Analysis of Naturalistic Driving Data to Assess Distraction and Drowsiness in Commercial Motor Vehicles

Rebecca Hammond
Susan Soccolich
04.13.2023

VIRGINIA TECH
TRANSPORTATION INSTITUTE
VIRGINIA TECH

Naturalistic Data and Study Background

- Naturalistic data provides a rich dataset allowing us to conduct many follow-on analysis and allows us to see the driver's environment without potential bias from an experimenter or in a controlled environment.
- Data were collected from 2013 – 2014 as part of a larger data collection effort involving tractor trailers and motorcoach vehicles.
- Contribute to the sparse body of research on CMV driver distraction and fatigue.

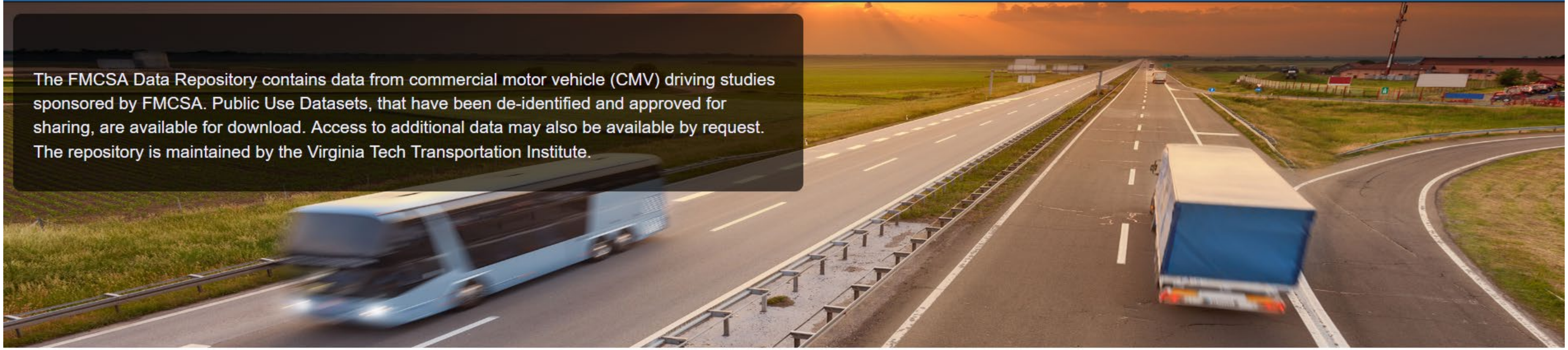
FMCSA Data Repository

- Provides access to FMCSA-funded datasets
- Public use datasets are available for download
  - No personally identifying information (PII)
  - No video
- Must register on website to access downloadable data
- Identifiable data is available in the secure data enclave at VTTI
  - Video
  - GPS
  - Contact us button on website or email me at rhammond@vtti.vt.edu

https://fmcsadatarepository.vtti.vt.edu/

FMCSA
Federal Motor Carrier Safety Administration

Data Repository | Data Access | Query Data | FAQs | Contact Us | Rebecca Hammond



What's Available on This Website

Driver Descriptions and Assessments
Summary graphs and detailed records of driver assessments are provided addressing driver demographic information.

Summary of Data Collected
Graphs and detailed records describe data collection progress and characteristics of trips collected during the studies.

Vehicle Descriptions
Summary graphs and detailed records describe the types of vehicles involved in the studies.

Custom Query Capability
Build custom queries to search for records matching criteria that span multiple datasets.

Public Use Datasets
Deidentified public use datasets available for download.

Study Background Information
Access an overview of the FMCSA Research Data Repository and data collection procedures.

Highlighted Datasets

Commercial Driver Safety Risk Factors

FMCSA
Federal Motor Carrier Safety Administration

Data Repository | **Data Access** | **Query Data** | **FAQs**

- About
- Dataverse
- Public Use Datasets

ublic Use Datasets

To access the de-identified, Public Use Dataset, click

Data Dictionaries and Final Reports are also include

Analysis of Distraction and Drowsiness in CMV Drivers

Analysis of Distraction and Drowsiness in CMV Drivers Overview

Analysis of Distraction and Drowsiness in CMV Drivers Public Use Datasets — All Data (XLSX) | All Events Raw Eyeglance (CSV) | Raw PERCLOS (ZIP)

Analysis of Distraction and Drowsiness in CMV Drivers Data Dictionary — Download

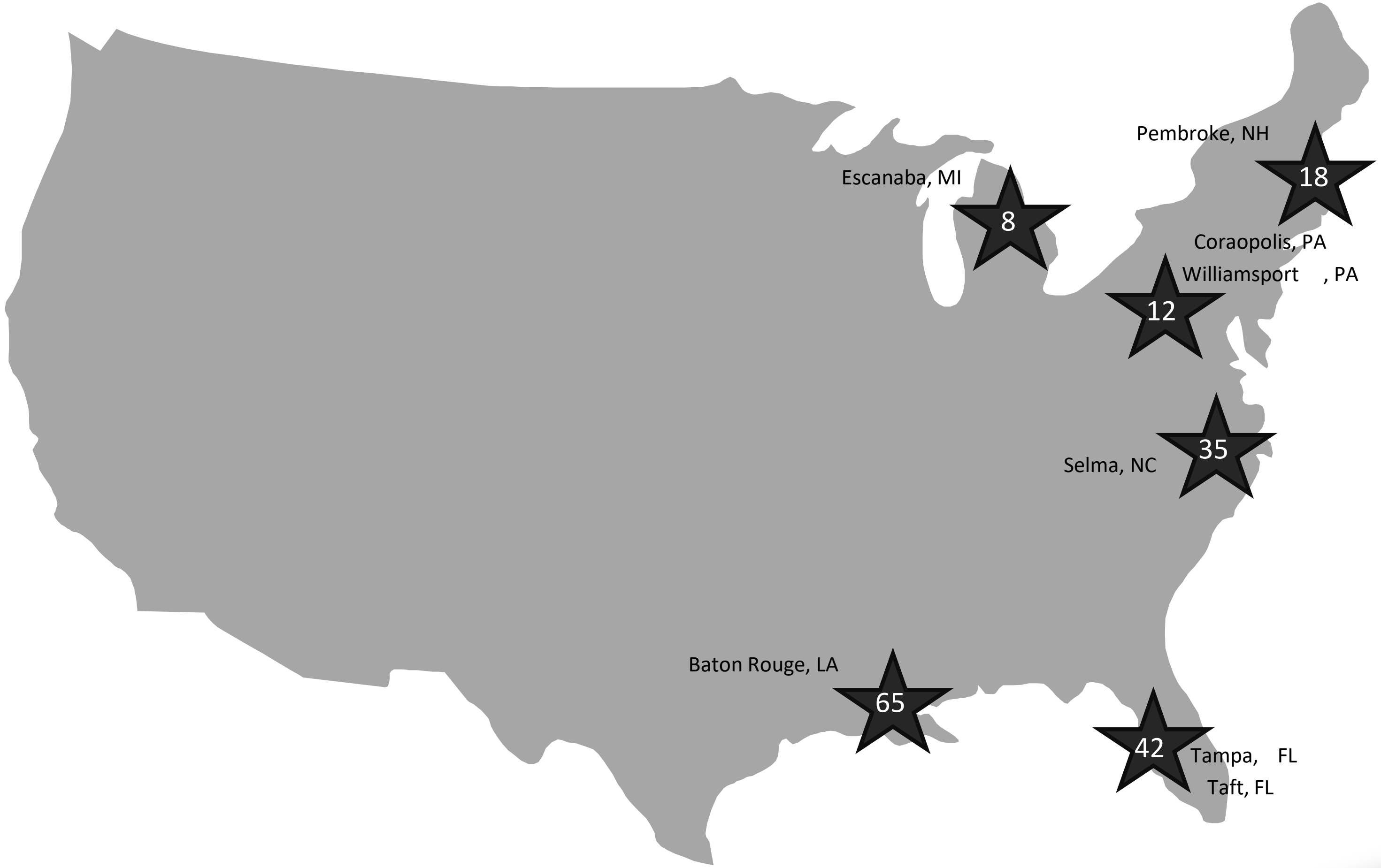
Analysis of Distraction and Drowsiness in CMV Drivers Final Report — Download

Data Collection

Data were collected continuously when truck is on and in motion during normal routes.

Sample Size

- 2.7 million miles of data
- 6 fleet locations
- 182 vehicles
- 172 drivers

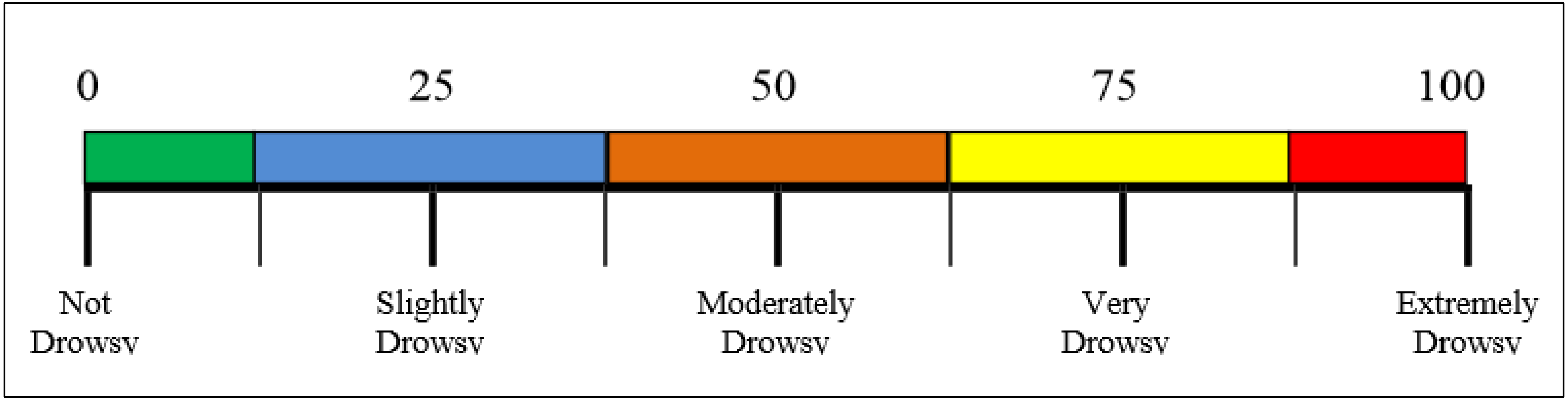


Data Collection and Data Reduction

- Kinematic (speed, accelerometers, GPS, network, etc.) and video from five cameras.
- Once data is collected, it is scanned for kinematic spikes, or triggers (e.g., hard brake, sharp steering, lane deviation) which are then reviewed for validity.
- Categorize as crash, near-crash, crash-relevant conflict and unintentional lane deviation (SCEs).
- Code each valid event with 60+ variables such as critical reason, reaction, environment, secondary tasks.

- Baseline epochs, or non conflict events created to use as a comparison.

Observer Rating of Drowsiness



- Subjective rating.
- One minute of video prior to the start of an SCE or baseline epoch.
- Analysts look for eye closures, yawning, decreased facial tone.
- Each event reviewed by three analysts, average is the final ORD value.

Manual PERCLOS

- “mathematically define proportion of a time interval that the eyes are 80 to 100 percent closed”
- 3 minutes of video are reviewed
- Objective measure

Analysis Results

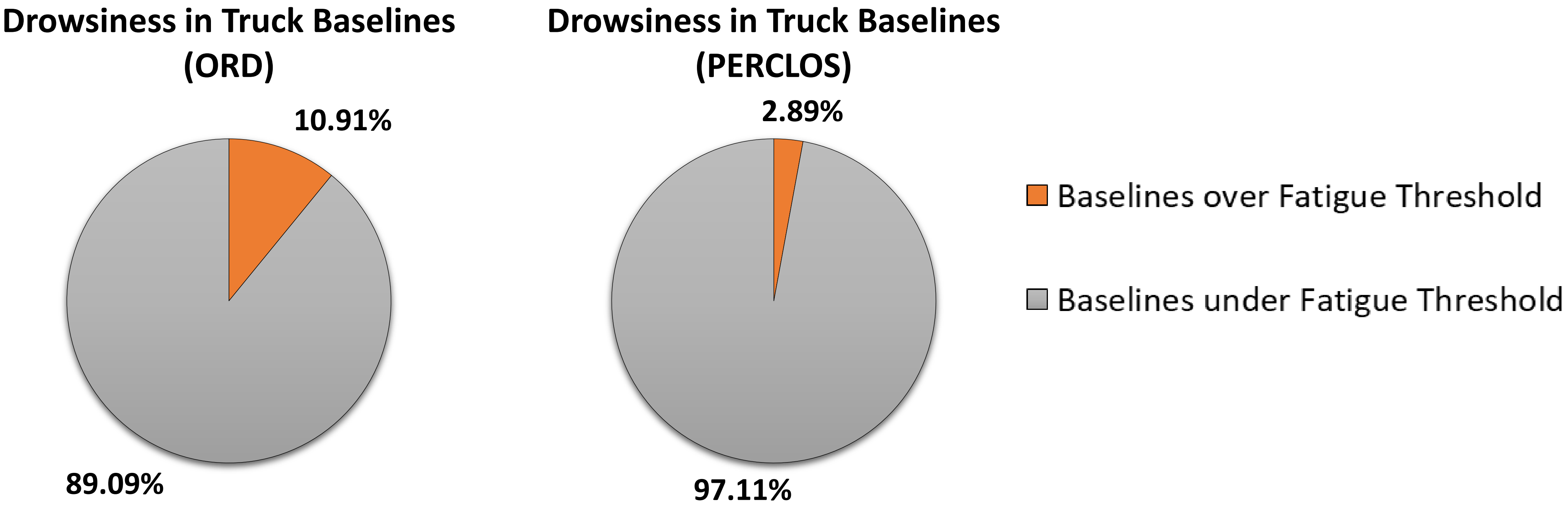
Risk of Cell Phone Use

- Data were collected after the nation-wide cell phone ban established in 2012. Did that ban reduce cell phone use?
- Identified instances of cell phone use, both hand-held and hands-free.
- Calculated odds ratios comparing cell phone use in SCEs to baseline epochs.
- Found that overall, cell phone use was down, when it did occur:
  - Hand-held cell phone use increased the risk of being involved in an SCE.
  - Hands-free cell phone use decreased the risk of being involved in an SCE (potentially helped drivers stay awake at night by talking to family or friends).

Cell Phone Task	All Faults OR	All Faults CI	V1 At Fault OR	V1 At Fault CI
All cell phone tasks	1.14	(0.93, 1.39)	1.40*	(1.13, 1.75)
Hand-held cell phone tasks	2.81*	(2.16, 3.66)	4.00*	(3.03, 5.27)
Hands-free cell phone tasks	0.51*	(0.38, 0.69)	0.46*	(0.33, 0.66)

Prevalence of Driver Drowsiness

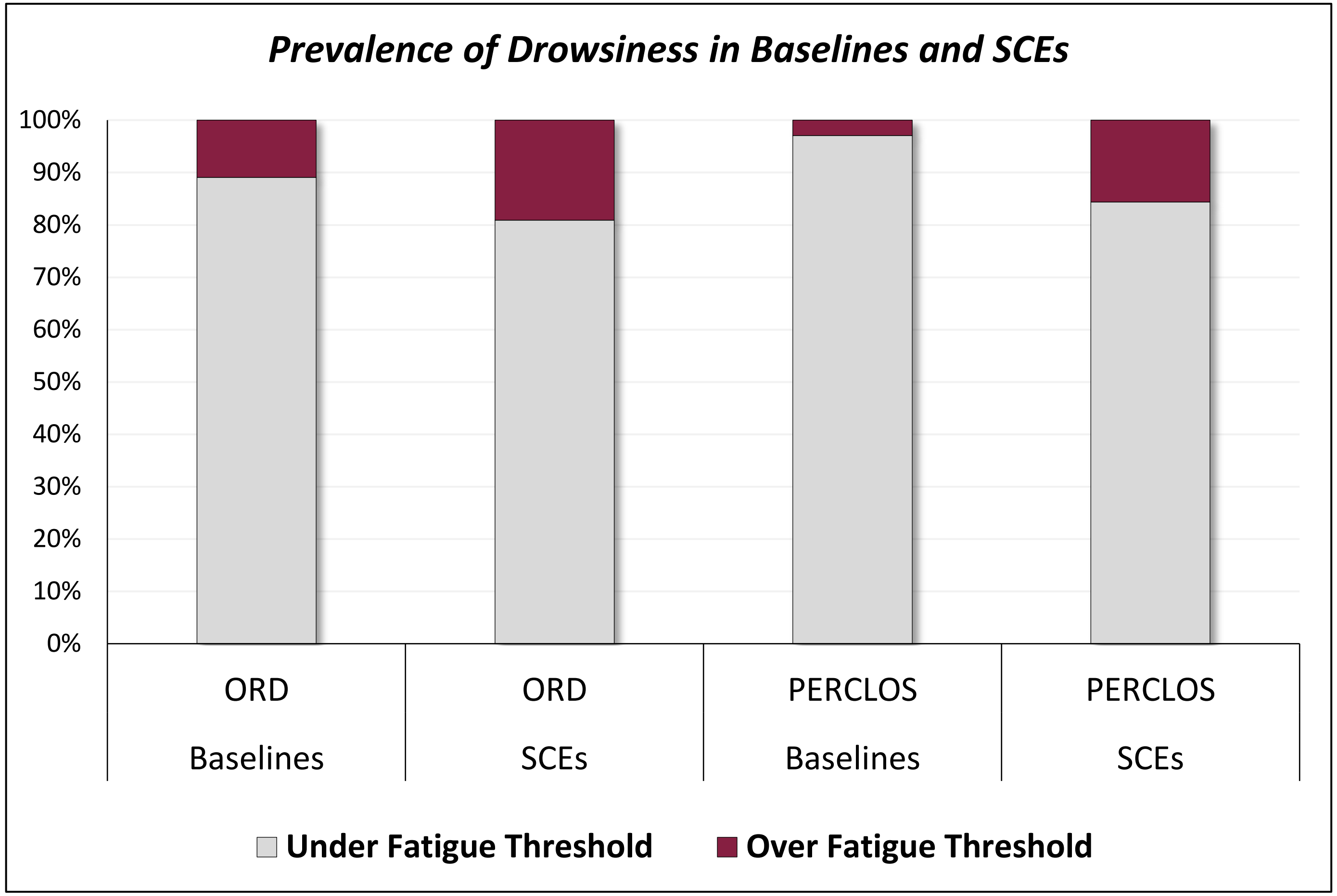
- Prevalence of driver drowsiness calculated from baselines
- Prevalence = # baselines with drowsiness over the fatigue threshold / total # baselines
- Valid drowsy measures were obtained from 1,962 baselines for ORD and 2,564 for PERCLOS measures
- ORD drowsy threshold included “very” and “extremely” drowsy; PERCLOS proportion threshold was 12%



Key finding: **Drowsiness observed in 3-11% of truck baselines.**

SCE Involvement while Drowsy

- Prevalence of driver drowsiness compared for baselines and SCEs
- Mixed-effect logistic regression model predicted the odds of being in an SCE while drowsy
- When drowsy, truck drivers were:
  - ~1.7 to 3.7x more likely to be involved in an *at-fault SCE*
  - ~1.3 to 2.9x more likely to be involved in an *all-fault SCE*



Key finding: **SCE risk grows significantly when drowsy for truck drivers.**

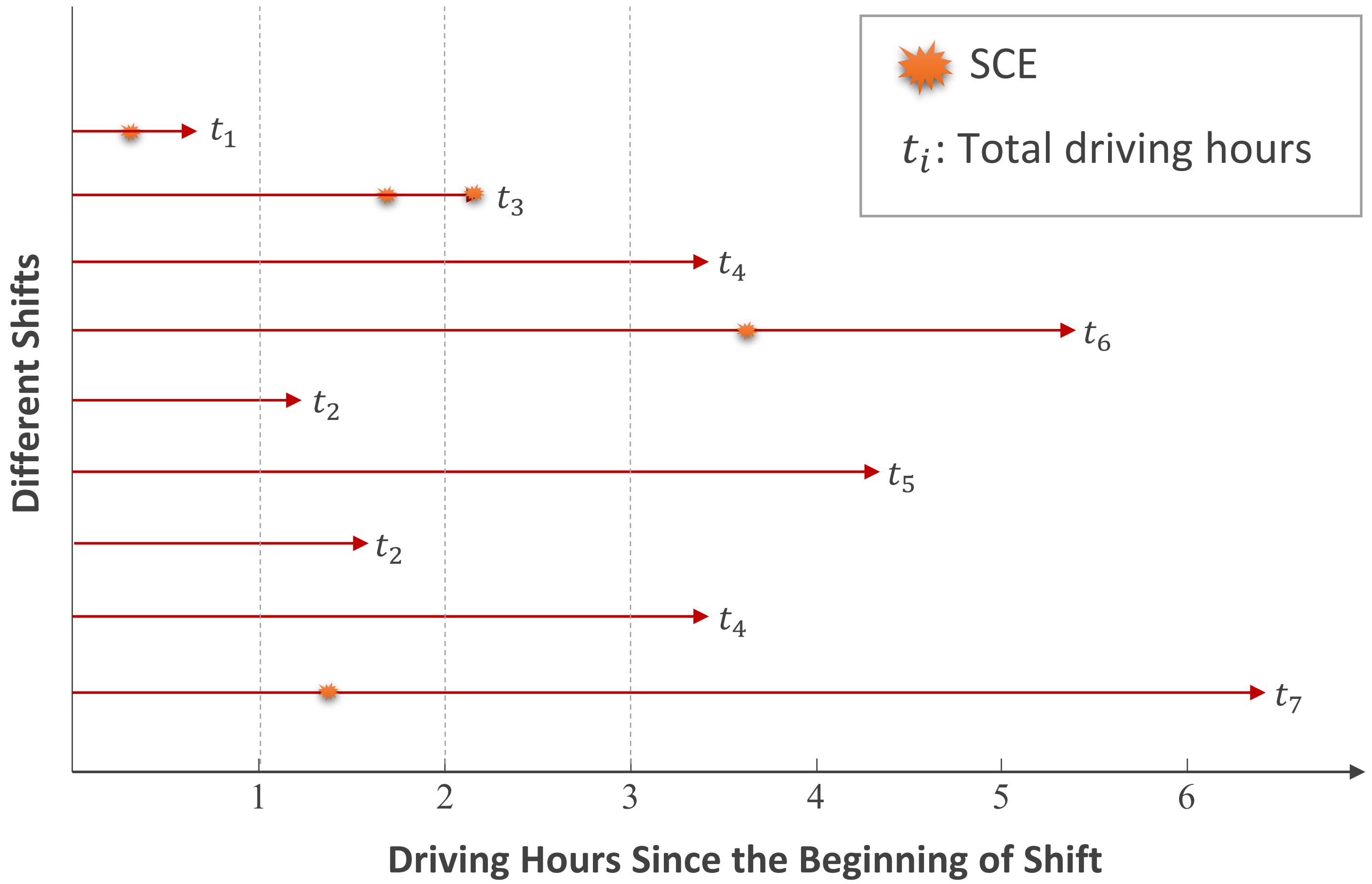
Prevalence of Driver Drowsiness During Secondary Tasks

- To assess whether drowsiness was more or less prevalent during secondary task engagement:
  - BLs and SCEs were categorized by presence of individual secondary tasks (i.e., baselines with task observed, baselines without task observed, etc.).
  - Within each subset, prevalence of driver drowsiness was calculated (using ORD and PERCLOS measures).
  - Drowsiness prevalence was compared in baselines with and without the secondary task using odds ratios.
  - This analysis was then repeated for SCEs.

Tasks Associated with Lower Drowsiness	Significant Subset	Tasks Associated with Higher Drowsiness	Significant Subset
Any secondary task (overall)	*BLs, SCEs (ORD)*	Removing/adjusting clothing	*SCEs (ORD, PERCLOS)*
Talking/singing	*SCEs (ORD)*		
External distraction	*SCEs (ORD, PERCLOS)*		
Eating	*SCEs (ORD)*		
Electronic dispatching device	*BLs (PERCLOS)*		
Hand-held browsing	*SCEs (ORD)*		
Hands-free call via headset/earpiece	*BLs, SCEs (ORD, PERCLOS)*		
Hands-free talk/listen	*BLs, SCEs (ORD, PERCLOS)*		

Key finding: **Certain tasks associated with significantly fewer observations of ORD drowsiness.**

SCE Rates by Driving Hour Since Shift Start



SCE rate for the i^{th} driving hour since shift start =

$$\frac{Number\ of\ SCEs\ in\ the\ i^{th} driving\ hour}{Total\ driving\ time\ in\ the\ i^{th} driving\ hour}$$

Examples:

1st driving hour:

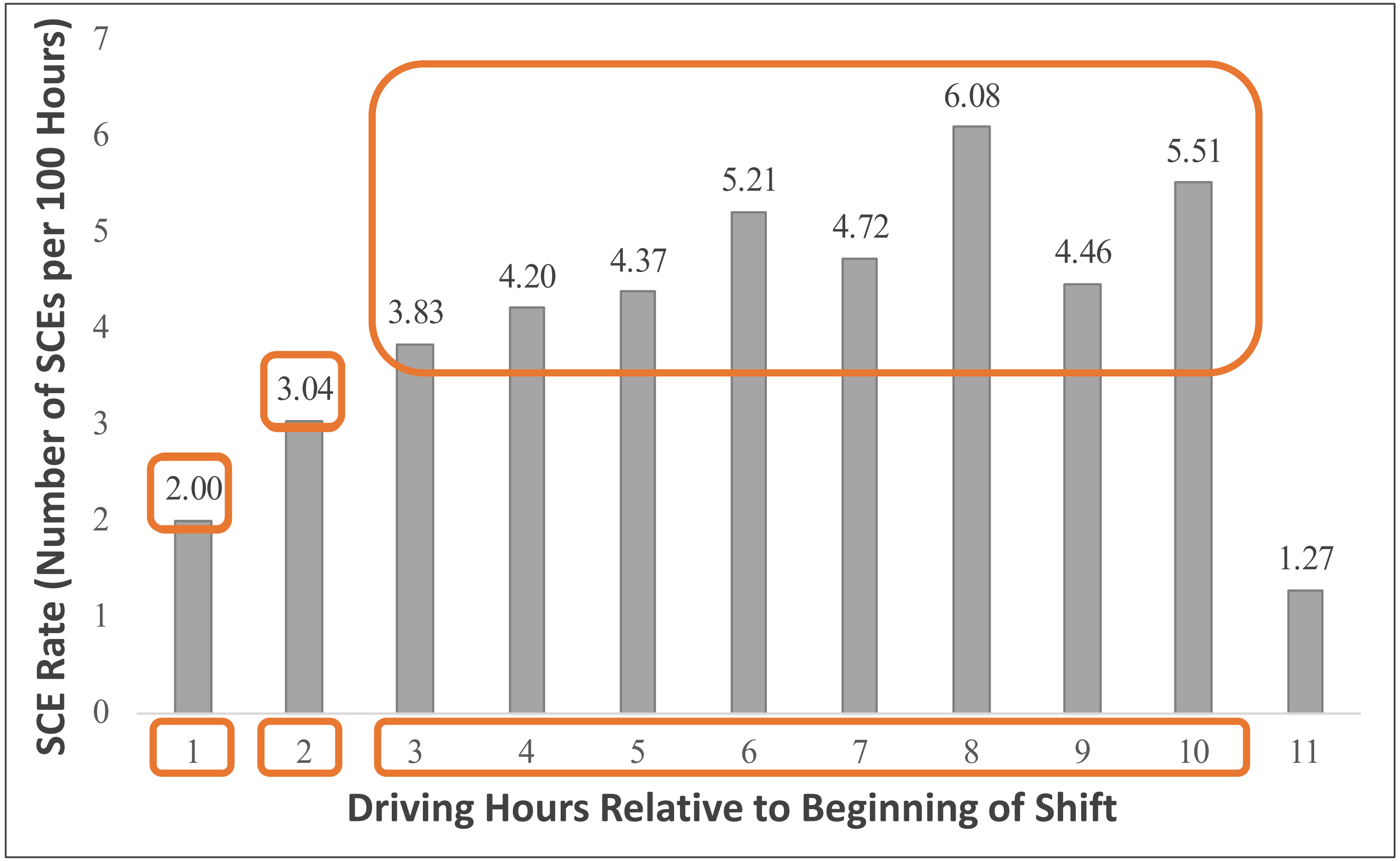
$$\frac{1SCE}{(t_1+1+1+1+1+1+1+1+1)hours}$$

2nd driving hour:

$$\frac{2SCEs}{(1+1+1+(t_2-1)+1+(t_3-1)+1+1)hours}$$

SCE Rates by Driving Hour Since Shift Start

- **Evaluated using:**
  - Mixed-effect Poisson model
  - Tukey multiple comparison
- **Significant comparisons:**
  - Hour 1 had a significantly lower SCE rate compared to Hours 2-10.
  - Hour 2 had a significantly lower SCE rate compared to Hours 3-10.
  - Hours 3-10 did not show a significantly different SCE rate.

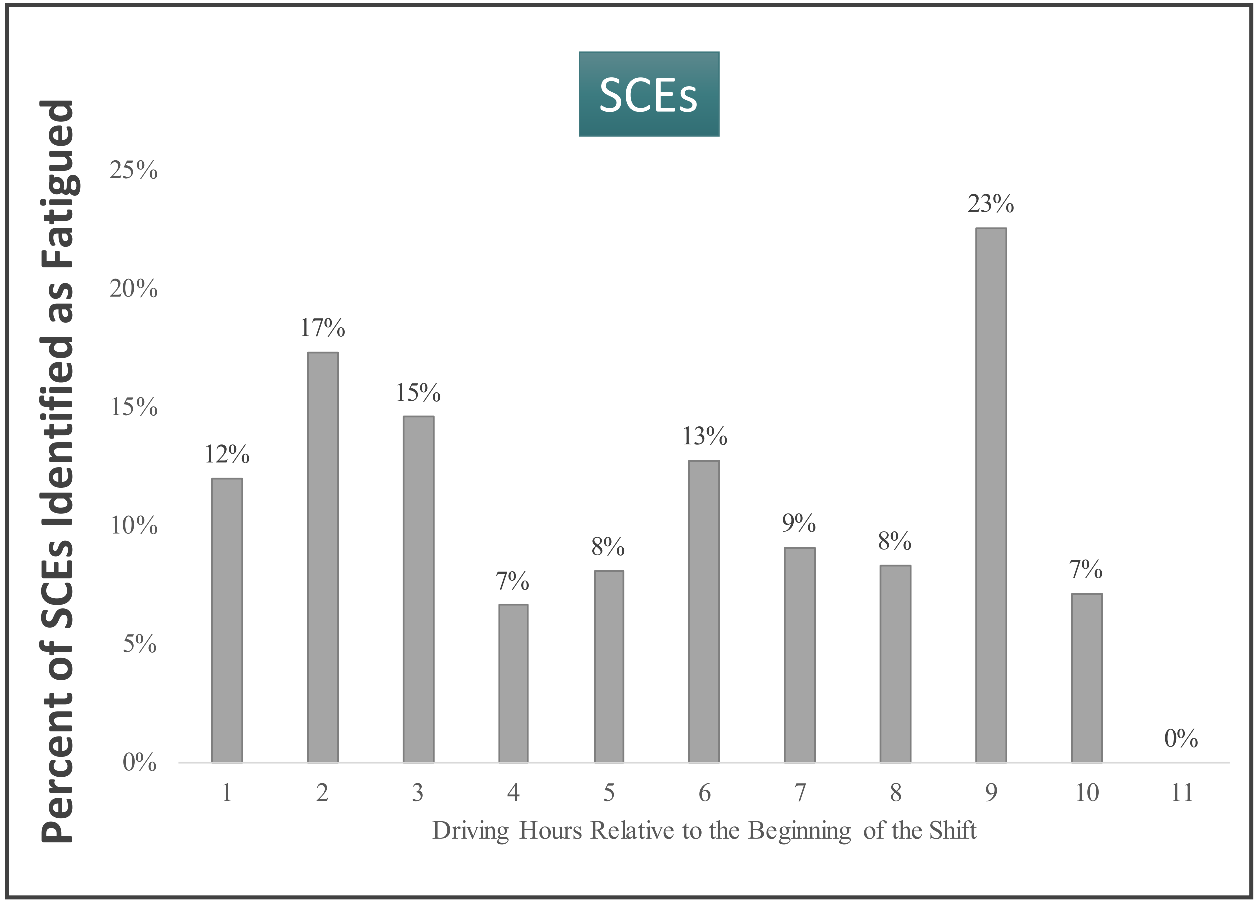


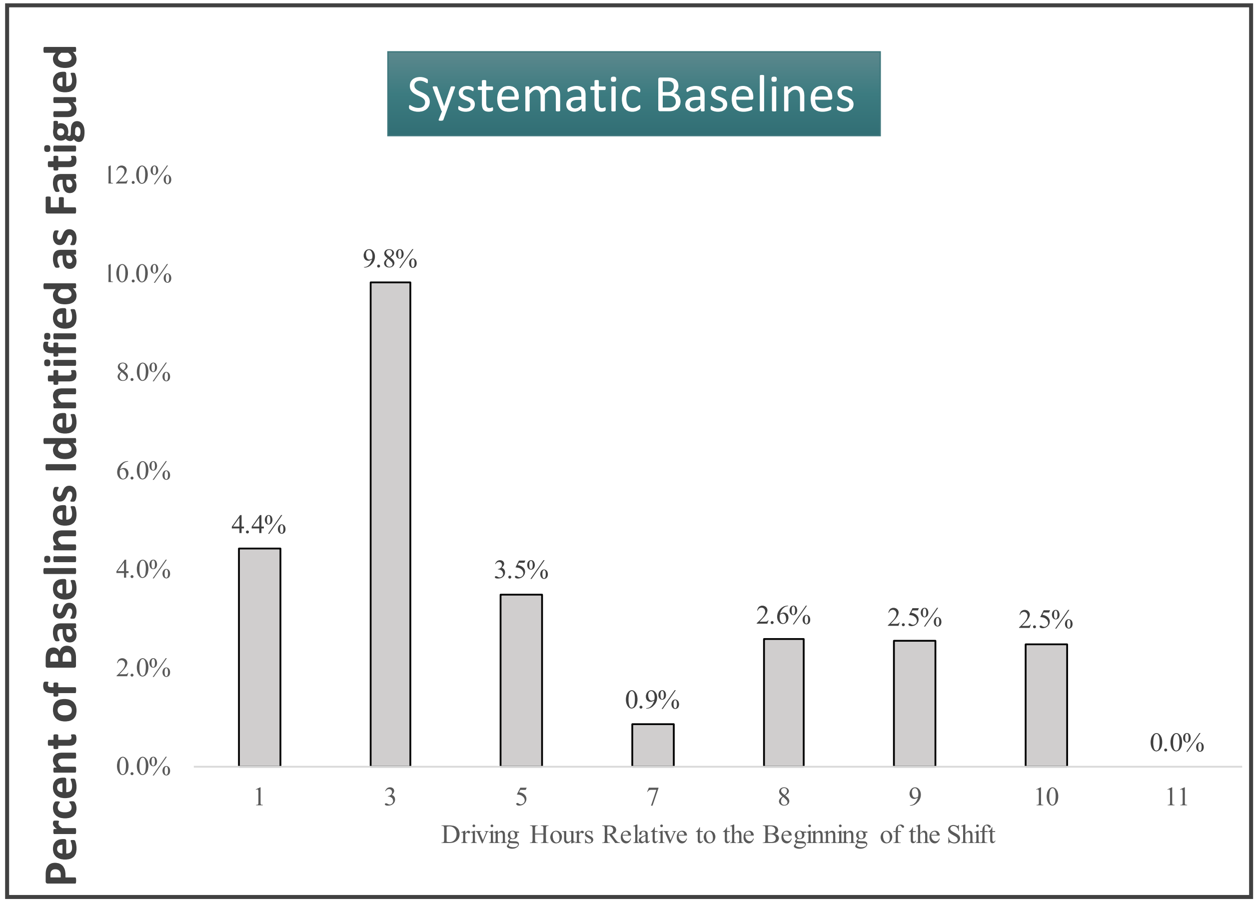
Key finding: **SCE rates show a pattern of increasing over driving time, with the 1st, 2nd, and 3rd-10th hour showing significant differences when compared.**

Prevalence of Driver Drowsiness by Driving Hour

- Objective
  - Evaluate temporal profile for both SCEs and systematically sampled BLs
- Analysis approach
  - Systematic sampled baselines:
    - Selected 200 shifts with driving time longer than 10 hours
    - Samples taken at 1, 3, 5, 7, 8, 9, 10, 11 hours
  - PERCLOS score used to determine drowsiness status
  - 2,325 valid SCEs and 929 valid systematic baselines
  - Mixed-effect Logistic regression models used to assess drowsiness prevalence by driving hour

Prevalence of Driver Drowsiness by Driving Hour





Key finding: **No clear pattern of drowsiness by driving hours, confirmed by modeling analysis.**

Summary of Study Findings

- Naturalistic driving data provides a unique opportunity to understand driver behavior in both safe and unsafe driving moments.
- The current study collected a large amount of naturalistic driving data from a diverse fleet set.
- Drowsiness was observed in as many as 1 in 10 baseline driving moments.
- Consistent with previous studies, drowsy driving increases SCE risk.
- SCE rates did not significantly increase after the 3rd driving hour.
- No clear pattern of drowsiness by driving hours, confirmed by modeling analysis.

Thank you

Rebecca Hammond rhammond@vtti.vt.edu

Susan Soccolich ssoccolich@vtti.vt.edu

Virginia Tech Transportation Institute

Division of Freight, Transit, and Heavy Vehicle Safety

Link to study report:
https://rosap.ntl.bts.gov/view/dot/57153/dot_57153_DS1.pdf