README and guidance

1 Overview

This document describes the replication files for "Congestion pricing with electric vehicle exemptions: Car-ownership effects and other behavioral adjustments" by Elisabeth T. Isaksen and Bjørn G. Johansen for the Journal of Environmental Economics and Management (JEEM).

The replication package constructs the figures and tables in the paper as well as the online appendix from a number of data sources using Stata 17 or 18.

The replication package consists of two parts. Part 1 contains the analysis based on Norwegian administrative data. All the do-files based on these data are group together. This analysis uses proprietary data made available for research, under specific and strict regulations, by Statistics Norway, and cannot be made publicly available.

Part 2 contains the analysis based on traffic and air pollution data. All the do-files based on these data are group together. The analysis on traffic is mainly based on proprietary data collected from a private road toll company (Ferde). These data cannot be made publicly available. Air pollution and weather data are available via public APIs. The raw data downloaded from the APIs are provided in the replication package.

This README-files documents part 2 of the analysis.

2 Data availability and provenance statements

Data on cars entering and exiting the toll cordons during 15-minute intervals in different cities are collected from the private road toll company Ferde. We are not allowed to publicly share these data.

We collected additional data on hourly traffic outside the Bergen toll cordon from the Norwegian Public Roads Administration. These data can be accessed via the following link: https://trafikkdata.atlas.vegvesen.no/. The downloaded raw data is provided in the replication package.

Data on air pollution can be accessed via NILU's public API: https://api.nilu.no/. The downloaded raw data is provided in the replication package.

Data on weather can be accessed via MetFrost's public API: https://frost.met.no/. The downloaded raw data is provided in the replication package.

3 Dataset list

The replication package presupposes the existence of the following datasets in the `STATA\data_in` folder:

Table 1. Data files used in the analysis.

Filename and data reference	**Description of data and variables included in analysis**
`QFS-NO017U-8505-1 – location.csv` *Source: Ferde (road toll company)*	Station meta data from Ferde. One file for each road toll point/cordon. "Location" refers to the location of the road toll collection, e.g., Bergen, or Kristiansand (22 locations in total). These datasets are not available in the folder.
`QFS-NO017U-8505-3 – location.csv` *Source: Ferde (road toll company)*	Traffic data from the company Ferde (15-minute intervals). One file for each road toll point/cordon. "Location" is e.g. Bergen, or Kristiansand (22 in total). These datasets are not available in the folder.
`indre_arna_2015.csv, indre_arna_2016_2018.csv, kråkenes_2014_2018.csv` *Source: https://trafikkdata.atlas.vegvesen.no/*	Publicly available data from two traffic sensors located outside the Bergen toll cordon. These datasets are available in the `STATA\data_in` folder.

The replication package presupposes the existence of the following datasets in the `Python\Metfrost\csv\` folder:

Table 2. Data files used in the analysis.

Filename and data reference	**Description of data and variables included in analysis**
`elements.csv, locations.csv, sources.csv` *Source: https://frost.met.no/*	Meta data downloaded from the Met Frost API. These datasets are available in the folder.

`MetFrost_air_temperature_SNXXX_YYYY.csv`
`MetFrost_MultipleVar_SNXXX_YYYY.csv`
`MetFrost_precipitation_SNXXXX_YYYY.csv`
Source: https://frost.met.no/
	Weather data downloaded from the Met Frost API for station SNXXX and year YYYY. These datasets are available in the folder.

Python scripts to download the relevant data from the API are available in `Python\Metfrost\`

The replication package presupposes the existence of the following datasets in the `Python\NILU\csv\` folder:

Table 3. Data files used in the analysis.

Filename and data reference	**Description of data and variables included in analysis**
`areas.csv, components.csv, stations.csv` *Source: https://api.nilu.no/*	Meta data downloaded from the NILU API. These datasets are available in the folder.

`historical_aq_YYYY_MXX_stationID_all.csv` *Source: https://api.nilu.no/*	Air pollution data from the NILU API for all stations for year YYYY and month MXX. Covers the period M1 2014 to M12 2018. These datasets are available in the folder.

Python scripts to download the relevant data from the API are available in `Python\NILU\`

4 Computational requirements

4.1 Hardware requirements

Estimations were done on a Windows 10 machine.

- Processor: Intel(R) Core(TM) i7-9700T CPU @ 2.00GHz 1.99 GHz
- Installed RAM: 16.0 GB

4.2 Software requirements

All code is run using Stata 17.

The following .ado-programs from SSC are used:

- leftalign
- estout
- geonear
- distinct
- reghdfe
- ftools
- coefplot

These packages can be installed by uncommenting the lines 48-54 in the script `STATA\dofiles\0_master.do`.

4.3 Runtime requirements

Given the aforementioned hardware, total run time (including results for the online appendix) is approximately one day.

5 Description of programs/code

The replication package has the following folder structure:

- STATA
  - ado
    - personal
    - plus
  - data_in
  - data_out

        - Ferde
        - Metfrost
        - NILU
        - SVV
    - dofiles
    - figures
        - longnames
    - ster
    - tables
- Python
    - Metfrost
        - csv
        - json
    - Metfrost
        - csv
        - json

All STATA code is located in the folder STATA\dofiles. This code assumes that the folders STATA\data_in, Python\Metfrost\csv\, and Python\NILU\csv\ contain all datasets described in Table 1. These are the only folders that are required to have any content before the code is run – all other folders are used for storing results created by the programs. Note that the most detailed documentation is in the form of in-line comments in the do-files themselves.

5.1 Description of code

DATA PREPARATION:

All tables and figures are constructed using a set of STATA programs. The results are reproduced by running the master do-file 0_master.do.do located in the STATA\dofiles folder. Provided that it is set up correctly (see the section "Instructions" below) it calls the other scripts to generate data sets and replicate tables and figures.

- The preamble of the master do-file sets paths and globals (lines 15-40).
- Step 1: Explains that you need to run the Python scripts located in Python\NILU and Python\Metfrost to download data from the public APIs.
- Step 2: Here you import data, calculate variables, and merge datasets. This step consist of 10 do-files. The do-files creates key datasets that are used for subsequent analyses. The main dataset used for the analysis of air pollution is called "OneBigFile_NILU_Metfrost.dta" (available in the replication package). The main dataset used for analysis of cars passing the toll cordons is called "Ferde_2021_passeringer_15min_AGG.dta" (not available in the replication package).

We encourage researchers interested in replicating the results to contact us if they require any assistance in putting the relevant datasets together.

ANALYSIS:

- Step 3: This step produced the figures and tables found in the main text in the paper.
- Step 4: This step produced the figures and tables found in the online appendix.

The files that are called consist of the following steps, documented by in-line code: (1) open one of the main datasets from STATA\data_out; (2) do the necessary sample selections and creation of

variables; (3) run any necessary regressions, (4) store the results of regressions as .ster-files in the folder STATA\ster; (5) create tables and figures by accessing (a) the edited dataset and/or (b) importing estimation results from STATA\ster; and (6) storing these tables and figures in the folders STATA\tables and STATA\figures.

OUTPUT GENERATION:

The tables and figures are named as in the published article. As an example, Figure 6 in the article consists of panels (a) and (b). These files are created by the scripts Figure6a.do and Figure 6b.do, and saved as Figure6a.pdf and Figure6b.pdf, etc. in the STATA\figures folder. Figures are stored as .pdf while tables are stored in .tex format.

By running 0_master.do from start to finish, the following tables and figures will be produced:

Table 4. List of output files produced by running the script 0_master.do.

Main text		Online Appendix	
Tables:	**Figures:**	**Tables:**	**Figures:**
	Figure 6	Table F.1	Figure A.1
	Figure 7	Table F.2	Figure B.3
	Figure 8	Table F.3	Figure F.1
		Table B.1	Figure F.2
			Figure F.3
			Figure F.4
			Figure F.5
			Figure F.7
			Figure F.8

5.2 Instructions

To re-run the code, do the following steps:

Step 1: Obtain access to 15-minute traffic data from Ferde. Download raw data on air pollution, weather, and traffic censors outside the Bergen toll cordon from public APIs - or use the provided raw data.

Step 2: Save the main folders Python and STATA at your preferred location, including all subfolders.

Step 3: Ensure that the folder from STATA\data_in, Python/Metfrost, and Pyton/NILU contains all files described in Table 1, obtained from Step 1.

Step 4: Copy the filepath to where the main folder STATA is stored.

Step 5: Open the master script STATA\dofiles\0_master.do and change the global "home" on line 20 so that it matches your folder path. Also ensure that the command at line 24 points to the PLUS folder where .ado-packages are installed.

Step 6: In case your PLUS folder does not contain the packages listed at lines 48-54, uncomment and run these lines to install the required packages.

Step 7 (optional): In case the data you have obtained have different formats or use different naming conventions, make the necessary changes to the files under step 2 in Section 5.1.

Step 8: Run the master script `STATA\dofiles\0_master.do`. When everything is finished, you will find all figures and table inputs (as presented in Table 2) in the folders `STATA\tables` and `STATA\figures`.

Optional: As long as the globals from the master file have been set correctly and the datasets are created (i.e. step 1 and 2 in the master file), it is possible to run subsequent do-files one by one, as each do-file starts by loading a fresh version of the dataset(s).