Hierarchical Bayes Mixed logit modelling for purchase car behaviour

GAETANO CARMECI[i], EVA VALERI[ii,*]

[i] DEAMS, Università di Trieste
[ii] European Commission, Joint Research Centre (JRC), Institute for Prospective Technological Studies (IPTS), Economics of Climate Change, Energy and Transport Unit, Edificio Expo. C/Inca Garcilaso 3, 41092 Seville

1. Introduction

Modelling car purchase decisions using discrete choice models (DCMs) has a long tradition. However, with the progress obtained with the electrification process of the conventional cars new interesting changes appear in the European market.

Hybrid, plug-in and battery electric vehicles are the new entrants in the car market with gradually growing market shares, and with the hydrogen fuel cell vehicles in the process of going from the concept stage to the manufacturing one. All these engine technologies, together with the already existing Compressed Natural Gas vehicles (bi-fuel CNGVs) and Liquefied Petroleum Gas Vehicles (bi-fuel LPGVs) ones, are what is called as Alternative Fuel Vehicles (AFVs).

Achieving substantial market penetration of AFVs requires large investments in infrastructure for cars and fuels production, and an expansion of the network of refuelling facilities (MacLean et al., 2004). So, an urgent challenge is to evaluate the willingness of individuals to adopt, buy and use low-emission cars, and their preferences towards specific car features (e.g. car size, refuelling distance, purchase price). This is important not only for both car producers and manufacturers and, more in general, the automotive sector, in guiding their car design developments, but also for policy makers, to implement suitable policy measures aimed to efficiently promote AFVs' expansion. Wider use of AFVs requires an improved understanding of consumer needs, attitudes and desires, as well as consumer willingness to change vehicle purchase and travel behaviour. In this direction, recently the European Commission funded the *Green eMotion*[1] project with the aim of: setting a framework for pan-European interoperable

---

* The views expressed are purely those of the author and may not in any circumstances be regarded as stating an official position of the European Commission

1 The *Green eMotion* project was officially launched by Siim Kallas, Vice President of the European Commission and Commissioner for Transport, at a high-level kick-off meeting in Brussels on 31st of March 2011. The project had a total budget of €42 million and was funded by the European Commission with €24 million.

electromobility which is commonly accepted, user-friendly and scalable; integrating smart grid developments, innovative ICT solutions and different types of EUs various urban mobility concepts; enabling an European wide market place for electromobility to allow for roaming; and, providing a unique knowledge base. The project was also aimed to estimate demand models to contribute understanding the reasons for the low penetration of the electric vehicles in three different countries in order to provide recommendations on how to boost the potential demand for electric vehicles (Cherchi et al., 2015).

Compared with other neighbouring countries, Italy appears to be lagging behind in the penetration of the AFVs, though having high air and noise pollution levels (with daily limits frequently overcome in the most industrialized regions), a strong economic dependence from oil imports, the highest index of motorization in the world after USA (60 million of people, 37 million of passenger cars). Only bi-fuel CNGVs and bi-fuel LPGVs have recently gained relevant market shares in some regions of the country. Consumer willingness to change travel behaviour and accept different types of vehicles and driving patterns is an important area of uncertainty (IEA, 2011). Public acceptance of AFVs is a key factor determining the ultimate success/failure of their technologies. On this, Valeri and Danielis (2015) calibrating a mixed error component logit model estimated the potential market shares of AFVs based on stated preference car choice data, and using a Monte Carlo simulation model, they evaluated under different scenarios, the Italians' reactions of potential policy measures aimed to boost the AFVs' diffusion. Cherchi et al. (2015) estimating hybrid choice models investigated the role of individuals' attitudes (such as environmental concern, technological interest and appreciation car feature, scepticism, and pro-environmental attitude) on car purchase decisions.

In this paper, we estimate Mixed Logit models under a Bayesian Hierarchical framework (called, HBML model). Our model permits us to take account of possible dependence of the car attribute random parameters on individual characteristics, like e.g. age and gender. Moreover, alternative-specific parameters and correlation across alternatives can be added straightforwardly to the model. Given the complexity of the models employed, the standard approach to obtain an approximation of the joint posterior distribution of both the model parameters and hyper-parameters is to use MCMC methods. The most commonly used are the Gibbs sampler and the Metropolis-Hastings (M-H) algorithm (as done for instance by Train 2002, Daziano 2015, Scaccia and Marcucci 2010). These methods, particularly Gibbs sampler, however, tend to be highly inefficient when applied to non-linear and hierarchical models as our HBML model. As a result, a huge number of sweeps of the MCMC algorithm is required to obtain a reliable approximation of the joint posterior distribution and marginals. To our knowledge, in the field of purchase behaviour for conventional and alternative fuel cars, this paper uses for the first time an alternative approach to Gibbs sampler and M-H algorithm, based on Hamiltonian Monte Carlo (HMC) methods (Duane et al., 1987; Neal, 1994, 2011). The HMC sampler accelerates both the convergence to the stationary distribution and the subsequent parameter exploration by exploiting the information coming from the gradient of the log probability function. Using a novel package of the R software, called *rstan*, we specify and estimate three flexible random parameter logit models through which we show individuals' sensitivity towards the tested attributes.

The paper is organised as follows: a synthetic review of the car purchase literature is presented in Section 2; the Stated Preference (SP) experiment and the methodology used are described in Section 3, while the econometric results are reported and discussed in Section 4. Section 5 proposes conclusions and future extensions.

2. Literature review of car choice

The first Bayesian applications to DCMs appeared at the beginning of the early nineties, applying the Bayesian approach to the conditional and nested logit models (Koop, Poirier, 1993, Poirier, 1996), using the Gibbs sampler (Gelfand, Smith, 1990) and data augmentation to perform Bayesian inference for the MNL probit model (Tanner, Wong, 1987) and binary and ordered choice models (Albert, Chib, 1993)[2].

Table 1 shows an overview of Bayesian DCM applications in the car choice literature.

Daziano and colleagues analysed Canadian consumers' choices for AFVs estimating hybrid choice models with a Bayesian approach (Daziano 2010, 2015, Daziano, Bolduc 2009, 2013a, 2013b). Using data from a survey conducted in 2002 by the Energy and Materials Research Group, Simon Fraser University (EMRG) and modelling the data in R language, they performed hybrid Kernel Gibbs models (Daziano 2010), to explain environmental preferences in a private vehicle choice context. Overall, they found that the latent variable of the environmental concern enters very significantly and positively into the choice model specification. In fact, reporting the highest effect for the hydrogen fuel cell vehicles, followed by AFVs, and then by hybrid vehicles, the environmental concern boosts the choice of alternative fuel technologies. Recently, Daziano (2015) implemented a structural choice model with a multinomial probit kernel and discrete effect indicators to study continuous latent segments of travel behavior and designing a vehicle purchase model. Exploiting five underlying latent attitudes to determine segments of pro-transit, pro-environment, pro-safety, cost-conscious, and pro-performance consumers, he found interesting results such as cost-conscious consumers appear as having a continuous sensitivity to changes in travel and fuel costs. This pattern of valuation of changes in fuel costs are reflected in an implicit discount rate of future energy savings – which is a measure of the energy paradox – that slightly increases with income. In addition, consumers that appreciate safety exhibit a lower probability of choosing not only hydrogen cars, but also hybrids.

Using 2002-2006 new midsize sedan aggregate sales data in US, Haaf (2014) applied a Bayesian approach to different Mixed Logit (ML) models (correlated mixed logit versus independent mixed logit) to calibrate private vehicles' market shares in a Matlab environment. The Bayesian estimation allowed having better select ASC forecasting method, considering them as model parameters (draws of predicted ASCs are made jointly with observed coefficient draws). He found that concerning estimate uncertainty, the share uncertainty from predictive ASCs is greater than uncertainty from observed coefficients.

Table 1 - Literature review of Bayesian DCM studies applied to the car choice

Authors	Year	Study context	Document type	Data type	Model type	Estimator type	Forecasting analysis	Software
Daziano, Bolduc	2009	Canada	Conference paper	SP	HCM	Logit mixture kernel	Yes	R
Daziano	2010	Canada	PhD thesis	SP	MNL, HCM	Gibbs sampler, discrete choice kernel	Yes	R
Daziano, Bolduc	2013a	Canada	Journal paper	SP	HCM	Gibbs sampler	No	R
Daziano, Bolduc	2013b	Canada	Journal paper	SP	HCM	Metropolis Hastings-within-Gibbs sampler	No	R
Haaf	2014	USA	PhD thesis	RP-SP	ML	Gibbs sampler	Yes	Matlab
Daziano	2015	Canada	Journal paper	SP	HCM, probit	Kernel sampler	Yes	R

Notes: SP = Stated Preference, RP = Revealed Preference, HCM = Hybrid Choice Model, MNL = Multinomial Logit Model

2 For an overview of the evolution of the Bayesian approach in the DCM literature see Brownstone (2000); and for a theoretical comparison between a hierarchical Bayes and maximum simulated likelihood for mixed logit models see Train (2001).

Overall, the literature survey reveals that at our knowledge the studies are very few and concentrated in Canada and USA. Although private car purchase choices via DCMs have a long tradition in the empirical literature around the world (e.g. Lave, Train, 1979), no studies on this specific research field with a Bayesian approach seem to be carried out in Europe and, even less, in Italy. For the latter, it seems that the studies about potential demand estimation of AFVs are carried out by Valeri and Danielis (2015) and Cherchi et al. (2015), but with the frequentist estimation approach. Moreover, all the reviewed studies are mainly concentrated on estimation of hybrid choice type of model calibrated with R software on SP data.

3. Survey design and data

The present survey was conducted in various Italian cities (Trieste, Bologna, Pesaro), collecting randomly car purchase choices with *face-to-face* interviews in the first half of the year 2013.

The labelled choice experiments contain seven car alternatives:

- Ford Fiesta (with a Diesel engine technology),
- VW Polo (with a Gasoline engine technology),
- Fiat Punto Evo (with a Bi-fuel – CNG engine technology),
- Natural Power Alfa Romeo Mito (with a Bi-fuel – LPG engine technology),
- Toyota Yaris (with a Hybrid – Gasoline engine technology),
- Peugeot iOn (with an Electric – Owned Battery engine technology),
- Renault Zoe (with an Electric – Leased Battery engine technology) (see Table 3).

These specific cars were chosen because they are very popular in Italy and representative of their fuel\ powertrain technology. They all belong to the same car segment (B segment), apart from the Peugeot iOn that belongs to the A segment.

Drawing from the literature, five attributes were included in the SP choice experiments: purchase price (€), annual operating cost (gasoline, insurance, tax, maintenance) (€), acceleration (seconds), range (kilometres), and refuelling distance (kilometres).

The selected attributes were set as follows for the choice experiments:

- Purchase price: -20%, Status Quo, +20%, +40%,
- Annual operating cost: -20%, Status Quo, +20%,
- Range: Status Quo, +20%, +40%,
- Acceleration: Status Quo, -10%, -20%, and
- Refuelling distance: Gasoline, Diesel and Hybrid cars (1 km, 5km, 10km); CNG and LPG cars (5km, 20km, 50km) and Electric cars (0km, 5km, 10km).

The Status Quo attributes for each car were set equal to the Italian average values as shown in Table 2. The purchase price is the prevailing market price in Italy in 2013. The annual operating cost is calculated by Rusich and Danielis (2013) assuming an annual driving distance of 10,000 km. The range and acceleration are the ones reported by the car manufacturers. The refuelling distance is estimated based on the current Italian fuel distribution network.

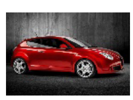
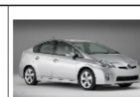
An efficient experimental design strategy was used with four waves in order to minimize the asymptotic standard error (e.g. Bliemer & Rose, 2010, 2011).

Due to time and budget constraints only 121 interviews could be collected in the first semester of the year 2013. Although the sample size is admittedly small, we decided not to carry out other interviews in 2013 and to devote more resources for the next years as new AFVs enter in the Italian car market and the consumers get acquainted to the new technologies.

Table 2 – Overview of the attributes-levels

Type of engine technology:	Purchase price	Annual operating cost	Range	Acceleration	Refuelling distance
	€	€	km	0-100 km in sec.	km
VW Polo (Gasoline):	11,900	2,081	900	13	1
Ford Fiesta (Diesel):	14,000	1,894	980	15	1
Fiat Punto Evo (Bi-fuel - CNG):	15,425	1,757	800	15	5
Natural Power Alfa Romeo Mito (Bi-fuel - LPG):	20,600	1,784	1,200	15	5
Toyota Yaris (Hybrid - gasoline):	18,650	1,920	1,000	13	1
Peugeot iOn (Electric – Owned Battery):	30,369	1,681	150	12	0
Renault Zoe (Electric – Leased Battery):	21,650	2,553	210	12	0

Table 3 - Choice task example

Car features:							
	Ford Fiesta (Diesel)	VW Polo (Gasoline)	Fiat Punto Evo (Bi-fuel - CNG)	Alfa Romeo Mito (Bi-fuel - LPG)	Toyota Yaris (Hybrid - gasoline)	Peugeot iOn (Electric – Own Battery)	Renault Zoe (Electric – Leased Battery)
Purchase price (€):	14,000	11,900	15,425	20,600	18,650	30,369	21,650
Range (km.):	980	900	800	1,200	1,000	150	210
Acceleration (0-100 km. in sec.):	15	13	15	15	13	12	12
Annual operating cost (€):	1,894	2,081	1,757	1,784	1,920	1,681	2,553
Refuelling (km.):	1	1	5	5	1	0	0
Which car would you buy?							

4. HIERARCHICAL BAYESIAN ML MODELLING

4.1 HBML MODELS SPECIFICATION

We estimated the following HBML model and two restricted variants of it. According to our specification, the $i-th$ interviewee, $i=1,\ldots,N$, with $N=121$, faces a choice among 7 alternatives in each of $J=12$ tasks and the person's utility from the different alternatives in the $j-th$ choice task is:

(1)

$$
\begin{aligned}
U_{ij,Gasoline.car} &= \beta_{PP,i}\text{PurchasePrice}_{ij,G} + \beta_{AOC,i}\text{AnnualOperatingCost}_{ij,G} + \beta_{R,i}\text{Range}_{ij,G} \\
&\quad + \beta_{A}\text{Acceleration}_{ij,G} + \beta_{RD}\text{RefuellingDistance}_{ij,G} + \varepsilon_{ij,G} = \eta_{ij,G} + \varepsilon_{ij,G} \\
U_{ij,Diesel.car} &= \alpha_{D,i} + \beta_{PP,i}\text{PurchasePrice}_{ij,D} + \beta_{AOC,i}\text{AnnualOperatingCost}_{ij,D} \\
&\quad + \beta_{R,i}\text{Range}_{ij,D} + \beta_{A}\text{Acceleration}_{ij,D} + \beta_{RD}\text{RefuellingDistance}_{ij,D} \\
&\quad + c_F\text{Female}_i + c_A\text{Age}_i + \varepsilon_{ij,D} = \eta_{ij,D} + \varepsilon_{ij,D} \\
U_{ij,CNG.car} &= \alpha_{CNG,i} + \beta_{PP,i}\text{PurchasePrice}_{ij,CNG} + \beta_{AOC,i}\text{AnnualOperatingCost}_{ij,CNG} \\
&\quad + \beta_{R,i}\text{Range}_{ij,CNG} + \beta_{A}\text{Acceleration}_{ij,CNG} + \beta_{RD}\text{RefuellingDistance}_{ij,CNG} \\
&\quad + c_F\text{Female}_i + c_A\text{Age}_i + \varepsilon_{ij,CNG} = \eta_{ij,CNG} + \varepsilon_{ij,CNG} \\
U_{ij,LPG.car} &= \alpha_{LPG,i} + \beta_{PP,i}\text{PurchasePrice}_{ij,LPG} + \beta_{AOC,i}\text{AnnualOperatingCost}_{ij,LPG} \\
&\quad + \beta_{R,i}\text{Range}_{ij,LPG} + \beta_{A}\text{Acceleration}_{ij,LPG} + \beta_{RD}\text{RefuellingDistance}_{ij,LPG} \\
&\quad + c_F\text{Female}_i + c_A\text{Age}_i + \varepsilon_{ij,LPG} = \eta_{ij,LPG} + \varepsilon_{ij,LPG} \\
U_{ij,Hybrid.car} &= \alpha_{H,i} + \beta_{PP,i}\text{PurchasePrice}_{ij,H} + \beta_{AOC,i}\text{AnnualOperatingCost}_{ij,H} \\
&\quad + \beta_{R,i}\text{Range}_{ij,H} + \beta_{A}\text{Acceleration}_{ij,H} + \beta_{RD}\text{RefuellingDistance}_{ij,H} \\
&\quad + c_F\text{Female}_i + c_A\text{Age}_i + \varepsilon_{ij,H} = \eta_{ij,H} + \varepsilon_{ij,H} \\
U_{ij,Electric.car-owned.battery} &= \alpha_{E-ob,i} + \beta_{PP,i}\text{PurchasePrice}_{ij,E-ob} + \beta_{AOC,i}\text{AnnualOperatingCost}_{ij,E-ob} \\
&\quad + \beta_{RE}\text{Range}_{ij,E-ob} + \beta_{A}\text{Acceleration}_{ij,E-ob} + \beta_{RD}\text{RefuellingDistance}_{ij,E-ob} \\
&\quad + c_F\text{Female}_i + c_A\text{Age}_i + \varepsilon_{ij,E-ob} = \eta_{ij,E-ob} + \varepsilon_{ij,E-ob} \\
U_{ij,Electric.car-leased.battery} &= \alpha_{E-lb,i} + \beta_{PP,i}\text{PurchasePrice}_{ij,E-lb} + \beta_{AOC,i}\text{AnnualOperatingCost}_{ij,E-lb} \\
&\quad + \beta_{RE}\text{Range}_{ij,E-lb} + \beta_{A}\text{Acceleration}_{ij,E-lb} + \beta_{RD}\text{RefuellingDistance}_{ij,E-lb} \\
&\quad + c_F\text{Female}_i + c_A\text{Age}_i + \varepsilon_{ij,E-lb} = \eta_{ij,E-lb} + \varepsilon_{ij,E-lb}
\end{aligned}
$$

where $\varepsilon_{ij,a} \sim i.i.d.$ extreme value, $\forall i=1,\ldots,N, \forall j=1,\ldots,J, \forall a \in \Omega$, with $\Omega=\{G,D,CNG,LPG,H,E-ob,E-lb\}$, and $\eta_{ij,a}$ represents the predicted utility of alternative a. Let $y_i=(y_{i1},y_{i2},\ldots,y_{iJ})$ be the interviewee's sequence of choices over the J

tasks, made under the assumption that the interviewee i chooses alternative $a^* \in \Omega$ in choice task j if $U_{ij,a^*} > U_{ij,a}, \forall a \neq a^*$. The likelihood for the entire sample $y = \left(y_1^{'}, ..., y_N^{'}\right)^{'}$ of the fixed and random parameters entering the predicted utility $\eta_{ij,a}, \forall a \in \Omega$, can be written as follows:

$$L(\lambda, \theta \mid y) = \prod_{i=1}^{N} L(\lambda, \theta_i \mid y_i) \tag{2}$$

where the vector λ contains all the fixed parameters, $\theta = \left(\theta_1^{'}, ..., \theta_i^{'}, ... \theta_N^{'}\right)^{'}$, with vector θ_i, $\forall i = 1, ..., N$, containing the random parameters characterizing the utility of i-th person, and $L(\lambda, \theta_i \mid y_i)$ is the product over the J tasks of standard logit formulas:

$$L(\lambda, \theta_i \mid y_i) = \prod_{j=1}^{J} \frac{e^{\eta_{ij, y_{ij}}}}{\sum_{a \in \Omega} e^{\eta_{ij,a}}} . \tag{3}$$

The HBML model is completed by specifying a hierarchical prior for the parameters. We assume that the random parameters on the car attributes Purchase Price, Annual Operating Cost and Range for non-Electric car are all linearly affected by the individual's gender and age[3]:

$$\begin{aligned} \beta_{PP,i} &= d_{0,PP} + d_{1,PP} Female_i + d_{2,PP} Age_i + u_{PP,i} \\ \beta_{AOC,i} &= d_{0,AOC} + d_{1,AOC} Female_i + d_{2,AOC} Age_i + u_{AOC,i} \\ \beta_{R,i} &= d_{0,R} + d_{1,R} Female_i + d_{2,R} Age_i + u_{R,i} \end{aligned} \tag{4}$$

where the error terms are mutually independent with $u_{PP,i} \sim i.i.d.N(0; \sigma_{PP}^2)$, $u_{AOC,i} \sim i.i.d.N(0; \sigma_{AOC}^2)$ and $u_{R,i} \sim i.i.d.N(0; \sigma_R^2)$. The hyper-parameters $d_{0,PP}, d_{1,PP}, d_{2,PP}$, $d_{0,AOC}$, $d_{1,AOC}$, $d_{2,AOC}$, $d_{0,R}, d_{1,R}$ and $d_{2,R}$ are assumed to be independently and identically distributed $N(0; 5^2)$. The error standard deviations, σ_{PP}, σ_{AOC} and σ_R, are assumed to be *i.i.d.* half-Cauchy (0, 2.5).

The Alternative Specific Constants (ASCs) are treated as random parameters and specified as follows:

$$\begin{aligned} &\alpha_{D,i} \mid \mu_D, \sigma_D \sim i.i.d.N(\mu_D; \sigma_D^2) \\ &\alpha_{CNG,i} \mid \mu_{CNG}, \sigma_{CNG} \sim i.i.d.N(\mu_{CNG}; \sigma_{CNG}^2) \\ &\alpha_{LPG,i} \mid \mu_{LPG}, \sigma_{LPG} \sim i.i.d.N(\mu_{LPG}; \sigma_{LPG}^2) \\ &\alpha_{H,i} \mid \mu_H, \sigma_H \sim i.i.d.N(\mu_H; \sigma_H^2) \\ &\left(\alpha_{E-ob,i}, \alpha_{E-lb,i}\right) \sim i.i.d.N^{(2)}\left(\left(\mu_{E-ob}, \mu_{E-lb}\right); \Sigma_E\right) \end{aligned} \tag{5}$$

3 The variable Age is specified as a dummy variable, assuming the value one if the respondent's age is greater than 29 years and zero otherwise.

Besides the Electric car constants that are assumed to be correlated, i.e. Σ_E is assumed to be not diagonal, all the other ASCs' are mutually independent. The prior distribution for the means $\mu_a, a \in \Omega$, standard deviations, $\sigma_D, \sigma_{CNG}, \sigma_{LPG}$, and σ_H, and covariance matrix Σ_E, are independent and we assume that the means are identically distributed $N(0;5^2)$. The covariance matrix Σ_E has an Inverse Wishart prior distribution with 6 d.f. and scale matrix $100I_2$, whereas the standard deviations are assumed to be *i.i.d.* half-Cauchy (0; 2.5). Finally, the fixed parameters are assumed to follow the independent prior distributions:

$$\begin{aligned} &\beta_A \sim N\left(0;10^2\right) \\ &\beta_{RD} \sim N\left(0;10^2\right) \\ &\beta_{RE} \sim N\left(0;10^2\right) \\ &c_F \sim N\left(0;10^2\right) \\ &c_A \sim N\left(0;10^2\right) \end{aligned} \tag{6}$$

The above model is labelled Model 2 and its estimation results are reported in Table 5. We also estimated two restricted versions of Model 2. In Model 1, we restricted the random parameters of the car attributes Purchase Price, Annual Operating Cost and Range of non-Electric car to be *i.i.d.* in the population, i.e. they were assumed to be no more dependent on the individual's gender and age. Therefore for this model in eq. (4) we set $d_{1,PP} = d_{2,PP} = 0$, $d_{1,AOC} = d_{2,AOC} = 0$ and $d_{1,R} = d_{2,R} = 0$. Moreover, the Electric ASC's were assumed to be independent, i.e. Σ_E was specified as a diagonal matrix, so that the Inverse Wishart prior distribution was replaced by two *i.i.d.* half-Cauchy (0; 2.5) priors for the standard deviations, σ_{E-ob} and σ_{E-lb}. The estimation results are reported in Table 4.

As Model 1, our preferred model, labelled Model 3, has the random parameters on the car attributes Purchase Price, Annual Operating Cost and Range of non-Electric car specified as *i.i.d.* in the population. However, like Model 2, we assumed the Electric ASCs' to be correlated. Moreover, we imposed the restriction that the two electric means were the same and equal to μ_E: $\mu_{E-ob} = \mu_{E-lb} = \mu_E$ (see Table 6 for the estimation results).

4.2 COMPUTATIONAL IMPLEMENTATION

Given the complexity of the models we presented, the standard approach to obtain an approximation of the joint posterior distribution of both parameters and hyper-parameters is to use MCMC methods. The most commonly used are the Gibbs sampler and the M-H algorithm (as done for instance by Train 2002, Scaccia and Marcucci 2010, Daziano 2015). These methods, particularly Gibbs sampler, however, tend to be highly inefficient when applied to non-linear and hierarchical models as our HBML model. As a result, a huge number of sweeps of the MCMC algorithm is required to obtain a reliable approximation of the joint posterior distribution and marginals. Of course, also the size of the burn-in sample has to be carefully monitored in order to avoid using draws from the transient phase of the Markov Chain. To our knowledge, in the field of purchase behaviour for conventional and alternative fuel cars, this paper uses for the first time an alternative approach to Gibbs sampler and M-H algorithm, based on HMC methods (Duane et al., 1987; Neal, 1994, 2011). The HMC sampler accelerates both convergence to the stationary distribution and subsequent parameter exploration by using the gradient of the log probability function in the *leapfrog* algorithm. Recently, these methods have been implemented by Andrew Gelman, Bob Carpenter and a group of researchers (see http://mc-stan.org/team/ for the all list) in a novel package of the R software (called *rstan*) (The Stan Development Team 2014, 2015). HMC methods have the ability to overcome some of the problems inherent in Gibbs sampling.

More specifically, the HMC methods implemented in Rstan use the No-U-Turn (NUTS) sampler (see Hoffman and Gelman, 2011, 2012, 2014). This sampler demonstrated to be able to efficiently solve the problem of tuning parameter of steps, a problem afflicting previous HMC algorithms. In fact, the Hamiltonian dynamics simulation requires not only the gradient of the log posterior but also two tuning parameters, the step size and the number of steps; moreover, it is very sensitive to how they are set. The step size parameter can be tuned during warmup based on Metropolis rejection rates, but the number of steps is not so easy to tune while maintaining detailed balance in the sampler. The NUTS sampler solves this problem by taking an ever increasing number of steps until the direction of the simulation turns around, then uses slice sampling to select a point on the simulated trajectory. While implemented HMC methods are more numerically intensive than Gibbs sampler and M-H algorithm (with a slow running time), they are highly more efficient. In our application, we obtained convergence after only 1,000 warmup draws and high values for the effective sample size of parameters.

4.3 THE HBML ESTIMATION RESULTS

In this section we report the results of the estimated models using *rstan* package (version 2.8.0) with R (version 3.2.2), based on the simulations of four mutually independent parallel chains, each one of length 4,000. We cautiously discarded the first 2,000 draws from each sample as burn-in period and retained the subsequent 2,000 draws, so that a total number of 8,000 draws were used for estimation.

As described in section 4.1, we estimated three HBML models (see Table 4, Table 5, and Table 6). Table 7 (in Annex) reports a brief description of the estimated parameters.

The content of each table is structured as follows: the second column, named *mean,* contains the MCMC estimates of the marginal posterior means for the parameters of interest, computed as the average of the saved simulations (i.e. the remaining simulations after warmup); while in the column *se_mean* the standard error of the MCMC estimator of the mean is reported (i.e. the precision of the estimated mean). The column named *sd* shows the MCMC estimate of the standard deviation of the posterior marginal distribution for the parameters of interest. Also the following estimated quantiles of the marginal posterior distribution are reported: 2.5%, 25%, 50%, 75%, and 97.5%. The effective sample size, reported as *n_eff,* represents a measure of the autocorrelation found in the Monte Carlo Markov chains. Without autocorrelation, the effective sample size is equal to the total number of simulated values for each parameter after warmup, i.e. the sample size; with autocorrelation, *n_eff* will be lower than the sample size. The larger the autocorrelation is, the lower the effective sample size will be, so that the ratio, *n_eff* over the sample size, measures the degree of mixing of Markov chains. The last column, *Rhat*, contains a statistic for assessing convergence of the chains to the same stationary distribution. Its value should be close to 1.0 when the chains have all converged to the same stationary distribution.

As described in section 4.1, the first model we estimated, called Model 1, corresponds to a standard random parameters mixed logit model. We set individual-specific random parameters for Purchase Price, Annual Operating Cost and Range of non-Electric car attributes. Individuals' socio-economic variables (female and age of respondents) enter as fixed parameters as well as the remaining attributes (Acceleration time, Refuelling Distance and Range of electric car). All random parameters (for both ASCs and slopes) are assumed to be *i.i.d.* Normally distributed in the population as well as mutually independent.

Results in Table 4 show that the sign of the population average effects of the car attributes, Purchase Price (-), Annual Operating Cost (-) and Range of non-Electric car (+), are in line with theory and significant[4] at 5% (see results for $d_{0,PP}$, $d_{0,AOC}$ and $d_{0,\mathrm{R}}$). Moreover, as expected, for the two Electric cars alternatives the attribute Range has not only a positive and significant effect on their choice (see β_{RE}), but also the magnitude of the estimated mean is more than six times the magnitude for non-Electric cars (see the mean of $d_{0,\mathrm{R}}$).

From Table 4, we can see that Refuelling Distance affects car choice both negatively and significantly: the posterior probability that β_{RD} is inside the (-0.03,-0.01) interval is equal to 0.95 while the posterior mean is -0.01. The only car attribute turning out to be not significant al 5% is Acceleration, although the posterior mean as expected is positive. All the aforementioned results are in line with the previous ones (Valeri, Danielis, 2015).

Individuals' socio-economic variables affect car choices of AFVs. Given the negativity and significance of the c_F posterior mean, it turns out that females tend to prefer the gasoline car alternative, holding other things constant, including age.

The sign of the relationship is reversed as far as the respondent's age is concerned: the posterior mean of c_A results to be 0.52 but slightly insignificant at 5%.

Controlling for these socio-economic effects, only the population means of ASCs for the two electric car alternatives turn out to be highly significant. With reference to the Gasoline car alternative, their population average effects are negative and their magnitudes dominate the other ASCs. Finally, their population means look very similar.

We considered two extensions of Model 1 along the following lines:

i) The individual-specific random parameters of the car attributes Purchase Price, Annual Operating Cost and Range of non-Electric car are assumed to be no more *i.i.d.* in the population, but dependent on both the gender and age of the respondent;

ii) The two ASCs of the Electric car alternatives are assumed to be correlated in the population.

Moreover, on the base of Model 1's results, we checked for the equality of the two population means of electric car ASCs.

The estimation results of Model 2 are reported in Table 5. Looking at the results, neither the variable female nor the age one seem to be able to explain the individual random parameters heterogeneity in the population as modelled in eq. (4). Moreover, the population covariance, $\sigma_{E-ob,E-lb}$, and the population correlation, rho_E, of the two Electric car ASCs are both positive

4 Given that the value zero is outside the reported 95% probability interval, computed from the approximated marginal posterior density.

and near significantly different from zero at 5%. Finally, the difference between their population means, $diff\_\mu_E$, is not significant at 5%.

Therefore, we estimated a third model, named Model 3, for which we set a common mean for the two ASCs of the Electric car alternatives and removed all the socio-economic interactions with the random parameters.

As we can see from Table 6, the common population mean of the two Electric car ASCs is negative and highly significant. As in Model 2, the population correlation coefficient is positive and near significant at 5%.

From a qualitatively viewpoint, all others results analysed above in depth, when commenting the results of Model 1, continue to be confirmed also by the other two models (Model 2 and 3).

5. Summary and future extensions

In this paper, we analysed the purchase behaviour for conventional and alternative fuel cars, using Italian stated preference discrete choice data, and we proposed modelling Multinomial Logit models under a Bayesian hierarchical framework. We specified a flexible Hierarchical Bayesian Mixed Logit model that permit us to take account of possible dependence of the car attribute random parameters on individual socio-economic characteristics, like age and gender. Moreover, alternative-specific and/or common parameters, as well as correlation across alternatives are easily included in the model. Instead of relying on traditional Gibbs Sampler or Metropolis-Hastings algorithm, we proposed for the first time in the field of purchase behaviour for conventional and alternative fuel cars, to use Hamiltonian Monte Carlo methods (Duane et al., 1987; Neal, 1994, 2011). The HMC sampler is more efficient than traditional MCMC methods, since it accelerates both convergence to the stationary distribution and subsequent parameter exploration by exploiting information coming from the gradient of the log probability function. We have thoroughly shown in the empirical application the usefulness of the proposed method.

In this first study, we assumed the normality of the population density of the random parameters. However, as a future extension we would like to consider discrete mixtures of normals (or other continuous distributions) for modelling the density of the random parameters, as done by Scaccia and Marcucci (2010) for public transport demand.

Table 4 – Results of the Model 1

Parameters	mean	se_mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
μ_D	0.43	0.01	0.25	-0.05	0.25	0.42	0.6	0.92	391	1
μ_{CNG}	0.24	0.01	0.27	-0.28	0.05	0.24	0.43	0.77	475	1
μ_{LPG}	-0.05	0.02	0.43	-0.91	-0.33	-0.03	0.25	0.76	541	1
μ_H	0.07	0.01	0.27	-0.47	-0.1	0.07	0.25	0.59	430	1
μ_{E-ob}	-2.35	0.07	0.84	-4.07	-2.92	-2.32	-1.76	-0.78	149	1.02
μ_{E-lb}	-2.2	0.1	1.05	-4.27	-2.89	-2.16	-1.48	-0.18	109	1.02
σ_D	1.13	0.01	0.17	0.81	1.01	1.13	1.24	1.49	902	1
σ_{CNG}	1.29	0.01	0.21	0.89	1.15	1.29	1.43	1.73	772	1.01
σ_{LPG}	2.2	0.02	0.41	1.49	1.92	2.17	2.45	3.09	601	1
σ_H	1.42	0.01	0.21	1.04	1.28	1.41	1.55	1.86	1162	1
σ_{E-ob}	2.08	0.04	0.59	0.97	1.68	2.05	2.44	3.3	257	1.01
σ_{E-lb}	1.67	0.03	0.5	0.82	1.3	1.62	1.98	2.77	241	1.02
$d_{0,PP}$	-0.42	0	0.03	-0.49	-0.44	-0.42	-0.4	-0.35	1641	1
$d_{0,\mathrm{AOC}}$	-2.62	0.01	0.26	-3.15	-2.79	-2.62	-2.45	-2.11	1758	1
$d_{0,\mathrm{R}}$	1.28	0.02	0.43	0.45	0.99	1.28	1.57	2.14	361	1
σ_{PP}	0.28	0	0.03	0.22	0.25	0.27	0.3	0.34	1268	1
σ_{AOC}	2.27	0.01	0.24	1.84	2.1	2.25	2.42	2.78	1423	1
σ_R	2.01	0.02	0.37	1.31	1.76	2	2.24	2.74	383	1
β_A	0.03	0	0.03	-0.03	0.01	0.03	0.05	0.09	1992	1
β_{RD}	-0.02	0	0	-0.03	-0.02	-0.02	-0.02	-0.01	8000	1
β_{RE}	7.43	0.25	3.42	0.96	5.11	7.42	9.68	14.29	186	1.01
c_A	0.52	0.01	0.28	-0.02	0.33	0.53	0.72	1.08	869	1
c_F	-0.84	0.01	0.29	-1.41	-1.04	-0.84	-0.64	-0.29	886	1

Table 5 – Results of the Model 2

Parameters	mean	se_mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
μ_D	0.44	0.01	0.25	-0.05	0.28	0.45	0.61	0.92	567	1.01
μ_{CNG}	0.24	0.01	0.27	-0.31	0.06	0.24	0.42	0.76	657	1
μ_{LPG}	0.04	0.02	0.43	-0.87	-0.23	0.06	0.34	0.85	533	1.01
μ_H	0.06	0.01	0.27	-0.47	-0.12	0.06	0.24	0.59	593	1
σ_D	1.15	0.01	0.18	0.83	1.03	1.14	1.26	1.51	962	1
σ_{CNG}	1.31	0.01	0.21	0.93	1.16	1.3	1.45	1.74	1133	1
σ_{LPG}	2.07	0.02	0.41	1.32	1.79	2.04	2.32	2.96	599	1.01
σ_H	1.48	0.01	0.21	1.1	1.33	1.47	1.61	1.91	1228	1
μ_{E-ob}	-2.97	0.06	0.97	-4.93	-3.63	-2.93	-2.29	-1.21	260	1.02
μ_{E-lb}	-3.05	0.07	1.15	-5.4	-3.79	-3.01	-2.26	-0.92	246	1.02
σ^2_{E-ob}	11.96	0.16	4.02	6.16	9.11	11.28	14	21.64	606	1.01
σ_{E-lb}	3.87	0.1	2.49	-0.18	2.16	3.56	5.23	9.72	662	1
$\sigma_{E-ob,\,E-lb}$	9.74	0.13	3.08	5.22	7.59	9.27	11.34	17.09	586	1.01
$d_{2,PP}$	-0.01	0	0.07	-0.14	-0.06	-0.01	0.04	0.13	4734	1
$d_{2,\mathrm{AOC}}$	-0.46	0.01	0.55	-1.55	-0.82	-0.46	-0.09	0.61	5066	1
$d_{2,\mathrm{R}}$	-0.55	0.01	0.68	-1.92	-1	-0.54	-0.09	0.76	2331	1
$d_{1,PP}$	-0.07	0	0.07	-0.21	-0.12	-0.07	-0.03	0.06	4687	1
$d_{1,\mathrm{AOC}}$	0.1	0.01	0.53	-0.94	-0.25	0.1	0.45	1.15	4348	1
$d_{1,\mathrm{R}}$	-0.48	0.01	0.64	-1.76	-0.9	-0.48	-0.06	0.74	2279	1
$d_{0,PP}$	-0.4	0	0.05	-0.5	-0.43	-0.4	-0.36	-0.3	3052	1
$d_{0,\mathrm{AOC}}$	-2.63	0.01	0.42	-3.48	-2.9	-2.62	-2.35	-1.83	3085	1
$d_{0,\mathrm{R}}$	1.73	0.02	0.56	0.65	1.35	1.72	2.11	2.84	902	1
σ_{PP}	0.29	0	0.03	0.23	0.27	0.29	0.31	0.36	1545	1
σ_{AOC}	2.41	0.01	0.25	1.96	2.24	2.4	2.58	2.95	1926	1
σ_R	2.17	0.02	0.45	1.31	1.86	2.16	2.47	3.05	593	1.01
β_A	0.03	0	0.03	-0.03	0.01	0.03	0.05	0.09	2347	1
β_{RD}	-0.02	0	0	-0.03	-0.02	-0.02	-0.02	-0.01	8000	1
β_{RE}	7.69	0.17	3.47	0.91	5.34	7.73	10	14.4	405	1.01
c_A	0.59	0.01	0.29	0.03	0.39	0.59	0.78	1.16	1205	1
c_F	-0.83	0.01	0.3	-1.42	-1.03	-0.83	-0.63	-0.25	1101	1
rho_E	0.34	0.01	0.18	-0.03	0.23	0.36	0.47	0.65	1218	1
$diff\_\mu_E$	0.01	0.03	0.76	-1.49	-0.49	0.01	0.51	1.48	918	1

Table 6 – Results of the Model 3

Parameters	mean	se_mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
μ_D	0.46	0.01	0.27	-0.08	0.28	0.46	0.64	0.97	353	1.01
μ_{CNG}	0.25	0.01	0.28	-0.33	0.06	0.26	0.44	0.79	446	1.01
μ_{LPG}	0.02	0.02	0.46	-0.96	-0.27	0.05	0.34	0.86	538	1
μ_H	0.08	0.01	0.29	-0.5	-0.11	0.09	0.28	0.64	394	1
σ_D	1.15	0	0.17	0.83	1.03	1.14	1.26	1.51	1251	1
σ_{CNG}	1.32	0.01	0.22	0.92	1.17	1.32	1.46	1.79	991	1
σ_{LPG}	2.13	0.02	0.42	1.41	1.83	2.09	2.39	3.01	678	1.01
σ_H	1.47	0.01	0.21	1.09	1.33	1.46	1.6	1.9	1304	1
$\mu_{E(ob\&lb)}$	-3.02	0.05	0.9	-4.82	-3.61	-2.99	-2.41	-1.34	323	1.01
σ^2_{E-ob}	11.51	0.1	3.47	6.13	9.03	11.01	13.44	19.61	1118	1
$\sigma_{E-ob,\,E-lb}$	3.65	0.07	2.37	-0.2	2.02	3.4	4.91	9.23	1282	1
σ^2_{E-lb}	9.41	0.08	2.73	5.19	7.45	9.02	10.99	15.67	1126	1
$d_{0,PP}$	-0.43	0	0.04	-0.5	-0.45	-0.43	-0.41	-0.37	1909	1
$d_{0,\mathrm{AOC}}$	-2.71	0.01	0.27	-3.25	-2.89	-2.71	-2.53	-2.2	2869	1
$d_{0,\mathrm{R}}$	1.33	0.02	0.44	0.45	1.03	1.33	1.62	2.2	698	1.01
σ_{PP}	0.29	0	0.03	0.23	0.26	0.28	0.31	0.36	1525	1
σ_{AOC}	2.37	0.01	0.25	1.92	2.19	2.36	2.53	2.89	1755	1
σ_R	2.07	0.03	0.47	1.14	1.76	2.08	2.39	2.97	342	1.02
β_A	0.03	0	0.03	-0.03	0.01	0.03	0.04	0.08	4036	1
β_{RD}	-0.02	0	0	-0.03	-0.02	-0.02	-0.02	-0.01	8000	1
β_{RE}	8.12	0.15	3.29	1.86	5.85	8.03	10.33	14.65	490	1
c_A	0.53	0.01	0.29	-0.04	0.33	0.53	0.73	1.11	646	1
c_F	-0.84	0.01	0.3	-1.42	-1.04	-0.84	-0.63	-0.25	833	1
rho_E	0.34	0	0.17	-0.02	0.23	0.35	0.46	0.64	1678	1

ACKNOWLEDGEMENTS

We thank Prof. Romeo Danielis to put at our disposal the data collected through the research project “Un Electric Car Club for the Friuli Venezia Giulia” in which the second author was involved.
We also thank Georgios Alaveras (JRC Seville), Francesco Pauli (University of Trieste), Matilde Trevisani (University of Trieste) for their useful advices about R software, and Michael Betancourt for teaching the “Bayesian Data Analysis” training course attended by the second author at IPTS (Seville), 15-19 June 2015, and for initial hints regarding our case study.

ANNEX

Table 7 – Description of the estimated parameters

Parameter	Parameter description
μ_D	Population mean of the ASC in the Diesel car alternative
μ_{CNG}	Population mean of the ASC in the CNG car alternative
μ_{LPG}	Population mean of the ASC in the LPG car alternative
μ_H	Population mean of the ASC in the Hybrid car alternative
μ_{E-ob}	Population mean of the ASC in the Electric car (with owned battery) alternative
μ_{E-lb}	Population mean of the ASC in the Electric car (with leased battery) alternative
$\mu_{E(ob\&lb)}$	Common mean of the ASC's in the two Electric cars' alternatives
σ_D	Population Standard deviation of the ASC in the Diesel car alternative
σ_{CNG}	Population Standard deviation of the ASC in the CNG car alternative
σ_{LPG}	Population Standard deviation of the ASC in the LPG car alternative
σ_H	Population Standard deviation of the ASC in the Hybrid car alternative
σ_{E-ob}	Population Standard deviation of the ASC in the Electric car (with owned battery) alternative
σ_{E-lb}	Population Standard deviation of the ASC in the Electric car (with leased battery) alternative
σ^2_{E-ob}	Population variance of the ASC in the Electric car (with owned battery) alternative
$\sigma_{E-ob,\,E-lb}$	Population covariance between the ASC's in the two Electric car alternatives
σ^2_{E-lb}	Population variance of the ASC in the Electric car (with leased battery) alternative
$d_{0,PP}$	Hyper-parameter (constant term) of the hierarchical model for Purchase Price random parameter
$d_{0,\mathrm{AOC}}$	Hyper-parameter (constant term) of the hierarchical model for the Annual Operating Cost random parameter
$d_{0,\mathrm{R}}$	Hyper-parameter (constant term) of the hierarchical model for the Range (non-Electric cars) random parameter
$d_{1,PP}$	Hyper-parameter (Female parameter) of the hierarchical model for the Purchase Price random parameter
$d_{1,\mathrm{AOC}}$	Hyper-parameter (Female parameter) of the hierarchical model for the Annual Operating Cost random parameter
$d_{1,\mathrm{R}}$	Hyper-parameter (Female parameter) of the hierarchical model for the Range (non-Electric cars) random parameter
$d_{2,PP}$	Hyper-parameter (Age parameter) of the hierarchical model for the Purchase Price random parameter
$d_{2,\mathrm{AOC}}$	Hyper-parameter (Age parameter) of the hierarchical model for the Annual Operating Cost random parameter
$d_{2,\mathrm{R}}$	Hyper-parameter (Age parameter) of the hierarchical model for the Range (non-Electric cars) random parameter
σ_{PP}	Population Standard deviation of the Purchase Price random parameter
σ_{AOC}	Population Standard deviation of the Annual Operating Cost random parameter
σ_R	Population Standard deviation of the Range (non-Electric cars) random parameter
β_A	Acceleration parameter (fixed parameter)
β_{RD}	Refuelling Distance parameter (fixed parameter)
β_{RE}	Range (electric cars) parameter (fixed parameter)
c_A	Coefficient of the respondent's Age (1 = age>29, 0 = otherwise) variable (fixed parameter)
c_F	Coefficient of the respondent's Gender (1 = female, 0 = otherwise) variable (fixed parameter)
rho_E	Population Correlation between the ASCs of the two Electric car alternatives
$diff\_\mu_E$	Difference of the population means between the ASCs of the two Electric car alternatives

References

ALBERT, J.H., CHIB, S. (1993) “Bayesian Analysis of Binary and Polychotomous Response Data“, *Journal of the American Statistical Association*, 88(422), pp. 669-679.

BLIEMER, M.C., ROSE, J.M. (2010) “Construction of experimental designs for mixed logit models allowing for correlation across choice observations“, *Transportation Research Part B: Methodological*, 44(6), pp. 720-734.

BLIEMER, M.C., ROSE, J.M. (2011) “Experimental design influences on stated choice outputs: An empirical study in air travel choice“, *Transportation Research Part A: Policy and Practice*, 45(1), pp. 63-79.

BROWNSTONE, D. (2001) “Discrete choice modeling for transportation“, in: HENSHER, D. (eds), *Travel Behaviour Research: The Leading Edge*. Pergamon, Amsterdam, pp. 97–124.

CHERCHI, E., MORRISSEY, P., O’MAHONY, M., WELDON, P., KELPIN, R., MANCA, F., MABIT, S., VALERI, E., CORCHERO, C. (2015) *Deliverable 9.1 - Consumers’ preferences and attitudes to, demand for, and use of electric vehicles (EV)*, *Green eMotion* project, version 5, available at: http://www.greenemotion-project.eu/upload/pdf/deliverables/D9_1-Consumers-preferences-and-attitudes_public.pdf.

DAZIANO, R.A. (2010) *A Bayesian approach to Hybrid Choice models*, PhD thesis, Sciences Sociales, Université Laval Québec.

DAZIANO, R.A. (2015) “Inference on mode preferences, vehicle purchases, and the energy paradox using a Bayesian structural choice model“, *Transportation Research Part B: Methodological*, 76(C), pp. 1-26.

DAZIANO, R.A., BOLDUC, D. (2009) “Canadian consumers’ perceptual and attitudinal responses towards green automobile technologies: An application of hybrid choice models“, *EAERE-FEEM-VIU European Summer School in Resources Environmental Economics: Economics, Transport and Environment*, Venice International University, Italy.

DAZIANO, R.A., BOLDUC, D. (2013a) “Incorporating pro-environmental preferences towards green automobile technologies through a Bayesian hybrid choice model”, *Transportmetrica A: Transport Science*, 9(1), pp. 74-106.

DAZIANO, R.A., BOLDUC, D. (2013b) “Covariance, identification, and finite-sample performance of the MSL and Bayes estimators of a logit model with latent attributes”, *Transportation*, 40(3), pp. 647-670.

DUANE, A., KENNEDY, A., PENDLETON, B., AND ROWETH, D. (1987) “Hybrid Monte Carlo“, *Physics Letters B*, 195(2), pp. 216-222.

GELFAND, A., SMITH A.F.M. (1990) “Sampling-based approaches to calculating marginal densities“, *Journal of the American Statistical Association*, Vol. 85, pp. 398-409.

HAAF, C.G. (2014) ”Vehicle Demand Forecasting with Discrete Choice Models: 2 Logit 2 Quit”, Dissertations, Paper 491.

HOFFMAN, M.D., GELMAN, A. (2011) The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo, http://arxiv.org/abs/1111.4246.

HOFFMAN, M.D., GELMAN, A. (2012) “The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo“, *Journal of Machine Learning Research*, In press.

HOFFMAN, M.D., GELMAN, A. (2014) “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo“, *Journal of Machine Learning Research*, 15, pp. 1593-1623.

KOOP, G., POIRIER D.J. (1993) “Bayesian analysis of logit models using natural conjugate priors“, *Journal of econometrics*, Vol. 56, pp. 323-340.

IEA, (2011), Technology Roadmap - Electric and plug-in hybrid electric vehicles (updated June 2011), available at: https://www.iea.org/publications/freepublications/publication/EV_PHEV_Roadmap.pdf

LAVE, C., TRAIN K.E. (1979) “A disaggregate model of auto-type choice“, *Transportation Research Part A*, 3(1), pp. 1-9.

MACLEAN, H.L., LAVE, L.B., GRIFFIN, M., (2004) “Alternative transport fuells for the future“, *International Journal of Vehicle Design*, 35 (1/2), pp. 27-49.

NEAL, R.M. (1994) “An improved acceptance procedure for the hybrid monte carlo algorithm“, *Journal of Computational Physics*, Vol. 111, pp. 194-203.

NEAL, R.M. (2011), “MCMC using Hamiltonian dynamics“. In Brooks, S., Gelman, A., Jones, G. L., and Meng, X.-L. (eds), Handbook of Markov Chain Monte Carlo, pp. 116-162. Chapman and Hall/CRC.

POIRIER, D.J. (1996) “A Bayesian Analysis of nested logit models“. *Journal of econometrics*, Vol. 75, pp. 163-181.

RUSICH, A., DANIELIS, R. (2013) “The private and social monetary costs and the energy consumption of a car. An estimate for seven cars with different vehicle technologies on sale in Italy“, Società Italiana di Economia dei Trasporti e della Logistica (SIET), Working paper n. 13.01.

SCACCIA, L., MARCUCCI, E. (2010) “Bayesian flexible modelling of mixed logit models“, Proceedings from the 19th International Conference on Computational Statistics (Lechevallier, Yves, Saporta, Gilbert eds.), Paris, France, August 22-27.

TANNER, M.A., WONG, W.H. (1987) “The calculation of posterior distributions by data augmentation (with discussion) “, *Journal of the American Statistical Association*, Vol. 82, pp. 528-50.

The Stan Development Team (2014) RStan getting started. https://github.com/stan-dev/rstan/wiki/RStan-Getting-Started.

The Stan Development Team (2015) Stan Modeling Language - User’s Guide and Reference Manual. Stan Version 2.8.0, Tuesday 8th September 2015, http://mc-stan.org/documentation/.

TRAIN, K.E. (2001) “A Comparison of hierarchical bayes and maximum simulated likelihood for mixed logit“, Working Paper, Department of Economics, University of California, Berkeley.

TRAIN, K.E. (2002) *Discrete choice methods with simulations*. Cambridge University Press, Cambridge.

VALERI, E., DANIELIS, R. (2015), "Simulating the market penetration of cars with alternative fuel\ powertrain technologies in Italy", *Transport Policy*, Vol. 37, pp. 44-56.