README FILE

"Replication package for: Abatement Strategies and the Cost of Environmental Regulation: Emission Standards on the European Car Market"

Data Availability Statements

The main data used in the paper are confidential, but may be obtained through Data Use Agreements with:

- Data containing information on vehicle prices, sales and attributes was purchased from Jato Analytics. The data panel containing information for 1998 to 2011 was purchased in several waves during the period 2006-2012 by the research unit of Frank Verboven, KU Leuven. Author Mathias Reynaert was working in this research unit when the project was initiated in 2015. Jato can be contacted at:

  Hunton House, Highbridge Estate, Oxford Road, Uxbridge, UB8 1LX

  +44 (0)20 8423 7100, enquiries@jato.com, jato.com

- Data containing information on fuel station visits in the Netherlands was purchased from Travelcard NV in July 2015. Travelcard NV can be contacted at:

  Travelcard P.J. Oudweg 4 Postbus 1324, 1300 BH Almere, +31(0)88 110 5000

  info@travelcard.nl, https://travelcard.nl/

- Data containing information on production sites, production quantity and characteristics is obtained from Pwc Autofacts:
  PwC Autofacts Germany,
  https://www.pwccn.com/en/industries/automotive/autofacts.html, Contact person German office: patrick.wiesner@pwc.com

To obtain the data a researcher can contact the data provider with a purchase request detailing the period, level of aggregation and variables that the researcher would like to purchase. The data provider usually offers a range of different options with different prices. Based on these options a price is negotiated and a purchase contract specifying the usage agreement between the academic institution of the researcher and the data provider is drafted. Finally, the data is delivered to the researcher. In my experience, the time between first contact and data delivery takes at least a month, sometimes much longer. The length of the procedure depends on administrative, budgetary and legal approvals in the academic institution of the researcher. The pricing can be dependent on timing of request and the local office of the data provider, but in general access to long detailed panels of prices and quantities on vehicle markets is expensive. IHS Markit is a competitor of Jato Analytics and provides similar data.

The following data is accessible for academic researchers:

- Data series by country year obtained from Thomas Reuters Refinitiv Eikon DataStream: gasoline and diesel fuel prices, GDP per capita, unit labor costs, and population. This data is often accessible through university libraries. See: https://www.refinitiv.com/en/products/datastream-macroeconomic-analysis
- Data series by country year on number of household members obtained from Eurostat: https://ec.europa.eu/eurostat/web/products-datasets/product?code=lfst_hhanwhtc
- Data for the USA vehicle attribute trends in Appendix Figure A1 is publicly available: https://www.epa.gov/automotive-trends/about-automotive-trends-data

Dataset list

Data file	Source	Notes	Provided
`version.dta`	Combined Jato, PwC, Datastream and Eurostat data	Confidential, used in stata_dofile.do, base dataset for versionA, versionB and Version C. Used for Table 1, 2, 3, A2, A3, A4 and A5 . Used for Figure 1, Figure A1.	No*
versionA	Generated in stata_dofile	Data 2011 used for Table A6	No
versionB	Generated in stata_dofile	Data 98-2007 used in demand and supply estimation Table 4, A6	No*
versionC	Generated in stata_dofile	Data 2007 for Table 5, 6, A7, A8, and Figure 2	
EPA_US	EPA	Used for Figure A1	Yes
CARLEVELgaps	TravelCard	Vehicle level gaps constructed in Sallee and Reynaert (forthcoming, American Economic Journal: Economic Policy). Used in Table 2, 3, A2, A3 and A4.	No
marketdata.dta		Marketdata contains the market level variables used in version.dta and versionA/B/C.	Yes
versionB_sample.dta		Gives a random sample of 10 data points with prices and quantities left blank.	Yes

We purchased the Jato Analytics data in waves over several years. At each step the PwC and macro variables from Eurostat and DataStream were added to construct a consistent panel. The file version.dta is a collapse to the engine version level of that panel, consistent with what we believe a researcher purchasing data from Jato now could receive as a single data file covering the same time period and merging the additional data. My programs begin with that data file. The data on on-road vehicle performance is discussed extensively in

Sallee and Reynaert (AEJ: Economic Policy), more specifically, I use vehicle level estimates estimated in Section G of that paper.

Computational requirements

Software Requirements

- Stata (code was last run with version 15)
- Matlab (code was run with Matlab Release 2018a) code uses 'Optimization Toolbox' and 'Statistics and Machine Learning Toolbox'
- Knitro Artelys (code was run with version 12), whenever Knitro is used a researcher could use Matlab optimizer. I add the code to this wherever applicable.

Memory and Runtime Requirements

The Stata code requires minutes on a standard desktop computer, the demand estimation 24 hours, the full set of simulations about 24 hours and the bootstrap 2 weeks. The code ran on a Windows Server computer that allowed Matlab to run in parallel with 14 workers. No specific memory or CPU requirements are needed to run the programs.

Description of programs

Folder codes:

- Stata:
  - Program 'stata_dofile.do' loads and prepares the data. It executes commands to generate Table 1, 2, 3, A2, A3, A4 and A5, Figure 1 and Figure A1.
  - Data input: `version.dta,` CARLEVELgaps.dta
  - Data output: versionA, versionB, versionC
  - Program usgraph.do loads EPA_US.dta to generate the US panel in Figure A1
- Matlab
  - createmfilesform.m and importfile.m loads the datasets versionA, versionB, versionC into Matlab
  - Subfolder Estimation contains the scripts for the demand estimation, the main file is named estimation.m
  - Subfolder Counterfactual contains the scripts for simulations, the main file is called Simulation_FOCS.m

Folder data

- Contains datasets EPA_US.dta, marketdata.dta and versionB_sample

Folder output

- Contains eps format for Figure 1 and Figure 2 and pdf format for Appendix Figure A1
- Contain excel file with all Tables as well as regression output written from the Stata and Matlab commands.

List of tables and programs

Figure/Table #	Program	Line Number	Note
Figure 1	dofile_stata.do	515-533	
Table 1	dofile_stata.do	509-511	
Table 2	dofile_stata.do	554-979	
Table 3	dofile_stata.do	982-1037	
Table 4	estimation.m	126-156	Logit Z2
		197-409	RC Logit Z2
	estimation_priceend.m	117-147	Logit Z1
		187-397	RC Logit Z1
	Simulation_FOCS.m	208-250	Marginal Costs Column 3
	Simulation_FOCS_NE.m	208-250	Marginal Costs Column 1 and 2
Table 5	Simulation_FOCS.m	450-733	
	Simulation_FOCS_BS.m	1-748	Computes Stand. Error on I-VII
Figure 2	Simulation_FOCS.m	838-852	
Table 6	Simulation_FOCS.m	666-703; 644-663; 607-618	Table combines results from Scenario IV, VI, III and IV with restriction
Figure A1	dofile_stata.do usgraph.do	481-506 1-24	EU Panel US Panel
Table A1	No code		
Table A2	dofile_stata.do	554-979	

Table A3	dofile_stata.do	554-979	
Table A4	dofile_stata.do	554-979	
Table A5	dofile_stata.do	414-479	
Table A6	estimation.m modelfit.m	416-493 1-64	Within Sample Out of Sample
Table A7	Simulation_FOCS_NE.m Simulation_FOCS.m	542-589 736-826	Column I Column II, III and IV
Table A8	Simulation_FOCS	556-589	

References

JATO ANALYTICS, 2012, Data on Vehicle Sales, Prices and Characteristics for EU Countries

1998-2011

TravelCardNV, 2015, Data on Fuel Station Visits Travelcard Holders 2004-2015

PwC Autofacts, 2012, Data on global production sites, quantities and capacities

Eurostat, 2012, Average number of persons per household by household composition, number of children and working status within households." URL https://ec.europa.eu/eurostat/web/

products-datasets/product?code=lfst_hhanwhtc.

Datastream International, 2012, Available: Datastream International." URL https://eikon.

thomsonreuters.com/index.html.

United States Environmental Protection Agency, 2012, Automotive Trends Data, URL https://www.epa.gov/automotive-trends/about-automotive-trends-data

Grigolon, Laura, Mathias Reynaert, and Frank Verboven, 2018, Consumer Valuation of Fuel Costs and Tax Policy: Evidence from the European Car Market, American Economic Journal: Economic Policy, 10 (3):193-225

Grigolon, Laura, Mathias Reynaert, and Frank Verboven, 2018, Replication data for: Consumer Valuation of Fuel Costs and Tax Policy: Evidence from the European Car Market, openicpsr-114687, https://www.openicpsr.org/openicpsr/project/114687/version/V1/view

Reynaert, Mathias and James M. Sallee, 2020, Who Benefits When Firms Game Corrective Policies?, American Economic Journal: Economic Policy, forthcoming

Reynaert, Mathias and James M. Sallee, 2020, Replication data for: Who Benefits When Firms Game Corrective Policies?, openicpsr-115161