Readme for replication package to reproduce the results of

Do Safety Inspections Improve Safety? Evidence from the Roadside Inspection Program for Commercial Vehicles

By Yuanning Liang

Date: 06.15.2025

The replication package includes the following components:

* A complete set of Stata do-files used to generate the final results presented in the paper.

* This read-me file, which provides detailed instructions on how other researchers can obtain the proprietary data via the listed website links. Due to legal and contractual restrictions, there are limitations on how the administrative data may be processed and stored. Additionally, the datasets used in this study are large in size—for example, the raw FMCSA data exceed 40GB even after compression. The event study datasets are even larger; the analysis file used to produce the baseline result in Figure 1 exceeds 100GB after compression. For these reasons, the processed datasets are not included in the replication package, both due to administrative constraints and concerns about file size.

* Instructions for submitting data requests and data dictionaries for each dataset are also provided, enabling readers to review the structure and content of the data before making a formal request.

*** Description of Stata-do-file ***

* 0_MCMIS_cleandata.do cleans variable names and irregular strings in raw data.

* 1_data_prep.do prepares analysis dataset for daily and monthly event studies. This do file also cleans FARS and TxDOT data and prepares their analysis datasets.

* 2_fig_tab.do produces figures and tables in the main text.

* 3_appendix.do produces figures and tables in the online appendix.

* The do files indicate which table/figure is being produced (numbering based on pre-print PDF)

* Stata version used: 18.0

*** Description of Source Data Files ***

1. Inspection, Crash, and Carrier company census data: From FMCSA MCMIS program DOT Open Data Catalog, or FOIA to make a formal request. A data request guide by FMCSA and file dictionaries are attached.
2. FARS crash data: Download data by year from ftp://ftp.nhtsa.dot.gov/fars/. A file list is attached.
3. Texas DOT crash data: TxDOT CRIS data extracted by registered users from https://www.txdot.gov/data-maps/crash-reports-records/crash-data-analysis-statistics.html. A CRIS data request guide is attached.
4. FHWA Traffic data sets:
    a. Travel Monitoring and Traffic Volume: from FHWA Office of Highway Policy Information at https://www.fhwa.dot.gov/policyinformation/travelmonitoring.cfm
    b. TMAS Class and Station Data: requested from FHWA 2012-2018 through FOIA
    c. VMT data: from FHWA Highway Statistics Series by year: Functional system travel - annual vehicle miles (https://www.fhwa.dot.gov/policyinformation/statistics.cfm)
5. Number of trucks and buses: BTS National transportation statistics: Table 1-11: Number of U.S. Aircraft, Vehicles, Vessels, and Other Conveyances https://www.bts.gov/archive/publications/national_transportation_statistics/table_01_11

5. Weather: GHCN-Daily from NOAA
(ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/)

*** Requests for Data Access ***

Most of the datasets listed above are proprietary and require approval from the relevant agencies prior to access and use. Researchers interested in obtaining these data should contact the respective offices responsible for maintaining the datasets directly.