



Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

**Evaluating the Impact of Driving Experience
and Previous Accident History on Motor Insurance
Premium**

By
Pricilla Moses

Student Number: 1818519
Module Number: LCDB7011
Module Name: Quantitative Research Methods
Module Lecturer: Dr. Helen Lee
Word Count: 8250

Abstract

This quantitative study evaluates the impact of driving experience and previous accident history on motor insurance premiums. The study is significant within the framework of general insurance practises and the operations of underwriters and actuaries. The analysis was conducted using SPSS software, and the relevant data were examined through one-way ANOVA and independent sample t-tests, supported by additional tools such as Levene's statistic, Tukey HSD, and Games–Howell. Driving experience and previous accident history were considered as the independent variables, while the motor insurance premium was treated as the dependent variable. The reliability of the dataset has been tested by Cronbach's alpha. Descriptive statistics were performed to understand the basic characteristics of the dataset after coding the variables appropriately. The study used a secondary dataset publicly available on www.kaggle.com, consisting of 500 responses. Furthermore, the contextualisation of risk theory and experience-pricing theory supported the study. Each statistical test was conducted to examine the relevance of the hypotheses developed. The findings of the study emphasise that driving experience is highly significant in determining motor insurance premiums and that risk-based insurance practises have a strong influence on the accuracy of premium determination. However, previous accident history has not shown a statistically significant effect; the impact was negligible, and this result did not align with the theoretical expectations of experience-pricing theory. Overall, the study provides academic contributions to both academia and industry, along with recommendations for improving existing premium calculation practises in underwriting systems.

Keywords: Motor insurance premium, Driving experience, Previous accident history, Cronbach's Alpha, One-way ANOVA, Independent Sample t-test, Games–Howell test, Risk theory, Experience-pricing theory

Table of Contents

Abstract	i
Table of Contents	ii
1. Introduction.....	1
1.1 Motor Insurance Industry.....	1
1.2 Problem Statement.....	2
1.3 Research Aim.....	3
1.4 Research Objectives.....	3
1.7 Significance and the Actionable Contribution	4
2. Literature Review.....	5
2.1 Motor Insurance Pricing	5
2.2 Theoretical Perspective: Risk Theory	6
2.3 Theoretical Perspective: Experience Pricing Theory (EPT).....	7
2.5 Variable Review: Previous Accident History on Insurance Premiums	9
2.6 Conceptual Framework.....	11
3. Methodology	13
3.1 Research Philosophy	13
3.2 Research Approach	13
3.5 Population and Sample Size.....	14
3.6 Variables	14
3.7 SPSS Tools and Statistical Tests in Data Analysis.....	14
3.8 Hypotheses Development	15
3.9 Ethical Consideration.....	16
4. Data Analysis and Results	17
4.1 Reliability testing.....	17
4.2 Descriptive Statistics.....	17

4.3	One-way ANOVA with post hoc tests.....	18
4.4	Independent Sample t-Test	20
5	Discussion	22
5.1	Interpretation of the Level of Driving Experience.....	22
5.2	Interpretation of Previous Accident History	23
5.3	Comparison of previous studies and the final outcome	23
5.4	Practical Implications.....	24
6	Recommendations & Conclusion	25
6.1	Contribution of the study	25
6.2	Future Recommendation	25
6.3	Limitations of the overall study	26
6.4	Conclusion	26
	Reflexive Report	28
	Bibliography	31
	Appendices.....	39
	Appendix 1: One-way ANOVA group descriptives	39
	Appendix 2: Tests of Homogeneity of Variance	39
	Appendix 3: ANOVA Effect Sizes	39
	Appendix 4: Multiple Comparison	40
	Appendix 5: Mean Plot	40
	Appendix 6: Group Statistics	41
	Appendix 7: Independent Sample Effect Sizes.....	41
	Appendix 8: Coded working of Raw data.....	41

Figure 1 Motor Insurance Market	1
Figure 2 Inverse Relationship Between Experience and Cost	8
Figure 3 Core Risk Matrix - Illustrated by the Author	9
Figure 4 Conceptual Framework designed by the Author	12

Abbreviations

AI	Artificial Intelligence
ANOVA	Analysis of Variance
BMS	Bonus-Malus System
CAGR	Compound Annual Growth Rate
CUE	Claim & Underwriting Exchange
EPT	Experience Pricing Theory
EV	Electric Vehicle
GLM	Generalised Linear Models
NCD	No Claim Discount
UBI	Usage Based Insurance

1. Introduction

1.1 Motor Insurance Industry

The global motor insurance segment is one of the most attractive insurance classes in general insurance practice. The high frequency of the mandatory purchasing class leads to massive annual industry growth, which is also affected by factors such as continuous technological innovation and challenges in the global vehicle market. The global motor insurance market is worth approximately USD 2.27 trillion in 2026, and it is acquiring a 6.65% CAGR (Mordor Intelligence, 2026). The market segment is mainly divided into two segments, such as the third-party insurance share, which is 42.37% while the comprehensive market share is 57.63% in 2025 (Mordor Intelligence, 2026).



Figure 1 Motor Insurance Market

Several key trends and drivers influence the global motor insurance market. One of the most influential trends, technology and telematics play a significant role in the market driven by the pace of digital transformation in AI, machine learning, and telematics, offering personalised risk and usage-based risk evaluation practises (Schwab, 2023). The global trend of EV usage is shaping the global motor insurance market with new risk assessments coupled with the repair cost evaluation (Total Systems PLC, 2026). Furthermore, regulators also have a major influence on the

dynamics of the motor insurance market despite global geographical relevance (Mordor Intelligence, 2026).

Premium pricing is the most significant and paramount of motor insurance practice as it may cause influences for the profitability margin, risk management strategies and the competitive advantage of the organisation (Mildenhall & Major, 2022). Motor insurance premium influence the bottom line of the company via several aspects; being the most compelling business class in every general insurance company, motor insurance acts as the primary cashflow intake supporting the financial stability and the solvency margin of the insurer (Ritho, et al., 2023), by funding claims, reserving adequate capitals to meet regulatory solvency ratios and protecting profit margins covering the operational costs and investment returns. Proper premium pricing strategies in motor insurance may influence on risk differentiation and anti selection of risk, attracting hazardous risk for a lower premium and also differentiation high cost for higher risk drivers and lower for lower risk drivers (Parodi, 2023). UBI pricing in motor insurance may reduce the risk in fraudulent and deterrence and risk reduction causing the behavioral modifications of drivers or telematics (Xie, 2023). Motor premium pricing also play crucial role in macroeconomic adaptations such as inflation hedging and EV transitions (Mehrotra & Chadha, 2025).

In calculation of motor insurance premium, driving experience and the previous claim history are vital components in risk evaluation. Underwriters and actuaries consider the driving experience as a proxy for the risk exposure (Isimoya, et al., 2022) and previous accident history may consider as a future risk predictor which helps measure frequency over severity and the discount of NCD and acts as a behavioral mirror (Xiong, et al., 2026). In this quantitative explanatory research it will find out how the driving experience and the previous accident history impact on the motor insurance premium.

1.2 Problem Statement

In motor insurance, drivers and the risk associated with them are considered the key factors in premium determinants. In practice, insurers calculate and justify different premium amounts for drivers based on their varying levels of experience and accident histories. This study focuses on evaluating the significant differences in premium amounts based on driving experience and previous accident history, using risk theory and EPT. The statistical analysis is conducted by using

SPSS, including the independent t-test and one-way ANOVA to examine the quantitative relationship of each variable.

1.3 Research Aim

The aim of this study is to evaluate the impact of the level of driving experience and the previous accident history effect on the premium amount by analysing the significant differences based on the risk theory and the EPT. The research is conducted by using the secondary data set which contains 500 cases using the appropriate SPSS tools.

1.4 Research Objectives

In order to evaluate the impact of driving experience and the previous accident history on the premium paid, the research is mainly focused on the below three research objectives:

1. To examine whether various levels of driving experience have a significant impact on motor insurance premiums.
2. To determine whether previous accident history significantly influences insurance premiums.
3. To evaluate how risk-based pricing and experience-based pricing practises explain variations in insurance premium costs.

1.5 Research Questions

This research is mainly focusing on evaluating the impact of driving experience and the previous accident history by driving the research on the below research questions:

1. Does driving experience significantly affect the insurance premium paid by drivers?
2. Does a previous accident significantly influence the insurance premium level?
3. How do risk-based pricing and experience-based pricing practises apply in insurance premiums among drivers?

1.6 Research Gap

Various research methodologies have focused on different geographical territories and factors, resulting in the publication of several empirical papers on the impacts of motor insurance premiums. However, limited studies have examined the combined impact of driving experience and previous accident history on motor insurance premiums using SPSS statistical evaluation,

specifically through one-way ANOVA and independent t-test methods. However, limited studies have examined the combined impact of driving experience and previous accident history on motor insurance premiums using SPSS statistical evaluation, specifically through one-way ANOVA and independent t-test methods.

1.7 Significance and the Actionable Contribution

This study will provide significant insights into how the driving experience and previous accident history affect motor insurance premiums and to what degree. Further, the finding of the study will guide insurers to reconsider the pricing factors and motor premium calculation. Simultaneously, academia also benefits from referring to the application of risk theory and the EPT in calculating motor insurance premiums. The findings of the study will also contribute to customer awareness regarding how a driver's driving experience and previous accident history affect the cost of motor premiums. The insurers may use the outcome of the study to strategise their premium pricing fundamentals, and academia will be guided to conduct a similar kind of research to evaluate the different factor impacts on the premium, such as the driver's age, year of manufacture of the vehicle, usage of the vehicle and the gender of the driver.

1.8 Scope of the Study

The study has been conducted by using the secondary data set available on a public source. The dataset of 500 cases contained several demographic variables such as driver's age, driver experience, previous accident history, annual mileage, car manufacture year and the insurance premium charged. The data has been gathered by a survey conducted for the UK drivers population (Kaggle, 2025).

2. Literature Review

2.1 Motor Insurance Pricing

Motor insurance pricing is a strategic process and a strict mechanical sequence of calculation to determine the risk premium paid by the customer (Parodi, 2023). Actuaries are practicing different risk evaluation formula to maintain the competitive edge and the long-term solvency margins of the insurance business (Ahmed, et al., 2022). In order to calculate the risk premium, the actuaries consider different criteria and the different aspects of the risk inculcated (Zhang, 2023) such as; scrutinising the pure risk portion (Kochenburger & Salve, 2023), applying the operational loadings and profit margins, and determinants of discounting factors (Worku, et al., 2024). In conventional insurance, static historical data and generalised linear models¹ (GLM) used to calculate motor insurance premiums (Rahmawati, et al., 2024). The model evaluates the historical data to determine future premium cost by multiplying baseline risk factor by various other demographic factors (Rahmawati, et al., 2023). According to the multiplicative structure² (Acclavio & Maieli, 2024), final premium is calculated by using the formula; Final premium = Base premium x Factor_{age} x Factor_{gender} x Factor_{vehicle group} x etc., (Coutts, 1984). Motor insurance premium is mainly be rated by three derivatives; the proposer (insured's demography), the risk (vehicle classification) and the exposure (territorial, behavioral or usage constraints). However, there have been identified several limitation since the conventional premium calculation practises. The nature of the proxy biasness punishes the safe drivers, even though if they fallen inside the proxy (Johnson, 2025). The static nature of the risk evaluation foresee the risk as it is for the entire annum and no opportunity for improved factors in between the policy period. Fraudulent vulnerability and lagging in system data update may have an impact in real time risk analysis process in the conventional premium pricing models (Ahmed, 2025). Underwriting principles play a vital role in bridging the gap between the actuarial formulae and the actual risk perception in a business strategy and it decides whether to accept the risk or the risk appetite aligned with the insurer's capacity as well. There are five fundamental underwriting principles practiced by global insurance industry in terms of pricing strategy (Alles, 2026). Firstly, the principle of utmost good faith or *Uberrimae Fidei* is the legal and ethical foundation of pricing to declare the every material fact related to the risk (Kaushish & Patnaik, 2025). Secondly, the principle of mutual benefit and

¹ GLM – a flexible extension of ordinary linear regression (OLR) that allows the dependent variable to follow different types of probability distributions in a line straight line.

² Variables combined by multiplications rather additions.

equity to ensure that the policy holder's due premium reflects the fair share of the risk contribution and the risk classification discipline (Sugden, 2024). Thirdly, the principle of adverse selection is mitigated by the underwriters by doing the required inspections and screening test before selecting the risk (Dionne, et al., 2024). Fourthly, underwriters main job is to select risk, classify and to decline any risk which is not fallen with the risk appetite of the organisation (Isimoya, et al., 2022). Fifthly, underwriters ensure adequate rate making and financial soundness by evaluating the combined ratios for solvency protection and long-term consistency (Maggioni & Turchetti, 2024).

2.2 Theoretical Perspective: Risk Theory

Risk theory was originally introduced by Filip Lundberg in 1903 as mathematical risk theory (Khairiati, et al., 2025). Lately it was extended to concepts such collective risk theory and ruin theory providing the root course for insurance analytics (Alcoforado, 2022). In 1930 the Swedish mathematician Harald Cramér refined the risk theory in Cramér-Lundberg Model of insurers probability of insolvency (Englund & Martin-Löf, 2001). Risk theory plays a pivotal role in insurance to quantify the risk on probability and managing portfolio from future financial losses. The risk theory models in insurance mainly focus on two areas; frequency and severity (Mildenhall & Major, 2022). Frequency evaluates how often claims occur and severity evaluates the individual claim cost (Wang, et al., 2024). These two key elements combinedly evaluate the total loss distribution in a risk portfolio to determine the capital gain or the future bankruptcy of the business in future (Fiveable, 2026). The primary action of motor insurance underwriting is classifying homogeneous groups from diversified drivers, and the risk theory is providing validity for the risk classification (EBSCO, 2021). Further explaining the homogeneity and the variation, risk theory emphasises the level of frequency and severity of a future expected claim (Boomhower, et al., 2024) and in the same time to illustrate the statistically significant relationship between frequency and the severity. In scrutinising the motor insurance risk, actuaries group the standard drivers according to the frequency, by poisson distribution³ model (Rahmawati, et al., 2023) and simultaneously the high variance of personal driving behavior, lead actuaries to apply negative binomial distribution (Zhang, et al., 2022). Risk theory ensures the equity of the premium by eliminating cross subsidisation (Supian, et al., 2022). The mathematical steps involved in the risk

³ Frequency of events happen in a fixed time period, area, or space – mostly used in insurance, healthcare, and financing.

theory converts the statistical data of a risk exposure into a financial value as a premium and it happens with the application of various mathematical steps and models. The collective risk model in the risk theory calculate the pure premium related to risk to ensure that right amount of money has been charged to cover a future claim cost (Springer Berlin Heidelberg, 2005). $S = X_1 + X_2 + \dots + X_n$, according to this formula, S represent the aggregate total claim cost in a selected risk, n represents the random number of claims and the X is the financial amount of the loss. Insurance premium cannot be charged only for the pure risk; it has to be loaded for the future claim cost and the volatility and the unpredictability of the claims. Insurers may load a percentage value to the premium by following expected value principle⁴ (Mwigilwa, 2024) and the loading depends on the variance, following the variance principle method⁵ (Moriah, et al., 2023).

2.3 Theoretical Perspective: Experience Pricing Theory (EPT)

Experince pricing or rating theory is grounded on the credibility theory (Whitney, 2024). EPT is a framwork that calculate premiums on the previous claim history or variable history rather than the demographic classifications (Badescu & Lin, 2024). The theory emphasise how much of a weight (credibilty) that an underwriter assign for the past claim history and records of the policyholder when scruitnising and risk assessment (Badescu & Lin, 2024). The fundamental credibilty formula is expressed as; if the credibility factor is represented by Z, $\text{Premium}_{\text{New}} = Z \cdot (\text{Individual past experience}) + (1-Z) \times (\text{Class base rate})$. As an example, if the driver has obtained the driving liscence not more than a one year his individual statistical data set is too small to analyse the risk on past experince where a vehicle fleet operated for more than fiver years have a rich content in the past histroty to analyse. In the context of motor insurance the EPT is applied in two key frameworks such as; the Bonus – Malus system⁶ (BMS) (Ieosanurak, 2023) and Prospective experince rating⁷. In here the bonus is represnting the no claim discount which is offered to the claim free policyholder. Malus represent the panelty step back to remove the applied NCD when a faults accident occur. The prospective experince rating is applied for mainly the commercial fleets such as delivery netwroks and haulage providers. Actuaries are calculating the mod factor or the experince modification factor by dividing the historical three years of losses by the expected losses (Parodi, 2023) and charge the discount or the premium loading accordingly. When

⁴ Specifically used in insurance to measure the probability and the average outcome of random variables.

⁵ Premium calculation method in insurance pricing.

⁶ Special pricing method used in motor insurance for premium calculation.

⁷ A method of premium calculation to determine current premium on prospecting future claims and losses.

comparing the EPT with conventional classic rating systems and UBI, EPT is placing in the middle featuring, main data focus on the historical events and losses, number of accidents counted as primary variables, calculation time retrospect at the renewal and delay data dependency considering least 1-5 years. However, despite the technical guidance provided by the EPT, features of the framework are being questioned by the novel risk calculation methods such as UBI for its own operational limitations. The lagging effect of EPT is unable to assess the immediate behavioral change of a driver and it has no capacity to assess the safe driver who hits an accident with a totally uncontrollable facts (random volatility).

2.4 Variable Review: Driving Experience on Insurance Premium

According to insurance acturies, the driving experience is considered as a proxy and one of the most credible risk factors in pricing (Xie, 2024). Based on the hazardous risk perception less driving experience has a high claim probability which leads to a significantly high premium (Clemente, et al., 2023). The inverse relationship between the experience vs. cost follows a steepy inverse curve (Joeveer & Kepp, 2023).

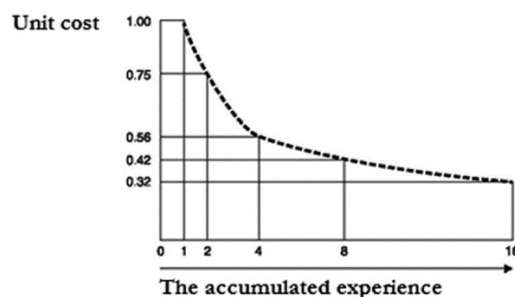


Figure 2 Inverse Relationship Between Experience and Cost

Source: (Reketye, et al., 2025)

Majority of the newly qualified drivers during their first year involve in statistically severe and frequent accidents (Rowe, et al., 2022) and this phase is named as novice premium peak (Xiong, et al., 2026). As the driver improves his defensive and maneuvering skills from the 2nd year there will be a drastical drop in cost which calls experience drop in the curve. After reaching the fifth year of the experience level, at the plateau, the curve starts getting flattened and the premium becomes steady and the premium will be evaluated on the previous claim history and other factors. Underwriters not generally penalise less experience drivers but investigate their loss history to calculate premium. According to the lack of hazardous perception, inexperience drivers take time

in their cognitive processing skills about the developing hazards in the road (Cao, et al., 2022). Statistical research publications show, that new drivers are overconfident in their judgments about the road conditions, stopping gaps and traffic conditions which lead for accidents (Hu, et al., 2026). Drivers age has an impact with their age as young age inexperience (double risk profile) and old age inexperience (mature novice profile). The double risk profiles will be charged for the biological immaturity and the low driving skills (Lee, et al., 2025) where mature novice profiles charge for lower baseline premium on statistical propensity of their reckless behavior (Isimoya, et al., 2022).

2.5 Variable Review: Previous Accident History on Insurance Premiums

Previous accident and claim history appeared to be similar in casual language, but in the technical aspect it is two different. The main difference lies between the capital financial movement and the physical occurrence of the event (Poufinas, 2023). The past accident history directly involve with the physical occurrence of the event regardless a financial pay out but it represent the raw physical exposure to the hazard (Lawrence, et al., 2023). Actuaries consider that past accident history is a mathematically accurate predictor about the future event frequency and the anticipating severity of the financial loss (Zheng, et al., 2025). Further it signals the insured's risk behavior and the environmental exposure that would cause for the future occurrence of claims. Actuaries consider previous accident history as a rating variable to predict future loss frequency vs. severity, as a probability shift and a look-back window (Meng, et al., 2022). According to actuarial logics, the frequency of the accident may stronger than the severity of the claim as it directly emphasis the moral hazard of the individual (Cevolini & Esposito, 2022).

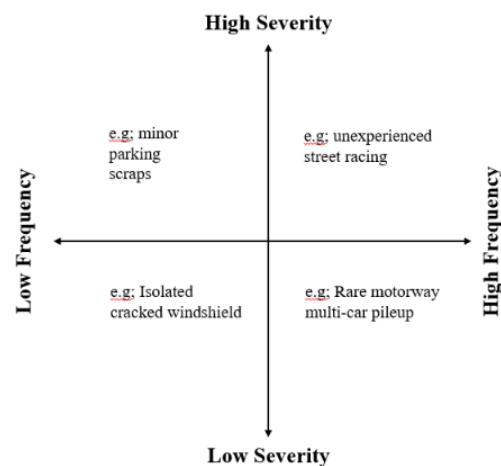


Figure 3 Core Risk Matrix - Illustrated by the Author

In the probability aspect, actuaries expect that an insured who lodged claims within past three to five years may have a tendency of filing new cases in future as habit as well (Denuit, et al., 2007). In the accident occurrence event though the insured is fault or non-fault the claim cost will have a direct impact in the renewal premium; if the insured is the fault he may be charged 20%-50% premium increase in the next renewal pricing (Engstrom, 2011). The accident history may influence the whole underwriting process by making impacts on NCD, compulsory excesses and the risk appetite (Aeron-Thomas, 2022). As the previous accident history has a direct impact on the renewal premium, there is a high vulnerability of non-disclosure fraud, and therefore it is to be centrally tracked. In the UK context it is monitored by Claims and Underwriting Exchange (CUE) (Gill, 2021). The pure premium formula is based on these two elements and it is being calculated as; $\text{Pure premium} = \text{Accident frequency} \times \text{Accident severity}$ (Manurung, 2025). According to the statistical researches the high frequency penalise the premium more than the severity and it is the single strongest indicator to predict the high severity catastrophic loss (Fackler, 2025). In an occurrence with a high frequency + low severity, underwriters may apply a minor surcharge and a compulsory excess (Selvakumar, et al., 2023). In an occurrence of a low frequency + high severity, underwriters may load only a portion to reflect the loss in the renewal premium, and further if there is a NCD, the premium will not be touched (Selvakumar, et al., 2023). In occurrence such as high frequency + high severity, the underwriters may decline the quoting (Selvakumar, et al., 2023).

2.5 Empirical Studies

There are several empirical papers that have conducted quantitative researches on the similar research interest, and they have emphasised that driver's characteristics have a direct impact on the determination of the insurance premium. The empirical study of (Henckaerts & Antonio, 2022) on "*the added value of dynamically updating motor insurance prices with telematics collected driving behavior data*", has conducted research by using novel data set to examine the driving behavior of young policyholders on motor insurance pricing on third-party portfolio. The research has done by using 400,000 participants for each year for entire period of three years. The research approach has been driven to analyse the UBI products and impact on the insurance premium which depend on the driver's habit and the style. The finding of the research has emphasised the improvement of risk classification via telematic tools and managers insights

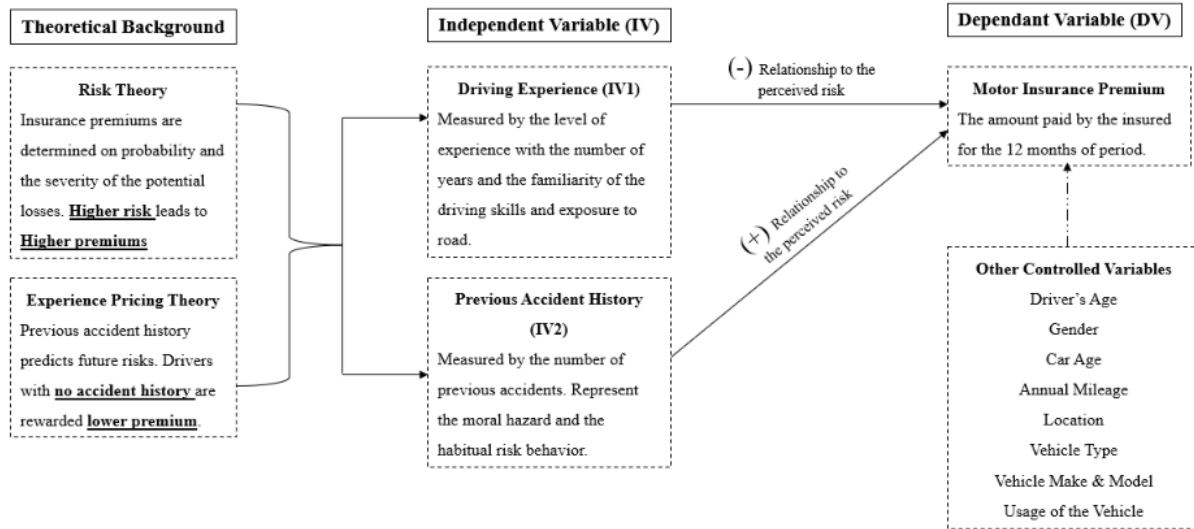
about the profit and rating structures under UBI pricing strategy. According to (Cevolini & Esposito, 2022), in the study of “*from actuarial to behavioural valuation; the impact of telematics on motor insurance*”, has examined the impact of digital tools on the motor insurance industry by the data extracted from telematic data regarding the drivers life styles. This qualitative study has been designed by conducting interviews with three main insurance companies. This study focuses on the proactivity rather waiting reactive for claims to mitigate the risks. This research paper is mainly supported by the risk theory. In the research publication of (Verschuren, 2022), on “*frequency-severity experience rating based on latent Markovian risk profiles*”, introduces novel rating on EPT. It explains that risk premium is improved to a credible premium with experience based pricing structure. This research is based on the quantitative literature review to prove the concept supported by Latent Markovian risk⁸ profiles. Therefore, considering the empirical papers already published the researcher has identified the gap of conducting a quantitative research to evaluate the impact of driving experience and the previous accident history on motor insurance premium guided by the risk theory and the EPT.

2.6 Conceptual Framework

In this conceptual framework, the researcher aims to illustrate the relationships between each variable and their effects on one another. The researcher has selected two fundamental theories, risk theory and EPT to evaluate how the driving experience and previous accident history impact the motor insurance premium. Both theories fundamentally influence each variable, with driving experience and previous accident history acting as independent variables, while the motor insurance premium serves as the dependent variable. Every behaviour of the independent variable can penalise or reward the dependent variable. Further, it illustrates how other factors also impact the dependent variable as controlled facts. Based on the theoretical fundamentals, the level of the driving experience (IV1) has a direct impact on the motor insurance premium (DV), due to the negative relationship of the effect to the perceived risk. Increasing driver's experience level decreases the perceived risk, resulting in a proportional decrease in the risk premium. The history of previous accidents (IV2) has a direct positive relationship to the perceived risk, which means when the number of previous accidents is being increased, the motor insurance premium is also being increased proportionally.

⁸ Derived from latent risk (hidden or unobserved) and markov process (future depends on the current state)

Evaluating the Impact of Driving Experience and the Previous Accident History on Motor Insurance Premium



Conceptual Framework Developed by the Researcher

Figure 4 Conceptual Framework designed by the Author

It also shows that IV1 and IV2 are not the only independent variables affecting the DV; other controlled variables include the driver's age, gender, vehicle age, mileage, location, type, make and model, and usage. However, in the framework it illustrates how the variables react to each other based on risk based principles and the experience-based pricing principles.

3. Methodology

3.1 Research Philosophy

This quantitative research of evaluating the impact of driving experience and the previous accident history on the motor insurance premium follows the discipline of positivism (Guba & Lincoln, 2005) and find the objective reality of how the variables influencing on each. The study will conduct a statistical review to identify how the independent variables are influencing on the dependent variable with the observed measurable data by developing hypotheses (Zumtavan, 2026) to investigate the relationship between variables. Positivism opens the screen to investigate objectively, how insurers apply risk based and experience based pricing for determining the motor premium by using the quantitative analysis.

3.2 Research Approach

This study is approached by the existing theories and knowledge such as risk theory and the EPT; hence the study adopts a deductive approach (Ali, 2024). The hypotheses developed on the theoretical background has been analysed by SPSS techniques. It also supports to accept or reject the developed hypotheses based on the existing empirical facts and findings.

3.3 Research Design

This study analyse numerical data and statistically investigating the relationship with the variables by using SPSS applications in a quantitative research design (Fischer, 2023). The cross-sectional data (Hunziker, 2024) set gathered at a single time point from various drivers with different experience level may enable to compare the relationship of the driving experience level and the previous accident occurrence on the premium saving levels. The data is analysed by using SPSS tools to compare the research hypotheses objectively. The reliability and validity of the data set is the first assessed by using Cronbach's Alpha (Ahmad, et al., 2024) test to verify that the selected data set is appropriate for the study purpose.

3.4 Data Collection Method

In order to conduct the research in cost effective and time efficient manner it has gathered information by secondary data collection method (Ajayi, 2023). The dataset with 500 cases have been extracted from Kaggle (www.kaggle.com), the publicly accessible online platform which provides averagely reliable dataset on academic, research purposes and machine learning purposes.

3.5 Population and Sample Size

The dataset contain 500 cases with motor policyholders, inspired by real word facts, which provides adequate amount of data (Wutich, 2024) for statistical analysis in quantitative data minimising sampling error (Harding, 2014). In order to feed SPSS, the dataset has been recoded into categories such as driving experince as five levels of experience ranging from level 1(L1) as the lowest to level 5 (L5) the highest experience level. The accident history has been recoded into two categories such as drivers without and with previous accident histroy as No=0 and Yes=1respectively. The motor insurance premium has been recoded according to the five premium saving levels.

3.6 Variables

In the study there are two independent variables (IV) (LeLAURIN & Wolery, 1992); driving experience (IV1) and the previous accident history (IV2). As discussed in section 3.5, these variables have been recoded to categorical data represnting the driver's experince level and the previous accident history as an event of occurrence. The level of driving experince has been analysed by using the one-way ANOVA to investigate whether insurance premium is differed on the level of experince groups. The previous accident history has been analysed by using independent sample t-Test to investigate wether the previous accident history can make a difference on the motor insurance premium. The motor insurance premium considered as the dependant variable (DV) (Bosch & Revilla, 2022) to invesitage the premium amount which is influenced by the driver's experience level and the frequency and severity over past accident history.

3.7 SPSS Tools and Statistical Tests in Data Analysis

To analyse the relationship between the level of driving experience and the occurence of the previous claim history on the level of premium saving would cost the policyholder. Therefore, the researcher has used the below-pointed SPSS analytical tests and tools to investigate the research objectives.

- A) Descriptive Statistics – the study has calculated the mean, minimum, maximum, standard deviation, and variance as its measures. It has also calculated the average experience level and the average premium saved by the policyholder. The minimum value represents the lowest boundary of the level experience and the premium saving range in the recorded data set. The maximum value represents the upper boundary of the level of experience and the

saved premium range of the given dataset. The standard deviation measures the variability of the level of experience and the range of saved premium from the mean values of the respective variables. The variance measured the squared deviation of data values from each mean value by searching the degree of the dispersion of motor insurance premiums.

- B) One-way ANOVA – This tool measures the significant differences in saved premium ranges across different experience level groups. This tool is more appropriate for analysing IV1, the level of driving experience, because it has been grouped into five different levels, allowing for measurement of its effect on the premium saving range. If the p-value is less than 0.05, it indicates a statistically significant difference in saved premium ranges across the respective driving experience levels, while a p-value greater than 0.05 indicates no significant difference. A Tukey HSD post hoc test has been conducted to analyse which group of driving experience level is most significant on each saved premium range.
- C) Independent Sample t-Test – This tool has been used to examine the significant difference in the saved premium range between the groups with previous accident history. The independent variable IV2, which represents previous accident history, consists of two groups: No (0) and Yes (1); therefore, the independent sample t-test is suitable for this analysis. If the p-value is greater than 0.05, it indicates there is no significant difference in the mean premium of the saved premium range between the groups with different previous accident histories; conversely, if the p-value is less than 0.05, it indicates that a significant difference exists between the groups.

3.8 Hypotheses Development

In order to evaluate the impact of driving experience and the previous accident history on the motor insurance premium the researcher has developed below hypothesis (Signoretta, et al., 2014) or statements predicting the relationship between variables. These hypothesis have been developed according to the testing tool.

A) Driving experience level and the motor insurance premium

H_0 – There is no significant difference in motor insurance premiums in the groups of different driving experience levels.

H_1 – There is a significant difference in motor insurance on the groups of different experience levels.

The one-way ANOVA was used to test the hypothesis since the data is categorical and has more than two groups

B) Previous accident history and the motor insurance premium

H_0 – There is no significant difference in motor insurance premium on the previous accident history

H_1 – There is a significant difference in motor insurance premium based on the previous accident history.

The hypothesis was tested to investigate the premium difference of motor insurance rates for drivers with and without accident history. Since there are two independent groups, such as with and without accident history, the independent t-test has been used.

3.9 Ethical Consideration

Ethical consideration is one of the most key criteria of a professional research and it represent the scientific ethics applied in the research work. Ethical considertaion in methodology is maintaining truthfulness, reporting data, privacy and confidentiality (Hasan, et al., 2021). In this research secondory data (Kaggle dataset) has been used and no direct involvement in human activitis. Further the dataset contains general numerical information and no personal information is gathered. The raw data has been used for necessary testing tools without manipulating on study purposes and honesty and the transparency have been maintained.

4. Data Analysis and Results

4.1 Reliability testing

As we gather secondary data from a source already publicly available, it is important to test the reliability of the dataset in order to evaluate the consistency and the stability of the dataset. In this quantitative study the reliability is tested by using the SPSS tool Cronbach's Alpha (Amirrudin, et al., 2021).

Case Processing Summary				Reliability Statistics		
		N	%			
Cases	Valid	500	100.0	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
	Excluded ^a	0	.0			
	Total	500	100.0			
a. Listwise deletion based on all variables in the procedure.				1.000	1.000	2

Two variables have been performed for the test: groups of driving experience levels and the levels of premium saving groups for 500 cases. In this study the reliability test performed as 1.000 which suggests the dataset is appropriate for the purpose of the study and the perfect match between the variables measured.

4.2 Descriptive Statistics

Three main recoded variables; driving experience (categorised levels of driving experiences), level of savings from the insurance premiums and the previous accident history show a high level of correlation among each.

Descriptive Statistics										
	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Variance Statistic	Skewness Statistic	Std. Error	Kurtosis Statistic	Std. Error
Level of SAVING frm Discounted IP	500	1	4	2.07	1.047	1.096	.489	.109	-1.043	.218
driverexp_code	500	1	4	2.07	1.047	1.096	.489	.109	-1.043	.218
PreviousAccidents_code2 (binary)	500	0	1	.85	.359	.129	-1.944	.109	1.788	.218
Valid N (listwise)	500									

A) Level of premium savings & level of driving experience

The above variables demonstrate statistically identical behaviour; therefore, the data interpretation is valued for both variables. The Likert scale ranges from 1 (minimum) to 4 (maximum), with premium savings categorised between £478.00 and £508.15, resulting in a premium difference of £30.15. Accordingly, the premium savings have been grouped, such as level 1= £500 or more, level 2= £491 – £500, level 3 = £481 – £490, level 4= £471- £480. The levels of driving experiences have been grouped as 1 = L1 {0-13year }, 2 = L2 {14 -20 years}, 3 = L3 {21-34 years}, 4 L4 {35 – 40 years}. A mean value of 2.07 indicates that the average score in the data set just passed the halfway point of the scale. In the dataset the majority of the premium savings are grouped between £491-£500 and the majority of the drivers have 14 to 20 years of driving experience compared to the available data. The spread is measured by Std. Deviation=1.047 and the Variance = 1.096, which interprets that the moderate amount of variation among 500 drivers is 0.1 away from the mean value. Skewness = 0.489 (slightly positive skew) interprets that higher end of the scale has a longer tail and the lower end is concentrated between 1 and 2. Kurtosis = -1.043 indicates that the distribution has a thinner tail and flatter peaks, meaning the responses are evenly distributed among the 1 to 4 scores.

B) Previous accident history (binary coding)

The range of previous accident history is binary as 0 or 1. Generally 0 represents drivers without accident history, and 1 represents drivers with accident history. The mean = 0.85 interprets that 85% of the responses (425 out of 500 drivers) have previous accident history, and only 15% of the responses have no previous accident history. The Std. Deviation = 0.359 and Variance = 0.129 represent that vast majority of the responses scattered around score 1.

4.3 One-way ANOVA with post hoc tests

In this test, the level of driving experience and the premium paid recoded information have been counted as variables.

A) Group descriptives

In the group descriptives, out of 500 numbers of drivers, the largest group represents L1 (n=199) and the smallest group L4 (n=59). The mean values indicate that the average score of the

premium paid has a consistent effect across different levels of driving experience, with scores of L1=1.03, L2=2.00, L3=3.00, and L4=3.92 [refer to appendix 1].

B) Homogeneity of variance

Based on the Levene Statistic measures on the “Based on mean” and “Based on trimmed mean” have significance (Sig.) if <0.001 and this happened when the group has no equal variance. Therefore, in post hoc evaluation the Tukey HSD has been ignored, and the Games-Howell test has been considered, which is specifically designed for unequal variance [refer to appendix 2].

C) ANOVA table

This table illustrates whether the mean value of each group has a statistically significant difference in the premium paid. The result shows $F(3, 496) = 5196.288$, Sig. $< .001$: Since the p-value is well below 0.05, this indicates that the result is highly statistically significant, meaning that the premium paid differs significantly based on the level of driving experience.

ANOVA					
PREMIUM PAID_CODE					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	515.293	3	171.764	5196.288	<.001
Within Groups	16.395	496	.033		
Total	531.688	499			

D) ANOVA effect size

ANOVA effect illustrates how much the variance in the premium paid code reflects the driver's experience level, resulting in an Eta-squared of .969 (with Omega-squared fixed-effect also at .969). This large effective size explains that 96.9% of the variance of the premium paid code can be reflected in the driver's experience level [refer appendix 3].

E) Post-hoc tests / multiple comparisons

As a result of the significant overall ANOVA outcome, the multiple comparison has been conducted to evaluate which pair of groups differentiated from one another. Since the variance is unequal, the Tukey HSD has been ignored, and the Games-Howell test has been considered. Refer to appendix 4 for the result illustration. In the results every single pair was statistically significant

($p < 0.001$) and among groups: L1 vs. L2 – L2 was significantly higher than L1 (mean difference = 0.970), L2 vs. L3 – L3 was significantly higher than L2 (mean difference = 1.000) and L3 vs. L4 – L3 was significantly higher than L3 (mean difference = 0.915). Please refer to appendix 5 for the means plot illustration.

4.4 Independent Sample t-Test

The level of premium savings has been compared between two groups: group 1 = drivers with no previous claim history and group 2 = drivers with accident history. According to the result, there are $n=76$ participants who have no previous accident history out of 500 drivers, and 424 have previous accident history.

A) Group statistics

ANOVA effect illustrates how much the variance in the premium paid code reflects the driver's experience level, resulting in an Eta-squared of .969 (with Omega-squared fixed-effect also at .969). However the mean value of the “with previous accident” results slightly higher and the independent sample t-Test has been tested for the deeper understanding [refer appendix 6].

B) t-Test table

The t-Test table consist of three segments; Levene's Test for Equality of Variances, t-Test for equality means and confidence interval respectively. In the Leven's test resulting Sig.=0.380 indicates that between the two groups the variance is equal. The $p\text{-value} > 0.05$ interprets that the null hypothesis has been failed to reject.

Independent Samples Test										
Levene's Test for Equality of Variances					t-test for Equality of Means					
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference
						One-Sided p	Two-Sided p			Lower Upper
Level of SAVING frm Discounted IP	Equal variances assumed	.771	.380	-1.431	498	.076	.153	-.186	.130	-.442 .069
	Equal variances not assumed			-1.541	110.925	.063	.126	-.186	.121	-.426 .053

The t-test with equal variance shows a t-statistic (t) of -1.431, degrees of freedom (df) of 498, and a two-sided p-value (Sig. 2-tailed) of 0.153. This result interprets the standard threshold for significance as $\alpha = 0.05$. And the two-sided p-value = 0.153 is > 0.05 , the difference between the two groups is not statistically significant. Even the one-side p value = 0.076 is not statistically significant. In the 95% confidence interval for the mean difference, which ranges from -0.442 to

0.069, the interval has passed zero; it further confirms that there is no true difference that exists between the two groups in the population.

C) Independent sample effect size

The table interpret how large the difference is, regardless of sample size, Cohen's $d = -0.178$ and according to standard conventions, an effect size around 0.2 is considered which is small A value of -0.178 indicates a minimal effect, reinforcing that any observed difference between the two groups is practically negligible [refer to appendix 7].

5 Discussion

5.1 Interpretation of the Level of Driving Experience

The results from the one-way ANOVA revealed that the level of driving experience and the saving of the insurance premium levels are highly significant. The mean values over four groups of driving experience behave significantly different, reporting the p-value is <0.001 . This information simply indicates that the level of driving experience has a major impact on insurance premium pricing. The descriptive statistics interpret the mean values of the premium saving over the level of driving experience, which has been progressively increased from L1 to L4. The lowest level of driving experience in the L1 {0-13 years} is related to the recoded group with the lowest savings on insurance premiums, and the highest driving experience level L4 {31-40 years} is related to the highest level of premium saving. Further, the result of the Game-Howell post-hoc analysis confirmed that every pair within the groups of driving experience levels is statistically significant. The overall findings of the one-way ANOVA indicate that insurers do practise systematically different pricing strategies according to the level of driving experience. The eta-squared value of 0.969 indicates a large effect, accounting for 96.9% of the variance in savings from recoded premium pricing. Finally, according to the dataset, the driving experience has a high impact on the premium pricing.

The findings endorse the risk theory, that insurance premiums reflect the probability and severity of future risks. By applying the risk theory, insurers are evaluating the characteristics of the insured which could be leading to future losses. The insurers predict that experienced drivers possess a high level of road familiarity, improved understanding of the hazard, improved driving dexterity and low accident probability. Accordingly, the findings direct insurers to practise a risk-based pricing strategy in calculating motor insurance premiums.

The final outcome reflects the relevance to the empirical findings as well. As (Henckaerts & Antonio, 2022) explains that driver's behavior and the driving-related risk has a high impact on the insurance pricing and the final finding of this study also align with the findings of this empirical paper. Futher (Cevolini & Esposito, 2022) emphasises that insurers do heavily depend on the behavior and experience based indicators when systemise the motor insurance premiums.

Therefore, the present findings demonstrate the relevance with the risk theory that the driving experience has a high relationship of determining the motor insurance premium and H_1 is accepted.

5.2 Interpretation of Previous Accident History

The significant differences between the two groups of “without accident history” and “with accident history” has been evaluated by using the independent sample t-test in concern of the insurance premium saving levels. The mean value of the group “with accident history” = 2.09 is slightly higher than the value of “without accident history”=1.91. However, the independent sample t-test readings $t = -1.431$, $df = 498$, and $p = 0.153$ statistically indicate that the significant difference between the two groups is low. The p-value of the significance is > 0.05 and it interprets that the null hypothesis (H_0) is unable to be rejected. According to the sample population, the study does not discover substantial evidence to support H_1 . Further, the 95% confidence interval ranging from -0.442 to 0.069 indicates that there is no proper population difference between the two groups, and the Cohen’s d effect size = -0.178 illustrates that the difference between the two groups is small and negligible; hence, the H_0 could be accepted.

The original work of this study does not support the EPT. According to the theory, the previous accident history has a direct impact on the future claim prediction. This final outcome may indicate or assume several possibilities. From the insurer's perspective, the dataset may be more reliable to make decisions upon the impact of the driving experience level on the premium pricing rather than considering the previous accident history. The recoded categories of premium saving levels may not capture the actual premium differentiation in the dataset, or the dataset may contain characteristics of more homogeneous accident history. Furthermore, the fact that 85% of the drivers reported having an accident history limited the variation between the two groups, which may reduce the t-test's ability to detect a statistically significant difference. Therefore, the final outcome of this study does not sufficiently support the conceptual idea of the experience pricing theory.

5.3 Comparison of previous studies and the final outcome

The outcome of the one-way ANOVA finds a statistically significant relationship between the level of driving experience and the savings of motor insurance premiums which is strongly consistent with the findings of (Henckaerts & Antonio, 2022) and (Cevolini & Esposito, 2022). The both studies emphasise the driving behaviour and experience have a direct impact on the

premium determination. Accordingly to the previous studies, insurers assess and classify drivers according to the characteristics of the behaviour and experience of the drivers.

The outcome of the previous accident history is slightly deviate from the the previous studies. According to (Verschuren, 2022) and several prior studies, argue that the insurance premium is penalised by the record of accident histroy. These findings align with the EPT to predict the future loss based on the history of frequency and the severity of the previous accident. The original finding of this study slightly deviate from the key theoretical and the similar empirical evidence due to various reasons. The unusual precence of 85% drivers with accident history in the dataset has a direct influence on the final finding. Apart of this the researcher believes there could be an effect of the coding structure and the sample distribution as well on the outcome of the previous accident histry. However, even the negligable influence of the relationship of previous accident history on the insurance premium interprets there is statistically significant relationship between two variables although it is not dominant on the factors of insurance premium pricing. Further, the researcher personally believes the financial impact on the differnce of the accident history and the claim histroy may have a influence on the true premium which may cause a slight variance of the significant differnce between two groups.

5.4 Practical Implications

The findings highlight some practical implications for insurers' practice of underwriting the motor premium. The final outcome of the relationship between the driving experience and the motor premium further emphasises, in the acturial or underwriting practice, the experience-based pricing upgrade of the accuracy of the risk rate. Further, insurers may use the true picture of the variance from low to high risk drivers to design their business development strategies, including marketing activities. It proves the risk-based pricing strategies are more appropriate for pricing motor premiums. Even though the study outcome of the previous accident history does not support or align with the current practical premium pricing strategies, insurers may still rely on the majority of the empirical evidence and the theoretical perspectives that positively support the concept. Furthermore, the test could be repeated with a different dataset to verify and validate the hypothesis. Furthermore, insurers may consider other demographic factors such as the driver's age, car make and model, annual mileage and previous claim history.

6 Recommendations & Conclusion

6.1 Contribution of the study

This quantitative statistical study contributes several key findings to the importance of academia and industry as well. The outcome strengthens both the existing empirical knowledge and the current practice of risk-based pricing strategies. At the same time, it encourages underwriters to consider and investigate the impact of various demographic characteristics of drivers, as these factors can influence premium pricing. The theoretical framework and the SPSS tools used in this study, including one-way ANOVA and independent sample t-tests, may be applied in future academic research to uncover additional relevant information regarding risk-based insurance and user-based insurance applications.

From the insurer's perspective, the outcome leads to thinking of the importance of the risk-based and the experience-based premium pricing and how influential it is in shaping the competitive advantage of the product. The level of driving experience indicates how insurers can focus their marketing activities and new product development on capturing untapped market segments based on the dynamics of years of driving experience. Further, regulators also may refer to the findings of this study to determine the regulatory influences on underwriting practices when imposing excesses and policy terms and conditions to maintain transparency and fairness in motor insurance pricing.

6.2 Future Recommendation

Primarily, there are several recommendations that could be introduced for insurance companies. Firstly, insurance companies should enhance their experience-based underwriting systems, specifically by refining the characteristics used in new risk classification models to analyse how driving experience influences premium pricing strategies. Secondly, to improve the pricing accuracy, insurers should improve predictive data analysis models that have been integrated with drivers' behaviour, telematics data and driving patterns. This exercise is expected to be effective in future risk assessment models and useful in developing resilience strategies. Thirdly, insurers may revisit their safer driving incentives to be competitive with the market by introducing more attractive NCD schemes, premium discounts and driver loyalty programmes.

Secondarily, there are several recommendations that could be introduced for the policymakers or the regulators. This study reveals that driving experience significantly influences motor insurance pricing, and the terms and conditions of motor policies should be fairer for both insurers and the insured. Therefore, regulators could closely monitor inappropriate premium discrimination and the loading of unfair excesses. Further, regulators may promote more technically advanced digitally operated risk evaluation systems to improve the accuracy of the premium pricing. Also, regulators and insurance companies can collaborate to develop high-quality risk evaluation models that identify accident-related risks and predict future anticipated losses accordingly.

6.3 Limitations of the overall study

Finally, the study has managed to achieve its research objectives and find out the reality for the research questions successfully. However, the study has several limitations that should be noted for future improvements. The large number of drivers, such as 500, would be sufficient according to the research sample rules, but the generalisability may fall into statistical pitfalls, causing hypersensitivity of data. Another specific concern is that the dataset obtained from Kaggle is not exactly appropriate for the study in terms of data quality, variable definitions and coding structures. Comparing to the available dataset and the possible variables, it has been tested only for the driving experience and the previous accident history, whereas other demographic variables such as the driver's age, car make and model, and the annual mileage also could be considered to test the impact on the motor insurance premiums. The premium paid had to be recoded and converted into categorical data otherwise, the variables were flowing in opposite directions, which was not supported for the reliability test, creating a negative indicator. The study could also have used regression analysis to identify the complexity of the relationship between variables.

6.4 Conclusion

This quantitative study was to evaluate the impact of the driving experience and the previous accident history on the motor insurance premium by using the statistical software SPSS. The study has been conducted using the secondary dataset publicly available on Kaggle.com, and the descriptive statistics, one-way ANOVA and independent sample t-test have been performed to investigate the scenario. The driving experience and previous accident history are considered dependent variables, and each has been analyzed under the hypothesis to determine their impact

using the dataset. The study has been supported by the context of risk theory and the EPT. From the findings it has been emphasised that the level of driving experience is highly significant in the premium determination and the previous accident history has no significant relationship. Therefore, the risk theory was supported by the final outcome, but the experience pricing theory was not supported as assumed. Overall the findings confirmed that the insurance pricing is strongly influenced by a risk-based classification system, and it highlights the importance of a data-driven underwriting system as well to increase the accuracy of the risk rating and the insights for the whole industry.

Reflexive Report

This industry-specific quantitative research has provided me with invaluable academic and practical learning experiences by evaluating the impact of the driving experience and previous accident history on the motor insurance premium. It allowed me to engage with a beneficial practice of SPSS software by using the various data analysis tools, statistical data interpretations and positivist research methodology related to the motor insurance market. Throughout the study, I had the opportunity to gain several insights related to quantitative research methods, the discipline of positivism, and the academic and practical challenges of conducting a quantitative study while maintaining the DBA (level 8) standard.

The primary objective of this exercise is to become familiarised with the statistical data analysis by using the SPSS tools while identifying the most appropriate testing tool. The principal learning outcome of this study is to prepare the secondary dataset to align with the research objectives through accurate coding, which was moderately rigorous. The original dataset available on Kaggle was a raw set of data with 1000 responses. Assuming that a large number of responses could lead to statistical issues related to hypersensitivity, the dataset was filtered down to 500 cases. The next challenging task was to recode the variable accurately with the correct categorical scale, as originally the variables operated in opposite directions, causing the reliability test, Conbach's Alpha, a negative reading. However converting the selected variables for the study the driving experience (IV1) was coded into 4 levels of driving experiences and transformed into categorical data and the previous accident history (IV2) was coded into binary variable as two groups derived as 'with accident history' and 'without accident history'. The dependant variable (DV), the premium paid, was coded as the levels of premium saved to match the scale direction by the relationship to each variable. For example, individuals with higher levels of experience perceive a lower risk in the insurance premium, and the adjustments for the variables were made carefully to ensure accurate alignment with the statistical interpretation.

. This process highlights the most important competency in the SPSS analysis, as even the smallest inconsistency in data may lead to misconduct of one-way ANOVA and independent sample t-test interpretation. Through this activity I learnt that SPSS analysis is not merely a number analysis but also a vast scope of logical reasoning and careful operation of variables.

At the beginning of this study, due to the knowledge gap I had in terms of SPSS tools such as one-way ANOVA and independent sample t-tests, the functions such as descriptive data analysis, effect size analysis, skewness and kurtosis, and homogeneity variance were conceptually dense. However, engaging with this study allowed me to better understand the main applications of each tool and their usability in the statistical data analysis and how to interpret the outcome of each test. At one point, the Levene's test for equality of variance became particularly important when unequal variances were identified among the groups in the one-way ANOVA analysis, and I learned that in such cases, the Tukey HSD post-hoc test should be replaced with the Games-Howell test. This made me realise the importance of selecting appropriate data analysis technique according to the characteristics of the dataset.

Initially the interpretation of the test results proved challenging for me, as I had a perception that quantitative data is full of numbers. As I progressed with the exercise, I realised that each number and result not only has a deeper meaning in its interpretation, which significantly impacts the final outcome of the findings, but also that I learned how to link these findings with the theoretical concepts of risk theory and experience pricing theory. I believe that quantitative research plays a crucial role in DBA studies. As the programme is structured to align more closely with the corporate world, aspects such as business decisions, operational performance analysis, financial analysis, risk management, and data-driven decision-making are fundamentally based on quantitative research and its outcomes. Furthermore, given my focus on the insurance sector, quantitative research is highly relevant in various areas, including risk evaluation, assessments, strategy development, performance management, customer experience evaluation, enterprise and risk management, and new product development, all of which rely on quantitative research findings. Therefore in my personal belief as a DBA student, the quantitative methodologies and the measurable objective findings may lead me to make evidence-based managerial decision-making that supports organisational strategies and policymaking.

Quantitative research allows one to analyse the data very closely and to understand better the idea of the patterns and the relationships of the variables applied. These kinds of numeric evidences are highly useful in the corporate sector to design pricing strategies, risk models and operational policies. Additionally, it is important to have a clear understanding of how to use a secondary dataset that aligns with our research objectives. The main limitation is that available

data may hinder the real scope of the research aims and objectives and the variable selection; adding new variables and giving definitions for them could be more challenging. In addition to drinking experience and accident history, demographic variables such as age, gender, vehicle type, vehicle usage, and past claim history may also significantly influence the motor insurance premium, as indicated by measurable findings in this quantitative study. Therefore, the quality of the outcome of the research may be unsubstantiated.

If primary data had been collected through the design of a comprehensive questionnaire, the responses would have been more relevant to the study and may have yielded more accurate findings. Furthermore, quantitative findings can only interpret the phenomenon to a certain extent, as numbers have limitations in their interpretability. I believe that for DBA-level programmes, it is essential to explore the deeper meanings of individual perceptions and other influencing factors. Therefore, I suggest that a pragmatic or mixed-method research methodology would be more suitable for a DBA-level thesis.

Overall, this study allowed me to improve my competencies in SPSS statistical analysis and enhance my skills in interpreting research findings related to the topic, while also supporting the relevant theories.. The study has also given me the freedom to maintain the originality of the work even when the previous accident history outcome was not aligned with the previous empirical results and the context of the experience pricing theory. I have learnt about the general challenges, limitations, advantages and disadvantages of a quantitative study and how to reinforce the quantitative methods with the scope of the DBA programme and in decision-making.

Bibliography

- Acclavio, M. & Maieli, R., 2024. Logic Programming with Multiplicative Structures. *arXiv preprint arXiv:2403.03032*.
- Aeron-Thomas, A., 2022. The role of the motor insurance industry in preventing and compensating road casualties. *Scoping Study Final Report*.
- Ahmad, N., Alias, F. A., Hamat, M. & Mohamed, S. A., 2024. Reliability analysis: application of cronbach's alpha in research instruments. *Pioneering the Future: Delving Into E-Learning's Landscape*, pp. 114--119.
- Ahmed, D., Yuantao, X. & Saeed Bhutta, U., 2022. Insurers' risk management as a business process: a prospective competitive advantage or not?. *European Journal of Management and Business Economics*, 31(3), pp. 345--366.
- Ahmmed, M. J., 2025. Systematic Review and Quantitative Evaluation of Advanced Machine Learning Frameworks for Credit Risk Assessment, Fraud Detection, And Dynamic Pricing in US Financial Systems. *International Journal of Business and Economics Insights*, 5(3), pp. 1329--1369.
- Ajayi, V. O., 2023. A review on primary sources of data and secondary sources of data. *Universal Journal of Arts, Humanities, Science and Technology Education Research RESEARCH ARTICLE* , 2(1).
- Alcoforado, R. G., 2022. Theoretical and practical advances in actuarial science: risk, ruin theory and life pensions. *PQDT-Global*.
- Ali, I. M., 2024. A guide for positivist research paradigm: From philosophy to methodology. *Ideology Journal*, 9(2).
- Alles, A., 2026. *Insurance Mastery: Essential Knowledge for Insurance Practitioners*. Colombo: s.n.
- Amirrudin, M., Nasution, K. & Supahar, S., 2021. Effect of variability on Cronbach alpha reliability in research practice. *Juornal Matematika, Statistika dan Komputasi*, 17(2), pp. 223--230.

- Badescu, S. { V. L. & Lin, X. S., 2024. Effective experience rating for large insurance portfolios via surrogate modeling. *Insurance: Mathematics and Economics*, Volume 18, pp. 25-43.
- Boomhower, J., Fowlie, M., Gellman, J. & Plantinga, A., 2024. *How are insurance markets adapting to climate change? risk selection and regulation in the market for homeowners insurance*. s.l.:National Bureau of Economic Research Cambridge, MA, USA.
- Bosch, O. J. & Revilla, M., 2022. Is tracking all that it takes?. *Exploring the validity of news media exposure measurements created with metered data*. Available at, Volume 10.
- Cao, S. et al., 2022. Hazard perception in driving: A systematic literature review. *Transportation research record*, 2676(12), pp. 666--690.
- Cevolini, A. & Esposito, E., 2022. From Actuarial to Behavioural Valuation. The impact of telematics on motor insurance. *Valuation Studies* , 9(1), pp. 109-139.
- Chelton, D. B., 1983. Effects of sampling errors in statistical estimation. *Deep Sea Research Part A. Oceanographic Research Papers*, 3(10), pp. 1083--1103.
- Clemente, C., Guerreiro, G. R. & Bravo, J. M., 2023. Modelling motor insurance claim frequency and severity using gradient boosting. *Risks*, 11(9), p. 163.
- Coutts, S. M., 1984. Motor insurance rating, an actuarial approach. *Journal of the Institute of Actuaries*, 111(1), p. 87--148.
- Denuit, M., Mar{\e}chal, X., Pitrebois, S. & Walhin, J.-F., 2007. *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*. s.l.:John Wiley & Sons.
- Dionne, G., Fombaron, N. & Mimra, W., 2024. Adverse selection in insurance. In: *Handbook of Insurance: Volume II*. s.l.:Springer, pp. 165--221.
- EBSCO, 2021. *Risk Theory and Credibility*. [Online]
Available at: <https://www.ebsco.com/research-starters/business-and-management/risk-theory-and-credibility>
[Accessed 16 May 2026].
- Englund, K. & Martin-Löf, A., 2001. Ernest Filip Oskar Lundberg. In: *Statisticians of the Centuries*. New York: Springer, pp. 308--311.

Engstrom, N. F., 2011. An Alternative Explanation for No-Fault's Demise. *DePaul L. Rev.*, Volume 61, p. 303.

Fackler, M., 2025. Premium rating without losses: How to estimate the loss frequency of loss-free risks. *European Actuarial Journal*, 12(1), pp. 275--316.

Fischer, H. E. a. B. W. J. a. N. K., 2023. Quantitative research designs and approaches. In: *Handbook of research on science education*. s.l.:Routledge, pp. 28-59.

Fiveable, 2026. *Fiveable*. (2026, March 4). 5.3 Aggregate loss distributions and stop-loss reinsurance – Actuarial Mathematics. *httaggregate-loss-distributions-stop-loss-reinsurance*. [Online]

Available at: <https://fiveable.me/actuarial-mathematics/unit-5/aggregate-loss-distributions-stop-loss-reinsurance/study-guide/z5N3FDreKMAxSkyr>
[Accessed 16 May 2026].

Gill, K. A., 2021. *Insurance fraud: causes, characteristics and prevention*. s.l.:University of Leicester (United Kingdom).

Guba, E. G. & Lincoln, Y. S., 2005. *Paradigmatic controversies, contradictions, and emerging confluences*. s.l.:Sage Publications.

Harding, B. a. T. C. a. C. D., 2014. Standard errors: A review and evaluation of standard error estimators using Monte Carlo simulations. *The Quantitative Methods for Psychology*, 10(2), pp. 107--123.

Hasan, N. et al., 2021. Ethical considerations in research. *Journal of Nursing Research, Patient Safety and Practise*, 1(1), pp. 1-4.

Henckaerts, R. & Antonio, K., 2022. The added value of dynamically updating motor insurance prices with telematics collected driving behavior data. *Insurance Mathematics and Economics*, 105(2).

Hunziker, S. a. B. M., 2024. Cross-sectional research design. In: *Research design in business and management: A practical guide for students and researchers*. s.l.:Springer, pp. 187--199.

Hu, W., Sun, L. & Cheng, L., 2026. Effects of confidence level and hazard type on the visual search patterns and hazard response times of young drivers. *Traffic Injury Prevention*, 27(3), pp. 365--371.

Ieosanurak, W. a. K. B. a. M. A., 2023. Claim modeling and insurance premium pricing under a bonus--malus system in motor insurance. *International Journal of Applied Mathematics and Computer Science*, 33(4), pp. 637--650.

Isimoya, O. A., Ajemunigbohun, S. S. & Osasona, A. V., 2022. MOTOR INSURANCE POLICIES'UNDERWRITING FACTORS: EXPLORATORY ANALYSIS FROM MOTOR INSURANCE PROVIDERS IN NIGERIA. *Economic \& Managerial Spectrum/Ekonomicko-mana{\v{z}}{\e}rske Spektrum*, 16(2).

Isimoya, O. A., Ajemunigbohun, S. S. & Osasona, A. V., 2022. MOTOR INSURANCE POLICIES'UNDERWRITING FACTORS: EXPLORATORY ANALYSIS FROM MOTOR INSURANCE PROVIDERS IN NIGERIA. *Economic \& Managerial Spectrum/Ekonomicko-mana{\v{z}}{\e}rske Spektrum*, 16(2).

Isimoya, O. A., Ajemunigbohun, S. S. & Osasona, A. V., 2022. MOTOR INSURANCE POLICIES'UNDERWRITING FACTORS: EXPLORATORY ANALYSIS FROM MOTOR INSURANCE PROVIDERS IN NIGERIA. *Economic \& Managerial Spectrum/Ekonomicko-mana{\v{z}}{\e}rske Spektrum*, 16(2).

Joeveer, K. & Kepp, K., 2023. What drives drivers? Switching, learning, and the impact of claims in car insurance. *Journal of Behavioral and Experimental Economics*, Volume 103, pp. 101-993.

Johnson, G. M., 2025. The hard proxy problem: proxies aren't intentional; they're intentional. *Philosophical Studies*, 182(5), pp. 1383--1411.

Kaggle, 2025. *car-insurance-premium*. [Online]
Available at: <https://www.kaggle.com/code/govindaramsriram/car-insurance-premium/input>
[Accessed 15 May 2026].

Kaushish, L. & Patnaik, P., 2025. Re-Evaluating the Principle of Utmost Good Faith in Indian Insurance Law: A Case for Legislative Reform. *LawFoyer Int'l J. Doctrinal Legal Rsch.*, 3(1).

- Khairiati, A., Budiarti, R. & Najib, M. K., 2025. Analysis of Ruin Probability in Insurance Companies Using the Cramer-Lundberg Model with Variations in Claim Distributions. *ZERO: Jurnal Sains, Matematika dan Terapan*, 9(2), pp. 386--395.
- Kochenburger, P. & Salve, P., 2023. An introduction to insurance regulation. In: *Research handbook on international insurance law and regulation*. s.l.:Edward Elgar Publishing, pp. 247--280.
- Lawrence, W. et al., 2023. Parametric insurance. In: *The global insurance market and change*. s.l.:Informa Law from Routledge, pp. 127--156.
- Lee, W.-H., Hsiao, C.-L. & Wen, K.-L., 2025. Promoting Sustainability in Motor Insurance: A Novel Approach To Evaluating Driving Behavior. *Information Systems Frontiers*, pp. 1-13.
- LeLAURIN, K. & Wolery, M., 1992. Research standards in early intervention: Defining, describing, and measuring the independent variable. *Journal of Early Intervention*, 16(3), pp. 275--287.
- Maggioni, M. & Turchetti, G., 2024. Solvency of an Insurance Undertaking. In: *Fundamentals of the Insurance Business*. s.l.:Springer, pp. 601--631.
- Manurung, T., 2025. Distribution Models of Claim Frequency and Claim Severity in Determining the Pure Premium of Car Insurance with the Application of a Deductible. *ZERO: Jurnal Sains, Matematika dan Terapan*, 9(2), pp. 662--670.
- Mehrotra, R. & Chadha, S., 2025. Macroeconomic Resilience to Socioeconomic Challenges and Corporate Stability in the Automotive Industry: Impact on Dividend Policy. *SocioEconomic Challenges*, 9(2), pp. 180--189.
- Meng, S., Gao, Y. & Huang, Y., 2022. Actuarial intelligence in auto insurance: Claim frequency modeling with driving behavior features and improved boosted trees. *Insurance: Mathematics and Economics*, Volume 106, pp. 115-127.
- Mildenhall, S. J. & Major, J. A., 2022. *Pricing insurance risk: Theory and practice*. s.l.:John Wiley & Sons.

Mildenhall, S. J. & Major, J. A., 2022. *Pricing insurance risk: Theory and practice*. s.l.:John Wiley & Sons.

Mordor Intelligence, 2026. *GLOBAL MOTOR INSURANCE MARKET*. [Online]
Available at: <https://www.mordorintelligence.com/industry-reports/global-motor-insurance-market>
[Accessed 14 May 2026].

Moriah, M., Vermet, F. & Charpentier, A., 2023. Measuring and mitigating biases in motor insurance pricing. *arXiv preprint arXiv:2311.11900*.

Mwigilwa, W., 2024. Non-zero sum stochastic differential game between two insurance companies under the variance premium principle. *Research Square*.

Parodi, P., 2023. *Pricing in general insurance*. s.l.:Chapman and Hall CRC.

Poufinas, T. a. G. P. a. P. T. a. Z. E., 2023. Machine learning in forecasting motor insurance claims. *Risks*, 11(9), p. 164.

Rahmawati, T., Susanti, D. & Riaman, R., 2023. Determining pure premium of motor vehicle insurance with generalized linear models (GLM). *International Journal of Quantitative Research and Modeling*, 4(4), pp. 207--214.

Rahmawati, T., Susanti, D. & Riaman, R., 2024. Determining pure premium of motor vehicle insurance with generalized linear models (GLM). *International Journal of Quantitative Research and Modeling*, 4(4), pp. 207--214.

Rekettye, G., Liu, J. & Danyi, P., 2025. Prices, Costs, and Profit. In: *Pricing for Strategic Managers: The Theory and Practice of Implementing Pricing Decisions*. s.l.:Springer Nature Switzerland, pp. 63--88.

Ritho, B. M., Simiyu, E. & Omagwa, J., 2023. The impact of loss ratio on the financial stability of insurance firms in Kenya. *Journal of Finance and Accounting*, 7(4), pp. 22--41.

Rowe, R. et al., 2022. Why are newly qualified motorists at high crash risk? Modelling driving behaviours across the first six months of driving. *Accident Analysis & Prevention*, Volume 177, p. 100.

Schwab, P.-N., 2023. *Into the minds*. [Online]

Available at: <https://www.intotheminds.com/blog/en/are-smart-boxes-the-future-of-car-insurance/>

[Accessed 14 May 2026].

Selvakumar, V., Satpathi, D. K., Kumar, P. & Haragopal, V., 2023. Modeling of Loss Severity of Motor Insurance Extreme Claims. *International Journal of Advances in Soft Computing \& Its Applications*, 15(3).

Signoretta, P., Chamberlain, J. M. & Hillier, J., 2014. A Picture Is Worth 10,000 Words': A Module to Test the 'Visualization Hypothesis' in Quantitative Methods Teaching. *Enhancing Learning in the Social Sciences*, 6(2), pp. 90-104.

Springer Berlin Heidelberg, 2005. Introduction. In: *A Course in Credibility Theory and its Applications*. Berlin, Heidelberg: Springer, pp. 1-14.

Sugden, R., 2024. Relational equality, mutual benefit and social insurance. *Oxford Open Economics*, 3(1), pp. i24--i29.

Supian, S., Mamat, M. & others, 2022. INSURANCE PREMIUM DETERMINATION MODEL AS NATURAL DISASTER MITIGATION EFFORT IN INDONESIA WITH A CROSS SUBSIDY SYSTEM.. *International Journal of Agricultural \& Statistical Sciences*, 18(2).

Total Systems PLC, 2026. *How Electric Vehicles are Impacting the Car Insurance Market*. [Online]

Available at: <https://www.totalsystems.co.uk/whitepapers/how-electric-vehicles-are-impacting-the-car-insurance-market/>

[Accessed 14 May 2026].

Verschuren, R. M., 2022. Frequency-severity experience rating based on latent Markovian risk profile. *Insurance: Mathematics and Economics*, Volume 107, pp. 379-392.

Wang, B., Chen, Y. & Li, Z., 2024. A novel Bayesian Pay-As-You-Drive insurance model with risk prediction and causal mapping. *Decision Analytics Journal*, Volume 13, pp. 100-522.

Whitney, A. W., 2024. The theory of experience rating. In: *The History of Actuarial Science Vol VII*. s.l.:Routledge, pp. 256--276.

Worku, A. T., Bayleyegn, Y. W. & Tafere, Z. B., 2024. Determinants of profitability of insurance companies in Ethiopia: evidence from insurance companies from 2011 to 2020 years. *Journal of Innovation and Entrepreneurship*, 13(1), p. 4.

Wutich, A. a. B. M. a. B. H. R., 2024. Sample sizes for 10 types of qualitative data analysis: an integrative review, empirical guidance, and next steps. *International Journal of Qualitative Methods*, Volume 23, pp. 1-14.

Xie, S., 2023. Analyzing the influence of Telematics-Based pricing strategies on traditional rating factors in auto insurance rate regulation. *Mathematics*, 12(19), p. 3150.

Xie, S., 2024. Estimating Risk Relativity of Driving Records using Generalized Additive Models: A Statistical Approach for Auto Insurance Rate Regulation. *Asia-Pacific Journal of Risk and Insurance*, 18(1), pp. 55--86.

Xiong, H., Guo, Y. & Zitikis, R., 2026. Beyond no-claims discount: a learning-embedded telematics-driven pricing system for dynamic premium adjustments. *Financial Innovation*, 12(1), p. 118.

Zhang, P., Pitt, D. & Wu, X., 2022. A new multivariate zero-inflated hurdle model with applications in automobile insurance. *ASTIN Bulletin: The Journal of the IAA*, 52(2), pp. 393--416.

Zhang, Z., 2023. Certainty equivalent, risk premium and asset pricing. In: *Fundamental Problems and Solutions in Finance*. s.l.:Springer, pp. 151--192.

Zheng, Y., Li, X. & Zhong, W., 2025. An empirical study on machine learning-based future loss prediction for insurance firms: evidence from China's property and casualty insurance sector. *Applied Economics*, pp. 1-14.

Zumitzavan, V., 2026. Hypotheses Development. In: *Social Science Methodologies for Management Research: From Philosophy to Practice in the AI Era*. s.l.:Springer, pp. 97-107.

Appendices

Appendix 1: One-way ANOVA group descriptives

Descriptives

PREMIUM PAID_CODE

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
L1- 0-10 years	199	1.03	.244	.017	1.00	1.06	1	3
L2- 11-20 years	128	2.00	.000	.000	2.00	2.00	2	2
L3- 21-30 years	114	3.00	.000	.000	3.00	3.00	3	3
L4- 31-40 years	59	3.92	.281	.037	3.84	3.99	3	4
Total	500	2.07	1.032	.046	1.98	2.16	1	4

Appendix 2: Tests of Homogeneity of Variance

Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
PREMIUM PAID_CODE	Based on Mean	14.388	3	496	<.001
	Based on Median	3.657	3	496	.012
	Based on Median and with adjusted df	3.657	3	252.027	.013
	Based on trimmed mean	6.875	3	496	<.001

Appendix 3: ANOVA Effect Sizes

ANOVA Effect Sizes^a

			95% Confidence Interval	
Point Estimate			Lower	Upper
PREMIUM PAID_CODE	Eta-squared	.969	.965	.972
	Epsilon-squared	.969	.964	.972
	Omega-squared Fixed-effect	.969	.964	.972
	Omega-squared Random-effect	.912	.900	.921

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

Appendix 4: Multiple Comparison

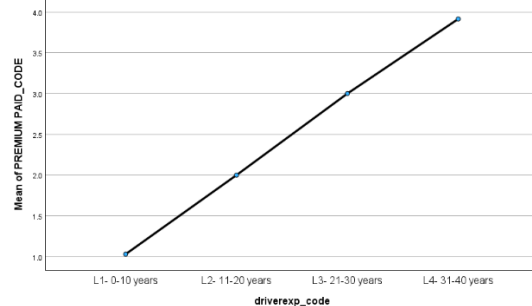
Multiple Comparisons

Dependent Variable: PREMIUM PAID_CODE

	(I) driverexp_code	(J) driverexp_code	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	L1- 0-10 years	L2- 11-20 years	-.970 [*]	.021	<.001	-1.02	-.92
		L3- 21-30 years	-1.970 [*]	.021	<.001	-2.02	-1.91
		L4- 31-40 years	-2.885 [*]	.027	<.001	-2.95	-2.82
	L2- 11-20 years	L1- 0-10 years	.970 [*]	.021	<.001	.92	1.02
		L3- 21-30 years	-1.000 [*]	.023	<.001	-1.06	-.94
		L4- 31-40 years	-1.915 [*]	.029	<.001	-1.99	-1.84
	L3- 21-30 years	L1- 0-10 years	1.970 [*]	.021	<.001	1.91	2.02
		L2- 11-20 years	1.000 [*]	.023	<.001	.94	1.06
		L4- 31-40 years	-.915 [*]	.029	<.001	-.99	-.84
	L4- 31-40 years	L1- 0-10 years	2.885 [*]	.027	<.001	2.82	2.95
		L2- 11-20 years	1.915 [*]	.029	<.001	1.84	1.99
		L3- 21-30 years	.915 [*]	.029	<.001	.84	.99
Games-Howell	L1- 0-10 years	L2- 11-20 years	-.970 [*]	.017	<.001	-1.01	-.92
		L3- 21-30 years	-1.970 [*]	.017	<.001	-2.01	-1.92
		L4- 31-40 years	-2.885 [*]	.040	<.001	-2.99	-2.78
	L2- 11-20 years	L1- 0-10 years	.970 [*]	.017	<.001	.92	1.01
		L3- 21-30 years	-1.000	.000	.	-1.00	-1.00
		L4- 31-40 years	-1.915 [*]	.037	<.001	-2.01	-1.82
	L3- 21-30 years	L1- 0-10 years	1.970 [*]	.017	<.001	1.92	2.01
		L2- 11-20 years	1.000	.000	.	1.00	1.00
		L4- 31-40 years	-.915 [*]	.037	<.001	-1.01	-.82
	L4- 31-40 years	L1- 0-10 years	2.885 [*]	.040	<.001	2.78	2.99
		L2- 11-20 years	1.915 [*]	.037	<.001	1.82	2.01
		L3- 21-30 years	.915 [*]	.037	<.001	.82	1.01

*. The mean difference is significant at the 0.05 level.

Appendix 5: Mean Plot



Appendix 6: Group Statistics

Group Statistics					
		PreviousAccidents_code2 (binary)	N	Mean	Std. Deviation
Level of SAVING frm Discounted IP	Without accidents		76	1.91	.955
	With accidents		424	2.09	1.061
					Std. Error Mean
					.110
					.052

Appendix 7: Independent Sample Effect Sizes

Independent Samples Effect Sizes					
		Standardizer ^a	Point Estimate	95% Confidence Interval	
				Lower	Upper
Level of SAVING frm Discounted IP	Cohen's d	1.046	-.178	-.423	.066
	Hedges' correction	1.047	-.178	-.422	.066
	Glass's delta	1.061	-.176	-.420	.069

a. The denominator used in estimating the effect sizes.

Cohen's d uses the pooled standard deviation.

Hedges' correction uses the pooled standard deviation, plus a correction factor.

Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

Appendix 8: Coded working of Raw data

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	DriverAge	DriverExp	PreviousAcc	AnnualMileage	CarManuf	CarAge	Insurance	Level of SAVING frm Discounted IP	carage_code	driverexp_code	manufacturing_cd	PREMIUM PAID_CQ	Previous accident	PreviousAccidents_code2 (binary)					
1																			
2	46		0	21	2025	0	486.15	2	2.00	2.00	2.00	2.00	4	0					
3	25	7	0	13	2005	20	495.55	1	1.00	1.00	1.00	1.00	4	0					
4	41	1	0	19	2007	18	494.25	1	1.00	1.00	1.00	1.00	4	0					
5	19	1	0	16	2010	15	498.20	1	1.00	1.00	1.00	1.00	4	0					
6	29	8	0	22	2025	0	492.90	1	1.00	1.00	1.00	1.00	4	0					
7	61	8	0	21	2024	1	486.55	1	1.00	1.00	1.00	1.00	4	0					
8	64	30	0	18	2011	14	480.50	3	3.00	3.00	3.00	3.00	4	0					
9	61	4	0	12	2006	19	489.10	1	1.00	1.00	1.00	1.00	4	0					
10	20	2	0	16	2025	0	496.20	1	1.00	1.00	1.00	1.00	4	0					
11	54	10	0	13	2003	22	489.05	1	1.00	1.00	1.00	1.00	4	0					
12	37	19	0	13	1990	35	491.05	2	2.00	2.00	2.00	2.00	4	0					
13	53	25	0	17	1995	30	485.75	3	3.00	3.00	3.00	3.00	4	0					
14	31	1	0	25	2012	13	496.05	1	1.00	1.00	1.00	1.00	4	0					
15	28	1	0	14	2011	14	496.20	1	1.00	1.00	1.00	1.00	4	0					
16	52	32	0	11	2015	10	481.55	4	4.00	4.00	4.00	4.00	4	0					
17	45	14	0	25	1997	28	490.85	2	2.00	2.00	2.00	2.00	4	0					
18	41	23	0	21	1994	31	489.05	3	3.00	3.00	3.00	3.00	4	0					
19	64	5	0	13	2011	14	487.75	1	1.00	1.00	1.00	1.00	4	0					
20	20	2	0	25	2003	22	498.85	1	1.00	1.00	1.00	1.00	4	0					
21	43	25	0	14	2020	5	485.10	3	3.00	3.00	3.00	3.00	4	0					
22	44	17	0	18	2019	6	487.60	2	2.00	2.00	2.00	2.00	4	0					
23	21	3	0	13	2017	8	496.35	1	1.00	1.00	1.00	1.00	4	0					
24	62	39	0	22	2015	10	478.00	4	4.00	4.00	4.00	4.00	4	0					
25	46	10	0	17	2014	11	489.75	1	1.00	1.00	1.00	1.00	4	0					
26	24	6	0	21	2007	18	496.25	1	1.00	1.00	1.00	1.00	4	0					