design Software'" (unpublished manuscript, Sept. 19, 2006).

[13]Raudenbush, "Statistical Analysis," 176. Spybrook et al., "Optimal Design," 10-11.

$$\mathrm{E}(MS_{treatment}) = N\sigma_{\mu}^{2} + \sigma_{Y}^{2} + \frac{NJ\delta^{2}}{4}$$
$$\mathrm{E}(MS_{cluster}) = N\sigma_{\mu}^{2} + \sigma_{Y}^{2}$$

$$F(1, J-2, \lambda) = \frac{E(MS_{condition})}{E(MS_{cluster})}\ , \lambda = \frac{\frac{NJ\delta^{2}}{4}}{N\sigma_{\mu}^{2} +\ \sigma_{Y}^{2}}$$

We estimate F() using estimates of σ_{μ}^{2} and σ_{Y}^{2} from the prior study's data, similar to the Monte Carlo analysis above, and using various values for N and J. We estimate power as $(1 - \hat{\beta}_k)\ =\ \Pr(\hat{F}\ >\ F_{.95})$, where $F_{.95}$ is the critical value of the F distribution with (1, J-2) degrees of freedom and noncentrality parameter $\lambda =\ \hat{\lambda}$, assuming α = .05.

Our analysis found that the field study had more power to detect overall effects than effects across multiple subgroups (see table 13).[14] The field study's approximate sample size of 100 drivers and an average of 25 PVTs had a 0.08 probability of detecting effects of -0.3 PVT lapses and a 0.42 probability of detecting effects of -1 PVT lapses for the overall population of drivers. Given these results, the study had low power to detect the overall effect of -0.33 lapses that it reported across all time periods, suggesting that the result may have been statistically unusual.[15] Our analysis found that increasing the sample size to approximately 2,000 to 4,000 drivers and 200 PVTs would have been required to detect an effect of -0.3 PVT lapses with conventionally adequate levels of power. However, samples as small as 250 drivers and 25 PVTs would be sufficient to detect effects of -1 PVT lapse.

[14]The required sample sizes are generally lower to detect overall effects of a given size, compared to detecting effects across multiple subpopulations, because the sample is not distributed among subpopulations.

[15]For the sensitivity analysis, we scaled the effect as the difference in PVT lapses between drivers taking 1 versus 2 night restarts (a positive value). For the power analysis, we scaled the effect as the opposite difference (a negative value), but this does not affect the power estimates.

Table 13: Power Analysis of Federal Motor Carrier Safety Administration's (FMCSA) Field Study for Overall Population

Assumptions			
Number of Psychomotor vigilance tests (PVT)	Number of Drivers	Effect Size	Estimated Power (Simple Test)
25	100	-0.3	0.08
200	100	-0.3	0.08
25	250	-0.3	0.13
200	250	-0.3	0.14
25	400	-0.3	0.19
200	400	-0.3	0.19
25	800	-0.3	0.32
200	800	-0.3	0.33
25	1000	-0.3	0.39
200	1000	-0.3	0.40
25	2000	-0.3	0.66
200	2000	-0.3	0.67
25	4000	-0.3	0.92
200	4000	-0.3	0.92
25	6000	-0.3	0.98
200	6000	-0.3	0.99
25	100	-1	0.42
200	100	-1	0.43
25	250	-1	0.80
200	250	-1	0.80
25	400	-1	0.94
200	400	-1	0.95
25	800	-1	1.00
200	800	-1	1.00
25	1000	-1	1.00
200	1000	-1	1.00
25	2000	-1	1.00
200	2000	-1	1.00
25	4000	-1	1.00
200	4000	-1	1.00
25	6000	-1	1.00

Source: GAO analysis of Federal Motor Carrier Safety Administration's Field Study on the Efficacy of the New Restart Provision for Hours of Service (Jan. 2014). | GAO-15-641

Note: Effect size is the mean difference in PVT lapses between drivers taking two-night versus one-night restarts. The power estimate is the probability of detecting a significant effect for the overall population of drivers, using samples of the sizes assumed.

Appendix III: GAO Identified Assumptions in the Hours of Service Rule's Regulatory Impact Analysis

As part of our review of the effects of the 2011 hours of service (HOS) rule, we were asked to identify the assumptions used by the Federal Motor Carrier Safety Administration (FMCSA) to estimate the costs and benefits of the rule. To do so, a GAO economist identified the key assumptions in the regulatory impact analysis (RIA) used by FMCSA to detail the costs and benefits of the 2011 HOS final rule.[1] This expert developed an initial list of key assumptions after reading the RIA. We then categorized these assumptions into five groups: which drivers were affected, how drivers were affected, economic costs, safety benefits, and health benefits. We also had a sixth category for assumptions that did not clearly fit into one of the aforementioned categories. FMCSA officials reviewed our list of assumptions, offered some corrections, and ultimately agreed that the list of assumptions accurately represented the key assumptions used in the RIA.

Below is a list of the assumptions we identified and FMCSA officials confirmed. These assumptions cover which drivers would be affected by the HOS rule and how they would be affected. The assumptions below also detail how operational changes (e.g. driver schedule changes) result in the economic costs and the safety and health benefits of the HOS rule that went into effect July 1, 2013. (See table 14).

[1]DOT, FMCSA, *2010-2011 Hours of Service Rule – Regulatory Impact Analysis*, RIN 2126-AB26 (December 2011).

Table 14: Key Assumptions Included in the Federal Motor Carrier Safety Administration's (FMCSA) Regulatory Impact Analysis for the 2011 Hours of Service (HOS) Rule

Group	Assumption
Which drivers affected	The HOS rule mostly affects drivers working 70 or more hours per week and the focus is on the truckload sector. • Truckload service is a type of For-Hire transportation service. • Night drivers within these sectors most likely to make schedule adjustments.
	Drivers are grouped into intensity groups based on their weekly on-duty hours, which are defined as follows: moderate drivers average 45 hours per week, high drivers average 60 hours per week, very high drivers average 70 hours per week, and extreme drivers average 80 hours per week. Overall, drivers that fall into the very high and extreme groups constitute 15 percent of drivers and incur the majority of effects from the HOS rule, while high Intensity drivers incur a small fraction of effects (see discussion of 30 minute rest break below). Only about 9 percent of the very high and extreme drivers are significantly affected, less than 3 percent of total driver workforce. In addition: • The 30-minute rest break affects the high, very high, and extreme intensity driver groups. • The 168-hour limit affects the very high and extreme intensity driver groups. • The two-night provision only affects a fraction of drivers from the very high and extreme intensity groups.
	Various assumptions are made on the percentage of days on which drivers use their 14th hour of on-duty time to determine the effect of the 30-minute rest break.
How drivers affected	Drivers with different work intensities reallocate time as necessary. Some time is reallocated within a driver's schedule to another day (except for drivers from the extreme intensity group) and time that cannot be reallocated is lost.
	Drivers from the extreme intensity group lose 8.70 hours per week, drivers from the very high intensity group lose 0.38 hours per week, and drivers from the high intensity group lose .05 hours per week. These effects stem from reduced work or driving time due to: • less daily on-duty time due to 30 minute rest break; and • less weekly on-duty time due to fewer restarts and/or longer restarts.
	Overlapping impacts of the rule provisions are accounted for to avoid double-counting the impact of rule provisions.
	The two-night provision has a minimal effect on most drives' schedules because: • drivers who end work week in late afternoon or evening can start again after 5am after a day and a half with two-night periods; • drivers who would usually work until 2 a.m. or 3 a.m., would only need to adjust by 1 or 2 hours and stop before 1 a.m.; and • only drivers in the very high and extreme intensity groups would lose a significant amount of time, and most night time drivers already take full weekends off for a greater than 34 hour restart.
Economic costs	The costs are driven by the extent rule-induced reduction in driving and on-duty time is lost to each driver.
	Reduced productivity for each driver intensity group is weighted by that groups' relative share of work hours, which is summed to get the total reduced industry productivity. This calculation is done for each aspect of the rule, including the 30-minute rest break and new restart provisions. Overlapping effects from each provision of the rule are accounted for.
	Estimate that a 1 percent reduction in productivity costs $356 million.
Safety benefits	Safety benefits are derived from reduced fatigue risk due to decreases in daily driving time and weekly work time and therefore estimated reduction in fatigue-related crashes.

Group	Assumption
	Fatigue is involved in 13 percent of crashes. To account for variation from this figure, scenarios are considered when fatigue is involved in 7 and 18 percent of crashes.
	Risk of a fatigue-related crash is higher in the 11th work hour. Shifting the 11th work hour to another driver with an average fatigue risk results in a 23 percent reduced risk of fatigue-related crashes.
	Reducing driving time leads to reallocation to: • other modes (e.g., rail), which are safer; • other drivers, who drive less; and • new drivers, with the same risk profile as current drivers. Therefore, any analysis of effects of reduced time needs to be on a net basis.
	The rule would have same relative effect on fatalities as on all crash damages caused by heavy trucks.
	Monetization of safety benefits is generated by estimating the reduction in large truck crashes due to the reduction in driver fatigue. This calculation accounts for costs saved from the average damages of large truck crashes and lives saved based on the value of a statistical life.
Health benefits	Health benefits accrue due to reduced work time and the associated increase in the opportunity to sleep.
	Each driver intensity group was assigned a low, medium, and high baseline level of sleep based on estimates of the relationship between hours worked and amount of sleep. Per night the expected gains in sleep by driver intensity are: • 0 minutes for drivers in the moderate intensity group; • 0.06 minutes for drivers in the high intensity group; • 1.44 minutes for drivers in the very high intensity group; and • 22.32 minutes for drivers in the extreme intensity group.
	Health benefits are obtained only for drivers who are not getting adequate sleep.
	Monetization of health benefits is based on the mortality characteristics of the new sleep profile achieved with the rule change based on the full value of a statistical life.
	Data on truck driver's age distribution and mortality rates—adjusted to account for the poorer health indicators in the truck driver population—used to calculate health benefits.
	There are various unquantified health benefits of the rule.
Other	The effect of the rule on congestion is not estimated or considered as part of the cost/benefit analysis.
	No technical changes will influence impact of the rule and the baseline represents the best assessment of the industry without a rule change.
	Data used in developing the rule (from various sources) is reliable.
	There will be full compliance with the rule. If there is less than full compliance, both benefits and costs will decline and will be unlikely to change the ratio of benefits to costs.
	Define local drivers as operating within a 100 mile radius. Over-the-road drivers operate beyond a 100 mile radius.
	Freight supply and demand are unaffected by rule.
	Standard discount rate used, taken from Office of Management and Budget's guidance, to determine the present value of monetary costs and benefits.

Source: GAO presentation of FMCSA assumptions. | GAO-15-641

Appendix IV: Analysis of Driver Schedule Data

To identify possible effects of the 2011 hours of service (HOS) rule, including which drivers were affected and how those drivers were affected, we purchased 3 years of drivers' schedule data from the American Transportation Research Institute (ATRI). The data contain information on daily on-duty hours for drivers that worked for 16 for-hire motor carriers from January 1, 2012, through December 31, 2014. Specifically, the dataset includes a driver identification number, the number of recorded on-duty hours for each driver for each day in the dataset, and the length of each restart the driver took. The 2012 dataset contains 57,096 unique driver records; the 2013 dataset contains 60,196 unique driver records, and the 2014 dataset contains 63,957 unique driver records. The data we used for our analysis of drivers' schedule data are not representative of the motor carrier industry, and therefore, our findings based on these data cannot be generalized to all carriers in that population.

Filtering the Data

The data we purchased was not filtered or processed by ATRI. We applied four data-quality filters to remove records that contained likely errors and to exclude drivers with minimal on-duty time that could potentially bias our results, specifically:

- We removed driver records with five or more null values in the on-duty-hours cells. This filter removed 40,622 records from the 2012 dataset, 42,496 records from the 2013 dataset, and 45,564 records from the 2014 dataset.
- We removed driver records with one or more instance of 18 or more hours of on-duty recorded in a single day. This filter removed 1,126 records from the 2012 dataset, 1,735 records from the 2013 dataset, and 1,520 records from the 2014 dataset.
- We removed driver records with less than 20 hours of total on-duty time logged during the year. This filter removed 32 records from the 2012 dataset, 34 records from the 2013 dataset, and 9 records from the 2014 dataset.
- We removed driver records that did not have a 0 value in the on-duty hours column on the day before a recorded restart with a value of 48 or greater. This filter removed 8 records from the 2012 dataset, 13 records from the 2013 dataset, and 61 records from the 2014 dataset.

After applying these filters, there were 15,308 unique driver records in the 2012 dataset, 15,918 unique driver records in the 2013 dataset, and 16,803 unique driver records in the 2014 dataset.

Defining and Calculating a Driver's Work Week

To compare our driver data to the Federal Motor Carrier Safety Administration's (FMCSA) estimates of drivers' weekly hours as well as to provide a meaningful denominator for hours worked (e.g. on-duty hours per work week), we defined and calculated a driver's work week. As described below, we considered several specifications in our analysis, including how many days to include in a work week, how to calculate a work week, whether to use the driver or the work week as the unit of analysis that is compared over time, and how to categorize weekly on-duty time.

Number of Days in a Work Week

As described in the background section, drivers may follow either a 70-hour over 8 days on-duty limit or a 60-hour over 7 days on-duty limit. According to ATRI, the majority of the drivers in the dataset we purchased likely follow the 70-hour on-duty limit. We calculated a work week using both 7 and 8 days to ensure we covered the range of possible work weeks. However, given that the majority of drivers in the dataset likely followed the 70-hour/8-day work week, we included these results in the body of this report. The number of days used to calculate a driver's work week has an effect on the total and average hours worked.

Calculation of a Work Week

We used a fixed 7 day and fixed 8 day work week in the analysis presented below and in our report. Before making this decision, we considered two ways of calculating a driver's work week: 1) a fixed work week and 2) a rolling work week. For drivers following the HOS rule, a work week is not based on a set week (e.g. Monday through Friday). Rather, drivers calculate their 60- or 70-hour on-duty limits over a rolling period of 7 or 8 consecutive days. To take this into account, we calculated a fixed work week and a rolling work week to approximate the weekly on-duty hours of drivers in our dataset. The fixed work week has 7 or 8 consecutive days that do not overlap. For example in a 90 day period, there are 12 7-day weeks and 11 8-day weeks.[1] The rolling work week

[1]The fixed weekly sums were calculated as follows. For 7-day fixed weekly sum, we added days 1 through days 7 (in this case January 1st through January 7th) to get the first weekly sum. Then we added days 8 through 14 to get the second weekly sum. We calculated weekly sums until we reached the end of the dataset. We used the same method to calculate the 8-day fixed weekly sum except 8 days were included instead of 7.

also has 7 or 8 consecutive days, but each subsequent week shifts the start of the week by only one day. Each rolling week overlaps with the previous week.[2] We tested these two calculation methods on a subset of data (January 1 to March 31, 2013, and 2014), and we found that the method we used to calculate a work week did not change the results of our analysis.

Weekly On-Duty Categories

To examine the assumptions and estimates FMCSA made in its regulatory impact analysis (RIA) for the 2011 rule, we categorized our results using weekly on-duty hour categories similar to those included in that document.[3] FMCSA defined four categories of drivers with the following parameters:

- “Moderate”—average weekly work time of 45 hours, includes weekly averages between 20 to 55 hours
- “High”—average weekly work time of 60 hours, includes weekly averages greater than 55 to 65 hours
- “Very High”—average weekly work time of 70 hours, includes weekly averages greater than 65 to 75 hours
- “Extreme”—average weekly work time of 80 hours, includes weekly averages greater than 75 hours

In our analysis, we split the Moderate category into two groups: those working less than 35 hours per week and those working more than 35 hours per week. We did this split to determine whether the behavior changes made by those at the low end of the Moderate category differed from those at the high end of the Moderate category. As a result, our analysis uses five weekly on-duty categories.

Unit of Analysis

To analyze the data, we considered two different units of analysis: a driver and an individual work week. Specifically, for the time periods we examined, we determined the number of drivers whose average work

[2]The rolling weekly sums were calculated as follows. For a 7-day rolling weekly sum, we added the on-duty hours for days 1 through days 7 (in this case January 1st through January 7th) to get the first weekly sum. Then we added the on-duty hours for days 2 through 8 to get the second weekly sum. We calculated weekly sums until we reached the end of the dataset. We used the same method to calculate the moving 8-day weekly sum except 8 days were included instead of 7.

[3]In the RIA, these weekly on-duty hour categories are referred to as “Driver Groups by Intensity of Schedule” or “Driver Intensity Groups.”

weeks fell into each of the on-duty hour categories described above. For example, a driver that averaged 57 on-duty hours per 8-day work week would be placed in the greater than 55- to 65-hour on-duty category. We also analyzed the data by the number of work weeks that fell into each of the on-duty categories.[4] For example, we identified 342,747 out of 959,529 total weeks with on-duty hours between 35 and 55 from January 1, 2012, to June 15, 2013 (see table 24 below).

Restart Length

To analyze drivers' use of restarts within our dataset, we followed the practice of both FMCSA and ATRI of excluding restarts longer than 72 hours (3 days). The rationale for this practice is that off-duty periods greater than 72 hours do not represent a normal or operational restart period by a commercial motor vehicle driver. While we provide descriptive information on all restarts in our dataset (see table below), we exclude restarts greater than 72.99 hours from our analysis of restart use by driver or driver groups (see tables 20, 21, and 25 below).

Data Analysis

We analyzed driver schedule data in two different ways. In the first section, we describe our analysis of the data by dividing the 3 years of data into: 1) seasonal datasets by creating 12 individual datasets composed of 75- to 90-day seasons and 2) before and after the rule datasets by creating two datasets of approximately 18 months before and after July 1, 2013. In the second section, we describe our statistical analysis of drivers' schedule data using a subset of the full dataset that covered the entire time frame—2012 to 2014—but were limited to drivers for whom we had data in all 3 years.

Analysis of the Dataset over Different Time Periods

Analysis of Seasonal Datasets

We divided each year of data into four seasons: winter, spring, summer, and fall. The spring, summer, and fall seasons were truncated to account for (1) the effective date of the 2011 HOS rule on July 1, 2013, and (2) the suspension of two of the restart provisions in the 2011 HOS rule.[5]

[4] This method of analysis is similar to the method used in FMCSA's 2007 Hours of Service study, which was referenced in the regulatory impact analysis.

[5] As noted in our report, Congress suspended enforcement of two of the HOS rule provisions, the 168-hour limit and the two-night provision, effective on December 16, 2014.

Specifically, data from June 16 to June 30 were removed to ensure that behavior changes in anticipation of the rule's going into effect did not bias the dataset. Similarly, data from July 1 to July 14 were removed to account for drivers' getting used to and understanding the rule. We also omitted data from December 16 through December 31 to account for the suspension of several restart provisions on December 16, 2014. To ensure we were comparing the same time periods in each season, all of these changes were applied across each year of data. (See table 15.)

Table 15: Description of Seasons Used in Analysis of Seasonal Datasets

Season	Dates
Winter	January 1 – March 31, 2012 January 1 – March 31, 2013 January 1 – March 31, 2014
Spring	April 1 – June 15, 2012 April 1 – June 15, 2013 April 1 – June 15, 2014
Summer	July 15 – September 30, 2012 July 15 – September 30, 2013 July 15 – September 30, 2014
Fall	October 1 – December 15, 2012 October 1 – December 15, 2013 October 1 – December 15, 2014

Source: GAO analysis of drivers' schedule data. | GAO-15-641

Dividing the data into seasons allowed us to isolate changes potentially due to seasonal variation and to show how drivers' behavior—averaged over a relatively short period of time (75 to 90 days)—changed. For example, several motor carriers and drivers told us that bad weather in January and February 2014 disrupted operations for many motor carriers. By separating the datasets into seasons, we were able to ensure that changes we observed from one period to another were consistent across all seasons and were not only due to seasonal variation.

We found that the percentage of drivers in our dataset working more than 55 hours per work week decreased after the 2011 HOS rule went into effect—between spring 2013 and summer 2013. In addition, our analysis suggests that the percentage of drivers working 55 or fewer hours per work week increased after the rule went into effect. Tables 16 through 22 show the results of our seasonal analysis of drivers' schedule data using an 8-day work week and a 7-day work week.

Table 16: Seasonal Comparison of Drivers by Average Weekly On-Duty Hours, per 8-Day Work Week for 16 For-Hire Motor Carriers (2012–2014)

		Average weekly on-duty hours										
		20 to less than 35 hours		35 to 55 hours		More than 55 to 65 hours		More than 65 to 75 hours		More than 75 hours		
Season	Year	Number of drivers	Percent of drivers	Number of drivers	Percent of drivers	Number of drivers	Percent of drivers	Number of drivers	Percent of drivers	Number of drivers	Percent of drivers	Total drivers[a]
Winter	2012	622	4.1	7890	51.5	4827	31.5	1,634	10.7	40	0.3	**15,308**
Spring	2012	650	4.2	7790	50.9	4847	31.7	1,625	10.6	57	0.4	**15,308**
Summer	2012	685	4.5	8040	52.5	4829	31.5	1,343	8.8	40	0.3	**15,308**
Fall	2012	664	4.3	7772	50.8	4931	32.2	1,534	10.0	40	0.3	**15,308**
Winter	2013	750	4.7	7825	49.2	5171	32.5	1,762	11.1	80	0.5	**15,918**
Spring	2013	702	4.4	7662	48.1	5287	33.2	1,820	11.4	69	0.4	**15,918**
2011 hours of service rule goes into effect												
Summer	2013	841	5.3	9110	57.2	4666	29.3	865	5.4	15	0.1	**15,918**
Fall	2013	864	5.4	9119	57.3	4657	29.3	831	5.2	10	0.1	**15,918**
Winter	2014	1003	6.0	10218	60.8	4532	27.0	658	3.9	3	0.0	**16,803**
Spring	2014	916	5.5	9441	56.2	5152	30.7	930	5.5	22	0.1	**16,803**
Summer	2014	953	5.7	9816	58.4	4914	29.2	805	4.8	15	0.1	**16,803**
Fall	2014	994	5.9	9995	59.5	4794	28.5	734	4.4	13	0.1	**16,803**

Source: GAO analysis of drivers' schedule data from 16 for-hire motor carriers. | GAO-15-641

[a]The number of drivers in each season does not add up to the total number of drivers. In each season some drivers who met our filter criteria did not work 20 or more hours per week.

Table 17:Seasonal Comparison of Drivers by Average Weekly On-Duty Hour, per 7-Day Work Week for 16 For-Hire Motor Carriers (2012–2014)

		Average weekly on-duty hours										
		20 to less than 35 hours		35 to 55 hours		More than 55 to 65 hours		More than 65 to 75 hours		More than 75 hours		
Season	Year	Number of drivers	Percent of drivers	Number of drivers	Percent of drivers	Number of drivers	Percent of drivers	Number of drivers	Percent of drivers	Number of drivers	Percent of drivers	Total drivers[a]
Winter	2012	1,196	7.8	11,325	74.0	2,386	15.6	62	0.4	0	0.0	**15,308**
Spring	2012	1,288	8.4	11,003	71.9	2,501	16.3	118	0.8	0	0.0	**15,308**
Summer	2012	1,309	8.6	11,249	73.5	2,250	14.7	74	0.5	0	0.0	**15,308**
Fall	2012	1,344	8.8	11,394	74.4	2,103	13.7	48	0.3	0	0.0	**15,308**
Winter	2013	1,461	9.2	11,661	73.3	2,312	14.5	74	0.5	0	0.0	**15,918**
Spring	2013	1,368	8.6	11,587	72.8	2,446	15.4	80	0.5	0	0.0	**15,918**

		Average weekly on-duty hours										
		20 to less than 35 hours		35 to 55 hours		More than 55 to 65 hours		More than 65 to 75 hours		More than 75 hours		
Season	**Year**	**Number of drivers**	**Percent of drivers**	**Number of drivers**	**Percent of drivers**	**Number of drivers**	**Percent of drivers**	**Number of drivers**	**Percent of drivers**	**Number of drivers**	**Percent of drivers**	**Total drivers[a]**
2011 hours of service rule goes into effect												
Summer	2013	1,701	10.7	12,417	78.0	1,298	8.2	22	0.1	0	0.0	**15,918**
Fall	2013	1,812	11.4	12,388	77.8	1,198	7.5	12	0.1	0	0.0	**15,918**
Winter	2014	1,959	11.7	13,276	79.0	1,081	6.4	7	0.0	0	0.0	**16,803**
Spring	2014	1,814	10.8	13,157	78.3	1,375	8.2	26	0.2	0	0.0	**16,803**
Summer	2014	1,988	11.8	13,218	78.7	1,212	7.2	14	0.1	0	0.0	**16,803**
Fall	2014	2,068	12.3	13,240	78.8	1,129	6.7	16	0.1	0	0.0	**16,803**

Source: GAO analysis of drivers' schedule data from 16 for-hire motor carriers | GAO-15-641

[a]The number of drivers in each season does not add up to the total number of drivers. In each season some drivers who met our filter criteria did not work 20 or more hours per week.

Consistent with our results using drivers as the unit of analysis, the overall percentage of weeks from our dataset in the more-than-55-hours-per-week categories decreased after the 2011 HOS rule went into effect. The overall percentage of weeks in the 55-or-fewer-hours categories increased after the rule went into effect. In both tables 18 and 19, the percentage of weeks with total on-duty hours greater than 65 hours per work week is higher than those found in tables 16 and 17. This finding suggests that while a driver may have an intense work schedule over one week, on average, drivers tend to have less intense work schedules over longer periods of time. Tables 18 and 19 show the total number of weeks that fell into our on-duty categories for drivers in our dataset using an 8- and 7-day work week.

Table 18: Seasonal Comparison of Weeks by Weekly On-Duty Hours, per 8-Day Work Week for 16 For-Hire Motor Carriers (2012–2014)

		Weekly on-duty hours										
		20 to less than 35 hours		35 to 55 hours		More than 55 to 65 hours		More than 65 to 75 hours		More than 75 hours		
Season	Year	Number of weeks	Percent of weeks	Number of weeks	Percent of weeks	Number of weeks	Percent of weeks	Number of weeks	Percent of weeks	Number of weeks	Percent of weeks	Total weeks[a]
Winter	2012	12,690	7.5	59,476	35.3	47,419	28.2	30,879	18.3	7,846	4.7	**168,388**
Spring	2012	9,816	7.1	48,298	35.1	38,281	27.8	25,002	18.1	6,961	5.1	**137,772**
Summer	2012	10,145	7.4	47,105	34.2	38,139	27.7	25,010	18.2	6,592	4.8	**137,772**
Fall	2012	11,800	8.6	44,165	32.1	37,999	27.6	26,704	19.4	7,344	5.3	**137,772**
Winter	2013	13,037	7.4	58,840	33.6	50,843	29.0	33,017	18.9	8,151	4.7	**175,098**
Spring	2013	9,934	6.9	46,951	32.8	41,991	29.3	27,852	19.4	6,658	4.6	**143,262**
2011 hours of service rule goes into effect												
Summer	2013	11,893	8.3	52,116	36.4	40,534	28.3	23,776	16.6	3,435	2.4	**143,262**
Fall	2013	12,663	8.8	54,269	37.9	40,848	28.5	21,872	15.3	2,979	2.1	**143,262**
Winter	2014	16,943	9.2	73,523	39.8	51,342	27.8	26,139	14.1	3,432	1.9	**184,833**
Spring	2014	12,393	8.2	57,543	38.1	44,652	29.5	23,628	15.6	3,005	2.0	**151,227**
Summer	2014	12,789	8.5	57,488	38.0	43,817	29.0	23,611	15.6	2,819	1.9	**151,227**
Fall	2014	14,074	9.3	60,310	39.9	42,614	28.2	22,208	14.7	2,592	1.7	**151,227**

Source: GAO analysis of drivers' schedule data from 16 for-hire motor carriers. | GAO-15-641

[a]The number of weeks in each season does not add up to the total number of week. In each season, some of the weeks that met our filter criteria did not have 20 or more on-duty hours.

Table 19: Seasonal Comparison of Weeks by Weekly On-Duty Hours, per 7-Day Work Week for 16 For-Hire Motor Carriers (2012–2014)

		Weekly on-duty hours										
		20 to less than 35 hours		35 to 55 hours		More than 55 to 65 hours		More than 65 to 75 hours		More than 75 hours		
Season	Year	Number of weeks	Percent of weeks	Number of weeks	Percent of weeks	Number of weeks	Percent of weeks	Number of weeks	Percent of weeks	Number of weeks	Percent of weeks	Total weeks[a]
Winter	2012	21,446	10.8	106,613	53.6	48,504	24.4	8,559	4.3	75	0.0	**199,004**
Spring	2012	15,413	10.1	79,550	52.0	28,720	25.3	7,394	4.8	63	0.0	**153,080**
Summer	2012	16,538	9.8	85,969	51.1	42,740	25.4	8,098	4.8	66	0.0	**168,388**
Fall	2012	17,734	11.6	76,961	50.3	38,120	24.9	7,290	4.8	107	0.1	**153,080**
Winter	2013	22,155	11.6	97,451	51.0	47,795	25.0	8,299	4.3	140	0.1	**191,016**
Spring	2013	16,455	10.3	80,989	50.9	41,232	25.9	7,403	4.7	96	0.1	**159,180**

		Weekly on-duty hours										
		20 to less than 35 hours		35 to 55 hours		More than 55 to 65 hours		More than 65 to 75 hours		More than 75 hours		
Season	**Year**	**Number of weeks**	**Percent of weeks**	**Number of weeks**	**Percent of weeks**	**Number of weeks**	**Percent of weeks**	**Number of weeks**	**Percent of weeks**	**Number of weeks**	**Percent of weeks**	**Total weeks[a]**
2011 hours of service rule goes into effect												
Summer	2013	20,352	11.6	96,299	55.0	38,164	21.8	3,908	2.2	32	0.0	**175,098**
Fall	2013	21,585	13.6	85,281	53.6	33,887	21.3	3,297	2.1	33	0.0	**159,180**
Winter	2014	28,624	14.2	110,126	54.6	40,214	19.9	3,971	2.0	39	0.0	**201,636**
Spring	2014	21,310	12.7	93,542	55.7	35,738	21.3	3,424	2.0	24	0.0	**168,030**
Summer	2014	23,275	12.6	102,064	55.2	39,259	21.2	3,689	2.0	29	0.0	**184,833**
Fall	2014	24,398	14.5	91,763	54.6	34,436	20.5	3,063	1.8	19	0.0	**168,030**

Source: GAO analysis of drivers' schedule data from 16 for-hire motor carriers. | GAO-15-641

[a]The number of weeks in each season does not add up to the total number of week. In each season, some of the weeks that met our filter criteria did not have 20 or more on-duty hours.

We found that restarts per driver remained relatively constant before and after the 2011 HOS rule went into effect for drivers in our dataset working between 35 to 75 hours per 8-day work week. We also found an increase of approximately one restart per driver for those working more than 75 hours per 8-day work week and a reduction of approximately one restart per driver for drivers working less than 35 hours per 8-day work week. Table 20 compares the restart use per driver across the on-duty hour categories.

Table 20: Seasonal Comparison of Restart Use per Driver by Average Weekly On-Duty Hours, Using 8-Day Work Week to Calculate Average Weekly On-Duty Hours for 16 For-Hire Motor Carriers (2012–2014)

		Average weekly on-duty hours					
Season	**Year**	**20 to less than 35 hours**	**35 to 55 hours**	**More than 55 to 65 hours**	**More than 65 to 75 hours**	**More than 75 hours**	**All drivers**
		Number of restarts taken per driver					
Winter	2012	7.3	8.1	10.8	12.1	12.2	**9.2**
Spring	2012	5.5	6.5	8.9	10.2	10.6	**7.5**
Summer	2012	5.9	6.8	9.0	10.3	11.1	**7.6**
Fall	2012	5.8	6.8	9.0	10.2	10.3	**7.6**
Winter	2013	7.5	8.0	10.4	12.1	12.4	**9.1**
Spring	2013	6.2	6.5	8.5	10.1	10.4	**7.4**
2011 hours-of-service rule goes into effect							
Summer	2013	5.1	6.4	8.4	10.3	11.3	**7.0**
Fall	2013	5.0	5.8	7.3	9.1	9.9	**6.2**
Winter	2014	6.3	7.7	9.4	11.6	13.0	**8.1**
Spring	2014	4.7	5.8	7.2	9.2	9.6	**6.3**
Summer	2014	5.0	6.3	7.8	9.9	10.7	**6.7**
Fall	2014	5.3	6.3	7.8	9.8	10.6	**6.7**

Source: GAO analysis of drivers' schedule data from 16 for-hire motor carriers. | GAO-15-641

Note: Excludes restarts greater than 72.99 hours.

We also analyzed restart use per driver per calendar week to understand how often drivers in our dataset use a restart. With this specification, we found that restart use after the 2011 HOS rule went into effect appears to vary by drivers' average weekly on-duty hours (see table 21). For drivers working on average more than 55 hours to 75 hours per 8-day work week, restart use per calendar week remained relatively constant before and after the 2011 HOS rule went into effect. For drivers working more than 75 hours per 8-day work week, restart use per calendar week appears to increase by almost 1. However, before the rule there were 69 drivers in this category and after the rule there were 15 drivers, therefore the change observed in Table 20 is based on a relatively small number of drivers. For drivers working 55 hours or less per 8-day work week, restart use per calendar week appears to decrease after the 2011 HOS rule went into effect.

Table 21:Seasonal Comparison of Restart Use per Driver per Calendar Week by Average Weekly On-Duty Hours—Using 8-Day Work Week to Calculate Average Weekly On-Duty Hours for 16 For-Hire Motor Carriers (2012–2014)

		Average weekly on-duty hours				
Season	**Year**	**20 to less than 35 hours**	**35 to 55 hours**	**More than 55 to 65 hours**	**More than 65 hours**	**All drivers**
		Number of restarts taken per driver per calendar week				
Winter	2012	0.56	0.62	0.83	0.93	**0.71**
Spring	2012	0.51	0.60	0.82	0.94	**0.69**
Summer	2012	0.53	0.61	0.81	0.92	**0.68**
Fall	2012	0.54	0.63	0.82	0.94	**0.70**
Winter	2013	0.58	0.62	0.81	0.94	**0.71**
Spring	2013	0.57	0.60	0.79	0.93	**0.69**
		2011 hours-of-service rule goes Into effect				
Summer	2013	0.46	0.57	0.75	0.92	**0.62**
Fall	2013	0.46	0.54	0.68	0.84	**0.58**
Winter	2014	0.49	0.60	0.73	0.91	**0.63**
Spring	2014	0.44	0.53	0.67	0.85	**0.58**
Summer	2014	0.45	0.57	0.70	0.89	**0.60**
Fall	2014	0.49	0.58	0.72	0.90	**0.62**

Source: GAO analysis of drivers' schedule data from 16 for-hire motor carriers. | GAO-15-641

Note: Excludes restarts greater than 72.99 hours.

We analyzed the length of restart periods to understand how many off-duty hours drivers in our dataset took. We found that before the 2011 HOS rule went into effect, about 8 percent of the restarts taken by drivers in our dataset were between 34 and 37 hours (see table 22). As a percentage of all restarts taken in each season, restarts with lengths of 34 to 37 hours decreased after the rule went into effect. The percentage of restarts with a length between 45 to 72.99 hours increased after the rule went into effect. As our dataset did not contain the start and end times of a driver's schedule or restarts, we are not able to analyze the extent to which there was a change in restarts with one- or two-night periods. For example, a driver working during daytime hours could take a 34 hour off-duty period that contains two nighttime periods (1 a.m. to 5 a.m.). In contrast, a driver working during nighttime hours may not have two nighttime periods (1 a.m. to 5 a.m.) when taking a 34-hour off-duty period. Therefore, we cannot determine the change in actual nights taken off-duty by looking at the length of a restart period on its own.

Table 22: Seasonal Comparison of Restart Length for 16 For-Hire Motor Carriers (2012–2014)

		Restart length										
		34 to 36.99 hours		37 to 44.99 hours		45 to 58.99 hours		59 to 72.99 hours		73 hours or more		
Season	Year	Number of restarts	Percent of restarts	Number of restarts	Percent of restarts	Number of restarts	Percent of restarts	Number of restarts	Percent of restarts	Number of restarts	Percent of restarts	Total restarts
Winter	2012	11,827	6.9	33,621	19.6	31,191	18.2	64,651	37.7	30,109	17.6	**171,399**
Spring	2012	10,012	7.1	27,306	19.2	25,563	18.0	52,514	37.0	26,487	18.7	**141,882**
Summer	2012	10,272	7.2	27,827	19.4	26,774	18.7	52,704	36.8	25,510	17.8	**143,087**
Fall	2012	10,834	7.5	29,997	20.8	25,353	17.6	51,604	35.8	26,519	18.4	**144,307**
Winter	2013	13,857	7.9	38,566	22.0	30,824	17.6	62,334	35.6	29,471	16.8	**175,052**
Spring	2013	11,553	7.9	30,232	20.7	24,535	16.8	52,886	36.3	26,588	18.2	**145,794**
2011 Hours of Service Rule Goes Into Effect												
Summer	2013	4,716	3.4	23,971	17.4	26,442	19.2	56,319	41.0	25,967	18.9	**137,415**
Fall	2013	4,525	3.6	23,356	18.6	22,726	18.1	49,495	39.5	25,356	20.2	**125,458**
Winter	2014	6,434	3.9	33,355	20.2	29,535	17.9	66,870	40.5	28,972	17.5	**165,166**
Spring	2014	4,858	3.7	24,097	18.2	22,311	16.8	54,550	41.1	26,847	20.2	**132,663**
Summer	2014	4,974	3.5	25,744	18.2	24,214	17.1	58,832	41.6	27,501	19.5	**141,265**
Fall	2014	5,051	3.6	26,705	19.0	22,624	16.1	59,597	42.4	26,513	18.9	**140,490**

Source: GAO analysis of drivers' schedule data from 16 for-hire motor carriers. | GAO-15-641

Analysis of Data before and after the HOS Rule

We also analyzed the data by dividing driver records into two periods: (1) before the rule went into effect—January 1, 2012, to June 15, 2013—and (2) after the rule went into effect—July 15, 2013, to December 15, 2014. As with the analysis by seasonal datasets, we applied filters to each year of data. To avoid double-counting drivers who appear across multiple years of data, we used the driver identification number to connect driver records with data in multiple years. The weekly on-duty hour average is calculated using the weeks that we have data for a driver. If a driver is only present in the 2012 data set, then that driver's average weekly on-duty hours are calculated based on only 2012 data. If a driver is present in both the 2012 and 2013 datasets, then that driver's average weekly on-duty hours are calculated using 2012 and 2013 (up to June 15, 2013) on-duty hours. The calculation for the post-rule dataset used the same method.

We compared the number and percentage of drivers in our dataset by average weekly on-duty hours before the rule went into effect and after the rule went into effect. As with the seasonal analysis, the percentage of drivers in our dataset working more than 55 hours per work week decreased, and the percentage of drivers in our dataset working 55 hours or less per 8-day work week increased. While this analysis shows similar trends in the data, there are differences in the number and percentage of drivers in each weekly on-duty hour group. Specifically, in the seasonal analysis, before the rule went into effect (winter 2012 through spring 2013) the percentage of drivers in our dataset working more than 65 hours per 8-day work week ranges from about 9 to close to 12 percent. As shown in table 23, approximately 8 percent of drivers in our dataset before the rule went into effect are categorized as working more than 65 hours per 8-day work week on average. This difference is likely due to the time frame over which the average weekly on-duty hours are calculated.

Table 23: Comparison of Drivers by Average Weekly On-Duty Hours before the 2011 Hours of Service (HOS) Rule Went into Effect and after, per 8-Day Work Week for 16 For-Hire Motor Carriers (2012–2014)

Average weekly on-duty hours	Before 2011 HOS rule went into effect		After 2011 HOS rule went into effect	
	Number of drivers	Percentage of drivers	Number of drivers	Percentage of drivers
20 to less than 35 hours	427	2.0	639	2.9
35 to 55 hours	11,737	54.9	14,659	65.7
More than 55 to 65 hours	7,444	34.8	6,359	28.5
More than 65 to 75 hours	1,758	8.2	663	3.0
More than 75 hours	16	0.1	2	0.0
Total drivers	**21,382**	**100**	**22,322**	**100**

Source: GAO analysis of drivers' schedule data from 16 for-hire motor carriers. | GAO-15-641

We also compared the number and percent of weeks before and after the rule went into effect categorized by total weekly on-duty hours (see table 24). Consistent with the seasonal analysis, the percent of 8-day work weeks with total on-duty hours of more than 65 also decreased.

Table 24: Comparison of Weeks By Weekly On-Duty Hours before the 2011 Hours of Service (HOS) Rule Went into Effect and after, per 8-Day Work Week for 16 For-Hire Motor Carriers (2012–2014)

Total weekly on-duty hours	Before 2011 HOS rule went into effect		After 2011 HOS rule went into effect	
	Number of weeks	Percentage of weeks	Number of weeks	Percentage of weeks
20 to less than 35 hours	73,471	7.7	95,929	9.5
35 to 55 hours	342,747	35.7	404,685	40.1
More than 55 to 65 hours	283,974	29.6	292,619	29.0
More than 65 to 75 hours	181,453	18.9	154,228	15.3

Total weekly on-duty hours	Before 2011 HOS rule went into effect		After 2011 HOS rule went into effect	
	Number of weeks	Percentage of weeks	Number of weeks	Percentage of weeks
More than 75 hours	44,789	4.7	19,730	2.0
Total weeks[a]	**959,529**	**100**	**1,008,537**	**100**

Source: GAO analysis of drivers' schedule data for 16 for-hire motor carriers. | GAO-15-641

[a]The number of weeks in each period (pre-rule and post-rule) does not add up to the total number of weeks. In each period, some of the weeks did not have 20 or more on-duty hours.

We also analyzed the length of restarts used by drivers' average weekly on-duty time. We found that before the rule went into effect the percentage of 34- to 36.99-hour restarts taken by drivers in our dataset in each on-duty time category ranges from about 6 to 11-percent (see table 25). After the HOS rule went into effect, the relative distribution of restarts by length taken in each weekly on-duty category changed. Specifically, the percentage of restarts lasting 34- to 36.99- hours decreased for all driver categories. The percentage of 34- to 36.99-hour restarts taken by drivers in our dataset working the longest hours decreased from close to 11 percent to about 9 percent. The percentage of 34- to 36.99-hour restarts taken by the other driver groups saw larger decreases. As with our seasonal analysis of restart length, our results do not allow us to state that the number of night time periods in a restart period changed after the 2011 HOS rule went into effect.

Table 25: Comparison of Restart Length by Weekly On-Duty Hours before the 2011 Hours-of-Service (HOS) Rule Went into Effect and after, per 8-day Work Week for 16 For-Hire Motor Carriers (2012–2014)

Average weekly on-duty hours	Restart length	Before 2011 HOS rule went into effect		After 2011 HOS rule went into effect	
		Number of restarts	Percentage of restarts	Number of restarts	Percentage of restarts
20 to less than 35 hours	34-36.99 hours	784	6.2	366	2.3
	37 - 44.99 hours	3,545	28.2	2,804	17.6
	45 - 58.99 hours	2,912	23.1	4,018	25.3
	59 - 72.99 hours	5,340	42.4	8,713	54.8
	Total Restarts	**12,581**	**100**	**15,901**	**100**
35 to 55 hours	34-36.99 hours	35,097	9.2	17,788	4.0
	37 - 44.99 hours	99,301	26.0	95,914	21.4
	45 - 58.99 hours	75,481	19.8	90,696	20.2
	59 - 72.99 hours	171,494	45.0	244,078	54.4
	Total Restarts	**381,373**	**100**	**448,476**	**100**

Average weekly on-duty hours	Restart length	Before 2011 HOS rule went into effect: Number of restarts	Before 2011 HOS rule went into effect: Percentage of restarts	After 2011 HOS rule went into effect: Number of restarts	After 2011 HOS rule went into effect: Percentage of restarts
More than 55 to 65 hours	34-36.99 hours	29,994	8.8	12,836	5.0
	37 - 44.99 hours	79,813	23.3	63,486	24.9
	45 - 58.99 hours	73,208	21.4	59,016	23.2
	59 - 72.99 hours	158,988	46.5	119,305	46.9
	Total Restarts	**342,003**	**100**	**254,643**	**100**
More than 65 hours	34-36.99 hours	9,198	10.7	2,610	8.5
	37 - 44.99 hours	22,210	25.7	10,768	34.9
	45 - 58.99 hours	26,600	30.8	9,182	29.7
	59 - 72.99 hours	28,313	32.8	8,316	26.9
	Total Restarts	**86,321**	**100**	**30,876**	**100**

Source: GAO analysis of drivers' schedule data from 16 for-hire carriers. | GAO-15-641

Note: Excludes restarts greater than 72.99 hours.

Statistical Analysis of Driver Schedule Data

We also analyzed whether there were differences in how the same set of drivers behaved before and after the 2011 HOS rule went into effect. To do this analysis, we developed several statistical regression models of hours worked and restart use, in order to further account for seasonal variation and differences across drivers who do not change over time, such as carrier. Using these models, we estimated whether the differences in hours worked and restart use before and after the 2011 HOS rule went into effect were statistically distinguishable from zero. We applied these models to a sample of 6,934 drivers who met our filters for the entire period between January 1, 2012, and December 31, 2014.

We estimated two sets of models, which made different assumptions about the distribution of the outcomes and the structure of the correlation of those outcomes within drivers over time. Both sets of models included season or month fixed effects and predicted hours worked per week; working 35 to 54, 55 to 64, or 65 or more hours per 8-day week; or the number of restarts per week.

First, we estimated linear models with driver fixed effects (implemented through a within-transformation) and standard errors that were robust to arbitrary forms of heteroskedasticity and driver-level autocorrelation. These adjustments corrected for heteroskedasticity caused by using linear models to analyze binary outcomes and for autocorrelation caused by analyzing repeated outcomes within drivers over time.

Second, we estimated generalized estimating equations (GEE) models, with outcomes within drivers having exchangeable or AR1 covariance structures. These covariance models were appropriate for the 3-year period, when the autocorrelation of driving schedules should be relatively consistent across days (exchangeable) or decline linearly (AR1). The models assumed the outcomes to be distributed normally, binomially, or negative binomially, depending on the scale, and used canonical link functions.

We summarize the results of our analysis as ranges across models in the estimated contrast between time periods before and after the policy change. Most of our findings were robust to these plausible model assumptions.

Results: Hours Worked per 8-Day Week

Drivers in our analysis sample worked approximately 1.1 to 2.5 fewer hours per 8-day week, on average, after the HOS rule was implemented than before, depending on model assumptions. Expressed as a proportion, the differences ranged from approximately 2.0 to 4.8 percent fewer hours. These differences were statistically distinguishable from zero at the 0.001 level.

Drivers in our analysis sample were approximately 24 to 29 percent less likely to work 65 hours or more per 8-day work week after the HOS rule was implemented, depending on model assumptions.[6] Similarly, drivers were approximately 12 percent more likely to work 35 to 54 hours per 8-day work week. All of these differences were statistically distinguishable from zero at the 0.001 level. The likelihood of working 55 to 64 hours per 8-day work week after the rule went into effect changed by less than 2 percent but the statistical significance and direction of the change varied according to model assumptions.

[6]These estimates are proportional differences in the odds of working at least 65 hours per week, derived from odds-ratios. When scaled as probabilities, the estimated differences ranged from -4.6 to -5.4 percentage points, with an overall rate of 19.8 percentage points. Estimates for the likelihood of working 35 to 54 and 55 to 64 hours per week are defined analogously.

Results: Number/Rate of Restarts

Drivers took approximately 6.1 to 6.5 percent fewer restarts per 8-day week, on average, after the HOS rule was implemented than before, depending on model assumptions and adjusting for variation in driving intensity (as a measure of exposure). These differences for the overall analysis population of drivers were statistically distinguishable from zero at the 0.001 level.

Appendix V: Biomathematical Models of Fatigue

GAO was asked to assess the safety impacts of the 2011 hours of service (HOS) rule, which are predicated on reducing driver fatigue. To assess whether the rule change could result in less fatigued drivers and potentially fewer fatigue-related crashes, we used a biomathematical fatigue model—the Fatigue Audit InterDyne™ (FAID) model—to assess the risk of driver fatigue for schedules that comply with the 2011 HOS rule and similar schedules that do not.[1]

The FAID model provides a fatigue score based on the start and end time of a work shift. The higher the FAID score, the higher the risk of fatigue. A standard work week of 40 hours, Monday to Friday, 9 a.m. to 5 p.m. results in a peak FAID score of 41. In contrast, a 40-hour work week from 11 p.m. to 7 a.m. results in a peak FAID score of 97. Research indicates that scores above 80 are comparable to the fatigue-related impairment found in individuals with a blood alcohol level of over 0.05 percent, above the legal limit in many countries.[2] The examples described below, both those based on hypothetical driver schedules and those based on interviews with motor carriers, differ significantly from a typical 9 a.m. to 5 p.m. 40-hour work week.

The schedules we modelled below often involve working long hours—60 or more hours per week—and working overnight. As a result, many of the peak fatigue scores shown below are well above the range considered safe. The purpose of our analysis is to show the relative difference in peak fatigue scores for schedules in operation after the rule went into effect and those in operation before the HOS rule went into effect.

We created different scenarios through which to test the risk of driver fatigue given rule-induced alterations to driver schedules. In some of these scenarios, we compare similar schedules in terms of the number and timing of weekly on-duty hours but that differ in terms of their compliance with the:

[1]We previously discussed the use of biomathematical fatigue models, including the FAID model, in GAO, *Freight Railroad Safety: Hours of Service Changes Have Increased Rest Time, but More Can Be Done to Address Fatigue Risks*, GAO-11-853 (Sept. 28, 2011).

[2]A. Fletcher, N. Lamond, C. van den Heuvel, and D. Dawson, "Predication of performance during sleep deprivation and alcohol intoxication by a quantitative model of work-related fatigue," *Sleep Research Online*, 5(2) (2003). The study found that a fatigue score of 80 is comparable to the impairment that would be observed in an individual with a blood alcohol concentration of 0.09 percent or greater.

- Two-night provision (two 1 a.m. to 5 a.m. periods in a restart)
- 168-hour limit (limits restarts to once every 168 hours)

We also created scenarios based on interviews with motor carriers and drivers. These interview-based scenarios include the schedules interviewees reported using before the 2011 HOS rule went into effect and schedules interviewees said they started using in order to comply with the 2011 HOS rule. These schedules demonstrate real world adaptations to the 2011 HOS rule as described to us. Our analysis is not generalizable to the entire industry, but rather is intended to illustrate how HOS schedule changes can affect the risk of driver fatigue.

Of the schedules we simulated, some did not require a change after the rule went into effect. For those that did require a schedule change, figures 8 through 11 below show the daily peak fatigue score when complying with the two-night provision and 168-hour limit in the 2011 HOS rule and the daily peak fatigue score for similar schedules that do not comply with one or both of those provisions in the 2011 HOS rule.

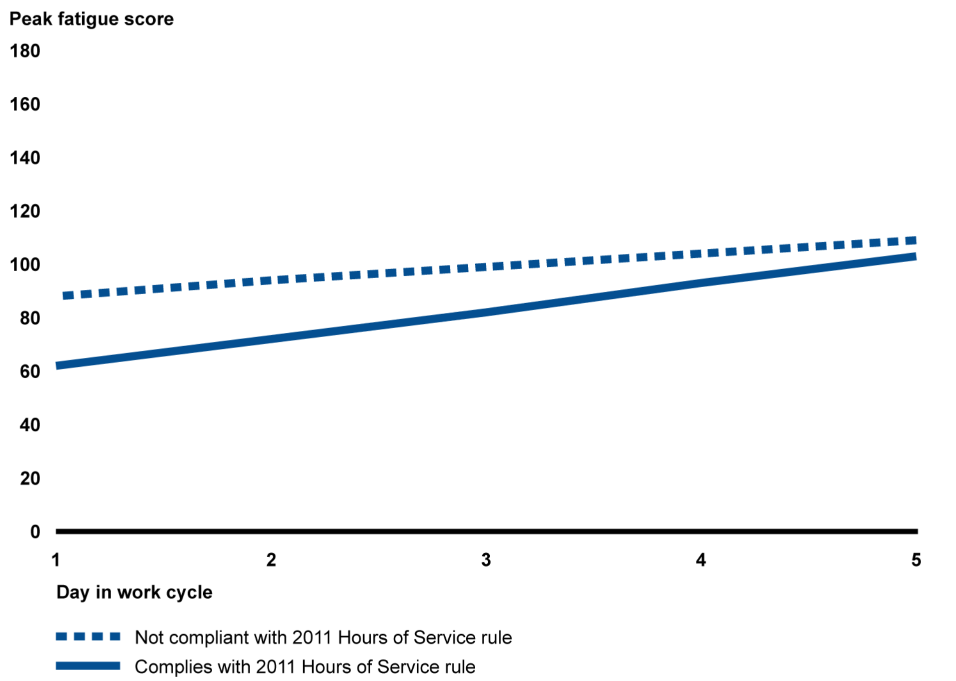
Simulated Schedules for Drivers Working Maximum Allowed Hours during Daytime Hours

Drivers can work up to 14 hours per day until reaching 70 hours—at which point they must either take a restart of at least 34 hours or remain off-duty until their consecutive 8-day sum of hours is less than 70. Therefore, we modelled a maximum daytime schedule with the following parameters:

- 14-hour shifts on-duty from 6 a.m. to 8 p.m. over five consecutive days

Under the previous HOS rule, a driver following this schedule would only need to take 34 hours off-duty to have a valid restart. However, after the rule went into effect and to comply with the 168-hour limit, a driver with this schedule must take 2 days off between work cycles. For example, the driver can work 14 hours between 6 a.m. and 8 p.m. on Monday through Friday, but must be off-duty Saturday and Sunday. This results in 58 hours off-duty. This schedule change—2 days off between work cycles instead of one day off—effectively lowers the average number of hours a driver can work. In terms of work time per 7-day period, the schedule that complies with the 168-hour limit averages 72 hours and the schedule that does not comply averages 82 hours. As shown in figure 8, the peak daily fatigue scores for the schedule that complies with the 2011 HOS rule (the schedule with 2 days off) are lower than the peak daily fatigue scores for the schedule that does not comply with the 2011 HOS rule. The peak daily fatigue scores for both of these schedules are high and above the levels generally considered safe on several days that are modelled.

Figure 8: Comparison of Peak Daily Fatigue Scores for Two Hypothetical Schedules Operating at the Maximum Allowed On-Duty Hours during Daytime Hours—One Schedule Complies with the 2011 Hours of Service Rule and the Other Does Not



Source: GAO analysis of simulated schedules using Fatigue Audit InterDyne (FAID) model. | GAO-15-641

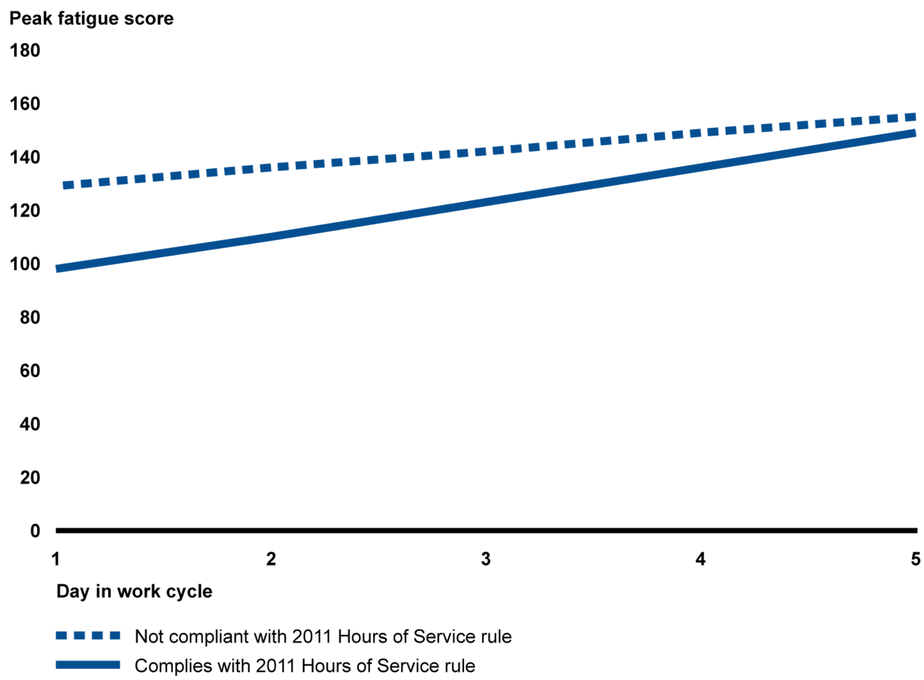
Simulated Schedules for Drivers Working Maximum Allowed Hours during Nighttime Hours

We also modelled a driver working the maximum allowed hours but during nighttime hours. The hypothetical schedule had the following parameters:

- 14-hour shifts on-duty from 8 p.m. to 10 a.m. over 5 consecutive days

Under the previous HOS rule, the driver only had to take off 34 hours between the end of the work shift at 10 a.m. on Saturday and the start of the next work shift at 8 p.m. on Sunday, but this schedule violates both the two-night provision and the 168-hour limit because the off-duty period does not contain two periods from 1 a.m. to 5 a.m. and begins less than 168 hours after the prior restart period began. Thus, to comply with both the two-night provision and the168-hour limit in the 2011 HOS rule, a driver with this schedule must take 2 days off between work cycles. For example, the driver can work a 14-hour shift between 8 p.m. and 10 a.m. on Monday through Friday nights, but must be off-duty Saturday and Sunday nights to comply with the 2011 HOS rule. This results in an off-duty period of 58 hours. As shown in figure 9, the peak daily fatigue scores for the schedule used after the rule went into effect (the schedule with two days off) are lower than the peak daily fatigue scores for the schedule in use before the rule went into effect. The peak daily fatigue scores for both of these schedules are high, above the levels generally considered safe.

Figure 9: Comparison of Peak Daily Fatigue Scores for Two Hypothetical Schedules Operating at the Maximum Allowed On-Duty Hours during Nighttime Hours—One Schedule Complies with the 2011 Hours of Service Rule and the Other Does Not



Source: GAO analysis of simulated schedules using Fatigue Audit InterDyne (FAID) model. | GAO-15-641

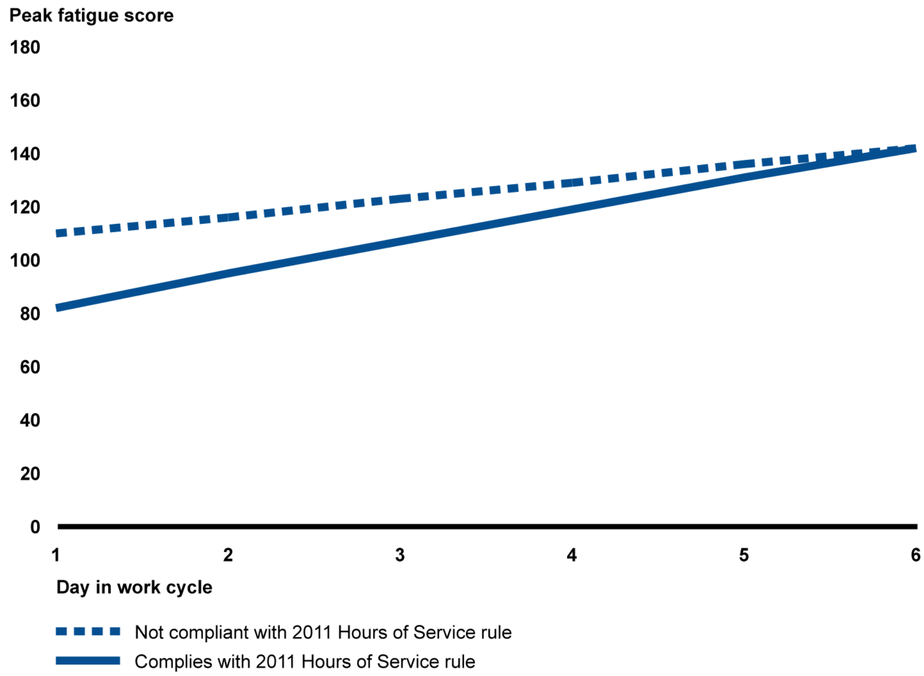
Simulated Schedules for Drivers Working 60- to 70-Hour Shifts during Nighttime Hours

Of the schedules we simulated, two additional hypothetical schedules needed to change to comply with the 2011 HOS rule. These hypothetical schedules had the following parameters:

- 10- to 12-hour on-duty shifts between 10 p.m. to 10 a.m. over 6 consecutive days (70 hours over a 6-day work cycle)
- 10-hour on-duty shifts from 12 a.m. to 10 a.m. over 6 consecutive days (60 hours over a 6-day work cycle)

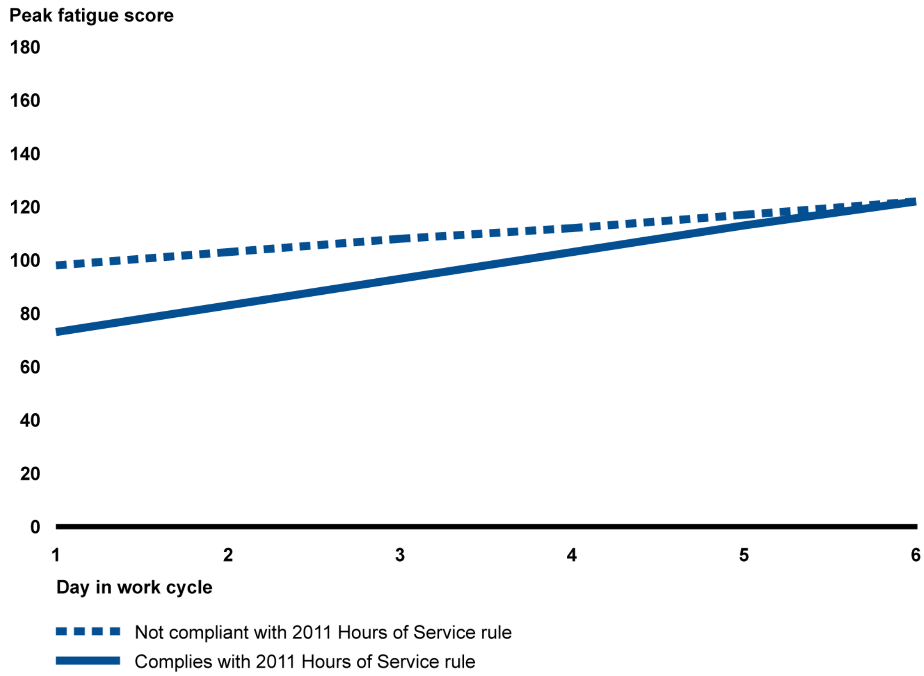
To comply with both the two-night provision and the168-hour limit in the 2011 HOS rule, drivers with these schedules must take 2 days off between work cycles instead of one day off as allowed under the prior HOS rule. For example, a driver can work a 10-hour shift between 12 a.m. and 10 a.m. on six consecutive nights but must be off-duty the next two nights. These two consecutive nights off allow this schedule to comply with the two-night provision, which requires two 1 a.m. to 5 a.m. periods in an off-duty period to qualify as a restart. As a result, a driver working 10-hour shifts on six consecutive nights (60 hours in a work cycle) must take an off-duty period of 52 hours. A driver working 10-to-12-hour shifts on six consecutive nights (70 hours in a work cycle) must take an off-duty period of 62 hours. As shown in figures 10 and 11, the peak daily fatigue scores for the schedules in use after the rule went into effect (the schedules with 2 days off) are lower than the peak daily fatigue scores for the schedules in use before the rule went into effect. The peak daily fatigue scores for the schedules shown in figures 10 and 11 are high, above the levels generally considered safe.

Figure 10: Comparison of Peak Daily Fatigue Scores for Two Hypothetical Schedules Operating at 70 On-Duty Hours during Nighttime Hours—One Schedule Complies with the 2011 Hours of Service Rule and the Other Does Not



Source: GAO analysis of simulated schedules using Fatigue Audit InterDyne (FAID) model. | GAO-15-641

Figure 11: Comparison of Peak Daily Fatigue Scores for Two Hypothetical Schedules Operating at 60 On-Duty Hours during Nighttime Hours—One Schedule Complies with the 2011 Hours of Service Rule and the Other Does Not



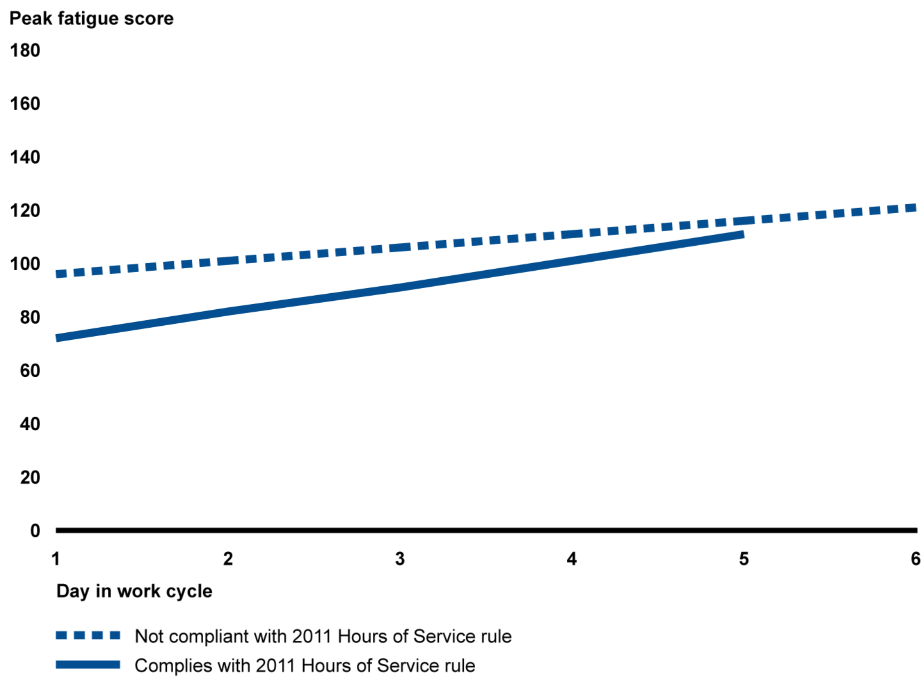
Source: GAO analysis of simulated schedules using Fatigue Audit InterDyne (FAID) model. | GAO-15-641

We also used the FAID model to estimate the risk of driver fatigue based on schedule changes described in our interviews with an industry stakeholder, motor carriers, and drivers. Figures 12 through 15 below show the daily peak-fatigue score for the schedules interviewees adopted in order to comply with the 2011 HOS rule and the daily peak-fatigue score for the schedules they used before the 2011 HOS rule went into effect.

Simulating Schedule Changes due to the Two-Night Provision Reported by Drivers

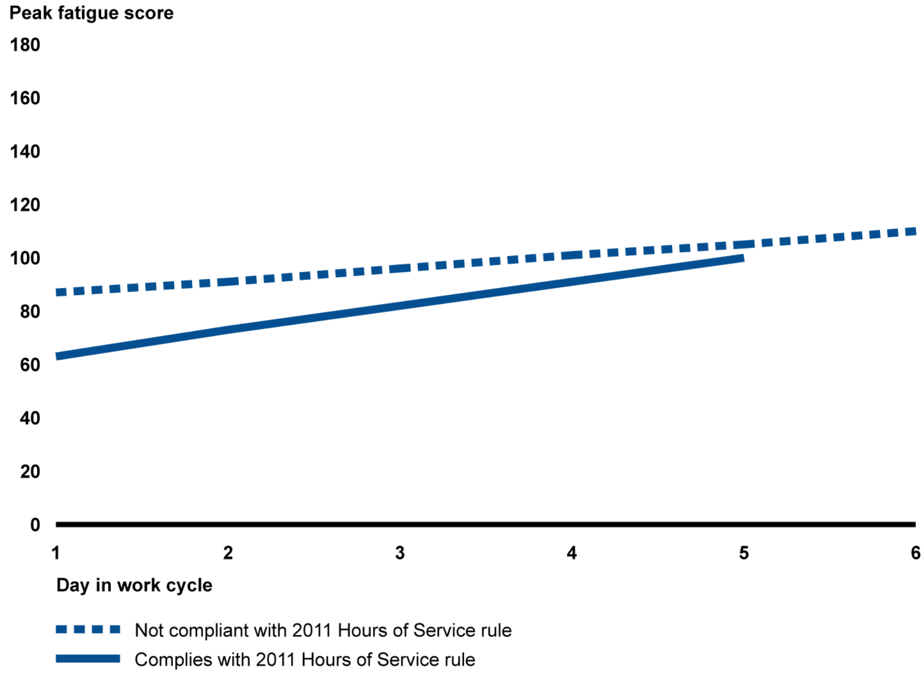
Representatives of a motor carrier and an industry association we spoke with reported changing the schedules of drivers working over-night or in the early morning hours to ensure that drivers could continue to take the restart and comply with the two-night provision in the 2011 HOS rule. In the first example, drivers worked six consecutive10-hour shifts per week from 12:30 a.m. to 10:30 a.m. before the rule went into effect. Over 7 days, their schedules averaged 61 on-duty hours. These drivers generally began and ended their shifts at the same time each day. After the rule went into effect, the motor carrier reduced the number of shifts per driver from six consecutive shifts per week to five consecutive shifts per week. This change was made to ensure drivers would have two 1 a.m. to 5 a.m. periods during their off-duty period and could still count the off-duty time as a restart. In the second example, drivers worked 10-hour shifts from 3:00 a.m. to 1:00 p.m. over six consecutive days. Similar to the first example, drivers working this schedule averaged 61 hours on-duty over 7 days. To ensure drivers could take a restart that complied with the two-night provision, the motor carrier reduced the total number of shifts per driver from six consecutive days to five consecutive days. As shown in figures 12 and 13, the schedules that comply with the 2011 HOS rule have lower peak-fatigue scores for each day a driver works this schedule. The lower peak-fatigue scores suggest a lower risk of fatigue for a driver working according to the schedule complying with the 2011 HOS rule than the driver working according to the schedule that does not comply with the rule. The peak daily-fatigue scores for the schedules shown in figures 12 and 13 are high, above the levels generally considered safe on many of the days modelled.

Figure 12: Schedule Change to Comply with Two-Night Provision—Peak Fatigue Scores of Driver Schedules with Overnight Shifts



Source: GAO analysis of simulated schedules using Fatigue Audit InterDyne (FAID) model. | GAO-15-641

Figure 13: Schedule Change to Comply with Two-Night Provision—Peak Fatigue Scores of Driver Schedules with Early Morning to Afternoon Shifts

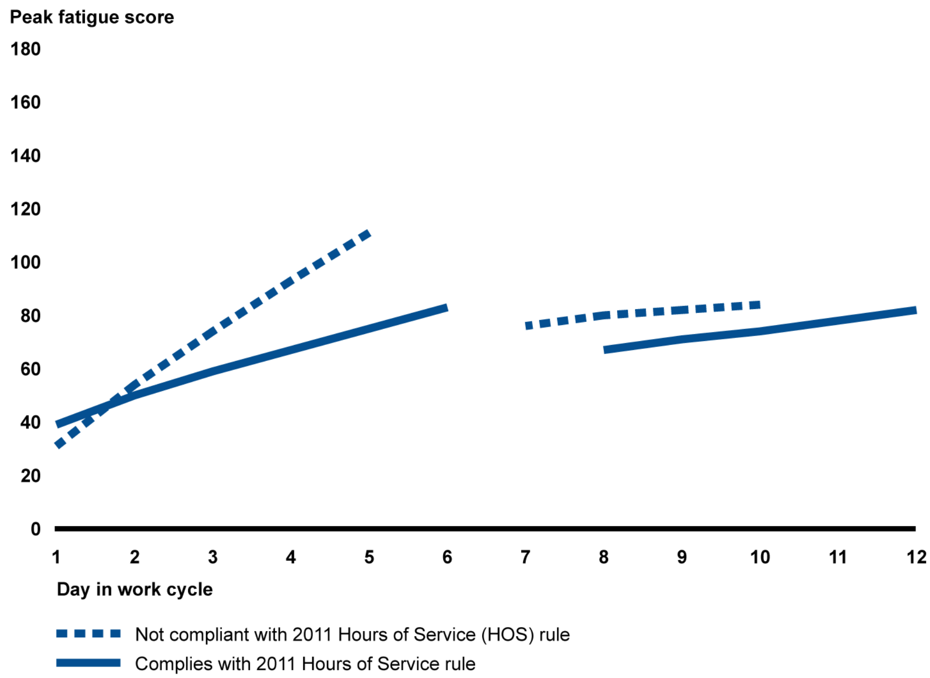


Source: GAO analysis of simulated schedules using Fatigue Audit InterDyne (FAID) model. | GAO-15-641

Simulating Schedule Changes due to the 168-Hour Limit Reported by Drivers

Representatives of other motor carriers we spoke with described schedule changes they made to ensure off-duty time complied with the 168-hour limit. For example, a representative of one motor carrier told us that a driver would typically work a long-haul route over 10 days. The driver would work between 10- and 14-hours per day for 5 days, reach the drop-off destination and go off-duty for 34 hours (a restart). Then the driver would drive home. Once home, the driver would take 3 days off which would count as a restart. To comply with the 168-hour limit and ensure the off-duty period would count as a restart, this schedule would have to change to ensure the off-duty periods began at least 7 days apart (168 hours). To do this, the driver could work 10- to 12-hour days for 6 days, take a 34-hour break on day 7, then drive home and take a shorter break, 2 days in this example. As shown in figure 14, complying with the 168-hour limit results in generally lower peak fatigue scores and therefore a lower risk of driver fatigue.

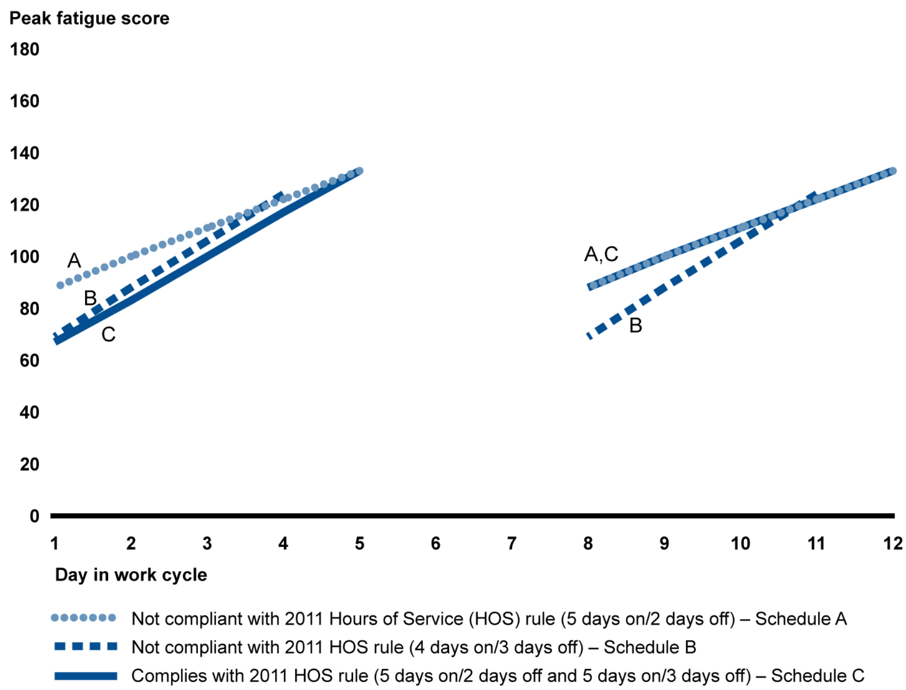
Figure 14: Schedule Change to Comply with 168-hour Limit—Peak Fatigue Scores of Driver Schedules with Differential Restart Use



Source: GAO analysis of simulated schedules using Fatigue Audit InterDyne (FAID) model. | GAO-15-641

In the second example, representatives of a motor carrier told us about short-haul drivers working overnight shifts whose schedules were changed to comply with the 168-hour limit. Before the rule went into effect, drivers could choose between two possible schedules. The first schedule (schedule A) had drivers working 12-hour shifts for 5 consecutive days, in our simulation from 8 p.m. to 8 a.m. Then drivers took the next 2 days off-duty and counted this time as a restart. The second schedule option (schedule B) had drivers working 13-hour shifts for 4 consecutive days, in our simulation from 8 p.m. to 9 a.m. These drivers then took the next 3 days off-duty and counted this time as a restart. Due to the 168-hour limit, the motor carrier switched drivers to the following schedule (schedule C): 12-hour shifts for 5 consecutive days with 2 days off followed by 12-hour shifts for 5 consecutive days with 3 days off. We compared the two schedules offered before the rule went into effect to the schedule offered after the rule was effective. As shown in figure 15, we found that schedule C, which complied with the 2011 HOS rule, had similar peak fatigue scores to Schedule A with the 5 days on-duty with 2 days off-duty schedule. Schedule C, the 2011 HOS compliant schedule, had slightly higher peak-fatigue scores than Schedule B, the 4 days on duty with 3-days off-duty schedule. This suggests that the schedule changes made by this carrier to comply with the 168-hour limit did not consistently lower the risk of fatigue faced by drivers working this schedule.

Figure 15: Schedule Change to Comply with the 168-Hour Limit—Peak Daily Fatigue Scores for Three Overnight Schedules with Differential Restart Use



Source: GAO analysis of simulated schedules using Fatigue Audit InterDyne (FAID) model. | GAO-15-641

Appendix VI: GAO Analysis of Vehicle Count Data

As was discussed in the body of our report, on July 1, 2013, the Federal Motor Carrier Safety Administration (FMCSA) began to enforce the 2011 hours of service (HOS) rule that made several key changes to the number of hours commercial drivers can work and drive per day and week. Specifically, the new rule altered when and how commercial drivers are permitted to "restart" the maximum number of hours they can work over a 7- or 8-day period, and required drivers to take a 30-minute, off-duty break during shifts that last more than 8 hours. FMCSA understood that when it designed this rule, it would result in changes to drivers' schedules. Specifically, FMCSA believed the rule would reduce driving time for drivers working 65 or more hours per week, but assumed that these hours would be shifted to other drivers or to other workdays rather than being eliminated altogether. While FMCSA asserted that total driving time for some individual drivers was likely to drop slightly due to the HOS rule, no attempt was made to quantify the effects of the rule on congestion. Stakeholders we spoke with were concerned that FMCSA did not adequately assess how the new HOS rule would impact traffic patterns, especially whether the rule would result in increased traffic congestion during morning hours, i.e., between 5 a.m. and 9 a.m.

To evaluate changes in commercial vehicle traffic before and after the 2011 HOS rule went into effect on July 1, 2013, we collected and analyzed data from the Federal Highway Administration's (FHWA) Weigh-in-Motion (WIM) and Classification database. This database includes information on non-commodity-carrying vehicles, including passenger vehicles, and commodity-carrying (i.e., commercial) motor vehicles through data sensors that are installed in roadways. These sensors allow states and FHWA to collect data on time and date, lane, speed, vehicle classification, and vehicle length, among other variables.

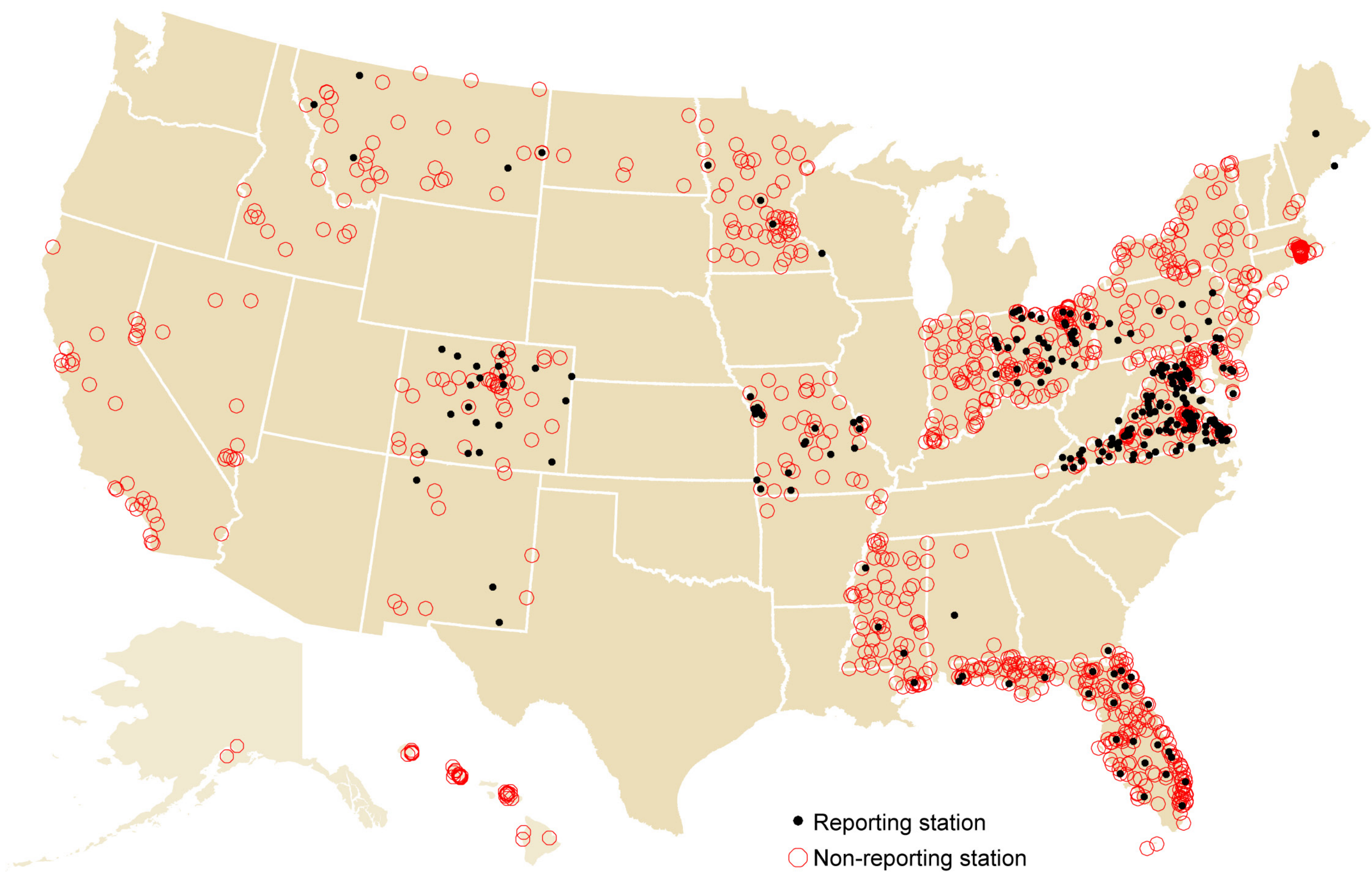
Our analysis of this data was limited to two time periods before and after the rule went into effect. Specifically, we chose to evaluate data from November through December 2012 and November through December 2013 for the following reasons:

- We wanted to allow sufficient time for carriers to adapt to the new rule, including preparing for the change prior to July 1, 2013 and adapting to the change after July 1, 2013.
- We wanted to control for any potential seasonal effects within the commercial carrier industry by including the same months prior to and after the HOS rule went into effect.

- We wanted to control for unusual events that could have impacted the commercial motor vehicle industry. For example, we excluded the time period from January and February 2014, which according to motor carriers and drivers we spoke with, had a large number of winter storms that severely impacted the industry. We also excluded the end of October 2012 to avoid capturing potential impacts on the industry from Superstorm Sandy.

We worked with FHWA officials to provide us with WIM data that fell into these timeframes. In order to ensure we were comparing data from similar locations between the two time periods we requested data only from stations that reported data for all days and hours between November through December 2012 and November through December 2013. Because not all stations reported data for every day or hour during these timeframes, our analysis was restricted to 324 stations from 14 states out of 987 total stations across the country (approximately 33 percent). (See figure 16).

Figure 16: Stations Reporting and Not Reporting Data for GAO's Analysis of Vehicle Counts



Source: GAO analysis. | GAO-15-641

We also asked FHWA officials to filter the data for us into variables most useful for our analysis. For example, although WIM data include information on direction of travel and specific lane of travel, we were primarily interested in total vehicle counts at each station, to capture traffic volume. As a result we requested vehicle count totals for each station by month, day, and hour. In addition, FHWA grouped the data into two distinct categories: (1) non-commodity-carrying vehicles, and (2) commodity-carrying vehicles.

We analyzed these data by combining individual station data and sorting them by hour of day before and after the HOS rule went into effect. We did this analysis for all days of the week, weekdays, and weekends to account for any change in traffic patterns caused by time of week. We also compared traffic patterns for non-commodity-carrying vehicles and commodity-carrying vehicles to determine whether one of those populations was behaving differently than the other and account for possible factors that might be influencing both groups, such as the economy.

Appendix VII: GAO Analysis of Crash Data

The Federal Motor Carrier Safety Administration (FMCSA) promulgated the 2011 hours of service (HOS) rule on the basis that it would improve safety due to a reduction in fatigue-related crashes. At the same time, however, some industry stakeholders, motor carriers, and drivers that we spoke with believe that the HOS rule has had some unintended effects that may actually lead to a decrease in safety, particularly because they anticipate increased congestion during certain early morning hours when roads tend to be congested. To analyze the safety effects of the 2011 HOS rule, we undertook two separate analyses to (1) assess whether there has been any apparent effect of the 2011 HOS rule on the numbers of motor carrier crashes and (2) whether the rule change appeared to affect the relative number of crashes that occurred after the rule's implementation between 5 a.m. and 9 a.m. We used monthly crash data from the Motor Carrier Management Information System (MCMIS) from 2008 through 2014, supplied by FMCSA, to address the first question, and individual crash data from the same source to address the second.[1]

Trends in the Numbers of Crashes

Table 26 shows information for each month from January 2008 to September 2014, on the numbers of total crashes involving motor carriers, as well as the number of crashes with injuries and the numbers of fatal crashes. As can be seen, the numbers of crashes show a good deal of variation over time, but there appears little by way of any patterned variation, or trend. For example, the monthly numbers of total crashes averaged 7,042, but ranged from a low of 5,593 in February 2009 to a high of 9,887 in January 2014. The monthly numbers of crashes with injuries and fatal crashes also showed a good deal of variation.

[1]FMCSA has authority to regulate motor carriers that travel between states or "interstate" motor carriers as well as all motor carriers that transport hazardous material. In addition, our report focused on large trucks transporting freight and based in the United States. Therefore, our dataset only included data from interstate and intrastate hazardous material vehicles that transport freight and are U.S.-based. We also filtered the data in several additional ways for data reliability purposes.

Table 26: Numbers of Total Crashes, Crashes with Injuries, and Crashes with Fatalities, by Month, from January 2008 to September 2014

Month/Year	Total crashes	Crashes with injuries	Crashes with fatalities
Jan - 2008	8,734	3,229	236
Feb - 2008	8,940	3,263	261
Mar - 2008	7,324	2,883	182
Apr - 2008	6,899	2,806	216
May - 2008	6,692	2,689	222
Jun - 2008	6,831	2,782	248
Jul - 2008	6,700	2,790	257
Aug - 2008	6,931	2,882	228
Sep - 2008	6,537	2,725	238
Oct - 2008	7,321	2,975	247
Nov - 2008	6,394	2,438	190
Dec - 2008	8,760	3,009	227
Jan - 2009	7,682	2,717	208
Feb - 2009	5,593	2,111	146
Mar - 2009	6,072	2,375	169
Apr - 2009	5,710	2,324	211
May - 2009	5,599	2,291	161
Jun - 2009	6,033	2,589	194
Jul - 2009	6,086	2,549	210
Aug - 2009	6,080	2,512	206
Sep - 2009	5,800	2,463	180
Oct - 2009	6,823	2,577	190
Nov - 2009	5,804	2,276	193
Dec - 2009	7,520	2,728	215
Jan - 2010	7,158	2,555	195
Feb - 2010	7,225	2,617	177
Mar - 2010	6,509	2,527	201
Apr - 2010	6,255	2,594	214
May - 2010	6,407	2,569	195
Jun - 2010	6,795	2,679	213
Jul - 2010	6,693	2,759	199
Aug - 2010	6,739	2,873	255
Sep - 2010	6,819	2,827	240
Oct - 2010	6,946	2,812	204

Month/Year	Total crashes	Crashes with injuries	Crashes with fatalities
Nov - 2010	6,938	2,649	243
Dec - 2010	7,950	2,884	223
Jan - 2011	7,829	2,794	203
Feb - 2011	7,554	2,611	173
Mar - 2011	6,588	2,543	207
Apr - 2011	6,367	2,484	186
May - 2011	6,455	2,531	197
Jun - 2011	6,831	2,749	226
Jul - 2011	6,361	2,571	208
Aug - 2011	7,021	2,936	258
Sep - 2011	6,861	2,673	244
Oct - 2011	7,148	2,845	226
Nov - 2011	7,152	2,694	211
Dec - 2011	6,932	2,543	221
Jan - 2012	7,138	2,618	195
Feb - 2012	6,175	2,364	201
Mar - 2012	6,876	2,666	232
Apr - 2012	6,364	2,458	212
May - 2012	6,973	2,802	221
Jun - 2012	7,129	2,801	211
Jul - 2012	6,613	2,513	219
Aug - 2012	7,254	2,781	257
Sep - 2012	6,578	2,580	238
Oct - 2012	7,530	2,957	253
Nov - 2012	6,986	2,599	227
Dec - 2012	7,665	2,682	205
Jan - 2013	7,988	2,874	222
Feb - 2013	7,192	2,478	179
Mar - 2013	7,432	2,683	206
Apr - 2013	6,945	2,597	192
May - 2013	7,348	2,783	246
Jun - 2013	7,051	2,738	248
Jul - 2013	7,169	2,851	204
Aug - 2013	7,398	2,893	224
Sep - 2013	7,283	2,948	227
Oct - 2013	7,882	2,925	235

Month/Year	Total crashes	Crashes with injuries	Crashes with fatalities
Nov - 2013	7,554	2,703	225
Dec - 2013	8,573	2,911	244
Jan - 2014	9,887	3,207	217
Feb - 2014	8,708	2,824	162
Mar - 2014	7,734	2,725	206
Apr - 2014	7,147	2,524	197
May - 2014	7,147	2,763	205
Jun - 2014	7,367	2,797	231
Jul - 2014	7,396	2,822	209
Aug - 2014	6,924	2,664	172
Sep - 2014	6,593	2,529	177
Total	**570,397**	**218,362**	**17,253**
Monthly Average(mean)	**7,042**	**2,696**	**213**

Source: GAO analysis of Motor Carrier Management Information System (MCMIS) data. | GAO-15-641

Note: This dataset only includes data from interstate and intrastate hazardous-material motor carriers that transport freight and are based in the United States.

Many potential factors can influence the number of crashes over time. For example, as the economy grows, it is reasonable to expect more goods and services are traded, resulting in more freight needing to be transported. Given the size of roads is unlikely to change much in the near term, more traffic may result in more congestion, and more congestion may result in more accidents. To assess whether the implementation of the 2011 HOS rule had an effect on crashes we developed two ordinary least squares regression models that controlled for trucking volume and seasonal variation, and in one case also controlled for the unusual winter weather in December 2013 to February 2014. (See table 27.)

Table 27: Models Estimating Differences in the Numbers of Total Crashes, Crashes with Injuries, and Fatal Crashes before and after the Rule Change, after Adjusting for Quarter and Trucking Output

	Model 1			Model 2		
	All crashes			All crashes		
Variable	Parameter estimate	standard error	P	Parameter estimate	standard error	P
Intercept	1,499.29	1,041.86	0.1543	1,584.57	953.81	0.1009
Q2	-761.05	170.99	<.0001	-632.93	159.84	0.0002
Q3	-780.65	172.62	<.0001	-610.86	163.76	0.0004
Q4	-103.26	178.16	0.5639	-42.25	163.79	0.7972
Rule Period	413.11	181.31	0.0255	116.30	182.24	0.5253
Trucking gross output	21.07	4.29	<.0001	21.55	3.93	<.0001
Dec 2013–Feb 2014				1345.12	341.30	0.0002
	Adjusted R-Square = 0.49			Adjusted R-Square = 0.57		
	Crashes with injuries			Crashes with injuries		
Variable	Parameter estimate	standard error	P	Parameter estimate	standard error	P
Intercept	911.24	331.69	0.0075	926.19	325.11	0.0057
Q2	-62.40	54.44	0.2553	-39.94	54.48	0.4658
Q3	13.30	54.96	0.8094	43.07	55.82	0.4428
Q4	43.65	56.72	0.4439	54.35	55.83	0.3335
Rule Period	-5.35	57.72	0.9264	-57.38	62.12	0.3586
Trucking gross output	6.90	1.37	<.0001	6.99	1.34	<.0001
Dec 2013–Feb 2014				235.83	116.33	0.0463
	Adjusted R-Square = 0.28			Adjusted R-Square = 0.31		
	Fatal crashes			Fatal crashes		
Variable	Parameter estimate	standard error	P	Parameter estimate	standard error	P
Intercept	28.14	40.06	0.4845	28.93	40.14	0.4733
Q2	12.78	6.57	0.0557	13.97	6.73	0.0413
Q3	24.55	6.64	0.0004	26.12	6.89	0.0003
Q4	23.52	6.85	0.0010	24.09	6.89	0.0008
Rule Period	-22.31	6.97	0.0020	-25.06	7.67	0.0016
Trucking gross output	0.76	0.17	<.0001	0.76	0.17	<.0001
Dec 2013–Feb 2014				12.47	14.36	0.3881
	Adjusted R-Square = 0.29			Adjusted R-Square = 0.29		

Source: GAO analysis of Motor Carrier Management Information System (MCMIS) data. | GAO-15-641

Model 1, fit separately for each of the three categories of crashes, regressed the numbers of crashes on 1) three quarterly dummy variables to assess for seasonal differences in the numbers of crashes across the four quarters of the year,[2] 2) a dummy variable (denoted “Rule Period”) to contrast the numbers of crashes in the months following the rule change with the numbers of crashes in the months preceding it, and 3) a linear covariate measuring “trucking gross output,” which is a quarterly measure used as a proxy for the numbers of motor carriers on the road and at risk of crashing.[3] The trucking gross output variable had a significant and positive relationship with the monthly numbers of all three groups of crashes (all crashes, injury crashes, and fatal crashes), and seasonal differences, as measured by quarter, were generally significant for total crashes and fatal crashes, but not for crashes with injuries. In model 1, the apparent effect of the rule change on crashes was significant and positive for total crashes, insignificant for crashes with injuries, and significant and negative for fatal crashes. However, our discussions with industry participants indicated that the winter from December 2013 through February 2014 was unusually harsh and made operations very difficult for the industry. These weather conditions may have been a contributing factor, irrespective of the rule change, to a short term rise in crashes during this time period. Because of this issue, we also tested an alternative iteration of this model: Model 2.

[2]The four quarters contrasted are the quarters of the calendar year; Q1 (the omitted or referent category) represents January–March, and Q2, Q3, and Q4 represent April–June, July–September, and October–December, respectively.

[3]To develop a proxy measure for variation in the level of truck traffic, we used data from the Bureau of Economic Analysis (BEA) on the contribution of the U.S. trucking sector to the Gross Domestic Product GDP of the United States. These data provide information on expenditures on trucking services in the U.S. economy. While annual data are available specifically for the trucking transport sector, quarterly data are provided only for a much broader sector—all transportation and warehousing. We first used the annual data to determine the share of expenditures in this broader sector that could be attributed to the more specific trucking sector. We then applied the annual shares to the quarterly data to estimate quarterly expenditures on trucking services. Because changes in sectoral spending over time can be due to changes in both the quantity and price of a good or service, we used the chain-type price index, also available from BEA, for the broad transportation sector to adjust the quarterly estimates for inflation. We did not have a monthly trucking gross output measure. As such, the quarterly output measure is the same for each month in each quarter, and thus cannot explain intra-quarter differences in crashes by month.

Model 2 estimates all these same effects as the first model for each of the three groups of crashes but includes an additional dummy variable to account for any independent effect on the numbers of crashes due to the unusually disruptive winter weather from December 2013 to February 2014. Adding this dummy variable had no effect on our estimate of the effect of the rule change on crashes with injuries, which remained insignificant, or in fatal injuries, which remained significant and negative. It did, however, reduce the size of the estimated impact of the rule change for total crashes and rendered the effect of the rule change on total crashes insignificant.

Crashes Occurring between 5 a.m. and 9 a.m.

As noted above, we were told by stakeholders that the two-nighttime provision of the HOS rule would potentially result in an increased traffic involving large trucks between 5:00 a.m. and 9 a.m. Presumably, because of this provision, drivers taking a restart would have to wait until at least 5:00 a.m. to begin their day when roads are more congested. Using individual crash data from MCMIS, we also investigated whether, among all crashes, there was any effect of the rule change on the likelihood of crashes occurring between 5 a.m. and 9 a.m. versus any other time of the day. For each of the three sets of crashes, we first examined simple two-way cross-classifications (shown in table 28) and fit a simple bivariate logistic regression model that regressed the logarithm of the odds on crashes occurring between 5 a.m. and 9 a.m. on a dummy variable denoted as “Rule Period” that contrasted crashes that occurred before and after the rule change. Table 28 provides information on the numbers and percentages of total crashes, crashes with injuries, and fatal crashes that occurred between 5 a.m. and 9 a.m., and at all other times of the day, before and after the rule change. As can be seen, between 19 percent and 20 percent of crashes occurred between 5 a.m. and 9 a.m., both before and after the rule change. The value of the likelihood-ratio chi-square, shown at the base of each table, indicates that there was no statistically significant difference in the occurrence of crashes in the 5 a.m. to 9 a.m. time.

Table 28. Numbers of Total Crashes, Crashes with Injuries, and Fatal Crashes That Occurred between 5 a.m. and 9 a.m. and at Other Times, before and after the Rule Change

	Time of crash (total crashes)				
Rule period	**Not 5 a.m.–9 a.m.**	**5 a.m.–9 a.m.**	**Total**	**Odds on 5 a.m.–9 a.m.**	**Odds ratio**
Before rule change	364,866	90,769	455,635	0.2488	
	80.1%	19.9%	100.0%		
After rule change	91882	22880	114762	0.2490	1.001
	80.1%	19.9%	100.0%		
Total	**456,748**	**113,649**	**570,397**		
	80.1%	**19.9%**	**100.0%**		
	Likelihood ratio chi-square = 0.014 with 1df, P - 0.91				

	Time of crash (crashes with injuries)				
Rule period	**Not 5 a.m.–9 a.m.**	**5 a.m.–9 a.m.**	**Total**	**Odds on 5 a.m.–9 a.m.**	**Odds ratio**
Before rule change	141,706	34,570	176,276	0.2440	
	80.4%	19.6%	100.0%		
After rule change	33760	8326	42086	0.2466	1.011
	80.2%	19.8%	100.0%		
Total	**175,466**	**42,896**	**218,362**		
	80.4%	**19.6%**	**100.0%**		
	Likelihood ratio chi-square = 0.637 with 1df, P - 0.43				

	Time of crash (fatal crashes)				
Rule period	**Not 5 a.m.–9 a.m.**	**5 a.m.–9 a.m.**	**Total**	**Odds on 5 a.m.–9 a.m.**	**Odds ratio**
Before rule change	11,416	2,702	14,118	0.2367	
	80.9%	19.1%	100.0%		
After rule change	2518	617	3135	0.2450	1.035
	80.3%	19.7%	100.0%		
Total	**13,934**	**3,319**	**17,253**		
	80.8%	**19.2%**	**100.0%**		
	Likelihood ratio chi-square = 0.484 with 1df, P - 0.49				

Source: GAO analysis of Motor Carrier Management Information System (MCMIS) data. | GAO-15-641

Also shown in table 28 is an alternative method of estimating the likelihood of crashes occurring between 5 a.m. and 9 a.m. by calculating the differences in the likelihoods of crashes in the 5 a.m. to 9 a.m. time frame and calculating odds and odds ratios, which are shown in the last two columns of table 28. The odds on crashes occurring between 5 a.m.

and 9 a.m. are calculated by taking the number (or percentage) of crashes that occurred in that interval and dividing it by the number (or percentage) of crashes that occurred at other times. With respect to total crashes, for example, the odds on crashes occurring between 5 a.m. and 9 a.m. before the rule change were 90,769 ÷ 364,866 = .2488, and the odds of crashes occurring between 5 a.m. and 9 a.m. after the rule change were 22,880 ÷ 91,882 = .2490. While somewhat different and less traditional than percentages, the odds have a fairly direct and simple interpretation—in this case they imply that in both periods there were roughly 25 crashes in that interval for every 100 that occurred at all other times of the day. The odds of crashes occurring between 5 a.m. and 9 a.m. are very similar for crashes with injuries and fatal crashes. The odds ratios in the final column are calculated by taking the odds of crashes occurring in the designated time frame after the rule change and dividing by the odds of crashes at that time before the rule change. For total crashes, for example, the odds ratio is calculated as follows: (.2490 ÷ 0.2488) = 1.001. This ratio is only slightly different than 1, indicating that the likelihood of crashes occurring between 5 a.m. and 9 a.m. was nearly identical before and after the rule. The odds ratios for crashes with injuries and fatal crashes produce similar conclusions.

One advantage of odds ratios is that unlike percentage differences, they can be adjusted using multivariate models (logistic regression models) so that they reflect the net effect of the rule change, after adjusting for other characteristics that might have differed, in this case, before and after the rule change. Table 29 shows that results of fitting bivariate (or unadjusted) logistic regression models (in the first row), and the multivariate regression models (in the remaining rows) to estimate the effect of the rule change on the likelihood of crashes occurring between 5 a.m. and 9 a.m. for total crashes, crashes with injuries, and fatal crashes. The bivariate, or unadjusted models, reproduce the same odds ratios we derived from the observed data in the two-way tables in table 28. We then fit multivariate models that regressed those same odds on the rule period variable while simultaneously controlling for a number of potentially confounding factors, including road, weather, and lighting conditions, trucking gross volume, and whether the crash occurred during a weekday or on the weekend. The coefficients for the multivariate model shown in table 29 are exponentiated odds ratios, which indicate the size and significance of the differences in the odds on crashes occurring between 5 a.m. and 9 a.m. as opposed to some other time of the day, both as a result of the rule change and as a result of the factors included in the model. The multivariate models reveal that a number of the control variables affected the likelihood of crashes occurring from 5 a.m. to 9 a.m., including different road, weather, and light conditions. Crashes were more likely to occur between 5 a.m. and 9 a.m., for example, when roads were icy rather than dry (by factors ranging from 2.2 for total crashes to 3.6 for fatal crashes), when conditions were foggy rather than clear (by factors ranging from 4.4 for total crashes to 5.0 for crashes with injuries and fatal crashes), and during dawn rather than daylight hours (by factors exceeding 40 in all cases). However, even with these factors controlled, the rule change appeared to have no significant effect on the likelihood of total crashes occurring between 5 a.m. and 9 a.m. Odds ratios in the multivariate models estimating the differences in the likelihood of crashes occurring before and after the rule change were very nearly 1.0, and ranged from 0.999 for total crashes to 1.025 for fatal crashes.

Table 29: Odds Ratios from Models Estimating the Effects of the Rule Change on the Likelihood of Crashes Occurring between 5 a.m. and 9 a.m., before and after Taking Account of Other Factors

		Odds ratio estimates		
Model	**Effect**	**Total crashes**	**Total crashes**	**Total crashes**
1	rule_period (bivariate model)	1.001	1.011	1.036
2	rule_period (multivariate model)	0.999	1.005	1.025
	road 2. WET vs 1. DRY	1.263[a]	1.296[a]	1.180
	road 3. WATER(STANDING, MOVING) vs 1. DRY	1.090	1.142	0.718
	road 4. SNOW vs 1. DRY	1.676[a]	1.782[a]	1.781[a]
	road 5. SLUSH vs 1. DRY	1.549[a]	1.517[a]	2.594[a]
	road 6. ICE vs 1. DRY	2.177[a]	2.283[a]	3.587[a]
	road 7. SAND,MUD,DIRT,OIL OR GRAVEL vs 1. DRY	0.949	1.015	1.021
	road 8. OTHER vs 1. DRY	1.605[a]	1.678[a]	2.069
	road 9. UNKNOWN vs 1. DRY	1.085	1.116	0.855
	weather 2. RAIN vs 1. NO ADVERSE CONDITIONS	0.809[a]	0.781[a]	0.891
	weather 3. SLEET, HAIL vs 1. NO ADVERSE CONDITIONS	0.719[a]	0.754[a]	0.583[a]
	weather 4. SNOW vs 1. NO ADVERSE CONDITIONS	0.707[a]	0.646[a]	0.732
	weather 5. FOG vs 1. NO ADVERSE CONDITIONS	4.354[a]	5.008[a]	5.006[a]
	weather 6. BLOWING SAND, SOIL, DIRT, OR SNOW vs 1. NO ADVERSE CONDITIONS	0.560[a]	0.436[a]	0.292[a]
	weather 7. SEVERE CROSSWINDS vs 1. NO ADVERSE CONDITIONS	0.478[a]	0.415[a]	0.608
	weather 8. OTHER vs 1. NO ADVERSE CONDITIONS	1.011	1.027	1.250[a]
	weather 9. UNKNOWN vs 1. NO ADVERSE CONDITIONS	0.963	1.056	1.186
	light 2. DARK - NOT LIGHTED vs 1. DAYLIGHT	0.876[a]	0.966[a]	1.059
	light 3. DARK - LIGHTED vs 1. DAYLIGHT	0.744[a]	0.772[a]	0.836
	light 4. DARK - UNKNOWN ROADWAY LIGHTING vs 1. DAYLIGHT	1.370[a]	1.471[a]	1.483[a]
	light 5. DAWN vs 1. DAYLIGHT	40.195[a]	46.850[a]	73.944[a]
	light 6. DUSK vs 1. DAYLIGHT	0.379[a]	0.416[a]	0.271[a]
	light 8. OTHER vs 1. DAYLIGHT	1.152	1.231	2.280
	light 9. UNKNOWN vs 1. DAYLIGHT	0.864	0.907	0.867
	trucking_gross_output	1.0000	1.0010[a]	1.0010
	weekend	0.802[a]	0.781[a]	0.811[a]

Source: GAO analysis of Motor Carrier Management Information System (MCMIS) data. | GAO-15-641

[a]Coefficient is significant at the .05 level.

As previously discussed, there are many potential factors that could influence the number of crashes in any given month, including weather and road conditions at the time of a crash. Our analyses take into account some of these factors by controlling or adjusting for seasonal differences and freight volume, but there are undoubtedly other potentially confounding factors, and we would have liked to have had a more direct measure of the numbers of motor carriers on the road in each month and at risk of crashing. Moreover, our dataset only included 15 data points, or months, since the rule went into effect, and neither the significant or insignificant changes we find in this short term are guaranteed in the longer term. Without additional data over a longer period of time, we are unable to robustly determine whether the HOS rule had an impact on crashes.

Appendix VIII: Comments from the Department of Transportation

U.S. Department of Transportation
Office of the Secretary of Transportation

Assistant Secretary for Administration

1200 New Jersey Avenue, SE
Washington, DC 20590

July 16, 2015

Ms. Susan Fleming
Director, Physical Infrastructure Issues
U.S. Government Accountability Office
441 G Street, NW
Washington, DC 20548

Dear Ms. Fleming:

In carrying out its safety mandate, the Federal Motor Carrier Safety Administration's (FMCSA) partners with stakeholders to reduce bus and truck-related crashes. To support this effort, FMCSA issued its 2011 hours-of-service (HOS) rule. While these new provisions were in place, there was (1) a decrease in the frequency of drivers using long work schedules, (2) a lower risk of driver fatigue, and (3) a reduction in the number of commercial vehicle crashes involving fatalities. This GAO report recognizes these achievements and also reported no increase in the number of large truck crashes during the morning rush hours between 5:00 am and 9:00 am.

FMCSA's January 2014 HOS study found that each of the three measures the researchers used independently supported the conclusion that drivers taking a one-night restart were more fatigued than those taking a two-night restart. This effect is consistent with the purpose of the restart provision, which is to ensure that truck drivers working long and demanding schedules have adequate nighttime rest periods for restorative sleep. In conducting this study, FMCSA adhered to standard principles and practices of scientific research, including the use of an independent peer review panel which evaluated and agreed with the project's methodology, analysis, and findings.

In an effort to continue our work, FMCSA:

- Pursues research avenues on the impacts of the HOS provisions and on effective countermeasures for driver fatigue.
- Commissioned the National Academy of Sciences to convene a panel of experts to recommend appropriate methodologies and statistical approaches for studying commercial motor vehicle driver fatigue and health.
- Is a co-sponsor of the North American Fatigue Management Program (www.NAFMP.org), an on-line training tool to help carriers and drivers address and mitigate fatigue.

Upon preliminary review of this report, FMCSA agrees with the GAO recommendation. The Department will provide a detailed response to the recommendation within 60 days of the GAO report issuance.

We appreciate the opportunity to provide additional perspective on the GAO draft report. Please contact Patrick D. Nemons, Deputy Director of Audit Relations, at (202)366-4986 with any questions or if the GAO would like additional information.

Sincerely,

Jeff Marootian
Assistant Secretary for Administration

Appendix IX: GAO Contact and Staff Acknowledgments

GAO Contact

Susan A. Fleming, (202) 215-2834 or Flemings@gao.gov

Staff Acknowledgments

In addition to the individual named above, H. Brandon Haller, Assistant Director; Amy Abramowitz; Sarah Arnett; Russell Burnett; Matthew Cook; Melinda Cordero; Leia Dickerson; Colin Fallon; David Hooper; Hannah Laufe; Ethan Levy; Grant Mallie; Joshua Ormond; Anna Maria Ortiz; Jerry Sandau; Doug Sloane; and Jeff Tessin made key contributions to this report.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's website (http://www.gao.gov). Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to http://www.gao.gov and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, http://www.gao.gov/ordering.htm.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on Facebook, Flickr, Twitter, and YouTube.
Subscribe to our RSS Feeds or E-mail Updates.
Listen to our Podcasts and read The Watchblog.
Visit GAO on the web at www.gao.gov.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Website: http://www.gao.gov/fraudnet/fraudnet.htm
E-mail: fraudnet@gao.gov
Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Katherine Siggerud, Managing Director, siggerudk@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548