Published in Theory and decision, 2018, vol. 84, no. 2, pp. 277–303, which should be cited to refer to this work.
DOI : 10.1007/s11238-017-9631-y

Modeling purchases of new cars: an analysis of the 2014 French market

Anna Fernández Antolín * Matthieu de Lapparent * Michel Bierlaire *

*Transp-OR, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, {anna.fernandezantolin,matthieu.delapparent,michel.bierlaire}@epfl.ch

Abstract

This paper analyses and compares different policy scenarios as well as discusses price elasticities and willingness to pay and to accept using revealed preference data from the French new-car market in 2014 by means of a cross-nested logit (CNL) model. We focus particularly on electric and hybrid vehicles. We use interactions between the cost (both fixed and running costs) and the household income in order to analyze the sensitivity towards different policy scenarios per income level.

Results show that the willingness to pay and to accept obtained in our study are consistent with the real market conditions. We also find that the most effective scenario in order to increase the market shares of new sold electric vehicles is that of a major technological advance such as a decrease in price due to cheaper manufacturing costs and an increase in driving range, rather than a policy-based scenario. Also, the market segment that has more potential to increase the market shares of electric vehicle purchase is the middle-income level.

In the paper, we discuss how to overcome the difficulties of working with revealed preference data, and propose a new method to impute the attributes of the unchosen alternatives, based on the empirical distributions observed in the data.

Keywords: car-type choice, policy analysis, revealed preference data, cross-nested logit

1 Introduction

The automobile sector is of interest for both the public and the private sectors. Governments and other public actors need to understand the car market in order to have valid forecasts of energy consumption, emission levels and even tax revenue. By means of these forecasts they can also derive optimal policy measures to, for instance, promote the use of electric vehicles to reduce emissions.

It is also interesting for private companies. The interest from automobile firms is obvious, but the car market is linked to many other sectors such as those providing the raw materials (steel, chemicals, textiles) and those working with automobiles such as repair and mobility services. Moreover, according to the European Commission, "the EU is among the world's biggest producers of motor vehicles and the sector represents the largest private investor in research and development (R&D)" [1]

In order to satisfy the needs of these public and private actors it is important to model car ownership, which has many dimensions. Car ownership models can be classified based on several criteria according to de Jong *et al.* (2004) such as: **i)** the inclusion of supply and demand, **ii)** the aggregation level, **iii)** the time representation (dynamic or static), **iv)** the time horizon (long-term or short-term forecasts), **v)** the inclusion of car-use and other socioeconomic characteristics, and **vi)** the type of market (private or business cars) among others. In this paper we focus on the demand side of private cars, in a disaggregate and static framework where we include socioeconomic characteristics of the car buyers. The objective is to have long-term forecasts. This is known as *static disaggregate car-type choice models*.

Our goal is to use revealed preference data to estimate these type of models. This allows to have more realistic demand indicators compared to the ones obtained with stated preference data, such as predicted market shares under several scenarios, willingness to pay and to accept several car attributes and price elasticities. We are particularly interested in the demand for hybrid and electric vehicles. Using RP is more challenging. The main difficulties are to define the choice set and to define the attributes of the unchosen alternatives. The main contribution of the paper is the way how we define the attributes of the unchosen alternatives. We use the empirical distribution of the attributes and draw from them.

The remaining of the paper is structured as follows. Section 2 contains a brief literature review, which is followed by the description of the data used in the paper, and how it is aggregated into different choice alternatives in Section 3. In Section 4 we discuss the adopted methodological approach, the results of which are discussed in Section 5. The application of the model is discussed in Section 6. We finalize the paper with some concluding remarks and future research directions in Section 7.

2 Literature Review

As mentioned in Section 1, we focus on static disaggregate car-type choice models. The first study dealing with this type of models was performed by Lave & Train (1979). For a complete review of the literature on car ownership the reader is refered to de Jong *et al.* (2004) and more recently to Anowar *et al.* (2014).

[1] http://ec.europa.eu/growth/sectors/automotive

Although it is clear that a choice of a private car is a discrete choice, there does not seem to be consensus in the literature about the definition of the choice set. The two main approaches are defined below. The first approach considers that a car is characterized by its make, model, engine and vintage (Birkeland & Jordal-Jorgensen, 2001). Then, for a given year, there may be over 1,000 alternatives. In this case, sampling of alternatives is usually required for the estimation of the model, although recent developments in model estimation (Mai *et al.* , 2015) allow to estimate large scale MEV models.

The second approach prefers an aggregate representation. For example Page *et al.* (2000) characterize a car by its engine size and fuel type. They have nine alternatives for petrol and seven for diesel. It greatly simplifies the specification and estimation of the model. A similar aggregation of alternatives is used by Hess *et al.* (2012). This approach is also justifiable from a behavioral point of view, arguing that decision-makers do not explicitly consider large choice sets.

The most popular model in this context is logit (Wu *et al.* , 1999; Choo & Mokhtarian, 2004). However, the Independence from Irrelevant Alternatives (IIA) property of logit, may lead to counter-intuitive results when alternatives share unobserved attributes. It is likely to happen in car-type choice no matter which of the previous two approaches is chosen. Other models have been considered, such as mixtures of logit models, (Brownstone & Train, 1998; McFadden *et al.* , 2000; Potoglou, 2008), nested logit models (Berkovec & Rust, 1985; McCarthy & Tay, 1998; Mohammadian & Miller, 2002, 2003; Cao *et al.* , 2006) and cross nested logit models (Hess *et al.* , 2012).

The interest in electric and hybrid vehicles has risen in the past years, through the analysis of stated preferences data (Glerum *et al.* , 2014; Hackbarth & Madlener, 2016; Beck *et al.* , 2013, 2016; Daziano, 2013; Hackbarth & Madlener, 2013; Daziano & Achtnicht, 2014; Brownstone & Train, 1998; Train, 1980). Massiani (2014) describes some of the most important limitations of the stated preference surveys being used currently in the literature, and questions the policy recommendations that can be obtained from them. Studies carried out with revealed preference data are generally quite old, and do not focus on electric and alternative fuel vehicles, mainly because of limitation of these vehicles in revealed preference data. Some examples include Berry *et al.* (1995, 1998); Train (1986); Berkovec (1985); Train & Winston (2007). In these studies aggregation of alternatives is used. Berkovec & Rust (1985) instead use sampling of alternatives, where 14 out of 785 alternatives are sampled for estimation. The main difficulty when using revealed preference data is to impute the attributes of the unchosen alternatives, in particular if there is an aggregation of alternatives. The studies cited above deal with it by imputing mean values of the attributes for each unchosen alternative.

We fill the gap in the literature by proposing an alternative way of imputing the attributes of the unchosen alternatives, based on multiple imputation using the empirical distributions of the attributes for each alternative. To the best of our knowledge, it is also the first study of car-type choice to focus on electric and hybrid vehicles in the context of the whole market using revealed preference data. Due to the lack of variability in the autonomy (the range) of electric vehicles that are currently in the market, we need to take the value of the willingness to pay for range from the literature. By doing this we are able forecast the impact of an increase of the range on the market shares of electric vehicles. We validate our model with demand indicators such as market shares, elasticities and willingness to pay and to accept.

3 Data and data aggregation

We use a dataset reporting sales of new cars in France in 2014. Each observation corresponds to the purchase of a new car. The data includes –but is not limited to– disaggregate information about the new car and its purchase, the car's characteristics and information about the main driver and the main driver's household. The dataset reports over 40,000 purchases. However, after selecting the variables that we use in the model and removing the missing values for any of them, 18,804 observations are used for this study. It is left for future work to recover some of these missing values.

In our approach we decide to consider a car-type as a combination between a market segment and a fuel type. The market segments are *full*, *luxury*, *medium*, *multipurpose vehicles (MPV)*, *off-road* and *small*[2]. The fuel types considered are *diesel*, *petrol*, *electric* and *hybrid*. Table 1 gives some examples of cars belonging to each market segment.

Market segment	Car
Full	Ford Taurus Toyota Avalon Hyundai Grandeur
Luxury	Mercedes-Benz S-Class Audi A8 BMW 7 Series
Medium	Opel Astra Honda Civic Audi A3
Multipurpose vehicle	Renault Espace Volkswagen Sharan Mercedes-Benz Vito
Off-road	Land Rover Freelander Chevrolet Captiva BMW X5
Small	Opel Corsa Ford Fiesta Toyota Yaris

Table 1: Examples of cars that belong to each market segment.

Out of the 18,804 observations, only 657 report the purchase of a hybrid vehicle. Moreover, they are always combined with either diesel or petrol. Consequently, we consider *hybrid* as a market segment rather than as a fuel type. Therefore, we have a total of 15 alternatives summarized in

[2]These segments are derived from the European Commission's segmentation

Table 2, together with the number of observations corresponding to each alternative after removing any missing values. Note that we should have 21 alternatives (3 fuel types multiplied by 7 market segments). The six missing alternatives are the combinations of *electric* vehicles with any market segment except *small*. This is because these alternatives do not exist in the French market. The only exception is *electric luxury* vehicles, that do exist (e.g: Tesla), but represent a negligible part of the car market. For this reason, we decide to remove this alternative (*electric luxury*) from the analysis.

Alternative	Market segment	Fuel type	Nbr. of obs.
1	Full	Diesel	323
2	Luxury	Diesel	178
3	Medium	Diesel	2,226
4	MPV	Diesel	1,375
5	Off-road	Diesel	2,044
6	Small	Diesel	5,538
7	Hybrid	Diesel	161
8	Full	Petrol	68
9	Luxury	Petrol	54
10	Medium	Petrol	663
11	MPV	Petrol	310
12	Off-road	Petrol	265
13	Small	Petrol	5,037
14	Hybrid	Petrol	496
15	Small	Electric	66
Total			18,804

Table 2: List of alternatives in the choice set and number of observations (after removing missing values).

From this definition of alternatives, it is obvious that alternatives that share either fuel type or market segment share unobserved attributes. Figure 1 proposes a correlation structure derived from the multi-dimensional nature of the choice set presented in Table 2. This correlation structure is used in the cross nested logit model presented in Section 4.

We assume that our dataset is representative of the population of new car buyers from an exogenous point of view (i.e: from a socioeconomic point of view). It is important to note that it might not be representative of the whole French population, but this is not an issue. For the representativeness of the choices, we are able to replicate the real market shares by applying the correction of the alternative specific constants as described by Train (2009).

4 Methodological approach and model specification

This section contains the methodological approach used to estimate the choice model. We define the choice model used in Section 4.1 followed by the model specification, the nesting structure and the

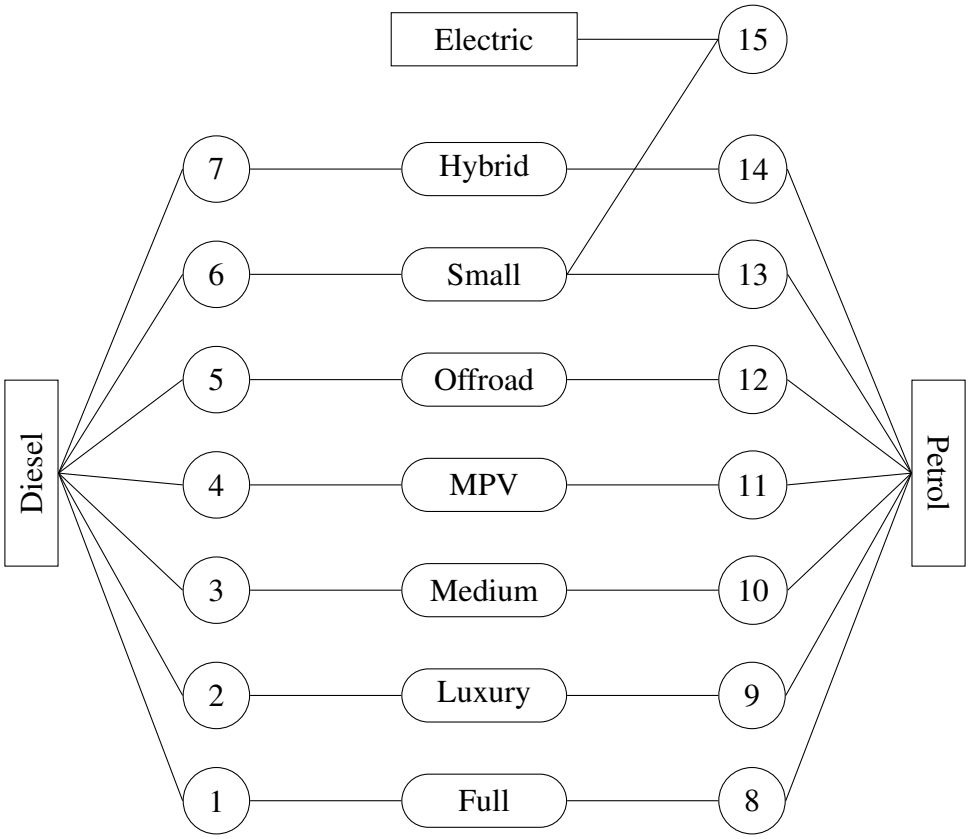


Figure 1: Cross nested structure.

definition of the variables used in Section 4.2. We finalize by describing how we import the parameter associated with range from the literature in Section 4.3.

4.1 Choice model

The choice models used in our study are both a logit model and a cross nested logit. We consider the logit model to be the benchmark in order to show the importance of accounting for the correlation across alternatives. We use the cross nested model in the application of the model.

The cross nested logit model allows to overcome the IIA property. Suppose that the choice set $\mathcal{C}$ is formed by M nests, $\mathcal{C}_1, \mathcal{C}_2, ..., \mathcal{C}_M$. The parameters α_{im} represent the degree of membership of alternative i to nest $\mathcal{C}_m$. For identification purposes it is bounded between 0 and 1 and the $\sum_{m=1}^{M} \alpha_{im} = 1$, $\forall i$. The expression of the choice probabilities for a cross nested logit model are:

$$P(i \mid \mathcal{C}_n) = \sum_{m=1}^{M} \frac{\left(\sum_{j \in \mathcal{C}_n} \alpha_{jm}^{\frac{\mu_m}{\mu}} e^{\mu_m V_{jn}}\right)^{\frac{\mu}{\mu_m}}}{\sum_{p=1}^{M} \left(\sum_{j \in \mathcal{C}_n} \alpha_{jp}^{\frac{\mu_p}{\mu}} e^{\mu_p V_{jn}}\right)^{\frac{\mu}{\mu_p}}} \cdot \frac{\alpha_{im}^{\frac{\mu_m}{\mu}} e^{\mu_m V_{in}}}{\sum_{j \in \mathcal{C}_n} \alpha_{jm}^{\frac{\mu_m}{\mu}} e^{\mu_m V_{jn}}}, \tag{1}$$

where $\mathcal{C}_n$ is the choice set of individual n, μ is the scale parameter of the model, normalized to 1, and μ_m, $m = 1, ..., M$ are the scale parameters of each nest, estimated from the data. For details on the normalization of the μ parameters, and a more detailed analysis of the cross-nested logit model, the reader is referred to Bierlaire (2006) and Abbe *et al.* (2007).

4.2 Definition of variables and model specification

We consider a logit and a cross nested logit model, with a linear in parameter specification for the utility functions. Table 3 shows the variables considered in the model. They are divided in (i) attributes of the recently purchased car and (ii) socioeconomic characteristics of the main driver of the car and/or her household. It is important to note that the car attributes are those reported by the individuals, and not catalog attributes.

	Variable	Definition
Attributes	price	Purchase price after discounts and government schemes [€]
	cons	Fuel consumption [l/100km]
	max_power	Engine power [bhp]
	range_EV	Reported average range achieved from a full charge [km]
Socioeconomic characteristics	agglomeration	1 if main driver lives either in a city or in the suburbs
	town_rural	1 if main driver lives either in a town, village or rural area
	university	1 if education level of main driver is at least a bachelor degree
	nbr. cars	Total number of cars in regular use in the household
	nbr. adults	Number of adults in the household (including main driver)
	nbr. child.	Number of children in the household (aged 18 or less)
	income	10 if annual gross household income ≤ 10 [k€]
		15 if annual gross household income $\in [10, 20)$ [k€]
		25 if annual gross household income $\in [20, 30)$ [k€]
		35 if annual gross household income $\in [30, 40)$ [k€]
		45 if annual gross household income $\in [40, 50)$ [k€]
		55 if annual gross household income $\in [50, 60)$ [k€]
		65 if annual gross household income $\in [60, 75)$ [k€]
		87.5 if annual gross household income $\in [75, 100)$ [k€]
		112.5 if annual gross household income $\in [100, 125)$ [k€]
		150 if annual gross household income $\in [125, 175)$ [k€]
		200 if annual gross household income ≥ 175 [k€]

Table 3: Definition of the variables used in the model.

Since the dataset consists of revealed preference choices, we have no direct access to the attributes of the unchosen alternatives and they have to be imputed. The state-of-the-art is to impute the attribute of an unchosen alternative as the mean of that attribute from the chosen alternatives (Train, 1986;

Berkovec, 1985). In other words, if an individual chose a *small petrol* car, the *max_power* of the *off road petrol* car is usually imputed as the mean *max_power* of the observed *off road petrol* cars. In this paper, instead, we perform multiple imputation (see, for example, Schafer (2000)), by considering the empirical distribution of each attribute for a given alternative. This distribution consists in the observed values of other people's chosen alternatives. Algorithm 1 shows how, where K is the number of multiple imputations, N is the set of respondents, $\mathcal{C}_n$ is the set of alternatives available to individual n (as discussed in Section 3) and Y is the set of cars [3]. We define $t : Y \to \mathcal{C}_n$ as the function that maps each car with its car-type such that $t(y) = i$ if car y belongs to alternative (car-type) i. Note that $t(\cdot)$ is surjective but not injective. This is, each car belongs to a car-type, and two different cars can belong to the same car-type. We estimate the model repeatedly with the different datasets $D_1, ..., D_k$ built as defined by Algorithm 1. We denote by $\widehat{\theta}^k$ the maximum likelihood estimates of the parameters obtained using dataset D_k. Therefore, we obtain a distribution of the model parameters rather than a point estimate.

Data: number of multiple imputations K, set of respondents N, set of alternatives $\mathcal{C}_n$, set of cars Y, vector of attributes for each car x_y
Result: Datasets $D_1, ..., D_k$ containing attributes of chosen and unchosen alternatives
begin
  for $k = 1 : K$ **do**
    for $n \in N$ **do**
      for $i \in \mathcal{C}_n$ **do**
        if *individual* n *chose alternative* i **then**
          attributes of alternative $i \leftarrow$ attributes of chosen car
        else
          select randomly (with equal probability) a car y such that $t(y) = i$
          attributes of alternative $i \leftarrow$ attributes of car y
        end
      end
    end
    $D_k \leftarrow$ attributes of chosen and unchosen alternatives
  end
end

Algorithm 1: Attributes of all alternatives.

Tables 4 and 5 show the model specification. Note that both price and fuel consumption are interacted with income. The fuel consumption is also multiplied by the mean fuel price (diesel or petrol), calculated for 2014 in France (Institut national de la statistique et des études économiques, 2016c). Petrol price is denoted as *pp* and diesel price is denoted as *pd* in the table. The rest of the variables appear linearly in the model. Note that some of them are rescaled for numerical reasons. Note also that the specification of the utility functions is the same for the logit, and the cross nested logit models.

[3] Here, a *car* is defined by a combination of make-model-type. The alternatives are car-types, as defined in Table 2.

Parameter	1	2	3	4	5	6	7	8
ASC_i	1 in the utility function of alternative $i \in (2, 15)$, 0 in the rest							
$\beta_{price_inc_i}$	$\frac{price_i \cdot 100}{income}$ in the utility function of alternative i, 0 in the rest							
β_{inc_full}	$\frac{income}{10000}$	0	0	0	0	0	0	$\frac{income}{10000}$
β_{inc_luxury}	0	$\frac{income}{10000}$	0	0	0	0	0	0
β_{inc_medium}	0	0	$\frac{income}{10000}$	0	0	0	0	0
β_{inc_MPV}	0	0	0	$\frac{income}{10000}$	0	0	0	0
$\beta_{inc_offroad}$	0	0	0	0	$\frac{income}{10000}$	0	0	0
β_{inc_hybrid}	0	0	0	0	0	0	$\frac{income}{10000}$	0
$\beta_{nbr_adults_small}$	0	0	0	0	0	nbr. adults	0	0
$\beta_{nbr_chilren_small}$	0	0	0	0	0	nbr. child.	0	0
$\beta_{nbr_cars_lux}$	0	nbr. cars	0	0	0	0	0	0
$\beta_{nbr_cars_hybrid}$	0	0	0	0	0	0	nbr. cars	0
$\beta_{university}$	0	0	0	0	0	0	university	0
$\beta_{town_rural_EV}$	0	0	0	0	0	0	0	0
$\beta_{town_rural_hybrid}$	0	0	0	0	0	0	town_rural	0
β_{conso_inc}	$\frac{cons_1 \cdot pd \cdot 100}{income}$	$\frac{cons_2 \cdot pd \cdot 100}{income}$	$\frac{cons_3 \cdot pd \cdot 100}{income}$	$\frac{cons_4 \cdot pd \cdot 100}{income}$	$\frac{cons_5 \cdot pd \cdot 100}{income}$	$\frac{cons_6 \cdot pd \cdot 100}{income}$	$\frac{cons_7 \cdot pd \cdot 100}{income}$	$\frac{cons_8 \cdot pp \cdot 100}{income}$
β_{max_power}	$\frac{max_power_1}{10}$	$\frac{max_power_2}{10}$	$\frac{max_power_3}{10}$	$\frac{max_power_4}{10}$	$\frac{max_power_5}{10}$	$\frac{max_power_6}{10}$	$\frac{max_power_7}{10}$	$\frac{max_power_8}{10}$
β_{range_EV}	0	0	0	0	0	0	0	0

Table 4: Model specification (part 1/2).

Parameter	9	10	11	12	13	14	15
ASC_i	1 in the utility function of alternative $i \in (2, 15)$, 0 in the rest						
$\beta_{price_inc_i}$	$\frac{price_i \cdot 100}{income}$ in the utility function of alternative i, 0 in the rest						
β_{inc_full}	0	0	0	0	0	0	0
β_{inc_luxury}	$\frac{income}{10000}$	0	0	0	0	0	0
β_{inc_medium}	0	$\frac{income}{10000}$	0	0	0	0	0
β_{inc_MPV}	0	0	$\frac{income}{10000}$	0	0	0	0
$\beta_{inc_offroad}$	0	0	0	$\frac{income}{10000}$	0	0	0
β_{inc_hybrid}	0	0	0	0	0	$\frac{income}{10000}$	0
$\beta_{nbr_adults_small}$	0	0	0	0	nbr. adults	0	0
$\beta_{nbr_chilren_small}$	0	0	0	0	nbr. child.	0	0
$\beta_{nbr_cars_lux}$	nbr. cars	0	0	0	0	0	0
$\beta_{nbr_cars_hybrid}$	0	0	0	0	0	nbr. cars	0
$\beta_{university}$	0	0	0	0	0	university	university
$\beta_{town_rural_EV}$	0	0	0	0	0	0	town_rural
$\beta_{town_rural_hybrid}$	0	0	0	0	0	town_rural	0
β_{conso_inc}	$\frac{cons_9 \cdot pp \cdot 100}{income}$	$\frac{cons_{10} \cdot pp \cdot 100}{income}$	$\frac{cons_{11} \cdot pp \cdot 100}{income}$	$\frac{cons_{12} \cdot pp \cdot 100}{income}$	$\frac{cons_{13} \cdot pp \cdot 100}{income}$	$\frac{cons_{14} \cdot pp \cdot 100}{income}$	0
β_{max_power}	$\frac{max_power_9}{10}$	$\frac{max_power_{10}}{10}$	$\frac{max_power_{11}}{10}$	$\frac{max_power_{12}}{10}$	$\frac{max_power_{13}}{10}$	$\frac{max_power_{14}}{10}$	$\frac{max_power_{15}}{10}$
β_{range_EV}	0	0	0	0	0	0	$\frac{range_EV}{100}$

Table 5: Model specification (part 2/2).

Nesting structure The nesting structure is defined as in Figure 1, where the numbered circles represent the 15 alternatives, the oval shape boxes represent the nest related to the market segment, and the rectangle boxes represent the nests related to the fuel type. We define one membership parameter, α_{MS}, that defines the membership to the *market segment* nests (the ones surrounded by a rounded rectangle). Then, $1 - \alpha_{MS}$ gives the membership to the *fuel type* nests. More general specifications were tested, but the resulting models were not identified. This is a strong assumption, and more investigation is left for future research. We define also five scale parameters, μ_k, $k \in \{\text{medium, offroad, small, diesel, petrol}\}$. μ_{electric} has to be normalized to one, because it only contains one alternative. μ_ℓ, $\ell \in \{\text{hybrid, MPV, luxury, full}\}$ are normalized to one because they reach the lower bound when we try to estimate them.

4.3 Parameter associated with range for electric vehicles

Due to the lack of variability in the range of electric vehicles in the data, the parameter $\beta_{\text{range_EV}}$ can not be estimated with sufficient precision. Since the willingness to pay for range is well studied in the literature based on stated preference data, and is known to be one of the determinants of electrical vehicle purchase, we import it from the literature and use it in our model. Dimitropoulos *et al.* (2013) perform a meta-analysis based on 129 willingness to pay estimates and find that consumers are willing to pay between 66 and 75US$ on average for a 1-mile increase in range, which is equivalent to between 30.8 and 35.0€/km [4]. For the results shown in Sections 5 and 6, we consider the value 34€/ km. We note $\text{WTP}(\text{range}_{\text{lit}}) = 34$.

From the definition of willingness to pay for range (since alternative 15 is the one related to electric vehicles) we obtain:

$$\text{WTP}(\text{range}_{15,n}) = -\frac{\frac{\partial V_{15,n}}{\partial \text{range}_{15,n}}}{\frac{\partial V_{15,n}}{\partial \text{price}_{in}}} = -\frac{\beta_{\text{range_EV}} \cdot \text{income}_n}{100 \cdot \beta_{\text{price_inc_15}}}, \tag{2}$$

and by equalizing $\text{WTP}(\text{range}_{\text{lit}}) = \text{WTP}(\text{range}_{15,n})$:

$$\beta_{\text{range_EV}} = -34 \cdot \frac{100 \cdot \beta_{\text{price_inc_15}}}{\text{income}_n}. \tag{3}$$

We define $\beta_{\text{range_EV}}$ as defined in Equation (3) and estimate all the parameters simultaneously.

5 Results

The estimation results for both the logit and the cross nested logit models are reported in Table 6. The reported parameters are the means of the parameters obtained with the 50 realizations of the multiple imputation method. The t-tests are computed from the standard errors, and the standard errors of each

[4] For the change in units, we consider the mean exchange rate between US$ and € for 2014 which is 1.33$/€ according to the European Central Bank.

parameter s_p are obtained from the empirical distribution of the estimates as follows:

$$s_p = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}(\widehat{\theta}_p^k - \bar{\widehat{\theta}}_p)^2}, \tag{4}$$

where K is the number of multiple imputations, $\widehat{\theta}_p^k$ is the value of the parameter at imputation k, and $\bar{\widehat{\theta}}_p = \frac{1}{K}\sum_{k=1}^{K}\widehat{\theta}_p^k$.

	Logit		CNL	
	mean param.	t-test [5]	mean param.	t-test [5]
ASC_2	-1.90	-57.8	-1.86	-38.3
ASC_3	2.44	119	2.20	18.3
ASC_4	1.87	79.6	1.78	21.3
ASC_5	1.94	85.9	1.88	18.1
ASC_6	3.64	136	2.99	23.4
ASC_7	0.207	8.09	0.295	8.83
ASC_8	-2.38	-112	-2.26	-54.1
ASC_9	-3.10	-79.3	-3.07	-37.4
ASC_{10}	1.20	61.1	1.43	13.7
ASC_{11}	0.465	21.1	0.417	11.3
ASC_{12}	-0.283	-13.4	0.389	2.20
ASC_{13}	3.44	127	2.81	19.0
ASC_{14}	1.23	45.8	1.14	21.8
ASC_{15}	0.0968	3.62	-0.0533	-0.225
β_{inc_full}	0.143	79.4	0.120	24.6
β_{inc_luxury}	0.204	107	0.179	33.6
β_{inc_medium}	0.0155	10.1	0.00875	4.55
β_{inc_MPV}	0.0605	41.9	0.0375	16.9
$\beta_{inc_offroad}$	0.117	73.0	0.0693	17.4
β_{inc_hybrid}	0.0939	53.3	0.0614	16.4
$\beta_{nbr_adults_small}$	-0.0911	-39.2	-0.0687	-10.2
$\beta_{nbr_chilren_small}$	-0.235	-196	-0.190	-14.6
$\beta_{nbr_cars_lux}$	0.291	85.8	0.295	20.9
$\beta_{nbr_cars_hybrid}$	-0.260	-131	-0.240	-43.2
$\beta_{university}$	0.180	52.8	0.175	34.6
$\beta_{town_rural_EV}$	0.556	130	0.546	44.5
$\beta_{town_rural_hybrid}$	-0.270	-70.2	-0.234	-20.4

[5]The reported t-tests are against zero for all parameters except for the μ parameters. For the μ parameters, the reported t-tests are against one.

$\beta_{price_inc_1}$	-0.128	-67.9	-0.119	-56.7
$\beta_{price_inc_2}$	-0.102	-56.4	-0.0950	-37.2
$\beta_{price_inc_3}$	-0.109	-54.4	-0.0805	-15.9
$\beta_{price_inc_4}$	-0.134	-65.0	-0.118	-24.4
$\beta_{price_inc_5}$	-0.116	-62.8	-0.0959	-21.7
$\beta_{price_inc_6}$	-0.107	-38.2	-0.0778	-18.4
$\beta_{price_inc_7}$	-0.169	-96.7	-0.168	-63.9
$\beta_{price_inc_8}$	-0.0717	-31.2	-0.0690	-31.6
$\beta_{price_inc_9}$	-0.136	-52.0	-0.122	-51.0
$\beta_{price_inc_10}$	-0.112	-39.4	-0.0819	-14.6
$\beta_{price_inc_11}$	-0.140	-62.8	-0.130	-63.2
$\beta_{price_inc_12}$	-0.0995	-37.6	-0.0981	-21.3
$\beta_{price_inc_13}$	-0.0948	-26.4	-0.0672	-15.2
$\beta_{price_inc_14}$	-0.146	-58.2	-0.131	-37.9
$\beta_{price_inc_15}$	-0.507	-75.8	-0.428	-67.6
β_{conso_inc}	-0.105	-8.74	-0.0718	-7.30
β_{max_power}	0.0565	43.1	0.0456	20.3
μ_{medium}	-	-	1.67	3.49
$\mu_{offroad}$	-	-	1.39	2.37
μ_{small}	-	-	1.91	2.38
μ_{diesel}	-	-	6.34	1.83
μ_{petrol}	-	-	3.25	5.68
α_{MS}	-	-	0.589	11.9

Table 6: Mean of the parameter estimates. Number of multiple imputations: K=50.

Unless pointed out, the following interpretations are valid for both the logit and the cross nested logit models.

Income The interactions between the income level and the market segment have the expected relative magnitudes. The normalized market segment is *small*. Therefore, the interpretation of the results is that people with larger income levels have a larger preference towards *luxury* vehicles, then *full*, followed by *offroad* and *hybrid*, with almost the same magnitude. The less preferred alternatives, all else being equal, for people with larger income are *MPV*, then *medium* and the less preferred is the reference level *small*.

Other socioeconomic characteristics We also model the effects of number of children, of adults and of vehicles in a household, the education level and the residence location.

From the negative values of $\beta_{nbr_adults_small}$ and $\beta_{nbr_children_small}$ we can conclude that

the more people live in a household (either adults or children), less likely it is no have a *small* vehicle compared to households with less people. Moreover, the number of children has a stronger effect in the decrease of the probability of buying a *small* vehicle than the number of adults.

From the estimation results we can also conclude that the larger the number of cars in a household, the more likely it is to buy a luxury car (since $\beta_{\texttt{nbr_cars_lux}} > 0$). Similarly, the larger the number of cars in a household, the less likely it is a hybrid one ($\beta_{\texttt{nbr_cars_hybrid}} < 0$). From the positive value of $\beta_{\text{university}}$ we can conclude that individuals who go to university are more likely to buy hybrid and pure electric vehicles compared to people that do not go to university. For the residence location, we find a surprising result: individuals living in towns or rural areas are more likely to buy an EV than those living in a city or in the suburbs ($\beta_{\text{town_rural_EV}} > 0$). For hybrid cars it is however the opposite: individuals living in cities and suburbs are more likely to buy one than those living in towns or rural areas ($\beta_{\text{town_rural_hybrid}} < 0$).

Price interacted with income Both pairwise t-test comparisons between the parameter estimates, and a likelihood ratio test reject the hypothesis of generic price parameters. All the price parameters are negative, as expected, and individuals are more sensitive to high prices for electric vehicles (alternative 15) than to any other alternative (since $\beta_{\text{price_inc_15}}$ is the largest parameter in absolute value).

Other attributes of the alternatives The fuel consumption is multiplied by 1.48 €/ℓ, that is the mean petrol price in France in 2014 for the petrol alternatives, and by 1.29 €/ℓ, that is the mean diesel price in France in 2014 for the diesel alternatives (Institut national de la statistique et des études économiques, 2016c). The variable is therefore a proxy to the running costs. We interact it with the household income analogously as we do for price (or the fixed cost). As expected $\beta_{\text{conso_inc}}$ is negative, meaning that all else being equal, individuals prefer cars with less fuel consumption. We also model the engine power, that has a positive effect. All else equal, individuals prefer vehicles with more power, as expected.

Nest and membership parameters The five reported nest parameters are significantly different from one, meaning that the alternatives that belong to them share unobserved attributes. However, the other five nest parameters are to be fixed to one. μ_{electric} is fixed to one because it only contains one alternative, so it can not be identified. The other four (μ_{full}, μ_{luxury}, μ_{MPV}, μ_{hybrid}) are fixed to one because when we try to estimate them they reach the lower boundary 1.

The membership parameter α_{MS} is between 0 and 1 and it represents how much (out of one) an alternative is explained by the market segment. The fact that it is larger than 0.5 means that an alternative *belongs more* to its corresponding market segment than to its corresponding fuel type. Note that we consider the same membership parameter to market segment and to fuel type for all the alternatives.

6 Application of the model

In this section we apply the model described above in order to obtain demand indicators such as price elasticities (Section 6.1), aggregate market shares under different policy scenarios (Section 6.2) and willingness to pay and to accept (Section 6.3). We consider only the cross-nested logit.

For the application of the model, instead of doing multiple imputation (as we do in the estimation process) we impute each attribute of an unchosen alternative by the mean value of each attribute for the chosen alternatives [6]. In other words, for an individual n that chose alternative i, the values of an attribute of the unchosen alternative j, x_{jn} are imputed as the average of attribute x of those individuals who chose alternative j, $\bar{x}_j$. Moreover, in order to replicate the correct market shares in the base case, we need to calibrate the alternative specific constants as described by Train (2009, p. 67).

6.1 Price elasticities

Let p_{in} be the current value of the price variable, and $p^+_{jn} = p_{jn} + \Delta p_{jn}$ the future value. Keeping all other variables at their current values, we denote $P_n(i)$ the choice probability of alternative i and $P^+_n(i) = P_n(i) + \Delta P_n(i)$ the choice probability involving p^+_{jn}. The disaggregate arc elasticity for individual n is defined as follows:

$$E^{\Delta P_n(i)}_{\Delta p_{jn}} = \frac{\Delta P_n(i)}{P_n(i)} \cdot \frac{p_{jn}}{\Delta p_{jn}}, \quad \forall i = 1, ..., 15, \forall j = 1, ..., 15, \tag{5}$$

If $i = j$ in Equation (5) then it is called the direct arc elasticity, and otherwise the cross arc elasticity. In our application, the alternative scenario is a decrease of 20% of alternative j, $\Delta p_{jn} = -0.2 \cdot p_{jn}$. Results for each pair $(i, j), i = 1, \ldots, 15$, $j = 1, \ldots, 15$ are shown in Table 7.

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-0.998	0.0109	0.0109	0.0109	0.0109	0.0109	0.0109	0.0109	0.0109	0.0109	0.0109	0.0109	0.0109	0.0109	0.0109
2	0.00412	-1.09	0.00412	0.00414	0.00414	0.00413	0.00412	0.00412	0.00412	0.00412	0.00412	0.00412	0.00412	0.00412	0.00412
3	0.0198	0.0198	-0.413	0.0198	0.0198	0.0199	0.0198	0.0198	0.0198	0.305	0.0198	0.0198	0.0198	0.0198	0.0198
4	0.102	0.102	0.102	-0.704	0.109	0.190	0.102	0.102	0.102	0.102	0.102	0.102	0.102	0.102	0.102
5	0.0540	0.0540	0.0540	0.0603	-0.686	0.0601	0.0540	0.0540	0.0540	0.0540	0.0540	0.364	0.0540	0.0540	0.0540
6	0.151	0.151	0.151	0.195	0.155	-0.309	0.151	0.151	0.151	0.151	0.151	0.151	0.252	0.151	0.155
7	0.00651	0.00651	0.00651	0.00651	0.00651	0.00651	-1.09	0.00651	0.00651	0.00651	0.00651	0.00651	0.00651	0.00651	0.00651
8	0.000747	0.000747	0.000747	0.000747	0.000747	0.000747	0.000747	-0.661	0.000747	0.000747	0.000747	0.000747	0.000747	0.000747	0.000747
9	0.00160	0.00160	0.00160	0.00160	0.00160	0.00160	0.00160	0.00160	-1.770345	0.00160	0.00165	0.00165	0.00184	0.00163	0.00160
10	0.00152	0.00152	0.0239	0.00152	0.00152	0.00152	0.00152	0.00152	0.00152	-0.534	0.00152	0.00152	0.00154	0.00152	0.00152
11	0.0199	0.0199	0.0199	0.0199	0.0199	0.0199	0.0199	0.0199	0.0199	0.0199	-0.765	0.0199	0.0217	0.0199	0.0199
12	0.00388	0.00388	0.00388	0.00388	0.0294	0.00388	0.00388	0.00388	0.00389	0.00388	0.00391	-0.785	0.00411	0.00390	0.00388
13	0.0863	0.0863	0.0863	0.0863	0.0863	0.143	0.0863	0.0863	0.0864	0.0872	0.0891	0.0877	-0.214	0.0879	0.0881
14	0.0140	0.0140	0.0140	0.0140	0.0140	0.0140	0.0140	0.0140	0.0140	0.0140	0.0141	0.0140	0.0149	-0.748	0.0140
15	0.00407	0.00407	0.00407	0.00407	0.00407	0.00426	0.00407	0.00407	0.00407	0.00407	0.00407	0.00407	0.00424	0.00407	-1.29

Table 7: Direct and cross arc elasticities for each pair of alternatives.

The diagonal values are negative, as should be, and the off-diagonal values are positive. Therefore, decreasing the price of an alternative i increases the probability of choosing alternative i and decreases the probabilities of choosing all other alternatives. Moreover, by means of the cross nested logit, we

[6] We also try multiple imputation, but the results do not change significantly, and considering it in this way saves time in the analysis.

get more realistic substitution patterns, compared to what we could obtain using a logit model. The ranges of the direct elasticities found are in line with what is reported by Train & Winston (2007) (between -1.7 and -3.2 depending on the model they use). Berry *et al.* (1998) report direct elasticities that are a lot larger in absolute value, going up to -126 for some vehicles. However, the cross elasticities reported by Berry *et al.* (1998) are close to what we find. For example, the cross elasticity between the Mazda 323 (that belongs to the *medium* market segment) and the Nissan Maxima (that belongs to the market segment *full*) is reported to be 0.056. We obtain a value of 0.0198.

As an illustration of the substitution patterns obtained thanks to the CNL specification, we analyze the elasticities related to alternative 3, the *medium diesel*. Figure 2 shows the values of the price arc elasticities obtained for alternative *medium diesel*, namely $E^{\Delta P_n(3)}_{\Delta p_{jn}}$, $j = 1, \ldots, 15$. We see that when the price of *medium diesel* is decreased, the largest arc cross elasticity is for alternative 10 (*medium petrol*). In other words, by making *medium diesel* cars cheaper, we attract medium-petrol buyers more than any of the other vehicle car-types. Due to the IIA property, this analysis could not be done with a logit model.

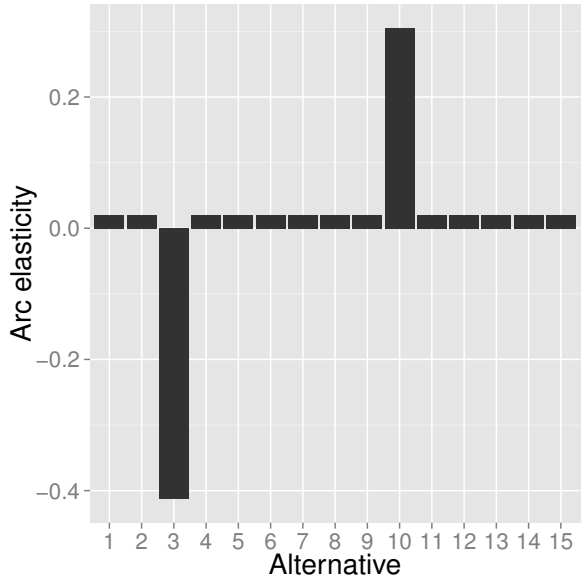


Figure 2: Price arc cross elasticities for medium diesel ($E^{\Delta P_n(3)}_{\Delta p_{jn}}$).

6.2 Comparing different future scenarios

For the forecasting exercise we consider three scenarios. The first one, denoted by *do nothing* scenario, corresponds to a foreseeable future where no specific policy is implemented. The second one, denoted the *tax* scenario, uses the same assumptions as the *do nothing* scenario, plus an increase of the registration tax for internal combustion vehicles of 10% and an increase in the fuel price. Finally, the *technological innovation* scenario uses the same assumptions as the *do nothing* scenario, plus a decrease in the price of electric vehicles of 15% and an increase of the range of these vehicles by 100%.

They are all considered to be related to a five-years horizon. The socioeconomic characteristics of new car buyers are assumed to remain unchanged.

The mean value of fuel consumption decreased from 6.95l/100km in 2010 to 6.49l/100km in 2015 (Comité des Constructeurs Français d'Automobiles, 2016). This represents a 7% decrease. We assume that the decrease in a five-years time horizon will be the same. For the price, in the *do nothing* scenario, motivated by the decrease of the bonus-malus in France from 2015 to 2016 from 4000€ to 1000€ for rechargeable hybrids and from 2000€ to 750€ for other hybrids (Ministère de l'environnement, de l'énergie et de la mer, 2016), we assume an increase of 2500€ of the price of all hybrid vehicles. Moreover, for the *tax* scenario we assume that an increase in the registration tax will render internal combustion vehicles 10% more expensive. For the *technological innovation* scenario we assume that an improvement in the manufacturing process will render electric vehicles 15% cheaper. For the electric vehicle's range in both the *do nothing* and the *tax* scenarios, we assume that within five years the range will increase of 50km for all vehicles. This is in line with the ranges for the Nissan leaf comparison between 2011 (117km) and 2016 (172km) (U.S. Department of Energy, 2016). For the *technological innovation* scenario we assume that the ranges for all electric vehicles on the market are doubled. Finally, for the fuel price, we assume that the petrol and diesel prices will be the same, as the French government has reported that they would like the difference between both prices to decrease (Sud Ouest, 2015). We assume that the taxes are constant in the *do nothing* and *technological innovation* scenarios, and use the forecast for the price of imported fuel (European Comission, 2011), resulting in 2.44€/l. For the *tax* scenario we use the same import price, and increase the taxes by 50% which results in 3€/l. These assumptions are summarized in Table 8.

	Do nothing scenario	**Tax scenario**	**Technological innovation scenario**
Max. power	-	-	-
Fuel cons.	7% decrease in fuel consumption (alt 1-14)	7% decrease in fuel consumption (alt- 1-14)	7% decrease in fuel consumption (alt. 1-14)
Price	- Hybrid vehicles (alt 7 and 14) 2500 € more expensive	- Hybrid vehicles (alt. 7 and 14) 2500 € more expensive - Internal combustion engine vehicles (all alt. except 7,14,15) 10% more expensive	- Hybrid vehicles (alt. 7 and 14) 2500 € more expensive - Electric vehicles (alt. 15) 15% cheaper
Range	+50km	+50km	100% increase
Fuel price	diesel=petrol=2.44€/ℓ	diesel=petrol=3€/ℓ	diesel=petrol=2.44€/ℓ

Table 8: Description of the different tested scenarios.

We compute the market shares of each alternative for each scenario. They are reported in Table 9. In order to interpret these results, we focus on the electric vehicle alternative and plot the market shares per income level and scenario. This is shown in Figure 3. There are 11 income levels as shown in Table 3 and they are labeled from *Income 1* for the lowest income level, to *Income 11* for the highest. Indeed, all the scenarios have an increase in the share of new sold electric vehicles, and the

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Do nothing	1.22	0.546	4.13	14.5	7.74	35.6	0.552	0.126	0.140	0.366	2.48	0.618	30.0	1.75	0.303
Tax	1.16	0.517	4.11	13.9	7.41	36.2	0.574	0.120	0.125	0.363	2.34	0.593	30.5	1.83	0.333
Techno. innov.	1.24	0.552	4.13	14.5	7.87	34.8	0.551	0.132	0.144	0.379	2.54	0.645	30.5	1.74	0.362
Base	1.24	0.554	4.17	14.6	7.93	35.3	0.598	0.130	0.143	0.366	2.50	0.627	29.8	1.84	0.274

Table 9: Predicted market shares for each alternative and scenario in percentages.

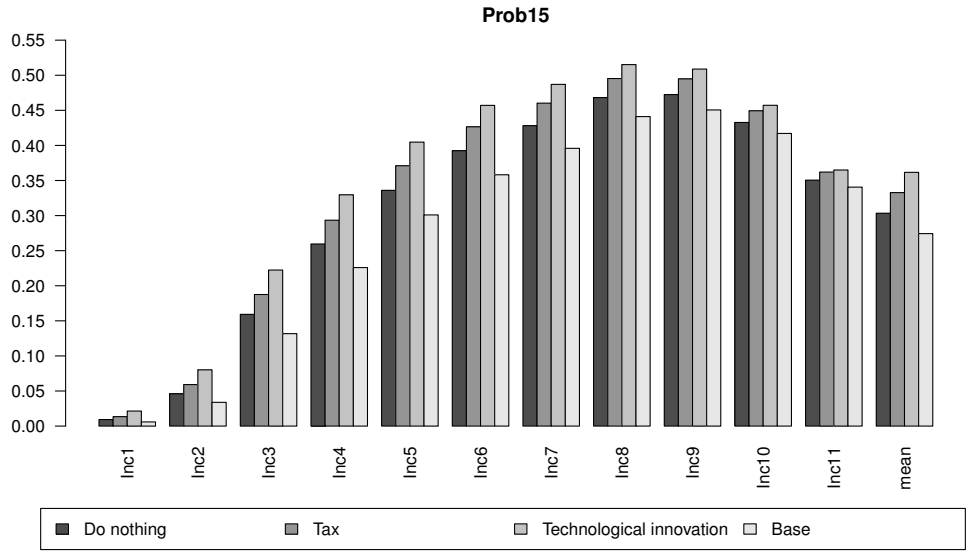
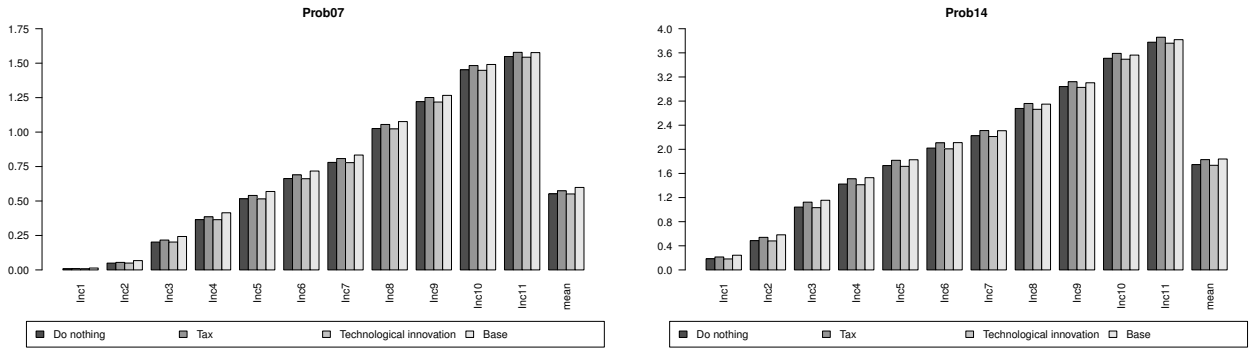


Figure 3: Market shares for the electric vehicle alternative for the base case and each of the scenarios, per income level

most effective scenario is the one with a major technological advance. It is also very interesting to note how the increase in market share is higher for medium income levels rather than low or high income levels. In other words, people with lower income levels can still not afford the electric vehicles, and people with higher income levels are less attracted by the improvements.

We repeat the analysis for hybrid vehicles. Figure 4(a) shows the hybrid diesel and Figure 4(b) shows the hybrid petrol. In both cases, the share is a growing function of the income. However, for the *hybrid diesel* alternative, the market shares in the defined scenarios actually decrease for all income levels. This indicates that without the subsidies given today, the sales of hybrid diesel cars would decrease. For the petrol case, with the *tax* scenario, the market shares increase slightly for all income levels, but they decrease in both the *do nothing* and the *technological innovation* scenarios. We can conclude that unless internal combustion engine (ICE) cars are made more expensive (like in the *tax* scenario), the combination of increasing the fuel price and decreasing the subsidies for hybrid vehicles does not allow the new sales market shares of these alternatives to increase.

It is important to note that these predicted market can only be calculated using revealed preference data for the sample enumeration. Indeed, the values of the attributes in stated preferences data are



(a) Hybrid diesel

(b) Hybrid petrol

Figure 4: Market shares for hybrid vehicles for the base case and each of the scenarios, per income level.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean
Fuel cons. [€/ (ℓ/1000km)]	7.76	9.74	11.50	7.83	9.65	11.90	5.51	15.39	8.73	12.97	8.17	10.83	15.82	8.14	-	10.28
Max power [€/ bhp]	24.41	30.66	36.18	24.64	30.38	37.46	17.35	42.21	23.96	35.57	22.40	29.71	43.39	22.32	6.80	28.50

Table 10: Willingness to accept for fuel consumption and willingness to pay for maximum power for each alternative.

engineered by the experimental design, and do not represent any market reality. Note also that the shares refer to the reference population, that is the set of people buying a new car during a given year.

6.3 Willingness to pay

We compute the marginal willingness to pay for an increase of maximum power and the willingness to accept an increase in fuel consumption, which have the following expressions

$$\begin{aligned} \text{WTP}(\text{max_power}_{\text{in}}) &= -\frac{\frac{\partial V_{\text{in}}}{\partial \text{max_power}_{\text{in}}}}{\frac{\partial V_{\text{in}}}{\partial \text{price}_{\text{in}}}} = -\frac{\beta_{\text{max_power}} \cdot \text{income}}{1000 \cdot \beta_{\text{price_inc_i}}}, \\ \text{WTA}(\text{cons}_{\text{in}}) &= \frac{\frac{\partial V_{\text{in}}}{\partial \text{cons}_{\text{in}}}}{\frac{\partial V_{\text{in}}}{\partial \text{price}_{\text{in}}}} = \frac{10 \cdot \text{price(fuel)} \cdot \beta_{\text{conso_inc}}}{\beta_{\text{price_inc_i}}}. \end{aligned} \tag{6}$$

Results are summarized per alternative in Table 10. By considering an individual who drives 13000km (which is the mean mileage in France in 2014 for private vehicles (Institut national de la statistique et des études économiques, 2016a)) and who keeps the car for 5.4 years (Institut national de la statistique et des études économiques, 2016b), she would be willing to pay 735€ more for a car that is 1 $\frac{\ell}{1000\text{km}}$ more efficient. For the maximum power our results show that an individual is willing to pay 2580€ more for a car that has 100bhp more of maximum power, if all else is equal. Both results are in line with what is observed in the new car market. This indicates that the new car market is close to equilibrium.

7 Conclusions and future work

In this paper we have developed a CNL model for car type choice for new buyers, using revealed preference data. A multiple imputation method has been applied for the attributes of the non chosen alternatives. We have used the estimated model to study the effects of several policy scenarios in the market shares of different car-types with a special focus on electric vehicles. The results are in line with the market share variations obtained in other models using revealed preference data (but not focusing on electric vehicles). We also computed price elasticities, that are in line with values found in the literature, and willingness to pay and to accept values, that are in line with what is observed in the new vehicle market.

By using revealed preference data we have encountered two major difficulties. The first relates to the definition of the choice set. We aggregated several make-model-type of cars into fifteen alternatives defined as a combination of a market segment (full, luxury, medium, MPV, offroad and small) and a fuel type (diesel, petrol, hybrid, electric). This definition of the alternatives makes it natural to estimate a cross nested logit model, since alternatives that share either market segment or fuel type share unobserved attributes. The results confirm that the cross nested logit model is better than the logit model both in terms of fit and in terms of realistic behavioral results.

The second major difficulty of using RP data related to the definition of the attributes of the unchosen alternatives. To the best of our knowledge, it is the first time that the empirical distributions of the attributes of the alternatives are used to impute the attributes of the unchosen alternatives. This is computationally slow, but much more realistic than considering the mean of the attributes of the chosen alternatives, as was done in the literature in the past.

This methodology, however, is not free of limitations. As for any choice experiment, we are not able to estimate a parameter if the attribute related to it has very little variability. In our data, this is the case for the range of electric vehicles. EVs represent a very small part of the 2014 car market in France, and therefore, the reported ranges do not allow to estimate the sensitivity to the autonomy of electric vehicles. In order to overcome this problem, we use the willingness to pay for range reported in the literature.

Out of the three tested scenarios, the most effective in order to increase the use of electric vehicles is the *technological innovation* scenario. The results we obtain are also divided per income levels. The income level that would increase the market share of EV the most is income level number 4, which corresponds to an annual gross household income of between 30,000 and 40,000 €.

Some improvements in the presented research would include weighting the alternatives by the amount of different number of vehicles that they contain, as done in Train (1986). Moreover, the variables related to the attributes of the alternatives could contain measurement errors, since they consist of reported values (instead of catalog values), which might cause endogeneity. Auxiliary models for the car attributes can be estimated and integrated with the choice model to solve this issue. Our framework allows for these auxiliary models to be included easily. Moreover, these auxiliary models would also allow to recover observations containing missing attributes. Finally, a future research direction would be to take into account the price endogeneity as in Berry *et al.* (1995). New results by Mai *et al.* (2015) show that it might be feasible to estimate the model over the full set of alternatives, without the need for either aggregation or sampling of alternatives. It would also be interesting to

estimate the same model over the full set of alternatives and compare the results obtained with the two approaches.

Acknowledgments

This study is financed by a research agreement with Nissan International SA, which is gratefully acknowledged.

8 Bibliography

Abbe, E., Bierlaire, M., , & Toledo, T. 2007. Normalization and correlation of cross-nested logit models. *Transportation Research Part B: Methodological*, **41**(7), 795–808.

Anowar, Sabreena, Eluru, Naveen, & Miranda-Moreno, Luis F. 2014. Alternative Modeling Approaches Used for Examining Automobile Ownership: A Comprehensive Review. *Transport Reviews*, **34**(4), 441–473.

Beck, Matthew J., Rose, John M., & Hensher, David A. 2013. Environmental attitudes and emissions charging: An example of policy implications for vehicle choice. *Transportation Research Part A: Policy and Practice*, **50**(Apr.), 171–182.

Beck, Matthew J., Rose, John M., & Greaves, Stephen P. 2016. I can't believe your attitude: a joint estimation of best worst attitudes and electric vehicle choice. *Transportation*, Jan., 1–20.

Berkovec, James. 1985. Forecasting automobile demand using disaggregate choice models. *Transportation Research Part B: Methodological*, **19**(4), 315–329.

Berkovec, James, & Rust, John. 1985. A nested logit model of automobile holdings for one vehicle households. *Transportation Research Part B: Methodological*, **19**(4), 275–285.

Berry, Steven, Levinsohn, James, & Pakes, Ariel. 1995. Automobile Prices in Market Equilibrium. *Econometrica*, **63**(4), 841.

Berry, Steven, Levinsohn, James, & Pakes, Ariel. 1998 (Mar.). *Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market*. Tech. rept. w6481. National Bureau of Economic Research, Cambridge, MA.

Bierlaire, M. 2006. A theoretical analysis of the cross-nested logit model. *Annals of Operations Research*, **144**(1), 287–300.

Birkeland, M. E., & Jordal-Jorgensen, J. 2001. Energy efficiency of passenger cars.

Brownstone, David, & Train, Kenneth. 1998. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, **89**(1–2), 109–129.

Cao, Xinyu, Mokhtarian, Patricia L., & Handy, Susan L. 2006. Neighborhood design and vehicle type choice: Evidence from Northern California. *Transportation Research Part D: Transport and Environment*, **11**(2), 133–145.

Choo, Sangho, & Mokhtarian, Patricia L. 2004. What type of vehicle do people drive? The role of attitude and lifestyle in influencing vehicle type choice. *Transportation Research Part A: Policy and Practice*, **38**(3), 201–222.

Comité des Constructeurs Français d'Automobiles. 2016. SOeS; MEDDE; ASFA, Kantar Worldpanel; TNS; Setra; CPDP.

Daziano, Ricardo A. 2013. Conditional-logit Bayes estimators for consumer valuation of electric vehicle driving range. *Resource and Energy Economics*, **35**(3), 429–450.

Daziano, Ricardo A., & Achtnicht, Martin. 2014. Forecasting Adoption of Ultra-Low-Emission Vehicles Using Bayes Estimates of a Multinomial Probit Model and the GHK Simulator. *Transportation Science*, **48**(4), 671–683.

de Jong, Gerard, Fox, James, Daly, Andrew, Pieters, Marits, & Smit, Remko. 2004. Comparison of car ownership models. *Transport Reviews*, **24**(4), 379–408.

Dimitropoulos, Alexandros, Rietveld, Piet, & van Ommeren, Jos N. 2013. Consumer valuation of changes in driving range: A meta-analysis. *Transportation Research Part A: Policy and Practice*, **55**(Sept.), 27–45.

European Comission. 2011. *Energy Roadmap 2050. Impact assessment and scenario analysis*.

Glerum, Aurélie, Stankovikj, Lidija, Thémans, MichaÃ«l, & Bierlaire, Michel. 2014. Forecasting the demand for electric vehicles: accounting for attitudes and perceptions. *Transportation Science*, **48**(4), 483–499.

Hackbarth, AndrÃ©, & Madlener, Reinhard. 2013. Consumer preferences for alternative fuel vehicles: A discrete choice analysis. *Transportation Research Part D: Transport and Environment*, **25**(Dec.), 5–17.

Hackbarth, AndrÃ©, & Madlener, Reinhard. 2016. Willingness-to-pay for alternative fuel vehicle characteristics: A stated choice study for Germany. *Transportation Research Part A: Policy and Practice*, **85**(Mar.), 89–111.

Hess, Stephane, Fowler, Mark, Adler, Thomas, & Bahreinian, Aniss. 2012. A joint model for vehicle type and fuel type choice: evidence from a cross-nested logit study. *Transportation*, **39**(3), 593–625.

Institut national de la statistique et des études économiques. 2016a. Accessed on 20.11.2016.

Institut national de la statistique et des études économiques. 2016b. Accessed on 20.11.2016.

Institut national de la statistique et des études économiques. 2016c. *Prix moyens à la consommation en métropole - Utilisation de véhicules, biens et services de loisirs*. Accessed on 21.11.2016.

Lave, Charles A., & Train, Kenneth. 1979. A disaggregate model of auto-type choice. *Transportation Research Part A: General*, **13**(1), 1–9.

Mai, Tien, Frejinger, Emma, Fosgerau, Mogens, & Bastin, Fabien. 2015 (June). *A Dynamic Programming Approach for Quickly Estimating Large MEV Models*. Tech. rept.

Massiani, Jérôme. 2014. Stated preference surveys for electric and alternative fuel vehicles: are we doing the right thing? *Transportation Letters*, **6**(3), 152–160.

McCarthy, Patrick S., & Tay, Richard S. 1998. New Vehicle Consumption and Fuel Efficiency: A Nested Logit Approach1. *Transportation Research Part E: Logistics and Transportation Review*, **34**(1), 39–51.

McFadden, Daniel, Train, Kenneth, *et al.* . 2000. Mixed MNL models for discrete response. *Journal of applied Econometrics*, **15**(5), 447–470.

Ministère de l'environnement, de l'énergie et de la mer. 2016. *Bonus-Malus : dfinitions et barèmes pour 2016*. Accessed on 21.11.2016.

Mohammadian, Abolfazl, & Miller, Eric. 2002. Nested Logit Models and Artificial Neural Networks for Predicting Household Automobile Choices: Comparison of Performance. *Transportation Research Record: Journal of the Transportation Research Board*, **1807**(Jan.), 92–100.

Mohammadian, Abolfazl, & Miller, Eric. 2003. Empirical Investigation of Household Vehicle Type Choice Decisions. *Transportation Research Record: Journal of the Transportation Research Board*, **1854**(Jan.), 99–106.

Page, Matthew, Whelan, Gerard, & Daly, Andrew. 2000. Modelling the factors which influence new car purchasing.

Potoglou, Dimitris. 2008. Vehicle-type choice and neighbourhood characteristics: An empirical study of Hamilton, Canada. *Transportation Research Part D: Transport and Environment*, **13**(3), 177–186.

Schafer, Joseph L. 2000. *Analysis of incomplete multivariate data*. 1. ed., 1. crc press reprint edn. Monographs on statistics and applied probability, no. 72. Boca Raton: Chapman & Hall/CRC. OCLC: 249266966.

Sud Ouest. 2015. *Prix du gazole et de l'essence : ce qui va changer pour les automobilistes*. Accessed on 21.11.2016.

Train, Kenneth. 1980. The potential market for non-gasoline-powered automobiles. *Transportation Research Part A: General*, **14**(5-6), 405–414.

Train, Kenneth. 1986. *Qualitative choice analysis: theory, econometrics, and an application to automobile demand*. MIT Press series in transportation studies, no. 10. Cambridge, Mass: MIT Press.

Train, Kenneth. 2009. *Discrete choice methods with simulation*. 2nd ed edn. Cambridge ; New York: Cambridge University Press. OCLC: ocn349248337.

Train, Kenneth E., & Winston, Clifford. 2007. Vehicle choice behavior and the declining market share of U.S. automakers. *International Economic Review*, **48**(4), 1469–1496.

U.S. Department of Energy. 2016. *Energy Efficiency & Renewable Energy*. Accessed on 21.11.2016.

Wu, Ge, Yamamoto, Toshiyuki, & Kitamura, Ryuichi. 1999. Vehicle Ownership Model That Incorporates the Causal Structure Underlying Attitudes Toward Vehicle Ownership. *Transportation Research Record: Journal of the Transportation Research Board*, **1676**(Jan.), 61–67.