Public opinion on Usage-Based Motor Insurance Schemes: a stated preference approach

ABSTRACT

This paper aims to investigate which parameters affect users' willingness to pay for alternative usage-based motor insurance pricing schemes such as Pay-as-you-drive (PAYD) and Pay-as-how-you-drive (PHYD). For that reason, a dedicated questionnaire was designed and administered to 100 participants including both revealed and stated preference questions and proposed scenarios regarding current and alternative insurance schemes. In order to account for unobserved heterogeneity, a mixed logit model was applied to analyze vehicle insurance choice. Candidate variables include the effect of driving characteristics, drivers' demographics and the price of vehicle insurance premiums. Two distinct mixed logit models were developed; one mixed logit model to investigate the factors influencing the choice of present insurance policy over PAYD and one for present insurance policy over PHYD. Results indicated that women and smartphone owners are more likely to choose a new insurance schemes. Kilometers and cost reduction were also found to affect similarly the choice for both Usage-Based-Motor Insurance (UBI.) Moreover, the higher the speed reduction imposed to the user, the lower the probability of the UBI scheme to choose it. It was also found that people over 40 years old are less likely to choose PHYD insurance.

Key-words: Insurance; willingness to pay; stated preference; discrete choice

1. INTRODUCTION

Usage-based motor insurance (UBI) schemes, such as Pay-as-you-drive (PAYD) and Pay-how-you-drive (PHYD), constitute new innovative concepts that have recently started to be globally commercialized. The core concept is based on the fact that drivers pay insurance premiums depending on their travel and driving behavior instead of a fixed price based on demographics and/or their driving experience only. In spite of having been only recently implemented, it appears to be a very promising practice with a potentially significant impact on traffic safety as well as on traffic congestion mitigation and pollution emissions reduction (Tselentis et al., 2017).

Insurance charging systems based on travel behavior are often called Pay-As-You-Drive (PAYD) Usage-Based Insurance schemes. Drivers' travel behavior can be defined as their strategic choices (whether on a real-time basis or not) concerning which type of road network they use and at what time they drive in order to fulfil their travel needs. These choices are directly linked to their exposure to crash risk through their mileage, the road network type chosen and the related traffic conditions, the period of time chosen to drive and the related weather conditions. In the primary form of PAYD, mileage was only incorporated in the models as a travel behaviour characteristic. This was concluded based on the fact that mileage and crash risk are much correlated. Indeed, many studies (Litman, 2005, Bordoff and Noel, 2008) in literature indicate a relationship between VMT (vehicle miles travelled) and crash risk. For instance, Edlin (2003) found that the elasticity of the number of crashes occurring with respect to VMT is approximately 1.7 which means that if mileage was reduced by 10%, crashes would be reduced by 17% while in other research the elasticity of crash risk was found to be around 1.2 (ICBC Research Services Data, 1998). More specifically, the authors claim that the 1981-1982 recession led to a 10% VMT and 12% insurance claims reduction in British Columbia. In support of the above, Ferreira and Minikel (2010) found that there is a high statistical significance between mileage and risk and that they are positively correlated.

Another PAYD insurance scheme is the Pay-at-the-Pump (PATP) method which was the early stage of the mileage-based insurance policy that appeared later. Considering that fuel consumption and mileage are somehow correlated, these two methods share many similar characteristics and the same conceptual basis. PATP is the second most influential method of UBI which considers fuel consumption as its main indicator instead of mileage. For example, Wenzel (1995) argued why insurance premiums should be estimated based on use. Claiming that VMT is a good predictor of crash costs, he proposed a travel behaviour-based system which was actually a per-gallon surcharge for consumers, a method similar to the PATP method. Wenzel also suggested that premiums should be the sum of a fixed amount based on location, vehicle safety characteristics and driving record, most of which are travel behaviour characteristics, plus a variable amount based on fuel consumption (per-gallon surcharge).

On the other hand, insurance charging systems based on Driving Behavior are often called Pay How You Drive (PHUD) Usage Based Insurance schemes. Driving behavior can be defined as drivers' operational choices at real time in handling the vehicle within the existing traffic conditions. These choices are directly linked to the probability of getting involved in a traffic accident, based on the way they are driving, e.g. by speeding, harsh braking, harsh accelerating, harsh cornering, being distracted by mobile phone, etc.. The main advantages of UBI schemes compared to the conventional ones so far are discussed in more details in Sugarman, (1994), Litman (2004a), Litman 2004b and Tselentis et al. (2017) and so on. For instance, Bolderdijk et al. (2011) found that speed violations of young drivers are significantly reduced with PAYD schemes. The potential financial benefits and incentives are likely to lead to reduce speeds as Todelo et al. (2008) state. Similarly, other studies found that PHYD (or pay-as-you-speed) can be very beneficial in road safety (Lahrmann et al., 2012).

During the last few decades traditional motor insurance has started to gradually transform into Usage-Based Insurance. The question, , to what extent is this new type of motor insurance going to be widely adopted and which indicators will be fully incorporated, remains though. According to Tselentis et al. (2017), UBI will play a key role in motor insurance market in the future and as a result it will strongly influence traffic safety in total. Figure 1 illustrates the types of insurance that currently exist in the marketplace as well as the intuition of the authors on how motor insurance future will be formed. Since the trend in innovative motor insurance revealed above is to implement schemes that progressively incorporate travel and behavioural factors the authors consider that future models will be in the form of Pay-As-How-You-Drive (PAHYD) including parameters from both PAYD and PHYD models.

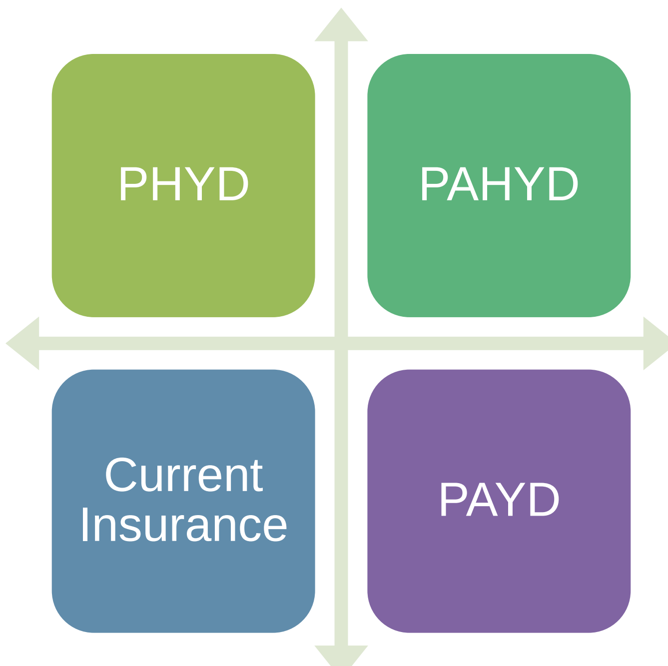

Figure 1: UBI and current Insurance policies. Source: Tselentis, D. I., Yannis, G., & Vlahogianni, E. I. (2017). Innovative motor insurance schemes: a review of current practices and emerging challenges. Accident Analysis & Prevention, 98, 139-148.

In order to estimate insurance premiums, the "Willingness to Pay" (WtP) methodology is examined, which is in fact the reflection of the individual estimate on how much money an individual is willing to pay (or sacrifice) so as to obtain certain benefits or even avoid costs (Persson and Cedervall, 1991). Apart from the opinion of each individual on the desired goods or services value in comparison to other desirable objects, the amount specified by the respondent also reflects the ability of people to pay. Individuals can judge their own wealth and therefore values and estimates derive from an oriented domination of the consumer. The existing income or wealth distribution is considered acceptable if the amount resulting from the WtP is adjusted by the individual's ability to pay (Persson, 1992).

When analyzing stated preferences in discrete choice situations, one common way is to apply (random parameters) mixed logit models (Brownstone, 2000). One reason for choosing this type of models is to account for unobserved heterogeneity and variations among observations. It is therefore important to apply such a methodology that allows for the influence of variables affecting users' preferences to vary across the sample. This is an important consideration raised by relatively recent research carried out by Brownstone and Train (1999), Train (1999a, 1999b), Revelt and Train (1997, 1999), McFadden and Train (2000), and Bhat (2001). The aforementioned studies have demonstrated the effectiveness of the mixed logit model that can explicitly account for such variations. Therefore, it is suggested that mixed logit models are superior to traditional logit models. Due to the effectiveness of the mixed logit model, it is also widely applied in other fields of transport, as for example in road safety (Gkritza and Mannering, 2008; Ben-Akiva et al., 2007).

In general, relevant literature on the field is very limited since the analysis of the Usage-Based Motor insurance schemes via willingness to pay is a novel subject and has only recently been starting to be explored. Consequently, the present paper aims to add to the current knowledge by being one of the first attempts to identify the parameters that affect users' willingness to pay for usage-based motor insurance, proposing alternative pricing methods such as PAYD and PHYD. More specifically, it is aimed to investigate and provide insight on the understanding of the impact of driving characteristics (driving style and driving needs), drivers' demographics (gender, age, marital status, income, etc.) and the specific characteristics of vehicle insurance premiums on vehicle insurance choice. In order to achieve the aims of the study, a mixed logit model is implemented.

The paper is structured as follows: Section 2 provides an illustration of the sample, the experiment and the choice situations. Section 3 is dedicated to a concise theoretical background of the mixed logit model, whilst Section 4 illustrates and discusses the findings of the models utilized for PAYD and for PHYD. Finally, the last section provides the main conclusions of the study as well as directions for further research.

2. METHODOLOGY

2.1 Discrete Choice Experiment

In order to identify users' preferences and the criteria influencing their choice, the two pricing methods (PAYD-PHYD) were evaluated by respondents using multiple choice and scaled questions. For most questions, a five levels scale was used (1-5) in which the significance of individual factors was evaluated as 1 = "not at all" to 5 = "very much".

The dedicated questionnaire was designed including both revealed preference questions about current vehicle and insurance type, as well as stated preference scenarios related to current and alternative insurance schemes. To increase the number of alternative tested scenarios, two different sheets were designed with four PAYD and eight scenarios PHYD each and each of the 100 respondents answered a single sheet. The questionnaire is structured in 4 sections and questions included:

• general respondent's driving data (years since license was obtained, vehicle make, current insurance cost etc.),
• driving behavior data
• alternative stated preference scenarios about the new insurance premium policies (PAYD and PHYD) and their benefits
• personal - demographic data to draw conclusions about the sample characteristics.

The required time for completion was 10-12 minutes and it was administered to drivers being stopped at a motorist's service station in the Attica region, Greece. The following quoted text was read in each respondent before the administration of the questionnaire:

"In the context of dealing with road accidents, consideration will be given to the future application of an alternative pricing policy based on the use and / or driving behavior of each user, as recorded by a smartphone or an in-vehicle device (On-Board-Diagnostics i.e. OBD). Monitored driving information will be confidentially disclosed to the insurance company that will evaluate the insurance premium annually. Information and further advices will also be provided to the driver via the Internet and / or a smartphone application. These insurance schemes are:

a) based on the use of the vehicle (annual mileage) i.e. the driver will be able to choose a specific annual mileage package based on his needs and pay lower premiums per annum than the current situation if it does not exceed the permitted mileage of the package (Pay-As-You-Drive - PAYD)

b) based on improved driving behavior (lower average speed, lower number of acceleration and braking events etc.) the driver will pay lower premiums (Pay-How-You-Drive - PAHD)

Along with lower premiums and better driving behavior, the driver will have lower accident risk and fuel costs (energy-efficient driving) and potentially additional rewards within the Loyalty Programs (gifts, etc.)."

As for the number of scenarios chosen, it was decided that for the proper implementation of the research the number of scenarios should be reduced. Based on the number of possible values that the variables of the stated preference questionnaire were designed to take, the number of different scenarios results to 16 for PAYD and 80 for PHYD. The number of different combinations in this study was reduced based on an orthogonal design analysis that was implemented, under the assumption that no correlations between typical alternatives exist. Occasionally, in stated preference surveys fractional factorial design can be used instead of full factorial design. Both these designs ensure orthogonality however, the full factorial design would include 16 out of 80 scenarios respectively, in contrast to the fractional comprising (usually much) fewer combinations and are guaranteed to meet some desirable statistical properties, such as the identification and accuracy (Tselentis et al., 2017).

Table 1 summarizes all alternative specific variables used in different scenarios used both for present insurance and the two new insurance schemes, PAYD and PHYD. Present insurance's values were chosen to be zero to facilitate the respondent by not being affected by changes both in new and present insurance schemes.

Regarding the PAYD and PHYD insurance schemes, it should be noted that the respondents were given different scenarios that arose from the orthogonal design in which variables used are in form of percentage reduction. For instance in PAYD schemes, percentage reduction in mileage and percentage reduction in insurance cost are used to counterbalance the reduction in driving distance and cost savings. In other words, respondents were asked to assess how much it would be worth for them to reduce their mileage in exchange for a reduction in their annual insurance fees. The introduction of these variables in this form in the scenarios intends to capture the exact willingness to pay of the respondents i.e. to quantify the percentage reduction drivers are willing to alter their mileage in order to switch to a new insurance scheme. This could not be captured if an absolute minimum mileage value was given in the scenarios tested instead since the most important to take into consideration is the percentage reduction for each respondent and not the absolute value by itself. The latter could not be easily interpreted in the analysis of the stated preference part of the questionnaire where the actual annual mileage of each driver are not taken into consideration. Finally, percentages are preferred over absolute values in order to render feasible the comparison between a) current and future insurance schemes and b) individuals. Regarding all variables used in the questionnaire, respondents were informed that their driving behavior would be recorded during the evaluation period and as a result the user could monitor the value of mileage and speed and therefore adapt his driving habits within the requested limits to gain the respective profit presented in each scenario.

TABLE 1 Descriptive Statistics for Alternative Specific Variables

ALTERNATIVE SPECIFIC VARIABLES	Abbreviation	Mean	St.deviation
PRESENT INSURANCE			
% reduction in mileage (current Insurance)	KM	0.00	0.00
% reduction in Insurance Cost (current Insurance)	COST	0.00	0.00
% reduction in Speed (current Insurance)	SPEED	0.00	0.00
PAYD INSURANCE*			
% reduction in mileage (PAYD Insurance)	KM	11.76	6.58
% reduction in Insurance Cost (PAYD Insurance)	COST	11.69	6.63
PHYD INSURANCE**			
% reduction in mileage (PHYD Insurance)	KM	6.25	9.61
% reduction in Insurance Cost (PHYD Insurance)	COST	11.43	6.78
% reduction in Speed (PHYD Insurance)	SPEED	11.47	6.80

*reduction is compared to the traditional scheme
**reduction is compared to the traditional scheme

As for PAYD and PHYD variables used, percentage change in Mileage allowed to be driven within the insured period and percentage change in Annual Insurance cost were chosen for PAYD as it accounts only for how much you drive. On the other hand, PHYD represents how you drive so percentage change in Average Vehicle Speed variable is also considered in addition to PAYD variables. As illustrated in Table 1, mileage and insurance cost variables for the PAYD scenarios range between -20% and -5% change with a mean and standard deviation of -11.76 % and -11.69% respectively. As for PHYD, mileage, cost and speed variables range between -20 and 5, -20 and -5 and -20 and -5 while their means and standard deviations are -6.25 and 9.61, -11.43 and 6.78 and -11.47 and 6.80 respectively. Generally, in PAYD, mileage and cost reduction intermediate levels used were -5%, -10%, -20% while intermediate levels used for PHYD were -5%, -10%, -20% for cost and speed reduction and +5%, 0%, -10%, -20% for mileage reduction.

The individual variables used in the models are shown in Table 2 and represent gender, age, whether the respondent is using a personal computer and a smartphone owner, the marriage status, income, occupation and education to name them by the order of appearance.

As for the dependent variable, it represents the choice of either present or usage-based Insurance i.e. PAYD/PHYD insurance schemes depending on the scenario answered. The choice of present insurance is represented by 0 whereas by 1 the choice of PAYD/ PHYD insurance.

It should be highlighted that individual variables are defined as all variables that characterize each individual such as age, gender, education etc. whereas alternative-specific variables are those variables that are used in stated preference questionnaire to test how a respondent's choice varies while their values are fluctuating.

The on-site survey took place in a Motorist Service Stations of a motorway in Attica, Greece. The interviews were made during a whole week both on weekdays and the weekend. The interviewers were randomly asking respondents to participate in the survey taking into account only whether or not the respondent is a holder of a valid driving licence as a screening question. No other screening questions such as age, years of active driving etc. were asked since the researchers' intention was to include younger drivers into the survey as well.

Regarding the sample characteristics, 100 respondents participated in the survey of which 45% were women, 53% married, 98% makes use of a PC and 78% is a smartphone owner. All individual specific variables tested in models developed are summarized in Table 2 along with their abbreviation and a few descriptive statistics such as mean, standard deviation, min and max values. The most important highlights are that:

• The majority of respondents were between 30-50 years old. That is also illustrated in figure 1 where it is shown how gender is distributed by age category. As it appears, 43% and 28% belong to the age category of 30-40 and 40-50 respectively.
• Most respondents' income was between 10,000 and 25,000 Euros.
• 45% was working in the public sector whereas 40% in private sector.
• 72% had pursued a degree after school.

Considering the sample characteristics illustrated in Table 2, one major remark is that the sample taken is a representative sample of the current motor insurance customer population. According to HMITN, the Greek population of drivers is similar to the one collected for the purpose of this research with a slight emphasis given on middle-age and younger drivers who form the future of motor insurance market in Greece. It has to be highlighted that the conducted research within this paper is aiming to identify the willingness to pay for alternative insurance schemes that do not exist in Greece at the moment but will probably exist in a decade. Therefore, it was considered preferable to administer the questionnaire to a less percentage of people whose age is more than 60 than the representative percentage of the Greek population of drivers. It should be noted that all respondents were at that moment insured with traditional motor insurance schemes and not with the new ones.

As for the definition of the variables, respondents were asked in the questionnaire to specify their main occupation i.e. the most profitable one for them as well as their level of education clarifying that this is the higher degree they hold. For both variables, only one answer is accepted so that they can be treated as categorical variables in the implemented analysis.

TABLE 2 Descriptive Statistics for Individual Specific Variables

INDIVIDUAL SPECIFIC VARIABLES	Abbreviation	Frequency
Gender = Female	GENDER_F	45
Age: 18-30 (reference category)	AGE1,2	11
Age: 30-40	AGE3	43
Age: 40-50	AGE4	28
Age: >50	AGE5	18
PC usage=yes	USAGE_PC	98
Smartphone Owner	SMARTPHONE	78
Married	MARRIED	53
Income <=10000 (reference category)	INCOME1	6
10000 < Income <= 25000	INCOME2	54
Income > 25000	INCOME3	40
Occupation: Public Sector	OCCU1	45
Occupation: Private Sector	OCCU2	24
Occupation: University Student	OCCU3	3
Occupation: Freelancer	OCCU4	9
Occupation: Enterpreneur	OCCU5	3
Occupation: Household	OCCU6	2
Occupation: Technician	OCCU7	0
Occupation: Pensioner (reference category)	OCCU8	7
Occupation: Unemployed	OCCU9	2
Occupation: Other	OCCU10	5
Education: Primary Education	EDU1	2
Education: Secondary Education (reference category)	EDU2	24
Education:Technological Educational Institute	EDU3	33
Education: University Degree	EDU4	11
Education: Postgraduate Degree	EDU5	24
Education: Ph.D.	EDU6	3
Education: Other	EDU7	3

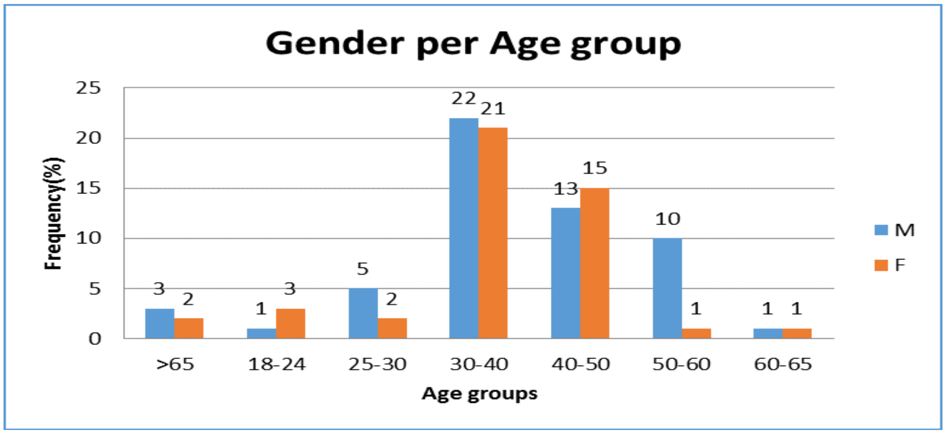

FIGURE 2 Gender distribution per age group.

When the preference on new motor insurance schemes is considered, (Figures 2 and 3), it is observed that the majority of the respondents are willing to switch to a new insurance policy. More specifically, in all education categories people seem to prefer a transition to UBI except from people with secondary education. The same applies to all age categories except for people between 50-60 years old, who answered that they would not switch to a usage-based insurance scheme.

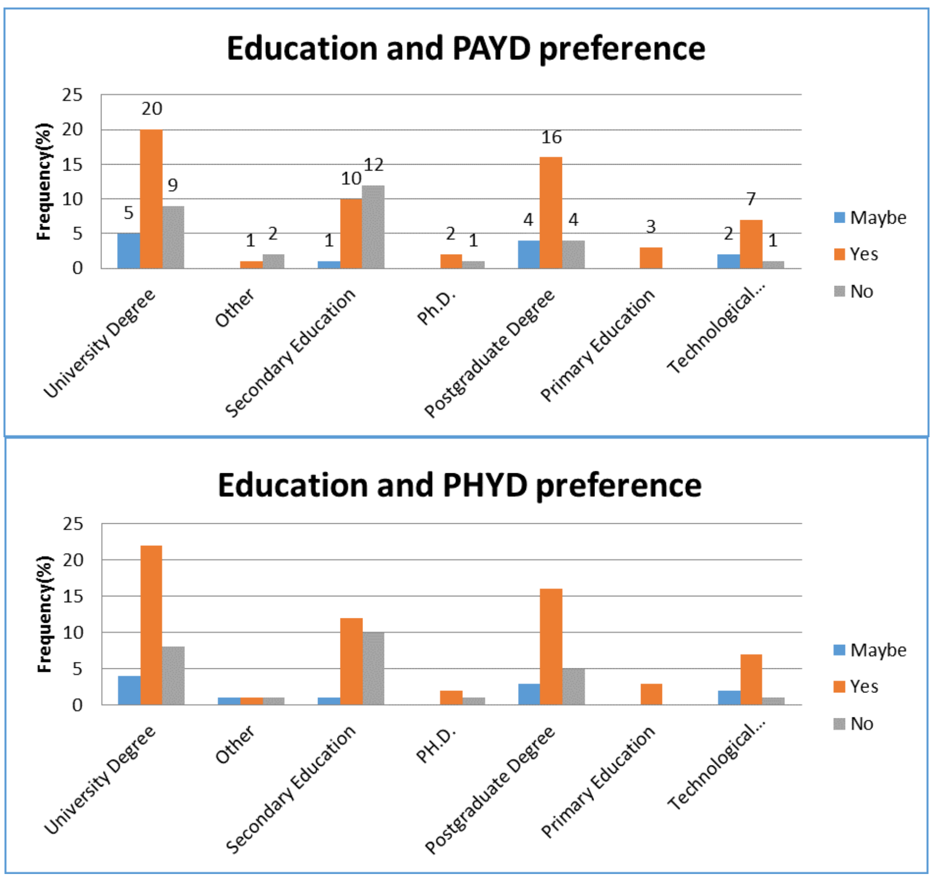

FIGURE 3 PAYD and PHYD preference distribution per education group.

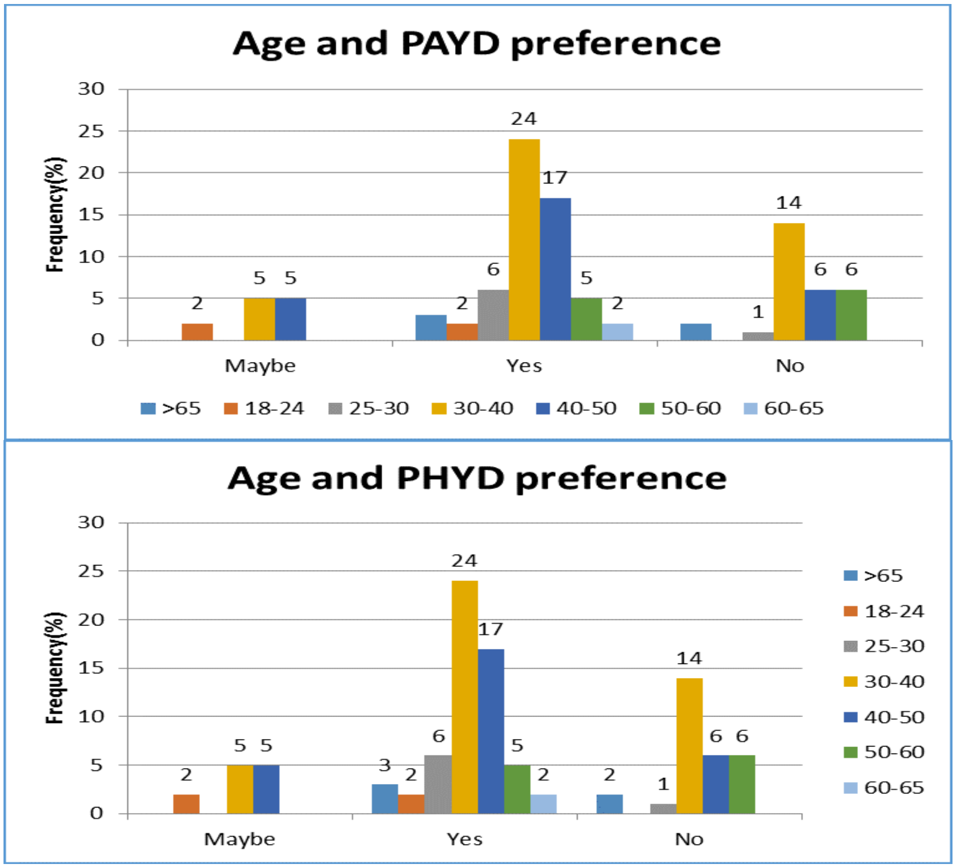

FIGURE 4 PAYD and PHYD preference distribution per age group.

2.2 Mixed Logit Models

The proposed methodology in order to analyze the stated preference questionnaire regarding Pay As You Drive (PAYD) and Pay How You Drive (PHYD) is the mixed logit model (random parameter model). Since the alternatives for each insurance scheme are two (the present insurance versus PAYD and present insurance versus PHYD), the binary logistic (fixed effects) model is initially considered appropriate.

However, the traditional fixed effects modeling approaches treat parameters as constant (fixed) across observations, meaning that the effect of any individual explanatory variable is the same for each observation or individual (Moore et al., 2011). Therefore, to account for unobserved heterogeneity, random-parameter models are applied assuming that the estimated parameters vary across observations. Train (1999) and Ben-Akiva et al. (2007) consider this model as a highly flexible model that can account for the standard logit limitations and at the same time allows for random variation across observations. In these models some parameters can be held fixed across observations while others are allowed to be random and follow a distribution (e.g. normal, lognormal, uniform, etc.).

Following Ben-Akiva et al. (2007) and Train (2009), a function determining discrete outcome probabilities is considered:

$$T_{in} = \beta_i X_{in} + \varepsilon_{in} \tag{1}$$

A mixed logit model is any model whose choice probabilities can be expressed in the form:

$$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta \tag{2}$$

where $L_{ni}(\beta)$ is the logit probability evaluated at parameters β:

$$L_{ni}(\beta) = \frac{e^{V_{ni(\beta)}}}{\sum_{j=1}^{J} e^{V_{nj(\beta)}}} \quad (3)$$

f(β) is a density function, $V_{ni}(\beta)$ is the observed portion of the utility, which depends on the parameters β. If utility is linear in β, then

$$V_{ni}(\beta) = \beta' x_{ni} \quad (4)$$

Then, the mixed logit probability takes the usual form:

$$P_{ni} = \int \left(\frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}}\right) f(\beta) d\beta \quad (5)$$

Mixed logit is a mixture of the logit function evaluated at different β's with f (β) as the mixing distribution. Estimation of the mixed logit model takes place by using simulation methods due to the difficulty in computing probabilities. More details about the mixed logit model can be found in Washington et al. (2003). Train (2009), provides a review of sampling techniques, but one of the most popular technique is considered to be the Halton draws (Washington et al., 2003), which were proposed by Halton (1960).

4. RESULTS

In this paper two distinct mixed logit models were developed; one mixed logit model in order to investigate which factors affect the choice of present insurance policy versus PAYD and one mixed logit model for present insurance policy versus PHYD. A common issue when fitting mixed logit models is the determination of which parameters should be random and which should be fixed (Moore et al., 2011). Moore et al. (2011) suggest starting with all possible independent variables and then gradually reduce them. For that reason, many different trials were conducted.

The next two subsections illustrate the proposed mixed logit models. In these models, 200 Halton draws were used. The parameters which were found to be random, were those whose standard deviations differ significantly from zero as Train (2009) and Milton et al. (2008) suggest. On the other hand, parameters whose standard deviations are not 95% statistically significant are considered as fixed across observations. It is noted that proposed random parameters followed the normal distribution. In order to present the performance of the model, goodness-of-fit measures such as log-likelihood and McFadden R^2 are calculated.

4.1. Pay As You Drive Scheme (PAYD)

The final model for the PAYD scheme is presented on Table 3. The model has an adequate fit in terms of likelihood ratio test (log-likelihood of empty model versus log-likelihood of the full model) and also McFadden R^2. More specifically, the likelihood ratio test was 61.19, and the McFadden R^2 was 0.212 indicating a reasonable fit of the model.

The variable "Km" and the variable "Cost" (which are alternative specific variables) as well as the constant term, were set to be random following the normal distribution across observations. However, only the standard deviation of the Km and the constant term were ultimately found to be statistically different from zero. Therefore, the cost variable is considered to be fixed across observations. The variable Km was found to have a mean value of 0.219 and a standard deviation 0. Therefore:

$$Z = \frac{0-0.219}{0.122} = -1.795.$$

According to the Z score table and the normal distribution function about 3% of observations are lower than zero. This means that in about 97% of observations, Km is associated with increased likelihood of selecting PAYD while only 3% of observations show a negative correlation. Therefore, in the vast majority of cases, it can be concluded that as the offered percentage reduction in driven mileage decreases, it is more likely that the drivers choose the PAYD policy.

Similarly, the constant term has a Z score of 1.044 (mean value = -1.179, s.d. = 1.129) means that about 86% percent of observations have a negative constant term.

The cost parameter were considered as fixed, therefore the negative sign of the beta coefficient (-0.158) denotes that as the cost reduction is lower, drivers are more likely to choose the present insurance. This happens because this variable expresses the percentage reduction in cost offered by the PAYD scheme. Therefore, a positive sign expresses an increase in offered cost reduction in PAYD scheme.

The positive value of the coefficient of SMARTPHONE variable (0.668), denotes that drivers who are more familiar with smartphones usage are more likely to choose PAYD scheme rather than the present insurance policy at a 90% level of confidence. The odds ratio was 1.95 meaning that drivers who are familiar with smartphones are about twice as likely to choose the PAYD scheme, than those who are not familiar with smartphones. Therefore, familiarity with technology is positively associated with acceptance of new alternative insurance policies as expected.

TABLE 3 Mixed logit Model Estimates (PAYD)

Variables	Estimate	Standard error	p-value	Conclusion	Odds ratio
Random parameters (normal distribution)					
Constant term	-1.179	0.529	0.026	95% significant	0.308
Standard deviation of constant term	1.129	0.491	0.022	95% significant	3.093
Km	0.219	0.051	<0.001	95% significant	1.245
Standard deviation of Km	0.122	0.045	0.006	95% significant	1.130
Cost*	-0.158	0.032	<0.001	95% significant	0.854
Standard deviation of Cost	-	-	-	non-significant	-
Fixed parameters					
SMARTPHONE	0.668	0.403	0.097	90% significant	1.950
Log-likelihood of the empty model	-259.279				
Log-likelihood of the full model	-203.500				
McFadden's pseudo R^2	0.212				

Cost variable was entered as fixed variable

4.2. Pay As How You Drive (PHYD)

The final model for the PHYD scheme is presented on Table 4. The model has an adequate fit in terms of likelihood ratio test values (log-likelihood of empty model versus log-likelihood of the full model) and values of McFadden R^2.

In this model, the constant term as well as the variables “Km”, “Cost” and “Speed” were set as random variables and be normally distributed. More specifically, Km has a mean value of 0.114 and a standard deviation of 0.061, Cost has a mean of -0.179 and standard deviation 0.065, while Speed has a mean value of 0.091 and 0.077. On the other hand, the constant term was found to have a mean value of -1.789 and standard deviation 1.197.

The interpretation of the random parameters is similar to the previous model by calculating the Z-scores and use the Z-tables, since all random parameters were normally distributed. Concerning Km, the calculated Z-values indicate that 97% of observations have a positive correlation with PHYD meaning that as the percentage change in km, tends from negative to zero (reduction is lower) the probability of selection of PHYD increases. Change in speed (variable Speed) has a similar interpretation, and results indicate that about 11% of observations have a negative association with PHYD while 89% have a positive association with PHYD. The mean value of the beta coefficient was found to be 0.091. This means that as the percentage reduction in speed tends to zero, the driver is more likely to choose the PHYD policy scheme.

On the contrary, variable Cost has a negative mean value as in the previous model, indicating that the percentage reduction in cost tends to be zero, the present policy is more probable to be selected by drivers. This is also supported by the Z score which indicates that about 99.7% of observations show a negative correlation of cost and PHYD.

The interpretation of the fixed parameters in this model is straightforward to a similar manner to the previous model. Age was found to be statistically significant for the PHYD scheme having an expected effect. More specifically, the beta coefficients of AGE4 and AGE5 have negative signs, indicating that drivers 40-50 years old and older than 50 years old are more likely to prefer the present insurance policy compared with younger drivers. More specifically, young drivers are almost 2.5 times and almost 3 times more probable to choose the PHYD policy, compared to drivers 40-50 years old and older than 50 years old respectively. Familiarity with smartphone use was found to be significant and expected, similar to the PAYD model. Its beta coefficient was 0.627, indicating that familiarity with smartphone and applications suggests high probability for drivers choose the PHYD scheme (similarly to the PAYD) compared to the present policy. In other words, the probability of PHYD selection by users familiar with smartphone use is 1.872 times higher than those who report low familiarity. Lastly, the beta coefficient of gender shows that female drivers would prefer the PHYD compared to male drivers (2.731 more likely than males). Therefore, female drivers are more willing to turn to new insurance policies in contrast to male drivers who are more tentative and prefer the traditional insurance policies. This could be attributed to the fact that female drivers are probably driving less frequently and are less likely to excess speed. Therefore, they would benefit from such alternative insurance policies.

TABLE 4 Mixed logit Model Estimates (PHYD)

Variables	Estimate	Standard error	p-value	Conclusion	Odds ratio
Random parameters (normal distribution)					
Constant term	-1.789	0.429	0.000	95% significant	0.167
Standard deviation of constant term	1.197	0.270	0.000	95% significant	-
Km	0.114	0.017	0.000	95% significant	1.121
Standard deviation of Km	0.061	0.027	0.022	95% significant	-
Cost	-0.179	0.025	0.000	95% significant	0.836
Standard deviation of Cost	0.065	0.025	0.009	95% significant	-
Speed	0.091	0.020	0.000	95% significant	1.095
Standard deviation of Speed	0.077	0.022	0.001	95% significant	-
Fixed parameters					
AGE4	-0.846	0.274	0.002	95% significant	0.429
AGE5	-1.176	0.433	0.007	95% significant	0.309
SMARTPHONE	0.627	0.309	0.042	95% significant	1.872
GENDER_F	1.005	0.244	0.000	95% significant	2.731
Log-likelihood of the empty model	513.250				
Log-likelihood of the full model	-416.500				
McFadden's pseudo R^2	0.216				

5. CONCLUSIONS

Within this paper, a methodological approach is proposed to identify the parameters that affect users' willingness to pay for alternative usage-based motor insurance pricing schemes such as PAYD and PHYD. Firstly, a dedicated questionnaire was designed and distributed to a random but representative sample of 100 participants in Attica region in Greece. In this questionnaire, specific scenarios were constructed in order to disclose respondents' preference towards insurance pricing schemes. It also included both revealed and stated preference questions regarding current and alternative insurance schemes.

The statistical analysis of the study consists of mixed logit models which are applied a) to account for unobserved heterogeneity and b) to assist in the better understanding of the effect of driving characteristics, drivers' demographics and the characteristics of vehicle insurance premiums on vehicle insurance choice. More specifically, two distinct mixed logit models were developed; one mixed logit model to investigate the factors influencing the choice of present insurance policy over PAYD and one for present insurance policy over PHYD.

To the best of our knowledge, the present study adds to current knowledge, as it is one of the very first times that a discrete choice experiment towards insurance policies is carried out. This is the core contribution of the study. Results indicated that female drivers and smartphone owners are more likely to choose a new insurance scheme as they are more familiar with new technologies. Kilometers and cost reduction were also found to affect the choice for both UBIs in a similar manner, i.e. the higher the kilometers reduction the lower the probability of the UBI scheme to be chosen and the higher the cost reduction the higher the probability of the UBI scheme to be chosen by a user. Moreover, as the speed reduction imposed to the user increases, the probability of choosing UBI scheme is reduced.

It was also found that people over 40 years old are less likely to choose PHYD insurance which is supported by descriptive statistics described in Data section. This is something expected, since older drivers show more familiarity with present insurance schemes.

Future research could extended by carrying out surveys in different countries and perhaps set up different scenarios, perhaps also including more parameters. Lastly, alternative models to account for unobserved heterogeneity could be utilized, for example the latent class model.

REFERENCES

Ben-Akiva M., D. McFadden, K. Train, J. Walker, C. Bhat, M, Bierlaire, A. Boersch-Supan, D. Brownstone, D. Bunch, A. Daly, A. DePalma, D. Gopinath, A. Karlstrom, and M. Munizaga. *Hybrid Choice Models: Progress and Challenges*. In: Y. Shiftan, K. Button, and P. Nijkamp, eds. Transportation Planning. Surrey, UK: Edward Elgar Publishing, Volume 13, Issue 3, 2007, pp. 163–75. (re-published: original first published in Marketing Letters, 2002).

Bhat C. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B*, Vol. 17, Issue 1, 2001, pp. 677–693.

Bolderdijk, J.W., Knockaert, J., Steg, E.M., Verhoef, E.T. (2011). Effects of Pay-As-You-Drive vehicle insurance on young drivers' speed choice: results of a Dutch field experiment. Accident Analysis and Prevention 43, 1181–1186.

Bordoff, J., & Noel, P. (2008). Pay-as-you-drive auto insurance: A simple way to reduce driving-related harms and increase equity. Hamilton Project Discussion Paper.

Brownstone D. and K. Train K. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, Vol. 89, 1999, pp. 109–129.

Brownstone D., D.S. Bunch and K. Train. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research Part B*, Vol. 34, Issue 5, 2000, pp. 315–338.

Edlin, Aaron S. 2003. "Per Mile Premiums for Auto Insurance." In Economics for an Imperfect World: Essays In Honor of Joseph Stiglitz, ed. Richard Arnott, Bruce Greenwald, Ravi Kanbur, and Barry Nalebuff. Cambridge: MIT Press.

Ferreira Jr, J., & Minikel, E. (2010). Pay-As-You-Drive Auto Insurance In Massachusetts: A Risk Assessment And Report On Consumer. Industry And Environmental Benefits, by the Department of Urban.

Gkritza K. and F.L. Mannering. Mixed logit analysis of safety-belt use in single- and multi-occupant vehicles. *Accident Analysis and Prevention*, Vol. 40, 2008, pp. 443–451.

Halton J. On the efficiency of evaluating certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, Vol. 2, pp. 84–90.

ICBC Research Services Data, 1998. Insurance Corporation of British Columbia.

Litman, T., 2004a. Pay-As-You-Drive Vehicle Insurance Converting Vehicle Insurance Premiums Into Use-Based Charges. Victoria Transport Policy Institute, Victoria.

Litman, T., 2004b. Pay-As-You-Drive Pricing for Insurance Affordability. Victoria Transport Policy Institute (www. vtpi.org)

Litman, T. (2005). Pay-as-you-drive pricing and insurance regulatory objectives. Journal of Insurance Regulation, 23(3), 35.

McFadden D. and K. Train (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, Vol. 15, 2000, pp. 447–470.

Milton J., V. Shankar V. and F. Mannering F. Highway accident severities and the mixed logit model: an exploratory empirical analysis. *Accident Analysis and Prevention*, Vol. 40, Issue 1, 2008, pp. 26–266.

Moore D., W. Shneider IV., P. Savolainen P. and M. Farzaneh. Mixed logit analysis of bicyclist injury severity resulting from motor vehicle crashes at intersection and non-intersection locations. *Accident Analysis and Prevention*, Vol. 43, Issue 3, 2011, pp. 621–630.

Persson U. and M. Cedervall. The Value of Risk Reduction: Results of a Swedish Sample Survey. *The Swedish Institute for Health Economics*, 1991, working paper.

Persson U. Three economic approaches to valuing benefits of safety measures, 1992, Lund University.

Revelt D. and K. Train. Customer specific taste parameters and mixed logit. *Working Paper*, 1999, University of California, Berkeley, Department of Economics.

Revelt D. and K. Train. Mixed logit with repeated choices: households' choice of appliance efficiency level. *Review of Economics and Statistics*, Vol. 80, Issue 4, 1997, pp. 647–657.

Sugarman, S.D., 1994. PAY-AT-THE-PUMP Auto Insurance: The California Vehicle Injury Plan (VIP) for Better Compensation, Fairer Funding, and Greater Safety Prepared for The Institute of Governmental Studies. University of California, Berkeley, California

Toledo, T., Musicant, O., Lotan, T. (2008). In-vehicle data recorders for monitoring and feedback on drivers' behavior. Transportation Research Part C, 16, 320-331.

Train K. Discrete choice methods with simulation (2nd ed.). Cambridge University Press Publication, 2009, Cambridge.

Train K. Halton sequences for mixed logit. *Working Paper*, University of California Berkley, Department of Economics, 1999b.

Train K. *Mixed logit models for recreation demand*. In: Kling, C., Herriges, J. (Eds.), Valuing the Environment Using Recreation Demand Models. Elgar Press, 1999a, New York.

Tselentis, D. I., Yannis, G., & Vlahogianni, E. I. (2017). Innovative motor insurance schemes: a review of current practices and emerging challenges. Accident Analysis & Prevention, 98, 139-148.

Washington S.P., M.G. Karlaftis and F.L. Mannering. *Statistical and econometric methods for transportation data analysis.* CRC Press LLC, 2003, Boca Raton, FL.

Wenzel, T. (1995). Analysis of national pay-as-you-drive insurance systems and other variable driving charges (No. LBL-37321). Energy Analysis Program, Energy and Environment Division, Lawrence Berkeley National Laboratory, University of California.

Hellenic Ministry of Infrastructure, Transport and Networks: http://www.yme.gr