README and guidance

1 Overview

This document describes the replication files for "Local Incentives and Electric Vehicle Adoption" by Askill H. Halse, Karen E. Hauge, Elisabeth T. Isaksen, Bjørn G. Johansen and Oddbjørn Raaum for the Journal of the Association of Environmental and Resource Economics (JAERE).

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 18. The analyses are carried out using proprietary data made available for research, under specific and strict regulations, by Statistics Norway.

2 Data availability and provenance statements

Statement about Rights

- **X** I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- ☐ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permission are documented in the LICENSE.txt file.

Summary of Availability

- ☐ All data **are** publicly available.
- ☐ Some data **cannot be made** publicly available.
- **X** **No data can be made** publicly available.

Except for a very limited set of supplementary data (the location of toll gates, fuel prices, electricity prices, consumer price indexes and changes in geographical classifications over time), no data can be made publicly available. See below for more details.

The information used in the analysis combines several Norwegian administrative registers (as described in the paper). The data use is subject to the European Union's General Data Protection Regulation (GDPR) per new Norwegian regulations from May 2018. The data used is on loan from Statistics Norway for a limited time-period, located at a secure server at the Frisch Centre and, due to confidentiality and security considerations, the data cannot be transferred outside the secure server or shared with others. Researchers interested in obtaining access to the register data employed in this paper are required to submit a written application to gain approval from Statistics Norway. The application must include a detailed description of the proposed project, its purpose, and its social contribution, as well as a description of the required datasets, variables, and analysis population. Applications can be submitted by researchers who are affiliated with Norwegian institutions accepted by Statistics Norway, or by researchers outside of Norway who collaborate with researchers affiliated with these institutions. The following link contains more information regarding access to data: https://www.ssb.no/en/data-til-forskning/utlan-av-data-til-forskere.

Any researchers interested in replicating the results of this paper are welcome to contact us for further information.

Some of the additional data sources for the online appendix, in particular related to input and output data from the national and regional transport models, are obtained from the Norwegian Public Roads Administration. We are not at liberty to share this information. However, other researchers might be able to obtain the same information by contacting the Norwegian Public Roads Administration directly. We are happy to cooperate in efforts to replicate our results, including seeking permission to access these data.

Finally, the location of toll gates, fuel prices, electricity prices, consumer price indexes and changes in geographical classifications over time are publicly available and can be downloaded using the links provided under "data references" (see below).

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 ***Data reference***	**Description of data and variables included in analysis**
`Househ_wp2.dta` *Source: Statistics Norway (2018a, c, d, f-j, m, o, s, 2019).*	A set of individual characteristics from various registers. Level of observations: (individual*year). Variables: Person ID, year, household ID, gender, #persons in household, age, basic statistical unit of residence, education code (NUS2000), income net of taxes, net wealth, employment dummy (labor income above one basic amount), retired dummy, home ownership dummy, second home ownership dummy, household type (composition of household members), number of adults in household, number of children below 18, number of children older than 18 not living with their parents. Sample: All individuals above 18 registered in Norway between 2005 and 2017.
`Work_grk_2005_2017.dta` *Source: Statistics Norway (2018a, b, e).*	Basic statistical units ('neighborhoods') of workplaces. Level of observations: (individual*year). Variables: Individual ID, year, workplace ID, basic statistical unit of workplace. Sample: All individuals' workplaces within Norwegian borders where location information is recorded in the registers between 2005 and 2017.
`Distances.dta` *Source: Halse et al (2022).*	Characteristics of trips between basic statistical units ('neighborhoods') in Norway along the road network. Level of observations: (basic statistical unit*basic statistical unit). Variables: basic statistical unit of origin, basic statistical unit of destination, distance,

	distance with bus lane, time, road toll by year. Sample: all origin-destination pairs in Norway where at least one individual commutes.
`Househ_select.dta` *Source: Statistics Norway (2018n).*	A set of all individual IDs and household IDs by year. Used to merge household ID to car ownership data, which has car ownership recorded by individual ID.
`Car_stock_end_of_year.dta` *Source: Statistics Norway (2018q).*	A dataset containing information about the car fleet, where car ownership is measured at December 31 each year. Level of observations: (car id*year). Variables: Car id, owner id, year, first registration date, fuel type, type of car transaction (new, second-hand or import). Sample: All passenger cars for private use that are part of the fleet Dec 31 each year.
`Kollektivtransport_grk` *Source: NPRA (2020a).*	A dataset containing level of service data for public transport from the Norwegian transport models. Level of observations: (origin basic statistical unit*destination basic statistical unit). Variables: origin, destination, time spent walking to/from stations, onboard time, waiting time, number of boardings, average cost per trip. Sample: All origin-destination pairs in Norway.
`Link_grkrets_regtype1.dta` *Source: NPRA (2020b).*	Type of geographic area. Level of observations: (basic statistical unit). Areas are split between the groups “Oslo”, “suburbs of Oslo”, “Bergen, Trondheim or Stavanger”, “suburbs of Bergen, Trondheim or Stavanger”, “other large cities”, “small cities” and “rural areas”. Sample: All of Norway.
`Driving_periods.dta` *Source: Statistics Norway (2018n, q, r).*	Odometer readings from cars at vehicle inspections (EU controls), including selected information about car owners. Level of observations: (car ID*driving period between EU controls*car owner). Note that driving periods, i.e. time between EU controls, can be split into several sub-periods if the car changes owner. Variables: Car ID, Driving period ID, driving sub-period ID, household ID, kilometers per day, start and end dates of driving period and sub-period, number of cars owned by the household, first registration date, diesel dummy, electric dummy, number of EU controls the car has been to, ID of other car owned by the household (if the household owns several, the newest one is selected). The same time-invariant car attributes for the other car. Sample: All EU controls for passenger vehicles between 2005 and 2017.
`Fuel_prices_taxes.xlsx`	National fuel prices and taxes. Level of observations: (month*year). Variables: year, month, gasoline price,

Source: Statistics Norway (2020a, b).	gasoline tax, diesel price, diesel tax, kWh price, kWh tax. Sample: 2005-2017.
`KPI.dta` *Source: Statistics Norway (2020c).*	List of consumer price indexes, where year 2017 is normalized to 100. Level of observations: (year). Variable: Consumer price index.
`Losdata_viken.DBF` *Source: NPRA (2020c).*	A dataset containing level of service data for cars from the transport model, used to create a measure of congestion. Level of observations: (origin basic statistical unit*destination basic statistical unit). Variables: origin ID, destination ID, driving time outside rush hours, driving time during morning rush, driving time during afternoon rush, distance. Sample: Oslo and Viken county.
`Lenker_viken_2020.DBF` *Source: NPRA (2020d).*	A dataset containing information from the transport models. Used to create Figure A5. Level of observations: (road link). Variables: driving time outside rush hours, driving time during morning rush, driving time during afternoon rush, distance, public transport lane dummy. Sample: All road links within Oslo and Viken.
`Grk_link_to_grk_number_in_2014.dta` *Source: Statistics Norway (2020d).*	A mapping between basic statistical unit IDs in 2014 and IDs used during subsequent years, in case the basic statistical unit ID has changed.

See in-line comments in the scripts `STATA\dofiles\1_CreateMainDataset.do` and `STATA\dofiles\2_CreateDrivingData.do` for more detailed information regarding each data file. Sources for each file are listed under "Data references" below.

Data references

Halse, Askill Harkjerr; Johansen, Bjørn Gjerde; Sand, Øystein (2022). Road tolls in Norway. Can be accessed at DOI: https://doi.org/10.18710/M2D2XP (last accessed 2022).

NPRA (2020a). Level of service data for public transport from the regional and national Norwegian transport models (`kollektivtransport_grk`). Norwegian Public Roads Administration, Oslo (last accessed: 2020).

NPRA (2020b). Codes for mapping basic statistical units to aggregate geographic areas (`Link_grkrets_regtype1`). Norwegian Public Roads Administration, Oslo (last accessed: 2020).

NPRA (2020c). Level of service data for cars from the regional transport model for the Oslo/Viken area (`Losdata_viken`). Norwegian Public Roads Administration, Oslo (last accessed: 2020).

NPRA (2020d). Network data on the road link level from the regional transport model for the Oslo/Viken area (`lenker_viken`). Norwegian Public Roads Administration, Oslo (last accessed: 2020).

Statistics Norway. (2018a). Wage information until 2015 (atmlto2005- atmlto2005; 10 separate datasets). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018b). Firm locations (bed_for_grk). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018c). Constant person characteristics (befolkning). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018d). Variable person characteristics (befolkning_tp). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018e). Firms and enterprises (bof2005-bof2015; 11 separate datasets). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018f). Living conditions (bofor2015-bofor2016; 2 separate datasets). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018g). NUS code for highest ongoing education (bu_igang). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018h). Information regarding completed education (f_utd_demografi). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018i). Marital status (famsam). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018j). Family type (famtyp). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018k). Driver's license status (fk2005-fk2016; 12 separate datasets). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018l). Changes in registered residences (flytting). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018m). Basic statistical unit of residence (grunnkrets). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018n). Household specific demographics (hush_fam2005-hush_fam2017; 13 separate datasets). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018o). Income from all sources (inntekt). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018p). Odometer readings from EU controls (kjl2005-kjl2017; 13 separate datasets). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018q). The central motor vehicle register (mreg_g2005-mreg_g2017; 13 separate datasets). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018r). Active roles in enterprises and organizations (roller). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2018s). Wealth and debt (skattestat). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway. (2019). Ownership of secondary home (`sek_fritid_g2012-sek_fritid_g2012`; 5 separate datasets). Statistics Norway, Oslo/Kongsvinger (last accessed: 2018)

Statistics Norway (2020a). Fuel prices. Downloaded from the Statbank (Table 09654, available at: https://www.ssb.no/en/statbank/table/09654). Statistics Norway, Oslo/Kongsvinger (last accessed: 2020)

Statistics Norway (2020b). Electricity prices. Downloaded from the Statbank (Table 09387, available at: https://www.ssb.no/statbank/table/09387/). Statistics Norway, Oslo/Kongsvinger (last accessed: 2020)

Statistics Norway (2020c). Consumer price index. Downloaded from the Statbank (Table 08184, available at https://www.ssb.no/statbank/table/08184/). Statistics Norway, Oslo/Kongsvinger (last accessed: 2020)

Statistics Norway (2020d). Changes in regional classifications over time. Downloaded from Statistics Norway (available at: https://www.ssb.no/metadata/alle-endringer-i-de-regionale-inndelingene). Statistics Norway, Oslo/Kongsvinger (last accessed: 2020)

4 Computational requirements

4.1 Hardware requirements

Estimations were done on a Windows 10 virtual machine.

- CPU speed: 2.10 GHz, 40 virtual processors
- Installed memory (RAM): 160 GB

We do not recommend attempting to replicate the results without access to a relatively fast computer/server, due to the computationally intensive regressions.

The largest .dta dataset that is read into memory is approximately 10 GB. The full list of data under "data references" will take up approximately 100 GB of hard drive space.

4.2 Software requirements

All code is run using Stata (code was last run using Stata 18, but will work using Stata 17 as well).

The following .ado-programs from SSC are used:

- estout
- reghdfe
- egenmore

These packages can be installed by uncommenting the lines 14-17 in the script `STATA\dofiles\0_Masterfile.do`.

4.3 Runtime requirements

Given the aforementioned hardware, total run time (including results for the online appendix) is approximately 57.5 hours.

5 Description of programs/code

The replication package has the following folder structure:

- STATA
  - data
    - in
    - out
  - dofiles
  - output
    - figures
    - ster
    - tables

All code is located in the folder STATA\dofiles. This code assumes that the folder STATA\data\in contains 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_Masterfile.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 first step of the master do-file sets paths (lines 11-28) and globals (lines 30-76). These globals control which explanatory/control variables that are used in the various regressions.
- The second part merges the datasets located in the folder STATA\data\in together. At line 82, the file STATA\dofiles\1_CreateMainDataset.do is called to create the main dataset on the household-year-level. At line 95, the file STATA\dofiles\2_CreateDrivingData.do is called to create a dataset of driving measurements at the car-periodic vehicle inspection-household level. These scripts produce two datasets stored in the folder STATA\data\out that are used for all subsequent analyses.

As explained above, we are not at liberty to share any of the data located in the folder STATA\data\in. Researchers wanting to replicate the results must apply for these data themselves, as explained in the "Data availability and provenance statements" above. It is unlikely that researchers will be able to obtain the exact same files as we have utilized, since Statistics Norway's process of creating datasets and naming variables might differ across commissions and caseworkers. Therefore, the abovementioned scripts for creating the datasets will likely have to be modified. However, the scripts should contain enough information in the form of in-line comments for this process to be straight-forward. We encourage researchers interested in replicating the results to contact us if they require any assistance in putting these datasets together.

ANALYSIS:

- The third step of the master do-file call separate files to replicate the tables (lines 103-110) and figures (lines 112-118) from the paper. Appendix figures and tables are replicated in the same manner by calling separate do-files below, from line 120 and onwards.

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 (in some cases this includes merging additional data to the main datasets from `STATA\data\in`); (3) run any necessary regressions, (4) store the results of these regressions as .ster-files in the folder `STATA\output\ster`; (5) create tables and figures by accessing (a) the edited dataset and/or (b) importing estimation results from `STATA\output\ster`; and (6) storing these tables and figures in the folders `STATA\output\tables` and `STATA\output\figures`.

OUTPUT GENERATION:

The tables and figures are named as in the published article. As an example, Figure 3 in the article consists of panels (a) through (d). These files are created by the script `Figure3.do`, and saved as `Figure3a.pdf`, `Figure3b.pdf`, etc. in the `STATA\output\figures` folder. Figures are stored as .pdf and .png, while tables are stored in .tex format.

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

Table 2. List of output files produced by running the script 0_Masterfile.do.

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

Comments, output to main text:

- Note that Table 6 is not produced by code, put together manually. However, the second column of Table 6 includes detailed enough information about both the source of each number as well as the formula for calculation that replication of this table should be straightforward.

Comments, output to online appendix:

- Figure A.1 is not produced by code but created manually in QGIS by importing the coordinates of toll gate locations as a layer. This figure is therefore not a part of the replication package. Data on toll gate level, including their coordinates, is publicly available and can be downloaded by using the following DOI: https://doi.org/10.18710/M2D2XP. Additional documentation of the toll gate data can be found in the report: Sand, Øystein; Bjørn Gjerde Johansen; Askill Harkjerr Halse; Svein Olav Sæter. 2022. Road Tolls in Norway, 2005-2021. TØI Report 1903/2022. url: https://www.toi.no/publications/road-tolls-in-norway-2005-2021-article37699-29.html (last accessed 2022).
- Tables A.2 and A3 are created by hand. The do-file `TableA2_A3.do` reproduces all numbers in the tables by means of the `xtsum` command and prints them to screen, but do not store the results in a `.tex` file.
- Table D.2 is created by hand. The do-file `TableD2.do` reproduces all regression results from Table D.2 and prints it to screen, where coefficients are listed row-wise and four different regressions with different outcome variables are listed column-wise. Table D.2 is a transposed version of this table.

5.2 Instructions

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

Step 1: Obtain access to register data from Statistics Norway (see elsewhere in this document and https://www.ssb.no/en/data-til-forskning/utlan-av-data-til-forskere for more information). Provided that the request is granted (i.e. is in accordance with the Statistics Act of Norway), obtaining data will take approximately 0.5-1 year and cost between 50,000 and 300,000 NOK.

Step 2: Save the main folder `STATA` at your preferred location, including all subfolders. This location must be on a secure server that satisfies requirements for storage of indirectly identifiable personal data.

Step 3: Ensure that the folder from `STATA\data\in` 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_Masterfile.do` and change the global "`home`" on line 20 so that it matches your folder path. Also ensure that the command at line 11 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 13-16, uncomment and run these lines to install the required packages.

Step 7 (optional): In case the data you have obtained from Statistics Norway have different formats or use different naming conventions, make the necessary changes to the files `STATA\dofiles\1_CreateMainDataset.do` and `STATA\dofiles\2_CreateDrivingData.do`.

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

Optional: As long as the globals from the master file have been set correctly and the datasets are created (i.e. lines 1-95 of the master file is ran), it is possible to run subsequent input files one by one, since each input file starts by loading a fresh version of the dataset(s).