

Supporting Information to:

Historical Penetration Patterns of Automobile Electronic Control Systems and Implications for Critical Raw Materials Recycling

Eliette Restrepo^{1,2,*}, Amund N. Løvik¹, Rolf Widmer¹, Patrick Wäger¹ and Daniel B. Müller²

¹ Empa, Swiss Federal Laboratories for Material Science and Technology; St. Gallen, CH-9014, Switzerland; amund.loevik@empa.ch (A.N.L.); rolf.widmer@empa.ch (R.W.); patrick.waeger@empa.ch (P.W.)

² Industrial Ecology Programme and Department of Energy and Process Engineering, Norwegian University of Science and Technology–NTNU; Trondheim, NO-7491, Norway; daniel.mueller@ntnu.no (D.B.M)

* Correspondence: eliette.restrepo@empa.ch

This supplement contains:

12 Pages

6 Figures

4 Tables

Contents

1.	Software.....	2
2.	Model variables	3
3.	Additional dataset distribution plots	4
4.	Association analysis.....	6
5.	Average historical penetration of automobile electronic control systems	7
6.	Additional results of Logistic GAM	9
7.	References.....	12

Figures

Figure S1.	Distribution of the dataset by additional car characteristics.....	4
Figure S2.	Association (correlation) among input variables.....	6
Figure S3.	Relationship between car characteristics and the presence of the electronic stability program (ESP) in new cars	9
Figure S4.	Relationship between selected car characteristics and the penetration of the navigation system (Navi) in new cars.....	10
Figure S5.	Average predicted penetration of the electronic stability program (ESP) in the validation set	11
Figure S6.	Average predicted penetration of the navigation system (Navi) in the validation set.....	11

Tables

Table S1.	R software packages used in the statistical analysis.....	2
Table S2.	Description of car characteristics considered in this study.....	3
Table S3.	Average penetration of the investigated ECS by production year of the cars.....	7
Table S4.	Average penetration of additional ECS by production year of the cars.....	8

1. Software

All the statistical analysis was carried out in R [1]. The R packages used are listed in Table S1 below.

Table S1. R software packages used in the statistical analysis.

Task	R package and reference
Plots: Histograms, historical penetrations, and variable importance results	cowplot[2] easyGgplot2[3] ggplot2 [4] gridExtra[5] grid[6] reshape2[7]
Computing the average penetration of ECS	dplyr[8] summarySE (function)[9] officer[10] flextable[11] magrittr[12]
Random forest model	ranger[13]
Logistic GAM	mgcv[14] mgcViz[15]

2. Model variables

Table S2 below lists the 14 car characteristics considered in the study (input variables), including their unit, categories (or value ranges), model variable name and data type.

The output variables (ECS considered) are listed in Table 1 of the methods section in the main manuscript. All output variables were of dichotomous data type (categorical with only two categories – one and zero)

Table S2. Description of car characteristics considered in this study.

Characteristic	Unit	Range/categories	Variable name	Data type
Price	CHF	[7 948–702 240]	Price	integer
Production year	a	[2001–2015]	Years	integer
Engine displacement	cm ³	[0-8285]	Displacement	integer
Mass	kg	[350 - 3050]	Mass	integer
Fuel type	n.a.	Electric Hybrid electric-diesel Hybrid electric-gasoline Diesel Gasoline	Fuel	categorical
Engine/Motor Power	kW	[4 - 552]	Power	integer
Body type	n.a.	Convertible Coupe Sedan Sports car Station wagon SUV Van	Body	categorical
Drivetrain type	n.a.	All wheel Front wheel Rear wheel	Drive	categorical
Transmission type	n.a.	Automatic Manual Semi-automatic Stepless	Transmission	categorical
Number of seats	n.a.	[1-9]	Seats	integer
Class	n.a.	Mini Small Low-medium Upper-medium Upper Luxury	Class	categorical
Number of doors	n.a.	[2-5]	Doors	integer
Number of engine cylinders	n.a.	0,2,3,4,5,6,8,10,12	Cylinders	integer
Number of gears	n.a.	[1-8]	Gears	integer

n.a. : does not apply.

3. Additional dataset distribution plots

Figure S1 below presents the distribution of the cars by the remaining eight characteristics considered.

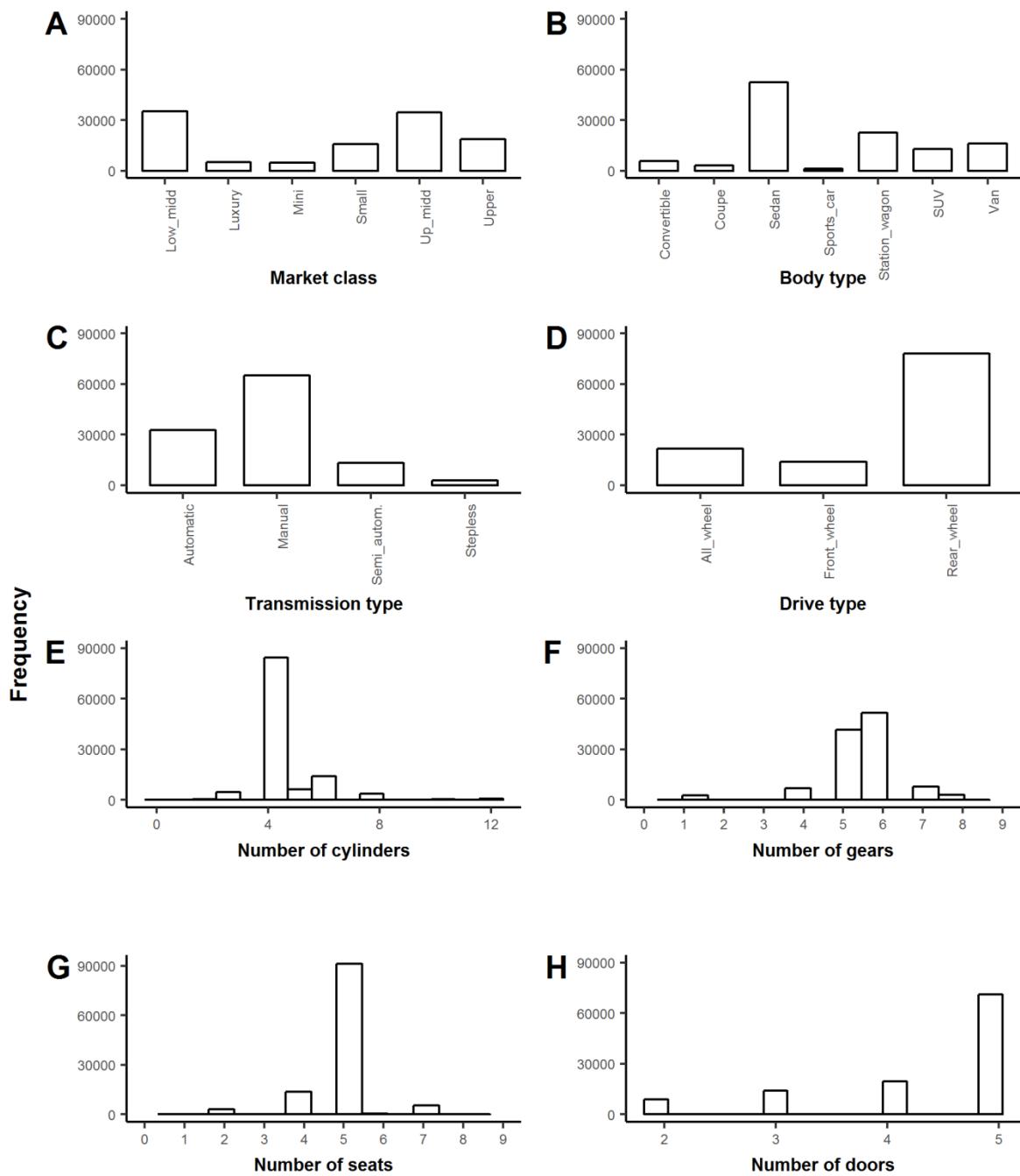


Figure S1. Distribution of the dataset by additional car characteristics: A) market class, B) body type, C) transmission type, D) drive type, E) number of cylinders, F) number of gears, G) number of seats and H) number of doors.

4. Association analysis

We analyzed the association among the input variables in order to discard redundant input variables from the statistical analysis. Specifically, we computed the *Goodman and Kruskal's tau measure of association* ($\tau(x,y)$) which indicates, with a value ranging from zero (0) to one (1) whether two input variables x and y are independent or perfectly predictable from each other. This measure was preferred due to its suitability for associations between numerical and categorical variables [16].

The correlogram in Figure S2 below presents the association between all pairs of input variables. The scale on the right side indicates the level of association: 0 (white) means the variables are independent; 1(bright red) means the variables are redundant. For example, the number of engine cylinders ("Cylinders") exhibited a slight ability to explain the size of the car's engine ($\tau(x,y) = 0.39$). Likewise, the car class ("Class") slightly explained variations in the car's mass ($\tau(x,y) = 0.39$). As all associations were weak ($\tau(x,y) < 0.5$) the variables were considered to be independent.



Figure S2. Association (correlation) among input variables. The association is measured by the Goodman and Kruskal's tau measure of association ($\tau(x,y)$) which ranges between 0 (no association) and 1(perfect association).

5. Average historical penetration of automobile electronic control systems

Data for the average penetration of established and emerging ECS displayed in Figure 3, Figure 4 and Figure 6 in the main manuscript is presented in Tables S3 below. Table S4 presents additional ECS not analyzed in detail in this study.

Table S3. Average penetration of the investigated ECS by production year of the cars.

Year	CC	ESP	ABS	ABS or ESP	Navi	stop_start	auto_pilot	jam_assist	park_pilot
2001	0.28	0.35	0.59	0.93	0.34	-	-	-	0.03
2002	0.30	0.39	0.55	0.94	0.40	-	-	-	0.03
2003	0.29	0.44	0.52	0.96	0.42	-	-	-	0.03
2004	0.31	0.49	0.49	0.98	0.45	-	-	-	0.02
2005	0.30	0.55	0.43	0.99	0.47	-	-	-	0.01
2006	0.30	0.56	0.42	0.99	0.47	-	-	-	0.00
2007	0.33	0.60	0.39	0.99	0.49	0.001	-	-	0.00
2008	0.33	0.64	0.34	0.99	0.52	0.004	-	-	0.02
2009	0.40	0.60	0.38	0.99	0.53	0.02	-	0.002	0.02
2010	0.39	0.66	0.33	0.99	0.50	0.05	-	0.004	0.02
2011	0.37	0.72	0.27	0.99	0.47	0.14	-	0.003	0.03
2012	0.35	0.74	0.26	0.99	0.48	0.25	0.004	0.01	0.03
2013	0.38	0.74	0.24	0.98	0.50	0.33	0.01	0.02	0.04
2014	0.36	0.79	0.19	0.98	0.46	0.40	0.02	0.03	0.03
2015	0.38	0.76	0.21	0.97	0.45	0.49	0.02	0.04	0.03

CC: Cruise control, ESP: Electronic stability program, ABS: anti-lock braking system, navi: Navigation system, stop_start: Stop-start system.

Table S4. Average penetration of additional ECS by production year of the cars.

Year	EPS	TPMS	FCW	EBS	DAD	EDR	LCA	LDW	LKA	DSR	PA	PDC	ADAS	ACC
2001	0.04	0.03	-	0.22	-	-	-	0.12	-	-	-	0.12	-	-
2002	0.05	0.07	-	0.32	-	-	-	0.15	-	-	-	0.15	-	0.003
2003	0.07	0.09	-	0.43	-	-	-	0.13	-	-	-	0.13	-	0.004
2004	0.09	0.11	-	0.54	-	-	0.001	0.14	0.002	-	-	0.14	-	0.01
2005	0.13	0.15	-	0.62	-	-	0.002	0.16	0.01	0.001	-	0.16	-	0.02
2006	0.16	0.15	-	0.62	-	-	0.002	0.17	0.01	0.01	-	0.17	-	0.02
2007	0.16	0.18	-	0.64	-	-	0.003	0.22	0.02	0.01	-	0.21	-	0.04
2008	0.18	0.20	-	0.63	-	-	0.003	0.23	0.03	0.01	-	0.23	-	0.04
2009	0.15	0.20	-	0.66	0.002	0.002	0.01	0.22	0.02	0.02	-	0.22	-	0.03
2010	0.15	0.19	-	0.66	0.004	0.004	0.01	0.21	0.03	0.01	-	0.21	-	0.03
2011	0.15	0.25	-	0.69	0.003	0.004	0.02	0.20	0.05	0.01	-	0.20	0.002	0.03
2012	0.15	0.31	0.004	0.69	0.01	0.01	0.02	0.23	0.07	0.02	0.01	0.22	0.002	0.03
2013	0.15	0.43	0.01	0.73	0.02	0.02	0.03	0.26	0.07	0.03	0.01	0.24	0.002	0.03
2014	0.16	0.59	0.02	0.67	0.03	0.04	0.05	0.26	0.08	0.04	0.01	0.24	0.002	0.04
2015	0.14	0.72	0.03	0.67	0.04	0.04	0.06	0.27	0.09	0.04	0.01	0.26	0.003	0.03

EPS: Electronic power steering, TPMS: Tire pressure monitoring system, FCW: Front collision warning, EBS: Emergency braking system, DAD: Drowsiness and attention detection, EDR: Event data recorder, LCA: Lane change assist, LDW: Lane departure warning, LKA: Lane keeping assist, DSR: Driver steering recommendation, PDC: Park-distance control, PA: Basic park assist, ADAS: Advance driver assistance system, ACC: Adaptive Cruise Control

6. Additional results of Logistic GAM

Figure S3 and Figure S4 show the Logistic GAM for the ESP and the Navi, respectively. The corresponding predictions in the validation set are shown in Figure S5 and Figure S6, respectively.

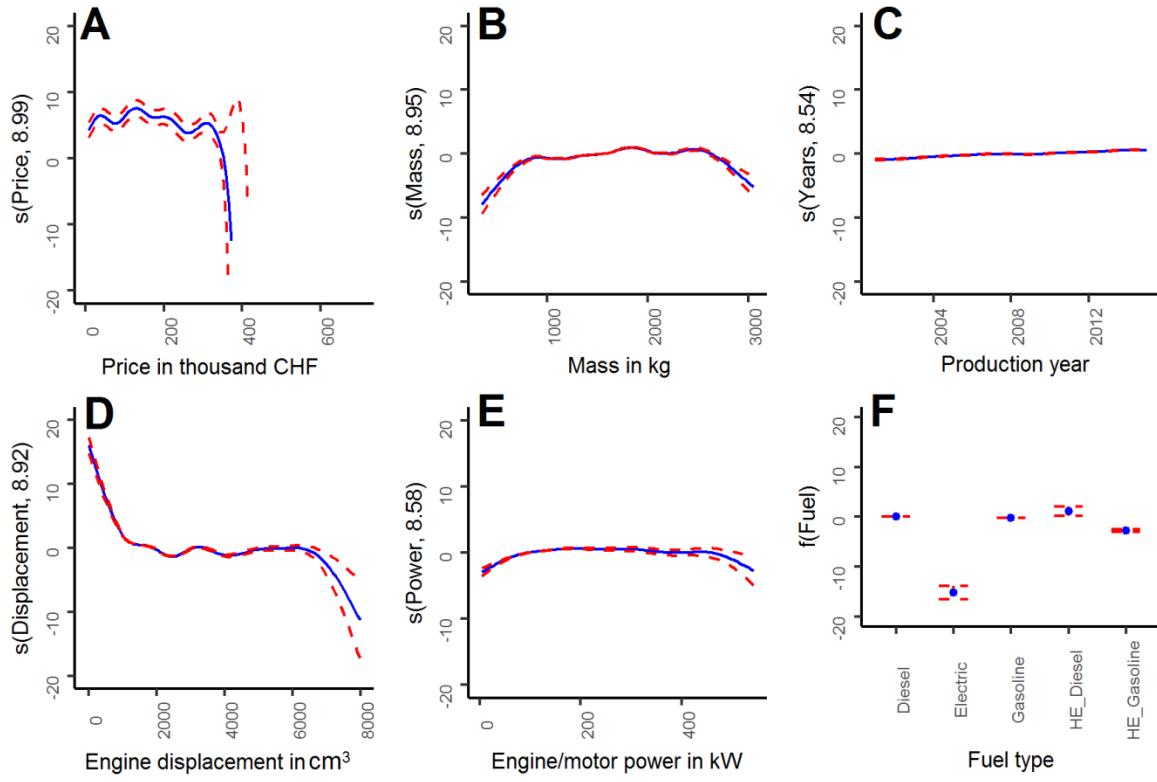


Figure S3. Relationship between car characteristics and the presence of the electronic stability program (ESP) in new cars. The blue lines represent the smoothing splines fit to each predictor after a Logistic GAM: (A) price, (B) mass, (C) production year, (D) engine displacement, (E) engine/motor power, (F) fuel type. The red dotted lines correspond to the confidence interval of the estimated function. The scale of the y-axis corresponds to the model transformation of the response variable and it is not given in the units of the response; the value in parenthesis indicates the degrees of freedom of each function.

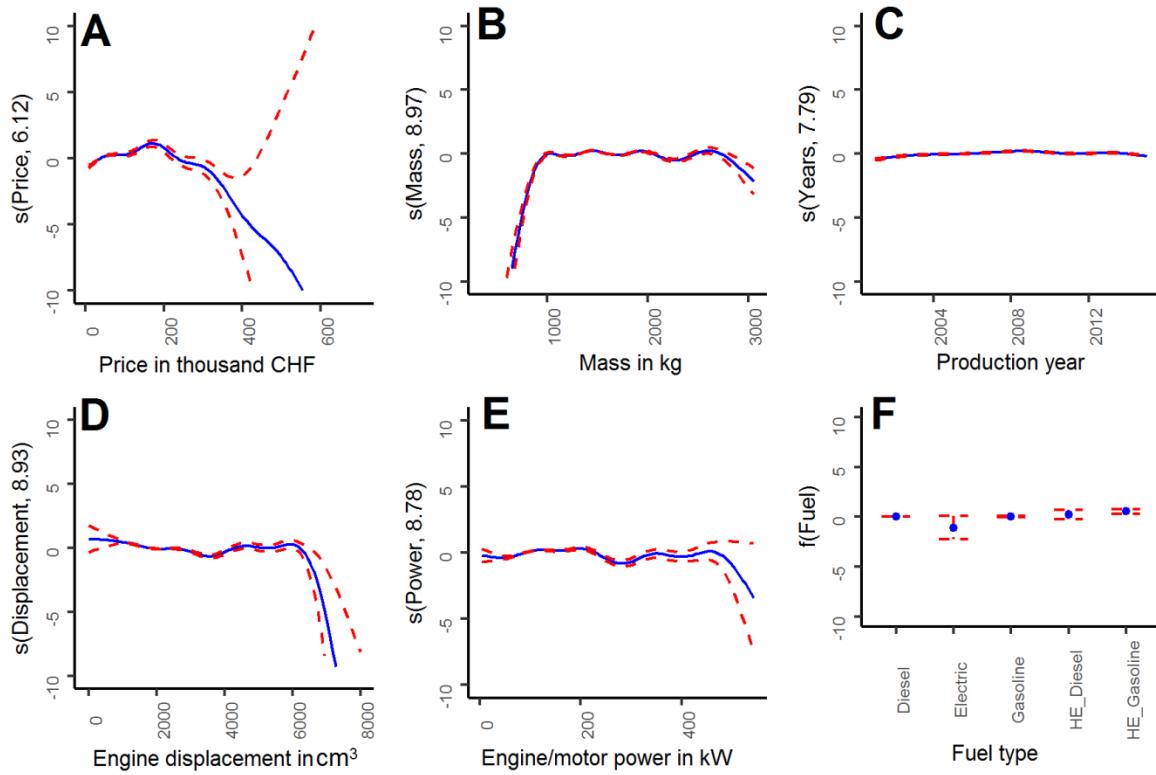


Figure S4. Relationship between selected car characteristics and the penetration of the navigation system (Navi) in new cars. The blue lines represent the smoothing splines fit to each predictor after a Logistic GAM: (A) price, (B) mass, (C) production year, (D) engine displacement, (E) engine/motor power, (F) fuel type. The red dotted lines correspond to the confidence interval of the estimated function. The scale of the y-axis corresponds to the model transformation of the response variable and it is not given in the units of the response; the value in parenthesis indicates the degrees of freedom of each function.

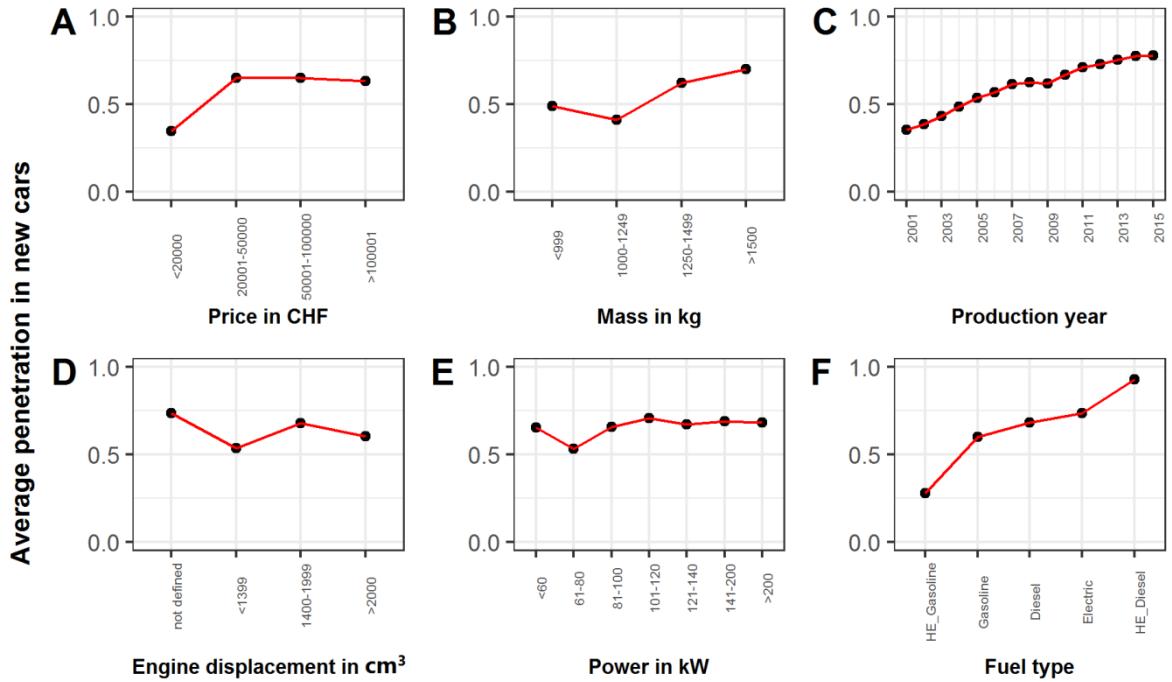


Figure S5. Average predicted penetration of the electronic stability program (ESP) in the validation set by the car's: (A) price, (B) mass, (C) production year, (D) engine displacement, (E) power, (F) fuel type.

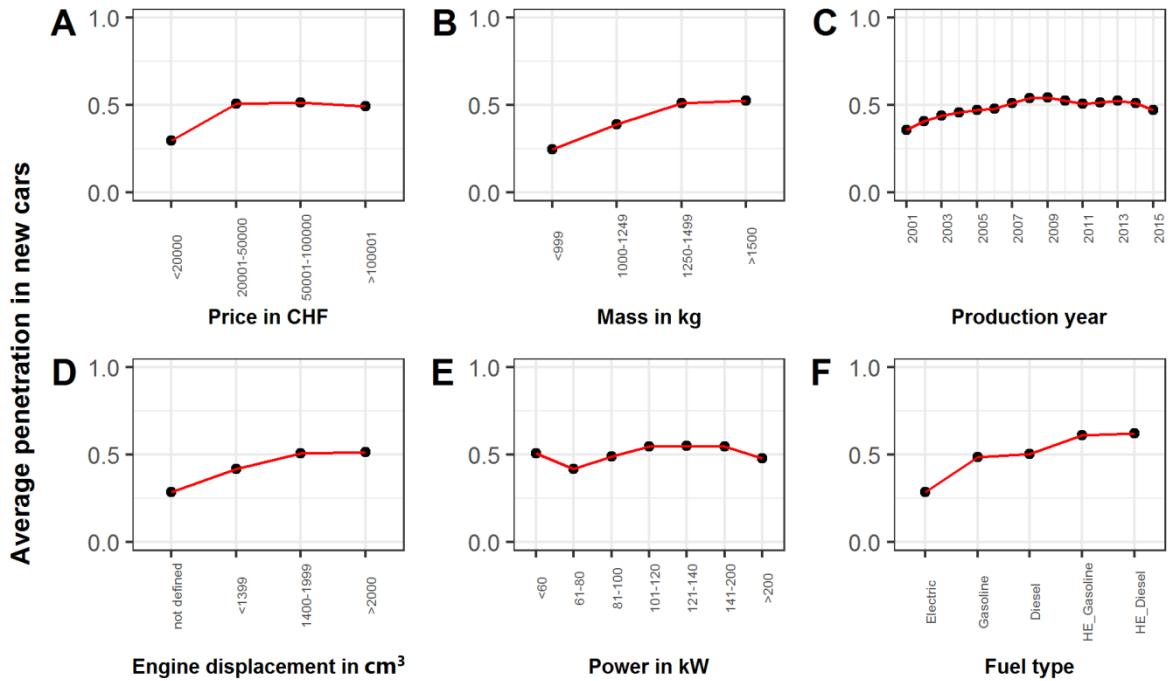


Figure S6. Average predicted penetration of the navigation system (Navi) in the validation set by the car's: (A) price, (B) mass, (C) production year, (D) engine displacement, (E) power, (F) fuel type.

7. References

1. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
2. Claus O. Wilke (2018). cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 0.9.3. <https://CRAN.R-project.org/package=cowplot>.
3. Alboukadel Kassambara (2014). easyGgplot2: Perform and customize easily a plot with ggplot2. R package version 1.0.0.9000. <http://www.sthda.com>.
4. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
5. Baptiste Auguie (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. <https://CRAN.R-project.org/package=gridExtra>.
6. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
7. Hadley Wickham (2007). Reshaping Data with the reshape Package. *Journal of Statistical Software*, 21(12), 1-20. URL <http://www.jstatsoft.org/v21/i12/>.
8. Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.0.1. <https://CRAN.R-project.org/package=dplyr>.
9. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
10. D. Gohel, F. Hangler, L. Sander, J. Calder, and J. Harrold, *Package "officer."* R Foundation for Statistical Computing, 2018.
11. David Gohel (2018). officer: Manipulation of Microsoft Word and PowerPoint Documents. R package version 0.3.2. <https://CRAN.R-project.org/package=office>.
12. Stefan Milton Bache and Hadley Wickham (2014). magrittr: A Forward-Pipe Operator for R. R package version 1.5. <https://CRAN.R-project.org/package=magrittr>.
13. Marvin N. Wright, Andreas Ziegler (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, 77(1), 1-17. doi:10.18637/jss.v077.i01.
14. Wood, S.N. (2017) Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.
15. Fasiolo M., Nedellec R., Goude Y. and Wood S. N. (2018) mgcviz: visual tools for Generalized Additive Models. R package version 0.1.1.
16. Ron Pearson (2016). GoodmanKruskal: Association Analysis for Categorical Variables. R package version 0.0.2. <https://CRAN.R-project.org/package=GoodmanKruskal>.