

The Effect of the Introduction of a “Pay Per Use” Option within Motor TPL Insurance

Stefan Trappl¹

*FHWien University of Applied Sciences of WKW, Austria
stefan.trappl@fh-wien.ac.at*

Karl Zehetner

*FHWien University of Applied Sciences of WKW, Austria
karl.zehetner@fh-wien.ac.at*

Robert Pichler

*FHWien University of Applied Sciences of WKW, Austria
robert.pichler@fh-wien.ac.at*

Received: 17 September 2013 / Revised: 25 October 2013 / Accepted: 21 March 2014 / Published online: 19 May 2014

ABSTRACT

In this paper the effects of the introduction of the so called “pay per use” -insurance products are examined. These products collect data of mileage of kilometers driven by policy holders. As a result of this data, policy holders can get a refund on the insurance-premium paid. Since there is a positive correlation between mileage and the risk of causing an accident the refund is granted to low-mileage drivers, so in theory the “pay per use” product is more attractive to low-mileage drivers than to long-distance drivers. The authors examine empirical evidence to find out whether or not it is mainly low-mileage-drivers who choose the “pay per use” product. Secondly, the authors examine whether there are other significant differences between characteristics of “pay per use” policy-holders and “traditional” policy-holders. Therefore a random sample of 4,000 car-insurance – clients (2,000 “pay per use” policy-holders and 2,000 “traditional” policy-holders) is reviewed. In addition the effects of the introduction of “pay per use” products are discussed, in case of a selection effect between low- and high -mileage drivers is observed.

JEL classification: D82, C12, G22

Keywords: insurance, pay per use, pay as you drive, adverse selection, selection effects

1. INTRODUCTION

According to the principle of individual equivalence it is very important for insurance companies to charge insurance premiums which are adequate to the individual risk of the persons insured. If insurance companies do not find risk factors to divide their clients according to risk-classes, they have to charge an equal premium to all their policy-holders (risk pooling). Typically

¹ Corresponding author: Stefan Trappl, Research Associate, FHWien University of Applied Sciences of WKW, 1180 Vienna, Währinger Gürtel 97, Austria; e-mail: stefan.trappl@fh-wien.ac.at; phone: +43 1 476 77-5817.

in this case the problems of asymmetric information, primarily adverse selection, appear. The problem of adverse selection was first described by George Akerlof (1970). A few years later Rothschild and Stiglitz (1976) provided a theoretical framework towards adverse selection. They came up with an opportunity for insurance companies to set incentives that policy holders choose an insurance contract in accordance with their individual risk.

It is very difficult for insurance companies to find risk factors for third-party motor-vehicle liability insurance policies defining the individual risk of their policy holders.

At the beginning of the automobile-era the insurance-premium depended only on the braking distance of the insured car. Mehring (1962) recommended a differentiation between car-related and individual-related risk factors. Indeed, most insurance companies use specifications of the car and characteristics of the policy holder but also a merit-rating system to calculate the individual insurance premium. However the calculation of individual insurance premiums is not very successful and the principle of individual equivalence is not strictly adhered to.

Since 2000, insurance companies have been trying to establish modern car-insurance products. So-called “pay per use” or “pay as you drive”, usage-based insurance products have been established in the insurance market and these products are meanwhile offered in almost all developed countries in the world. Via GPS, usage-based insurance-products record several types of information about mileage, road type and travelling time. With this information policy-holders signal to the insurance companies their driving behavior. Consequently this information is used to adjust individual insurance premiums in relation to the individual risk of a policy-holder so that the premiums are more in accordance with the principle of individual equivalence.

The aim of this paper is to ascertain if the introduction of usage-based insurance products has an impact on problems of asymmetric information. The main purpose of this paper is to ascertain empirically if the choice of a certain product (“pay per use”, vs. “traditional”) is influenced by specific criterions. The authors compared a “traditional” sample with a “pay per use” sample, to find out if there is a significant difference in the “pay per use”-risk-factor mileage. Secondly, the authors examined whether there are differences in the “traditional” risk factors like engine power or the age or the sex of policy-holders.

According to theory, the “pay per use” product should be chosen primarily by the low-mileage driver, while the high-mileage driver should avoid the “pay per use” product, because the insurance premium depends on mileage. On the other hand there should be low differences in the “traditional” risk factors between the two samples.

1.1. A brief overview of the third-party liability insurance market in Austria

There are about 4,250,000 passenger-cars in Austria. Consequently the mandatory third-party liability insurance is one of the most important insurance products in Austria. The gross volume of €1.8 billion, which is about 11.3% of all insurances contracted over all classes (Statistik Austria, 2007). Because of the deregulation of the insurance market in the 1990s and of the rising competition, many insurance companies tried to increase their market shares by offering generous premium discounts to new clients. For example: many new clients were placed in the best merit-rating system level (OEAMTC, 2009). Consequently the loss ratio of insurance companies rose to 86% in the year 2000 (VVO, 2004). Which means €86 € out of €100 was paid as compensation to third parties. A mere €14 was available to cover business operating expenses. As a result, nearly all insurance companies made a loss in the third- party liability insurance class. Thus most insurance companies tried to create alternative products, or other marketing strategies.

1.2. Premium differentiation

1.2.1. *Traditional third party liability insurance products*

The question about the fair premium of insurance products is as old as insurance itself. According to the principle of individual equivalence, every insured should pay a premium according to his or her individual risk. This is very important due to the problems of asymmetric information, especially in terms of adverse selection (Karten, 1994). For third-party liability insurance it is very difficult to define risk factors. There are several established individual-related risk factors such as age, sex, or residency (Mehring, 1962). Also car-related risk factors such as engine power, age of the car or its make (Mehring, 1962) are taken in consideration.

Another instrument for premium differentiation is the so called merit-rating system (bonus-malus system). If an accident happened in the past, a penalty for a bad loss-experience is charged. On the other hand, if no accident occurs a better insurance level class could be achieved. In fact without an accident, one level per year could be gained, and three levels could be lost in case of an accident. This instrument should divide individuals with a high loss-probability (“high risks”) from individuals with a low loss-probability (“low risks”) (Karten, 1994). Since 1994, the merit-rating system is no longer laid down in Austrian law (Verordnung des BMF, 1994). However most insurance companies still use this system, but as mentioned, they are often very generous in the classification of new clients.

As a consequence premium differentiation for “traditional” insurance products can be very difficult and often does not lead to the desired affect that policy holders pay a premium according to their individual risk.

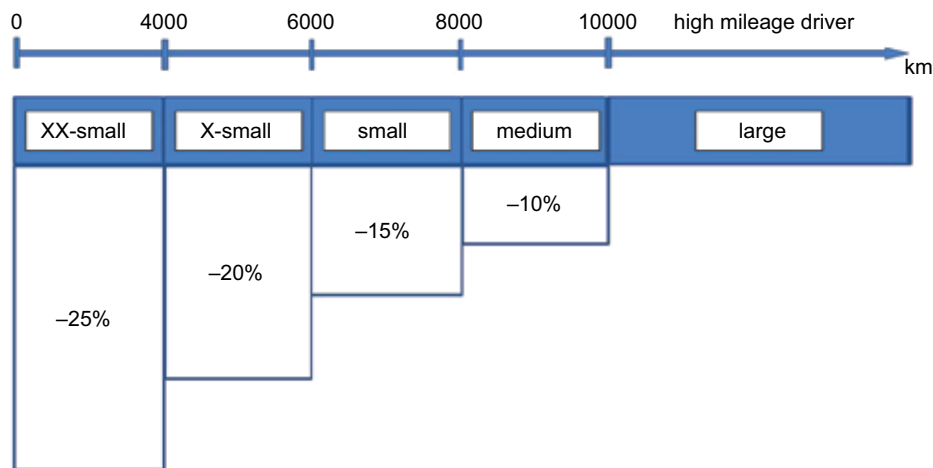
1.2.2. *The development of “pay per use”-insurance products*

Insurance companies face challenges with “traditional” insurance products as premiums to risk factors such as kilometers travelled (mileage), road type or average speed cannot be taken into account. But it is obvious that these relevant risk factors need to be considered. If a car is used only twice a year, there is a lower risk of causing an accident, than when hundreds of kilometers are driven per day.

As a consequence, in the 1990s the first theoretical ideas concerning “pay per use” insurance products were taken into consideration (Litman, 1997). Very important for the development of “pay per use” insurances was the advance of the “global positioning systems” (GPS), as almost all “pay per use” insurance products use GPS to keep track of the insured cars. In 2000 the American “Progressive Insurance Company” established the first usage-based car insurance product. This company also registered the widely used term “pay as you drive” as a trade mark. As a consequence the terms “pay per use insurance” or “usage-based insurance” are used to describe this sort of insurance product.

UNIQA is the only insurance company in Austria which offers a “pay per use” insurance product. It is called “Safe Line”. If you choose the product “Safe Line” you get a motor vehicle third-party liability insurance but, in addition, a black box is installed in your car. This black box records mileage and road type driven. These two additional risk factors are considered in the calculation of insurance premiums as follows: you get a premium reduction for the third-party liability insurance if you drive fewer than 10,000 km per year. Road type is also used as a risk factor. As driving on highways is statistically not as dangerous as driving on rural or urban roads, every kilometer driven on highways counts only for 0.8 real kilometers (UNIQA, 2012).

Figure 1
Premium Reduction Safe Line



Source: UNIQA (2009).

2. METHODS AND EMPIRICAL WORK

According to the descriptions above the research questions are:

- Is there a significant difference in the parameter values of the risk factors between “traditional” policy-holders and “pay per use” policy-holders?
- Do people with certain characteristics significantly more often choose the usage-based insurance contracts?

To answer these questions the authors analyzed a data set with information on 4,000 customers of the UNIQA insurance company. Of these, 2,000 customers chose a “traditional” insurance contract and 2,000 chose the “usage-based” insurance product “Safe Line”. From the research question mentioned above, the authors derived the following hypotheses:

- H_0 = there is no difference in the mileage per year between “pay per use” policy-holders and “traditional” policy-holders.
- H_0 = there is no difference in the other risk factors between “pay per use” policy-holders and “traditional” policy-holders.

The data set included information to the following risk factors:

- individual-related risk factors: age of the policy-holder, residency of the policy-holder, gender of the policy-holder,
- car-related risk factors: engine power, year of construction, conclusion of a Casco insurance, level in the merit-rating system, minimum guarantee sum,
- usage-based risk factors, mileage driven within one year. (These data are only available for the “pay per use” insurance contracts).

2.1. Applied statistical methods

As the mileage-data are only available for “pay per use” policy holders, the authors had to use two different statistical methods.

2.1.1. Logistic regression

Primarily, differences between “traditional” policy-holders and “pay per use” policy-holders with reference to the risk factors “age of the policy-holder”, “place of domicile of the policy-holder”, “gender of the policy-holder”, “engine power of the insured car”, “year of construction”, “conclusion of a Casco insurance”, “level in the merit-rating system”, and “minimum guarantee sum” are examined. For these risk factors data from all policy-holders are considered. As mentioned above the authors examine whether people with certain characteristics (= parameter values of the risk factors) typically prefer one or other of the two insurance contracts mentioned. Or, in other words: do certain characteristics significantly influence the decision whether a person choose a usage-based insurance or traditional insurance.

Questions in this vein are handled with diverse regression-analysis-methods. As the dependent variable has only two possible parameter values [$y = 1$ or $y = 0$], a binary regression method is applied. $y = 1$ means the event occurs: the usage-based insurance product is bought. $y = 0$ means the event doesn't occur: the usage-based insurance is not bought. The authors have chosen the logistic regression as it is a very robust statistical method and it is linked with fewer assumptions than other methods (Backhaus et. al., 2011).

The linear regression:

$$y = \alpha + \beta_1 * X_1 + \dots + \beta_j * X_j + \varepsilon \quad \text{Eq. (1)}$$

z = dependent variable

α = constant term

β_j = regression coefficient

ε = disturbance variable

In the case that the dependent variable y can only have two possible parameter values, the linear regression is not compatible. Due to linearity the model could deliver parameter values beyond the defined area. That means although only parameter values between 0 and 1 are reasonable, the model could deliver values between $-\infty$ and $+\infty$.

To handle this problem the logistic-regression transforms the linear regression-function into an s-shaped function. This s-shaped function can be interpreted as a probability-function.

To calculate the probability of the occurrence of event $y = 1$ $P(y = 1)$ variable z needs to be created.

$$y_k = 1 \quad \text{if } z > 0 \quad \text{Eq. (2)}$$

$$y_k = 0 \quad \text{if } z < 0 \quad \text{Eq. (3)}$$

$$z = \alpha + \beta_1 * X_1 + \dots + \beta_j * X_j + \varepsilon \quad \text{Eq. (4)}$$

The linear regression delivers z -values. However these z -values cannot be interpreted as probability values yet. Therefore the linear function needs to be transformed into a none-linear form.

In the first step, the exponential-function is applied on the linear regression-function:

$$p_1 = e^z \quad \text{Eq. (5)}$$

p_1 = first transformation – step to the probability – function

e = Euler's number

z = linear regression function.

The new function is limited downwards and can never take negative values. (The z-values still can have negative values).

In the second step, the function is limited upwards. Therefore the new function is divided by one plus the new function.

$$p = \frac{e^z}{1 + e^z} \quad \text{Eq. (6)}$$

p_1 = probability – function.

This function p is called the logistic function. With this function the occurrence-probability of buying the usage-based insurance product $P(y = 1)$ can be calculated.

The finalized regression-approach can be demonstrated as follows:

$$p(y = 1) = \frac{e^z}{1 + e^z} \quad \text{Eq. (7)}$$

$$z = \alpha + \beta_1 * X_1 + \dots + \beta_j * X_j + \varepsilon$$

The z-values are called logits. As the logits are the results of the linear regression, they can take values between $-\infty$ und $+\infty$. The linear regression is transformed into an s-shaped function therefore the β -Values cannot be interpreted as it is usually done in the linear regression. (An increase in the X_i -Value by 1, increases the event-probability by the β -value).

To interpret the regression-coefficients β , odds are used. Odds can be derived from the logistic regression-function:

$$p(y = 1) = \frac{e^z}{1 + e^z} \quad \text{Eq. (8)}$$

$$e^z = \frac{p(y = 1)}{(1 - p(y = 1))} \quad \text{Eq. (9)}$$

This formula is called the odds ratio. With the odds ratio the regression-coefficients can be interpreted as follows: An increase in the β -value by 1 increases the odds ratio (in favor of the event $y = 1$) by e^β .

2.1.2. One-sample t-test

Another focus is to examine whether there is a difference between average mileage of “traditional” policy-holders and “pay per use” policy-holders. As only the mileage-data from the “pay per use”-customers are known, the mileage of these people cannot be compared with the mileage of the “traditional” policy-holders. However the “pay per use” mileage-data can be contrasted with the average mileage of Austrian drivers.

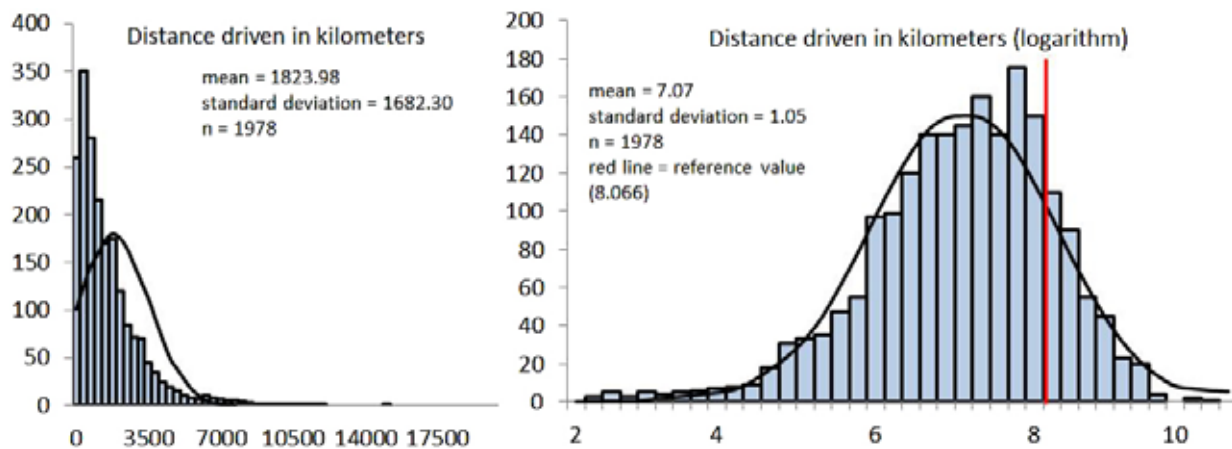
As the “pay per use” mileage data is available for three months, the mileage of an average Austrian car driven within three months needs to be calculated. Per year an Austrian car is driven 13,497 km on the average (VCÖ, 2010). During the relevant three months 23.6% of the annual car rides are undertaken (Asfinag, 2010). So during the comparison period $13,497 * 0.236 = 3,185.29$ km were driven with an average Austrian car.

According to this, the random sample of “pay per use” customers is compared with the arithmetical mean of the kilometers driven by an average policy-holder. In a next step, the data-set is tested for normality of the distribution to see if a normal distribution exists. As the data can only have positive values the natural logarithm is applied to the data to get a symmetric distribution. As you can see in Figure 2, the logarithmic data is optical alike a normal distribution.

Although the Kolmogorov-Smirnov-test and the Shapiro-Wilk test demonstrates, that there is no perfect normal-distribution, the one-sample t-test is applied, because the test is very robust against violations of the assumption of normal distribution (Kähler, 2011).

Figure 2

Data distribution and data log-distribution



Source: Authors' calculation.

2.1. Analysis of the quantity and quality of the data set

2.1.1. Quantity analysis

The analysis of the quantity of the data set entails the calculation of the effect size of the statistical methods. The sample contains of 4,000 data-sets which can be taken for the logistic regression respectively as a sample of 2,000 data-sets for which the one-sample t-test is available. The authors calculated the strength of the magnitude of the relationship between the independent and the dependent variables. As a very big sample, containing 4,000 data-sets, is available the level of significance should be 1% (two-sided) and the power (β -error) should be 99%. According to Cohen (1988) very small effects with an effect size smaller than 0.2 can also be identified with a dataset of this size. That means that even small effects of an independent variable on the dependent variable are significant.

Moreover the effect-size for the one-sample t-test is calculated. Again, the level of significance is 1% (two-sided) and the power (β -error) is 99%. The sample size for the one-sample t-test is now 2,000. The resulting effect size is 0.1, hence very small effects can be identified in this calculation. On the one hand it is very helpful to be able to identify small effects, but on the other hand we need to verify whether the interpretation of these small effects is useful.

2.1.2. Quality analysis

The sample studied was drawn from the total insurance policies of the UNIQA insurance company. Hence the sum of the UNIQA policies is the statistical population. The target-population is the sum of all policies in Austria. As big differences in the composition of the insurance portfolios of the different insurance companies do not exist, the results are meaningful for the entire insurance market in Austria.

The sample has been taken randomly by the UNIQA insurance group. Ex post, the sample was compared to see if it is representative concerning the available variables. As there were no significant differences in the distribution of the variables in the random-sample and the population, the random-sample is representative for the population.

3. EXECUTION OF THE ANALYSIS IN SPSS

3.1. Execution of the logistic regression

First the authors ran the logistic regression. The regression comprised 4,000 cases. As can be seen in Table 1 the binary dependent variable was classified as 1 = “pay per use” insurance and 0 = “traditional” insurance. Moreover the other non-metric but categorical variables were classified.

Table 1
Classifying of variables

Codification dependent variable				
original value		internal value		
traditional		0		
pay per use		1		
Codification categorical variables				
		observed frequency	Codification	
			(1)	(2)
Casco Insurance	no	1465	1.000	0.000
	partly	580	0.000	1.000
	full	1955	0.000	0.000
Insurance sum third party liability	10 m	2781	1.000	0.000
	15 m	1191	0.000	1.000
	others	28	0.000	0.000

Source: Authors' calculation.

The execution of the logistic regression brought the output given in Table 2.

Table 2
Regression coefficients

Step 1*	regression coefficient B	standard error	Wald-test	degrees of freedom	significance	exp(B)	95,0% confidence interval for exp(B)	
							lower bound	upper bound
KW	.010	.001	59.631	1	0.000	1.010	1.008	1.013
BAU	-.059	.010	33.095	1	0.000	.943	.924	.962
GEB	-.024	.003	70.330	1	0.000	.971	.971	.982
WO	.965	.086	126.677	1	0.000	2.624	2.218	3.104
SEX	-.233	.084	7.725	1	.005	.792	.672	.934
KAS			443.038	2	0.000			
KAS(1)	1.731	.132	170.854	1	0.000	5.649	4.357	7.323
KAS(2)	2.582	.123	442.987	1	0.000	13.228	10,401,	16.824
BM_alt	.086	.006	215.384	1	0.000	1.090	1.077	1.102
constant	113.188	20.552	30.332	1	0.000	1,435 E49		

* Included variables in step 1: KW – engine performance; BAU – year of construction; GEB – age of policy holder; WO – residency of the policy holder; SEX – gender of policy holder; KAS – conclusion of a full casco insurance; KAS (1) – conclusion of partial casco insurance; KAS(2) – no conclusion of casco insurance BM_alt – level in the merit-rating system.

Source: Authors' calculation.

The authors carried out the logistic regression with the risk factors “age of the policy-holder (GEB)”, “place of domicile of the policy-holder (WO)”, “gender of the policy-holder (SEX)”, “engine power of the car insured (KW)”, “year of construction (BAU)”, “conclusion of a Casco insurance (KAS)”, “level in the merit-rating system (BM_{alt})”, and “minimum guarantee sum (VS)”. As can be seen in Table 2, all available variables are within the regression and have a positive or negative but always significant effect on the probability that a person will choose the “pay per use” product ($y = 1$). Only the “minimum guarantee sum” has no contribution to this decision and fell out of the regression.

The logistic regression approach according to table 2 is:

$$\text{logit}(y) = \ln \left(\frac{p_i}{1 - p_i} \right) = z = 113.188 + \quad \text{Eq. (10)}$$

$$+ 0.010KW - 0.05BAU - 0.024WO - 0.233SEX + 1.731KAS(1) + 2.582KAS(2) + 0.086BM_{alt}$$

As all variables have a significant input on the decision between “pay per use” and “traditional” insurance, a closer look is taken on each individual variable.

3.1.1. Engine power of the insured car (KW)

The regression-coefficient is 0.010. The odds ratio therefore is $e^{0.010} = 1.01$. As the regression coefficient is positive, an increase in engine power leads to an increase in the probability that the “pay per use” product is chosen. But the effect is very small.

The range between the highest KW-value and the lowest KW-value is from 4 kw to 386 kw. But more than 50% of all cars in the sample are between 60 kw and 100 kw. Thus the odds that a person with a 100 kw-car chooses “pay per use” insurance in comparison to a person with a 60 kw-car are 1.49:1.

3.1.2. Year of construction (BAU)

The regression-coefficient is -0.059 . The odds ratio therefor is $e^{-0.059} = 0.942$. The range between the highest BAU-value and the lowest BAU-value is from 1978 to 2010. More than 50% of all cars were produced between 2000 and 2010. The odds that a person with a car produced in 2000 chooses “pay per use” insurance in comparison to a person with car produced in 2010 are 1:1.8.

3.1.3. Age of the policy-holder (GEB)

The variable GEB shows the age (not the birth-date) of the policy-holder. The regression-coefficient is -0.024 . Because of the negative sign the higher the age, the lower the probability that the policy holder will chose the “pay per use” product. The odds ratio is 0.976 which means that the probability that a person chose the “pay per use” product decreases by 2.4% per year. In the case that the difference adds up to 10 years this probability decreases by 27.12% (odds ratio 1:1.27).

3.1.4. Residency of the policy-holder (WO)

The variable WO is binary coded. Cars, insured in urban areas (more than 39,000 inhabitants), were coded as “Stadt” (1) cars insured in rural areas (below 39,000 inhabitants) were coded as “Land” (0). This variable has a high value of 0.965 which means, that the odds that people

in urban areas chose the “pay per use” product in comparison to persons from rural areas are 2.624:1. This is the first variable which shows a big difference in the two groups. The descriptive statistics also reflects this result. While nearly 75% of the 1,350 urban inhabitants chose the “pay per use” product, fewer than 50% of the 2,650 rural inhabitants chose this product. These figures indicate a clear selection-effect between rural and urban areas. The main reason for this disparity probably is that urban inhabitants drive fewer kilometers and that there is probably a strong correlation between mileage and residency.

3.1.5. *Gender of the policy-holder (SEX)*²

Similar to the variable “residency” the variable “gender of the policy-holder” was also binary coded (female = 0, male = 1). The descriptive statistic shows that 64.4% of all policy holders are men. The regression coefficient is -0.233 which means that the “pay per use” product is more attractive to women. In fact the odds that women chose the “pay per use” product are 1.26:1 in comparison to men. One explanation could be that men drive more kilometers per year than women, and therefore the “traditional” product is more attractive to men than for women.

3.1.6. *Conclusion of a Casco insurance (KAS)*

This variable can only take the values “no additional insurance” (0 = reference category), “part insurance cover” (1) and “own damage claim” (2). It is very interesting that there were huge differences in the three groups. According to the regression coefficient the odds that a “part-insured” person will chose the “pay per use” product instead of the “traditional” product are 5.649:1. The odds that an “own-damage-claim” insured person will chose the “pay per use” product instead of the “traditional” product are 13.228:1. Ex post, the authors found that financial incentives were offered to customers if they choose the “pay per use” product and in addition “part insurance cover” or “own damage claim”.

3.1.7. *Level in the merit-rating system (BM_alt)*

The UNIQA Insurance Company still uses, as do almost all insurance companies, a merit-rating system. This system is very similar to the former merit rating-system required by law in Austria till 1994. The regression-coefficient 0.086 is very low. The resulting odds-ratio is 1.090 which indicates that the variable “level in the merit-rating system” has a positive effect on the decision that a person chose the “pay per use” product. The higher the level in the merit-rating system, the higher is the probability that the person will chose the “pay per use” product. As the effect is very small and almost 75% of all policy holders are in level -2 , -1 and 0 , the variable (BM_alt) has practical no effect. The significant effect results from the fact that almost all policy holders (about 90%), who are in very high levels of the merit-rating system, chose the “pay per use” product. The authors assume that insurance brokers granted discounts to such policy-holders, if they declared their willingness to be “supervised” by the “pay per use” product. The authors did not find empirical evidence for this assumption.

3.1.8. *Minimum guarantee sum (VS)*

The variable “minimum guarantee sum” is the only variable, which was removed from the regression, as it had no significant impact on the dependent variable.

² The authors would like to mention that it is forbidden by law to charge different insurance premiums for men and women, as the European Court of Justice stated in 2011 that “The use of gender as a risk factor is a discrimination”. The data used originate from 2009. As this is a theoretical paper we use gender as a possible risk factor, although we know that this would be forbidden in insurance practice.

3.2. Results

Hence the results of the logistic regression indicate that there are selection effects caused by the introduction of “pay per use” products.

Therefore the null hypotheses: $H_0 =$ *There is no difference in the described risk factors between “pay per use” policy-holders and “traditional” policy-holders*, has to be rejected for all variables except for “minimum guarantee sum”.

3.3. Quality criteria of the statistical method

3.3.1. The interpretation of the quality criteria of the logistic regression

To check the quality of the logistic regression primarily the pseudo- R^2 -statistics are applied.

Table 3

Quality criteria for the model

Quality criteria for the model			
Step	-2 Log-Likelihood	Cox & Snell R-square	Nagelkerkes R-square
1	4005,713*	.319	.426
2	4007,509*	.319	.426

* Iteration stopped because change in parameter-calculation was lower than 0.001.

Source: Authors' calculation.

As shown in Table 3 the “Cox & Snell R^2 ” is 0.319 and “Nagelkerkes R^2 ” is 0.426. According to Backhaus et al (2011) these values are “acceptable” for the “Cox & Snell R^2 ”-value, and “good” according to “Nagelkerkes R^2 ”.

Table 4

Reference value

	acceptable	good	very good
McFaddens R^2	> 0,2	> 0,4	
Cox und Snell R^2	> 0,2	> 0,4	
Nagelkerke R^2	> 0,2	> 0,4	> 0,5

Source: Authors' calculation.

3.3.2. The interpretation of the quality criteria of the different variables

The second quality verification focuses on the impact of the different variables. As mentioned above all variables except “minimum guarantee sum” have significant regression-coefficient-values.

Table 5
Likelihood-quotient test (modelling if term excluded)

Variable	Log-Likelihood of the model	Change in -2 Log-Likelihood	degrees of freedom	Change in significance
step 2 KW	-2036.138	64.767	1	0.000
BAU	-2020.761	34.012	1	0.000
GEB	-2039.971	72.433	1	0.000
WO	-2069.310	131.111	1	0.000
SEX	-2007.623	7.736	1	.005
KAS	-2284.965	562.420	2	0.000
BM_alt	-2172.486	337.463	1	0.000

Variables included in step 2: KW – engine performance; BAU – year of construction; GEB – age of policy holder; WO – residency of the policy holder; SEX – gender of policy holder; KAS – conclusion of a full casco insurance; KAS(1) – conclusion of partial casco insurance; KAS(2) – no conclusion of casco insurance BM_alt – level in the merit-rating system.

Source: Authors' calculation.

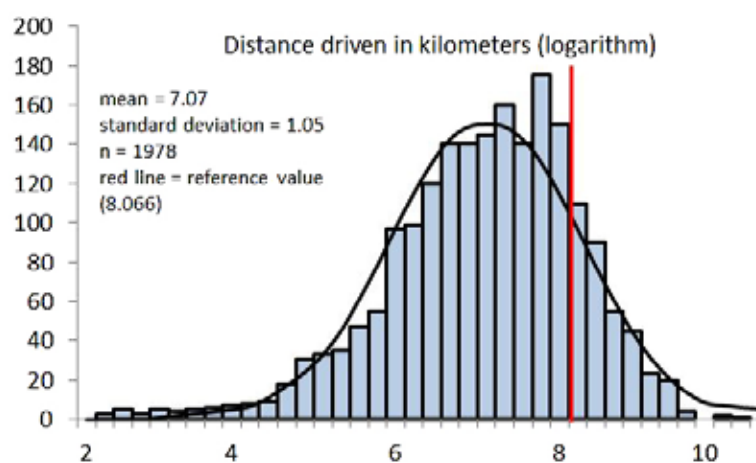
As a next step, a closer look is taken to see if the different variables contribute a positive value to the explanation of the variance in the model. This can be done with the “likelihood-quotient test”. As can be seen in the third column in Table 5 each variable has a positive contribution to the explanation of the variance of the model.

3.4. Execution of the one sample t-test

In a second step, the variable mileage is examined. As has been already mentioned, another statistical method has to be applied to examine whether there are differences in the mileage of “pay per use” policy-holders and “traditional” policy-holders. Hence the null-hypotheses therefore is: $H_0 = \text{There is no difference in the mileage per year between “pay per use” policy-holders and “traditional” policy-holders.}$

Figure 3

One-sample t-test, reference value



Source: Authors' calculation.

Table 6
One-sample t-test

Reference-value = 8.066; mean = 7,070						
LNDIST*	T-value	degrees of freedom	significance (two-sided)	mean difference	99.99% confidence interval	
					lower bound	upper bound
	−42189	1977	0.000	−0.99612	−1.0882 (6.9778)	−0.9041 (7.1619)

* LNDIST – mileage driven (logarithmised).

Source: Authors' calculation.

As demonstrated in Table 6 the test is highly significant.

3.5. Results

This means that the null hypotheses needs to be rejected as there is a difference in the mean of the “traditional” policy-holder mileage and the mileage of the “pay per use” policy-holder. In fact the mean of the “pay per use” policy-holder is 1,682.30 km, while the mean of an average Austrian car is 1,823.96 km. There is a difference of 141.66 km, consequently “traditional” policy-holders drive 8.42% more than “pay per use” policy holders.

4. DISCUSSION OF RESULTS

The results above show that there are selection effects between “pay per use” policy-holders and “traditional” policy-holders. In this paper it is empirically proven that low-mileage-drivers prefer “pay per use” insurance products as they can demonstrate their lower risk compared to high-mileage driver. [We have not thus far mentioned the correlation between mileage and the risk of causing an accident. It seems to be certain that low-mileage drivers have a lower risk of having an accident than high-mileage driver. Amongst others, Litman (2005) has shown a positive correlation between mileage and crash rates in empirically studies.]

According to Rothschild and Stiglitz (1976) there cannot exist a pooling-equilibrium in which all insured people choose the same insurance-contract. Under certain circumstances there can be a separating-equilibrium in which high-risk customers and low-risk customers choose the insurance contract which is adequate to their individual risk. The result, that it is possible that there is no equilibrium in a competitive insurance market, is caused by adverse selection as described in the famous paper of Rothschild and Stiglitz in 1976.

As we mentioned above, insurance companies use several characteristics to classify their clients into different risk-categories. But what is the effect of new risk categories? In the past policy-holders were pooled concerning the risk category, mileage, as there was no possibility of distinguishing between low-mileage drivers and high-mileage drivers. Insurance companies used other characteristics to divide their clients into risk categories.

What is more, third-party liability insurance is mandatory for passenger cars in Europe as a consequence insurance markets are not always perfectly competitive. But adverse selection is primarily a problem if there is a perfectly competitive market. These effects diminished the adverse-selection effects on insurance markets in the past. Especially in the third-party liability insurance markets, competition increased and the pressure to premium differentiation on insurance companies has been in evidence for the last few decades (Cohen 2005; Karten, 1994).

The new “pay per use” product can be used to indicate low risks for the insurance company. As described first by Spence (1973) the insurance company will reward this signal with a lower insurance-premium. As a consequence, all policy-holders will choose the “pay per use” -product if, and only if, the costs for the signaling are lower than the remuneration the insurance-company grants for this signal. High-mileage drivers will still choose the “traditional” product, but as there is no “subsidization” by the low-mileage drivers any longer, the price for the “traditional” product will rise. As a consequence, the “pay per use” -product will be attractive for other clients too. That is the same effect as Akerlof (1973) described when “lemons” drive “plums” out of the market. In theory “pay per use”- insurance might force traditional” insurances out of the market.

5. CONCLUSIONS

Does this theory go along with our empirical research? That would be the case if people made their insurance-decision in a rational way. According to this theory, low-mileage drivers should have chosen the “pay per use” product, while high-mileage drivers should have chosen “traditional” insurance products.

As illustrated in section 3.4, the authors found that people who drive fewer kilometers per year have a higher probability of choosing the “pay per use” product, than people who drive more kilometers per year. This confirms the theory mentioned above.

A second pre-condition which has to be fulfilled for the confirmation of the theory that “pay per use” products drive “traditional” products out of the market is the following: there should be no other independent variables which influence the decision between “pay per use” products and “traditional” products.

According to the results of the empirical analyses, it can be said in summary that almost all the independent variables reviewed have an impact on the dependent variable. But this fact is not really in conflict with our theory because in most cases (“engine power”, “year of construction”, “age of the policy-holder”, “gender of the policy-holder” and “level in the merit-rating system”) the impacts of the variables were very low. Big impacts (“conclusion of Casco insurance”, “residency of the policy-holder”) had different reasons. The “residency of the policy-holder” correlates strongly with the “mileage of the policy-holders”. The “conclusion of Casco insurance” has an immediate effect on the insurance-premium.

As a result, the authors see their theory fortified. That means the authors think that in the long run “pay per use” products will become more attractive to a certain group of car holders and, as a consequence, “pay per use” products might drive out “traditional” products from the market. But there are a few limitations to this theory which should be subject of further research.

Our results suggest venues for further research. It is possible that signaling has a negative effect on the personal benefit. As there are economic constraints against buying “pay per use” insurance product (the policy-holders are forced to give lots of information to the insurance company). If the policy-holder’s negative effects of the necessary transfer of information to the insurance company are greater than the positive effects of the money saved by choosing the “pay per use” product, the policy-holder will choose the “traditional” product even if he or she is a high mileage driver. Hence the negative effects of signaling should be empirically investigated.

References

- Akerlof G. (1970) The Market for “Lemons”: Quality Uncertainty and the Market mechanism, *Quarterly Journal of Economics*, Vol. 84, Issue 3, pp. 488–500.
- Backhaus K. et al (2011) *Multivariate Analysemethoden*, Berlin, Heidelberg, New York: Springer.

- Cohen A. (2005) Asymmetric information and learning in the automobile insurance market, *The Review of Economics and Statistics*, Vol. 87, Issue 2, pp. 197–207.
- Cohen J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, New York: Lawrence Erlbaum Associates.
- Kähler W. (2011) *Statistische Datenanalyse*, Wiesbaden: Vieweg & Teubner.
- Karten W. (1994) Das Einzelrisiko und seine Kalkulation, in: W. Große et al., *Versicherungszyklopädie*, Wiesbaden: Gabler.
- Litman T. (1997) Distance-based insurance as a TDM strategy, *Transportation Quarterly*, Vol. 51, Issue 3, pp. 119–137.
- Litman T. (2005) Pay-As-You-Drive Pricing and Insurance Regulatory Objectives, *Journal of Insurance Regulation*, Vol. 23, Issue 3, pp. 23–53.
- Mehring J. (1962) Die Schadenstruktur in der Kraftfahrt-Haftpflichtversicherung von Personenwagen, *Blätter der DGVM*, Vol. 6, Issue 1, pp. 23–41.
- Rothschild M., Stiglitz J. (1976) Equilibrium in competitive Insurance Markets: An Essay on the Economics of imperfect Information, *Quarterly Journal of Economics*, Vol. 90, Issue 4, pp. 630–649.
- Spence M. (1973) Job Market Signaling, *Quarterly Journal of Economics*, Vol. 87, Issue 3, pp. 355–374.