

**Gordon Institute  
of Business Science**  
University of Pretoria

**Financial behaviours of customers as determinants for risk  
aversion and insurance consumption in South Africa**

**Adriaan Botha**

**10569449**

**A research project submitted to the Gordon Institute of Business Science,  
University of Pretoria partial fulfilment of the requirements for the degree of  
Master of Business Administration**

**6 November 2017**

## **Abstract**

The aim of this research is to determine if there are significant relationships between individual level financial behaviours of customers and the demand for life insurance. The research addresses the gap in the academic literature on the understanding of which financial behaviours of individuals may be useful in determining risk aversion behaviours as assessed by the demand for life insurance.

South African life insurance data is used to develop three logistic regression models that predict take-up, lapse and cancellations of insurance respectively. Ten predictor variables were developed to measure the effect of income, savings and debt on the propensity to take-up, lapse or cancel life insurance.

The results showed that income, savings and debt were significant predictor variables and provide evidence that these measures may be useful to understanding customer preferences concerning insurance demand. The results show an increase in insurance consumption among low income consumers which is a finding unique to the South African context. The results also confirm that low income customers are at risk of both lapsing and cancelling their life insurance. Low levels of savings and debt may indicate an increase in the demand for life insurance but are also associated with increased risk of lapse.

## **Keywords**

Life insurance, Income, Savings, Debt, Insurance Propensity, Take-up, Lapse, Cancel, Logistic Regression Analysis

## Declaration

I declare that this research project is my own work. It is submitted in partial fulfilment of the requirements for the degree of Master of Business Administration at the Gordon Institute of Business Science, University of Pretoria. It has not been submitted before for any degree or examination in any other University. I further declare that I have obtained the necessary authorisation and consent to carry out this research.

Name:           Adriaan Johannes Jacobus Botha

Signature:

---

Date:            06 November 2017

## Table of Contents

Abstract.....	ii
Keywords .....	ii
Declaration.....	iii
1 Introduction to the Research Problem .....	1
1.1 Research Title.....	1
1.2 Research Problem .....	1
1.2.1 The link between the propensity for insurance consumption and individual-level financial behaviours of consumer .....	1
1.2.2 The role of risk aversion.....	2
1.2.3 Understanding insurance customers .....	4
1.2.4 Life insurance in South Africa.....	5
1.3 Purpose Statement .....	6
1.3.1 Data driven decision making .....	6
1.4 Field of Study.....	9
1.4.1 Key definitions.....	10
2 Literature Review .....	13
2.1 Introduction .....	13
2.2 Uncertain Lifetime and Life Insurance .....	15
2.3 Factors of insurance take-up.....	18
2.4 Lapses .....	18
2.5 Cancellations .....	22
2.5.1 Policy Termination.....	22
2.5.2 Policy Cancellation.....	22
2.6 Risk Aversion .....	25
2.6.1 Risk Aversion Behaviour .....	26
2.6.2 The Risk-Aversion Function .....	27
2.6.3 Relationships between functions of demand and risk aversion.....	29

2.6.4	Prudence and Temperance .....	30
2.7	Education .....	32
2.8	Financial sophistication .....	35
2.8.1	Income .....	35
2.8.2	Savings .....	36
2.8.3	Debt .....	40
2.9	Business Analytics .....	42
2.10	Research gap.....	43
3	Research Questions.....	47
3.1	Core research question .....	47
3.2	Sub research questions.....	47
3.2.1	Research question 1 - Income:.....	48
3.2.2	Research question 2 - Savings:.....	48
3.2.3	Research question 3 - Debt:.....	48
3.3	Hypotheses .....	49
4	Research methodology .....	50
4.1	Introduction .....	50
4.2	Philosophy .....	50
4.3	Approach.....	51
4.3.1	Target response variable.....	51
4.4	Strategy .....	52
4.5	Choices .....	52
4.6	Time horizon .....	52
4.7	Definitions .....	53
4.8	Techniques .....	54
4.9	Population .....	55
4.10	Unit of analysis.....	55
4.11	Sampling method .....	55

4.11.1	Individual customers .....	56
4.12	Sample size .....	56
4.13	Measurement Instrument .....	57
4.14	Data gathering process .....	57
4.14.1	Exclusions.....	58
4.15	Analysis approach.....	59
4.15.1	Logistic Regression.....	60
4.15.2	Background on statistical methods.....	62
4.15.3	Variables.....	64
4.15.4	Logistic regression model build-up.....	68
5	Results.....	69
5.1	Take-up.....	69
5.1.1	Graphical representation of take-up predictor variables.....	72
5.2	Lapse .....	79
5.2.1	Graphical representation of lapse predictor variables.....	81
5.3	Cancel.....	85
5.3.1	Graphical representation of the cancellation predictor variables.....	86
6	Discussion of Results .....	89
6.1	Introduction .....	89
6.2	Discussion of research question one: The influence of income on the demand for life insurance products.....	89
6.2.1	Discussion of income in the take-up model .....	90
6.2.2	Discussion of income in the lapse model.....	91
6.2.3	Discussion of income in the cancel model.....	92
6.2.4	Conclusions on income .....	93
6.3	Discussion of research question two: The influence of savings on the demand for insurance products .....	93
6.3.1	Discussion of savings in the take-up model.....	93

6.3.2	Discussion of savings in the lapse model .....	94
6.3.3	Discussion of savings in the cancel model .....	95
6.3.4	Conclusions on savings.....	95
6.4	Discussion of research question three: The influence of debt on the demand for insurance products .....	96
6.4.1	Discussion of debt in the take-up model.....	96
6.4.2	Discussion of debt in the lapse model .....	96
6.4.3	Discussion of debt in the cancel model .....	97
6.4.4	Conclusions on debt.....	97
6.5	Concerns.....	98
7	Conclusion .....	99
7.1	Introduction .....	99
7.2	Principal findings .....	99
7.3	Implications for management .....	103
7.4	Limitations of the research .....	103
7.5	Suggestions for future research .....	104
7.5.1	Incorporating transactional data .....	104
7.5.2	Extend observation period.....	104
7.5.3	More variables .....	105
7.5.4	Extend outcome period .....	105
7.5.5	Post-model adjustments.....	105
8	Consistency Matrix.....	106
9	References.....	107
10	Appendix .....	119
10.1	Appendix A.....	119
10.2	Appendix B.....	126

## Table of Figures

Figure 1: Graph showing South African GDP growth over the last 10 years. Source: World Bank national accounts data, and OECD National Accounts data files.....	44
Figure 3: Time horizon of research.....	
Figure 4: The effect of average credit turnover on take-up .....	69
Figure 5: The effect of average cheque balance on take-up, including customers with a negative account balance.....	70
Figure 6: The effect of average cheque balance on take-up, excluding customers with a negative account balance.....	70
Figure 7: The effect of savings on take-up.....	71
Figure 8: The effect of amount of total debt on take-up .....	71
Figure 9: The effect of the ratio of average credit turnover to income .....	72
Figure 10: The effect of the ratio of average credit turnover to income on take-up .....	72
Figure 11: The effect of the ratio of average cheque account balance to income on take-up.....	73
Figure 12: The effect of the ratio of total debt to income on take-up .....	74
Figure 13: The effect of the ratio of total debt to average cheque account balance on take-up.....	74
Figure 14: The effect of income on take-up .....	75
Figure 15: the effect of credit turnover on lapse.....	78
Figure 16: The effect of average cheque balance on lapse .....	79
Figure 17: The effect of savings on lapse .....	79
Figure 18: The effect of total debt on lapse .....	80
Figure 19: The effect of the ratio of credit turnover to income on lapse.....	80
Figure 20: The effect of the ratio of cheque balance to income on lapse .....	81
Figure 21: The effect of total debt to average cheque account balance on lapse.....	81
Figure 22: The effect of average credit turnover on cancellations, this was the only significant variable in the logistic regression model .....	83
Figure 23: The effect of total debt on cancellations .....	84



Figure 24: The effect of the ratio of debt to average cheque balance on cancellations 84

Figure 25: The effect of average income on cancellations .....85

Figure 26: The effect of savings on cancellations .....85

## Table of tables

Table 1: Key foundations and definitions in insurance .....	9
Table 2: The various consumer situations described by Yaari, (1965) .....	17
Table 3: Potential research hypotheses for the South African life insurance context ...	47
Table 4: The set-up of the research questions in relation to financial behaviours of customers and insurance demand.....	47
Table 5: Key data and methodology definitions used.....	48
Table 4: Research Methodology Choices .....	50
Table 6: Table showing the relative sample sizes for the different logistic regression models .....	53
Table 7: Data measurement and source.....	53
Table 8: Exclusions made to the research data .....	55
Table 9: A summary of the independent variables and their definitions used in the study .....	57
Table 10: A summary of the dependent variables and their definitions used in the study. Note that the information on target variables for life insurance products were included as binary variables in the model. This provides an overview of which target variables were specified for the research.....	59
Table 11: Data transformation into binary response variables .....	66
Table 12: Significant and insignificant predictor variables for lapse .....	67
Table 13: Logistic Regression Predicting Likelihood of Take-up based on Average Credit Turnover, Average Cheque Account Balance, Savings, Total Debt, Ratio of Credit Turnover to Income, Ratio of Average Cheque Account Balance to Income and Ratio of Total Debt to Average Cheque Account Balance.....	68
Table 14: Significant and insignificant predictor variables for lapse .....	76
Table 15: Logistic Regression Predicting Likelihood of Lapse based on Average Credit Turnover, Average Cheque Account Balance, Average Savings and Investments Balance, Total Debt, Ratio of Credit Turnover to Income, Ratio of Average Cheque Account Balance to Income and Ratio of Total Debt to Average Cheque Account Balance.....	77
Table 16: Significant and insignificant predictor variables for cancellations. ....	82

Table 17: Logistic Regression Predicting Likelihood of Cancellation based on Average Credit Turnover. ....	83
Table 18: Summary of the direction of the relationships between individual level financial behaviours of individuals regarding insurance demand. ....	101
Table 18: Consistency Matrix .....	122

## Table of Equations

Equation 1 .....	15
Equation 2 .....	15
Equation 3 .....	15
Equation 4 .....	16
Equation 5 .....	16
Equation 6 .....	27
Equation 4 .....	28
Equation 5 .....	28
Equation 6 .....	28
Equation 7 .....	28
Equation 8 .....	29
Equation 9 .....	29
Equation 10 .....	37
Equation 11 .....	38
Equation 12 .....	63

# 1 Introduction to the Research Problem

## 1.1 Introduction

*“Risk aversion is at the heart of the demand for insurance” (Outreville, 2014, p. 159)*

### ***1.1.1 The link between the propensity for insurance consumption and individual-level financial behaviours of consumers***

Ever since Bernoulli's (1738) exposition on a new theory on the measurement of risk, authors have been concerned with measuring risk aversion. Arrow, (1965) and Pratt, (1964) showed that their functions for the measurement and strength of risk aversion were excellent, and their measures of risk aversion have subsequently been utilised in a broad array of empirical and theoretical research on behaviour under uncertainty. Outreville, (2014) surveyed the vast contribution of literature on the connection between risk aversion (both relative and absolute) and insurance consumption. He reports that the connection that the demand for insurance is a well established response to risk aversion. Barseghyan, Molinari, O'Donoghue, and Teitelbaum's (2013) reported on using insurance decisions to estimate risk aversion behaviour. They found that households are generally risk averse to some degree and therefore necessitate a premium to invest. They further affirm that households subsequently buy insurance cover at actuarially imbalanced rates to satisfy their risk aversion. Slovic (2016) stated that there are three distinct stages of insurance purchase decisions. Firstly, consumers must contemplate the risks to be a problem (stage 1) and must then be cognisant that insurance is a suitable coping mechanism for the identified risk (stage 2). Finally, they will collect and process information relevant to the insurance purchase (stage 3). Based on the above, one may assert that it is well understood that risk aversion and insurance consumption may be related. Taking the above concept one step further, the literature has expanded on the links between other established risk aversion concepts and the propensity for insurance consumption. Numerous studies, for example; Hammond et al., (1967), Mantis and Farmer (1968), Duker (1969), Anderson and Nevin (1975), Ferber and Lee (1980), Burnett and Palmer (1984), Bernheim (1991), Zeits (2003), Nakata and Sawada (2007), Millo and Carmeci (2011), Feyen, Lester, & Rocha (2011), have established relationships between the propensity of life insurance consumption and a multitude of socio-economic and economic demographic factors, such as age and marital status, as cited by Outreville, 2014. This also includes life expectancy, number of dependents, and education (Zietz, 2003; Liebenberg, Carson & Dumm, 2012). These relationships have been well established and researched over a

significant period of time (Lewis, 1989; Browne & Kim, 1993, Outreville, 2015). However, Chui and Kwok, (2008), argue that the effects of national culture on the consumption patterns of life insurance, are more important concepts to consider. In an African context there has been very little work done on the above concepts. Demographic factors such as level of attained education, number of dependents, and health expenditure have been reported to have a significant influence on life insurance consumption in an African context (Alhassan & Biekpe, 2016). Despite these studies, there is a still room to account for the individual-level financial behaviours of customers (such as income, savings, and debt levels) that may be used as predictor variables for the consumption of life insurance. These variables are not easily attained but are however observable to some firms, primarily in banking and financial services. This may, in part, be the reason why they have not been extensively investigated in the past. There has however, been extensive literature on the positive relationship between income and insurance consumption (Lin & Grace, 2007). In much the same way Mulholland, Finke and Huston, (2016) also report the increase in insurance demand with increasing financial sophistication. This research, however, believes that there is a need to address the unknown specific responses to risk aversion in the South African context. This is primarily because there are significant cultural responses to bequest planning among low income consumers in South Africa (Roth, 2000). This paper therefore argues in favour of testing the academic literature on insurance consumption, due to the unique nature of the South African context, especially with regard to the high levels of inequality and poverty (Leibbrandt, Woolard, Finn & Argent, 2010).

The aim of this research is to determine whether each of these individual indicators are in fact good indicators for the consumption of life insurance in the South African context. A review of the literature on the effect of income on insurance consumption, may begin with early discussions by Browne and Kim (1993), Outreville (1996) and Enz (2000) and later discussions by Dragos, (2014); Shi, Wand and Xing, (2015) and Lange, Schiller and Steinorth, (2017). The accumulation of debt is discussed by Athrea (2008); Frees and Sun (2010); Aron and Muelbauer, (2013) and recently by van Winssen and van Kleef (2016). Finally, early discussions on savings by Headen and Lee (1974) and later reports by Somerville, (2004) and Feyen et al. (2011), will be reviewed and argued in this paper. Much of the work done on these concepts are in fact quite old. In the interim, there have been good developments, making it easier for individuals to access life insurance. This is especially true for low income individuals. This further reinforces the need to seek out new links between these predictor variables and insurance consumption. This paper also argues that very little work has

been done on relating these predictor variables to insurance consumption in emerging economies, as highlighted by Shi et al. (2015) as well as Outreville (2015).

### **1.1.2 The Role of Risk Aversion**

Outreville, (2014) reports that the determining factors of risk attitudes of individuals, is in a state of elevated interest in the expanding subject of behavioural finance that concentrates on individual characteristics that form general investment and financial practices. He asserts that, even though there is extensive literature published about the determinants of the demand for insurance, there are still numerous topics that still necessitate further consideration. He evaluates some of the empirical literature concerned with risk behaviour and risk aversion, with special emphasis on insurance consumption. Outreville (2014) continues to emphasise that the empirical studies on risk aversion and risk behaviours, may be characterised as follows: Firstly, the magnitude and measurement of risk aversion, and secondly, the investigation of socio-demographic factors related to risk aversion. More is said on behavioural finance of customers in subsequent chapters. This research paper in part heeds the call for further research into factors that are related to risk aversion behaviour.

Much research on the subject of behavioural insurance, largely emphasises the situational riskiness, while some other research concentrates on the individual's inclination to take risks in such situations (Benartzi & Thaler, 2001; Barber & Odean, 2001). The conformist anthropological concept is that people are directed in their choices among risk-taking and risk-avoiding approaches by their culture (Ward & Zurbruegg, 2000). A reintroduced attentiveness to this area of study is connected to the efforts by Hofstede (1983) and Newman and Nollen (1996). Outreville (2014) relates that it is astonishing that this topic has continued to be unexplored for an extended period, bearing in mind the significance of the article published by Hofstede (1995).

An effectual use of the above understanding of customer choices, in either taking or avoiding risks, would be to design insurance products which are differentiated by the variable customer risk preferences. John (1993) reports that the true cost of managerial time squandered on less effective insurance product development, is unknown and this is substantiated by Stark, (2015). It has been long established, as Urban and Hauser, (1993) reported, that a significant portion of all new products eventually fail in the market, causing a sizeable financial loss. Importantly, this was confirmed in a South African insurance context by Oldenboom and Russel (2000). Even though, in financial services, the financial losses due to failed products may be low (Stark, 2015) there are still many concealed expenses to be considered. These

may include the consequence of failure on corporate image, the misuse of managerial effort and time, as well as the loss of used resources (some of which may be key constraints in the business). Moreover, some ineffective products may not be withdrawn immediately once introduced, and resources must be engrossed in supporting it for current users, even for several years after the product launch (Salunke, S., Weerawardena, & McColl-Kennedy, 2011). Up to 70% of all product development expenses may be invested into products that are eventually cancelled or fail (Oldenboom & Russel, 2000). Ernst, Hoyer, & Rübсаamen, (2010) reported that much of the outcome of new product success or failure is dependent on a product manager's control. This confirms the value of research into product development activities, since product developers who have access to sufficient information in their line of work, will be empowered to make better decisions about their products. The aim is that this research paper will provide relevant information to insurance product developers that may enable the development of segment tailored insurance products through a better understanding of which customers are inclined to buy, cancel or lapse insurance.

### ***1.1.3 Understanding Insurance Customers***

To illustrate the benefits of efficacious product development through the suitable analysis of data, the argument is presented by Wamba, Gunasekaran, Akter, Ren, Dubey, and Childe, (2017) on the effect of companies' big data analytics capabilities and its effect on the performance of a firm. Their research shows that big data analytics capabilities may effectively be leveraged as a source of competitive advantage (at least in the short term). Their research showed that 65% of the variance in the performance of the firm, was explained by the big data analytics capabilities and process-oriented dynamic capabilities. However, big data analytics capabilities had a greater effect on firm performance than process-orientated dynamic capabilities. Their research separated big data analytics capabilities into three constructs (infrastructure flexibility, management capabilities, and personnel expertise capability), and provides insight into the linkages between big data analytics capabilities and firm performance.

Côrte-Real, Oliveira, and Ruivo, (2017) assessed the business value of big data analytics in European firms. Their results indicated that big data analytics may offer business value at numerous points in the value chain. Côrte-Real et al., (2017) go on to state that big data analytics may generate organizational agility through efficient knowledge management. They also demonstrated that agility (to some extent) mediated the effect between knowledge assets and competitive advantages seen in firms. Their model goes so far as to suggest that 77.8% of the variation in competitive



advantage can be explained by big data analytics expertise assets (endogenous, exogenous knowledge management and knowledge sharing partners).

The current trend of interest in big data analytics and the apparent promised improvement in firm performance, exists coupled with a relative lack of understanding how financial behaviours, specifically the propensity of taking on debt and savings levels, may be attributable to insurance consumption. The argument therefore exists to better understand which financial behaviours lead to an increased propensity for insurance consumption. Big data (defined as the ability to cross-reference and analyse different types of data) given the technological advancements and ability to collect this data and analyse it, may provide the opportunity to understand customers better. In the context of this research; this capability will assist in the understanding of the factors leading to the propensity to take up, cancel, or lapse insurance products. It can be argued that this is important, considering the role which insurance plays as a financial intermediary in some emerging economies (Outreville, 2015).

#### **1.1.4 Life Insurance in South Africa**

There is a strong case for studying determinants for insurance demand in the South African context, since empirical studies on the factors of insurance demand have principally concentrated on the life sector in the United States of America (Outreville, 2014). The life sector in the United States is significantly different to the life sector in emerging economies, predominantly due to the different cultural factors (such as the relative significance of health and life insurance in various social security settings), differences in gross domestic product (GDP) and institutional environments as exposed by Outreville, (2013) and Dionne, (2013).

Extensive research has been dedicated to understanding the demand of life insurance in developed countries, however the extent of the academic literature's understanding of this demand in developing markets remains under-developed, according to Shi, Wang & Xing (2015). This vacuum is regrettable, since emerging markets offer excellent opportunities for growth for insurance companies (Chang, Lee, & Chang, 2014). Mayers and Smith, from as early as 1983, have emphasised the important role of life insurance in economies with less developed capital markets.

Studies from Africa, specifically Ghana, have shown that the ex-ante moral hazard in insurance is a widely acknowledged problem, but is also often trivialized without the proper empirical foundations (Yilma, Van Kempen, & De Hoop, 2012). The insurance demand behaviour of the emerging markets (China) have been studied (Shi et al. 2015), but have not yet been determined in the South African context. Therefore

providing further motivation to consider micro-economic evidence of the determinants of life insurance demand in South Africa.

This research is motivated in part through the statement made by Chang et al. (2014), that “insurance is essential for the development of the banking industry”. They also state that insurance nurtures trade and commerce between different countries and in this way engenders bank revenues. Equally, companies that offer insurance products have long-term premium incomes and therefore have longer term investments. Thus encourage the development of stock market exchanges and local bond markets. These activities significantly contribute towards promoting economic growth, which is sorely needed in the South African context (also see Figure 2).

The core research questions are proposed in chapter three below.

## **1.2 Research Motivation**

Emamgholipour, Arab, Mohajerzadeh (2017) assert that life insurance is a specific form of continuing investment; therefore, the purpose of procuring life insurance, is to protect present and future damages of the insured person. While insurance performs a critical role in economic fiscal and development (Outreville, 2013), in some countries, insurance (especially life insurance) remains undeveloped, with a low penetration rate, especially in Africa (Enz, 2000; Alhassan & Biekpe, 2016). Hence, the purpose of this research is to ascertain the individual-level financial influences that impinge on life insurance demand.

### **1.2.1 Data Driven Decision Making**

Shanks, Sharma, Seddon and Reynolds (2010) as well as Sharma, Mithas, and Kankanhalli, (2014) advocate that using data to improve executive decision-making, is a competitive advantage. Chen, Chiang and Storey (2012) propose business analytics and similar analytical expertise can assist companies to ‘better understand its business and markets’ as well as take advantage of opportunities accessible by rich data. In the same way, Lavallo, Lesser, Shockley, Hopkins, and Kruschwitz (2011) account that the most successful organisations often take decisions grounded in rigorous analysis at twice the rate of organisations which perform at a lower level. These successful organisations use analytic insights to shape strategies and everyday operations. Brynjolfsson, Hitt, and Kim, (2011) reported on 179 firms which applied business analytics to various applications in decision making. They found that firms using data driven decision making processes, were associated with higher productivity and market value.

The argument therefore exists that companies who rigorously analyse their data (and have strong analytical capabilities) may be able to elicit a competitive advantage over those companies who use different techniques. This investigation into the links between the pre-defined banking behaviours and insurance behaviours, may provide some assistance for insurance companies to develop better insurance products in the South African context. The findings of this paper should, at the least, provide some basis for insurance companies to begin their own analytics on their customer segments.

At best, the findings of this paper may deliver interesting insights to which South African customers are likely consumers of insurance products. This will also enable businesses to develop specifically tailored products that are appetising for consumers who are most likely to take out insurance.

Eling and Kiesenbauer (2014) considered a large dataset to resolve which characteristics of insurance policies affected lapse rates. They established that product characteristics (product type) as well as policyholder characteristics (age and gender) were central drivers of lapse rates. These characteristics are imperative to understand for the purposes of value and risk-based management practices, but are also useful for the purposes of designing tailored products that would better serve customers' needs. For this reason, this paper seeks to find the linkages between individual level financial behaviour of customers and their insurance behaviours.

The fundamental purpose of this research paper, is to address the lack of academic research in the field of understanding which consumer-level financial metrics drive insurance consumption. The specific financial metrics (and therefore also the key concepts) that this investigation will address are personal income, savings, and debt, and how the varying levels of these personal financial indicators (referred to as financial behaviours) may be linked to insurance consumption.

The methodology of this paper, therefore, will use a systematic approach to attain useful information which would assist strategic decisions involving product development in the life insurance sector. This systematic approach is expanded on in chapter 4 of this paper.

Businesses invest in analytics but often employ poor methodologies which are not supported by a sound theoretical framework (Shanks et al., 2010, Sivarajah, Kamal, Irani, & Weerakkody, 2017). Thus, a methodological framework, practically implemented to analyse data, creates value through the facilitation of enterprise decisions. This paper aims to provide academic insights for a suitable approach to

analysing a dataset that a business may be in possession of and build on the corpus of knowledge on insurance propensity factors.

Han, Kamber and Pei (2012), Storbacka, (1997) and Herath and Rao (2009) express that some firms, especially banks, have access to large datasets of customer information which are relatively complete, of a good quality, and are reliable. Banking information systems have access to consumer financial transactions, allowing them to study consumer's behaviours (Agarwal & Qian, 2014) so much so that organisations with rich access to information are increasingly placing importance on securely managing the information (Herath & Rao, 2009). This access to information facilitates systematic data analysis on their customers. The problem is that many financial service providers are not using this wealth of information to significantly improve their offerings, often falling prey to disruption on an unprecedented scale over the past few years. In fact, Van Der Boor, Oliveira and Veloso (2014) found in their quantitative analysis that, users developed innovative banking services prior to the solution being commercialised by a producer. Their claim was consistent to findings by other authors, in that users were oftentimes the service innovators long before official financial service providers were. (Skiba & Herstatt, 2009; Oliveira & Hippel, 2011). The purpose of this paper is also to discover linkages between banking information and insurance consumption.

By intelligently analysing specific and focussed parameters, and finding insightful correlations between financial factors (such as income level, proportion of income saved, and debt) and insurance consumption (delineated by insurance purchases, insurance cancellations and lapse behaviours), an insurance firm may be able to solve for product design questions. Thus offer more effective value to its customers. The topic is discussed in greater detail in the subsequent sections.

The aim is that this research will offer a unique perspective, since the data it analyses is a rare combination of both detailed individual-level financial data gained from banking behaviours, and insurance purchases, lapses, and cancellations.

### **1.3 Field of Study**

The topic of this paper deals with product development and the uptake of insurance products. The paper is focused on analysing individual-level financial information of customers from banking data and linking this to insurance take-up (purchases), lapses, and cancellations. Thereby finding financial indicators of insurance propensity. The research aims to determine associations and relationships, specifically between insurance consumption and the aforementioned financial information (primarily: income, savings, and debt), thereby offering valuable information which will be useful to

business strategists who are looking at developing appropriate products for specific markets. The income, debt, savings, and insurance consumption information is available in the banking and insurance data deployed in this research.

The study used data from the insurance and banking sectors, but the methodology and findings will be useful and could be applied to other financial services or other industries, such as telecommunications or consumer electronics as they relate to understanding the attributes of consumption which are applicable and common.

For this paper to further substantiate why research is needed to find further linkages on insurance propensity, it is argued that performance is systematically improved through proper understanding of data and analytics. Lavelle et al., (2011) published a survey of nearly 3,000 analysts, executives, and managers, working across multiple industries globally and found that the higher performance companies use five times more analytics than lower performing companies. They found a prevalent conviction that analytics offers value.

Lavelle et al., (2011) further reported that improvements in data and analytics, was one of the highest priorities within high performing organizations and more than 20% reported significant pressure to adopt the latest and most advanced analytics and information systems. This research therefore also hopes to provide a contribution towards further investigation into insurance propensity and general analysis of insurance data.

The paper aims to explore parameters of consumer-level financial behaviour found in datasets that affect demand, and in this manner, assists with market entry, product design and strategic decisions since linkages between which levels of savings, debt and income, may lead consumers to be prone to purchasing or discarding insurance cover. Consequences of individual's behaviour towards risk or uncertainty are appropriate not only for insurance providers but also to other areas of financial services (Outreville, 2014).

### **1.3.1 Key Definitions**

This research is bounded by the following key definitions:

**Table 1: Key foundations and definitions in insurance**

Foundations	Definitions	Sources
Absolute Risk	Absolute risk aversion is two times the risk premium per unit of variance for infinitesimal changes in risk.	Deschamps,

Aversion	It measures the premium which an individual may spend on purchasing insurance cover against a minor risk. If it is negative then it is indicative of risk loving behaviour. It is denoted $A$ and is given by Equation 8.	(1973)
Behavioural Finance	Behavioural finance is the study of financial markets using broader models than those based on Von Neumann–Morgenstern expected utility theory and arbitrage assumptions. It embraces theories of cognitive psychology and limits to arbitrage. Cognitive psychologists have detailed many patterns concerning how individuals behave including concepts of overconfidence heuristics, mental accounting, conservatism, framing and the disposition effect.	Outreville, (2015)
Capabilities	Capabilities are subclass of resources, which is firm specific, is often not transferrable and is deeply embedded in an organisation. Capabilities improve productivity of the firm	Makadok, (1999)
Competitive advantage	A competitive advantage generates a greater amount of economic value than a competitor in its given product market.	Peteraf and Barney, (2003)
Cross-selling	Advances in information technology, make it possible to collect information on customers, the information may then be useful in identifying customers most likely to purchase other products. It also assists database marketers in targeting individuals for the promotion of new products, increasing the efficiency of both creating and distributing products, and securing faster returns on investments.	Kamakura, Kossar and Wedel, (2004)
Indirect utility function	This is the individual's maximal attainable utility when confronted with a vector of income as well as	Jehle and Reny, (2011)

goods prices. It imitates the individual's preferences as well as market conditions. The function is indirect since it is usually thought about in terms of preferences according to consumption rather than charges. The indirect utility can be calculated from an individual's utility function expressed in vectors of quantities of purchasable items.

Insurance default risk	The probability that a policyholder will not be reimbursed totally (or in some instances even partially) by the insurer in case of a valid claim (loss).	Zimmer, Schade and Gründl, (2009)
Mortality Selection	The connection between policy survivorship and termination. It is the relationship amongst policy survivorship and termination contained in a life insurance portfolio. It is important to actuaries with regard to risk management responsibilities, reserving and pricing.	Valdez, Vadiveloo, and Dias, (2014)
Prudence	The “propensity to prepare and forearm oneself in the face of uncertainty, in contrast to “risk-aversion” which is how much one dislikes uncertainty and would turn away from uncertainty if possible.”	Kimball, (1990, p.54)
Resources	Resources are defined as tangible and intangible assets used by the firms to conceive of and implement its strategies.	Barney and Arian (2001)
Relative risk aversion	It is the behaviour of consumers to endeavour to diminish their given amount of uncertainty. Risk averion is divided into constant relative risk aversion, decreasing relative risk aversion and increasing relative risk aversion. It is valid in situations of risk aversion and risk taking. Relative risk aversion is denoted $R$ and is given by Equation 9	Outreville, (2015) Deschamps, (1973)

Temperance	Temperance is “the desire to moderate total exposure to risk”. It causes individuals to respond to risk that are unavoidable though the reduction of other risks even though those risks may be statistically independent of the unavoidable risk.	Kimball, (1992)
Voluntary Deductible	This is an amount, which is chosen by the policyholder, paid to the insurer to meet a part of the claim. The amount rests on the policyholder who selects the threshold according to their affordability and risk. A Voluntary Deductible results in a premium rebate to the insured.	Van Kleef, Van De Yen, & Van Vliet, (2006).

---

#### **1.4 Research Ethics**

This research sourced anonymised banking and insurance data from a South African bank. Every effort was made to maintain the confidentiality of the company as well as the data. The customer information, such as identity numbers, customer numbers, addresses, and policy numbers were completely removed from the dataset. The bank provided official consent to conduct the study by using their information and the Gordon Institute of Business Science (GIBS) granted ethical clearance.

#### **1.5 Summary**

This chapter presented an introduction to the research problem. The motivation for the research was presented in Section 1.2 and motivated the need for this investigation. The field of the study was provided in Section 1.3, and the ethical implications of the study was given in Section 1.4. Chapter 2 delivers a review of the academic literature related to the research.



## 2 Literature Review

### 2.1 Introduction

Greene, (1963) in his paper on the psychological attitudes towards risk and theory of insurance consumption, states that if one recognises the conjecture that attitudes toward risk has an influence on insurance consumption, then variables such as the utility for money, and preferred risk levels, should equally be of significance in explaining insurance purchasing behaviour, since they are seemingly related to risk attitudes. Yaari, (1964) and Hakansson (1969), forms the starting point for “nearly all theoretical and empirical work on the demand for life insurance” according to Outreville, (2014), who reviews much of the literature on the effect of risk behaviour and aversion on the demand for insurance. He goes on to cite Bernheim, (1991), who states that an individual will use a range of variables to characterise the conceivable results of the decision being made, in order to maximize their lifetime utility. Outreville, (2014, p.265), continues to assert that insurance demand is a function of a number of financial factors, such as expected income, total assets (wealth), expected rates of return for alternative choices, as well as some subjective discounting functions which consumers use to assess their choices. Very importantly; Schlesinger, (1981) and Szpiro, (1985) as cited by Outreville, (2014, p.101), emphasise that it is implicitly assumed that risk aversion levels have a direct influence on the latter discounting factors, and that “risk aversion is positively correlated with insurance consumption”.

Since different individuals respond to given situations in disparate ways, several psychological experiments have been conducted in an attempt to classify profiles of those individuals who are risk averse as well as risk-taking (MacCrimmon & Wehrung, 1986; Smidts, 1997; Montibeller, & von Winterfeldt, 2015). Kogan and Wallach, (1964) and later, Zou and Scholer, (2016) highlight that the variance in the behaviours of individuals confronting analogous risky scenarios, may well be partly explicated by factors such as family background, position, prior experience, education, and location.

The literature review of this research will focus on individual-level customer financial behaviours, which this paper defines as the amount of, and frequency with which customers take on debt, the amount of savings they keep, and income that they generate as opposed to macro-economic drivers of insurance propensity. This approach is taken since customers are a very relevant parameter in the banking and insurance sectors. Indeed, the consumer is often the most important element of any business (DeSarbo, Jedidi, & Sinha, 2001). In a line of reasoning to this, Peppers, Rogers and Kotler (2016 p.76) similarly state that “for a customer relationship strategy

to work, a company must establish a focus on the customer, a commitment to a genuine understanding of the customer, and a culture in which every employee believes that the customer comes first. In short, the company needs to have a customer strategy, one focused on ensuring that everything the company does is oriented toward building solid customer relationships”. This is also true for the financial services and insurance market as Yousefi, (2016) declares that this relationship is “crucial to count for the survival and profitability of the organisation”. It is therefore important to understand which individual-level financial behaviour of consumers may assist with the prediction of insurance consumption. This understanding may be crucial for strategy formulation and developing products that are both appetising and have value to existing and potential insurance consumers. Today's markets are progressively dissimilar and understanding the differences between customers is a key facet of operations strategy and product development according to Hill, (2000) and Slack and Lewis, (2015).

The literature review will explain the factors of insurance propensity, namely; insurance take-up, lapse, and cancellations. Each concept will be discussed separately in the following sections. Take-up, lapse and cancellations are indicative of insurance behaviour and risk aversion.

## **2.2 Uncertain Lifetime and Life Insurance**

As previously mentioned, Yaari's (1965) work on customer responses to uncertainty of lifetime, is the starting point of nearly all empirical and theoretical work on the demand for life insurance according to Outreville, (2014). Yaari, (1965) states reported that consumers who meaningfully plan ahead, certainly takes account of their uncertain lifetime duration. He begins his discussion with a Fisher-type study of allocation over a period of time. If an individual assumes that they will live  $T$  years and  $c$  represents any random consumption strategy which the individual may anticipate, then  $c$  is a real-valued function on the interval  $[0, T]$ . So, for every  $t$  on this interval, the measure  $c(t)$  defines the rate of consumption expenditure (for example in Rands), which would occur at a given time  $t$  if the strategy  $c$  was implemented. Presuming the individual's choices are characterised by a utility function  $V$  (for a reminder this function ( $V$ ) is further denoted as the Fisher utility function) then the measure  $V(c)$ , (which is a real number) is in fact the utility of the consumption strategy  $c$ . Yaari (1965) goes on to assert that; an assumption must be made on the form of  $V$ . In this case, assume that  $V$  is the following form:

$$V(c) = \int_0^T \alpha(t)g[c(t)]dt$$

**Equation 1**

Where  $g$  (a concave real-valued function on the half-line  $[0, \infty)$ ) is the utility related with the consumption rate at every  $t$  and  $\alpha$  (a non-negative real-valued function on  $[0, T]$ ) is a subjective function for discount. Due to the form of Equation 1 we can say that  $V$  is the individual's preferences independent of time.

Yaari, (1965) continues to derive the consumption plan  $c^*$  which effectively maximises the preference function subject to a constraint of relative wealth. If the individual's initial assets are assumed to be zero at time  $t$  then the function  $S(t)$  (the individual's net assets) is the form:

$$S(t) = \int_0^t \left\{ \exp \int_{\tau}^t j(x)dx \right\} \{m(\tau) - c(\tau)\}d\tau$$

**Equation 2**

Where  $j(\tau)$  is the expected interest rate at time  $\tau$  and  $m(\tau)$  is the rate of all other earnings at time  $\tau$ . It follows that the function  $S(t)$  is simply the flow of earnings over the flow of expenditures (due to consumption), compounded for every moment at the current interest rate.

If the wealth constraint  $S(t) \geq 0$ , the consumption plan  $c$  is bounded and measurable,  $c(t) \geq 0$  for all  $t$  in the interval  $[0, T]$  and  $\int_0^t \left\{ \exp \int_{\tau}^t j(x)dx \right\} \{m(\tau) - c(\tau)\}d\tau = 0$  then the Fisher problem can be summarised as: Discover an allowable strategy  $c^*$  so that  $V(c^*) \geq V(c)$  for all allowable strategies  $c$ . (Yaari, 1965).

If the problem has a solution and the optimal plan ( $c^*$ ) exists, then the optimal plan  $c^*$  is continuous on the interval  $[0, T]$ , differentiable where positive and complies with the following differential equation:

$$\dot{c}^* = - \left\{ j(t) + \frac{\dot{\alpha}(t)}{\alpha(t)} \right\} \frac{g'[c^*(t)]}{g''[c^*(t)]}$$

**Equation 3**

Where  $\dot{c}$  and  $\dot{\alpha}$  signifies differentiation with respect to time. The function  $-\frac{\dot{\alpha}}{\alpha}$  may be considered as the individual's subjective rate of discount. The above equation (Equation 3) tells us that the best consumption strategy is decreasing where the rate of subjective discount is larger than the rate of interest and increasing when the rate of

interest is greater than the rate of subjective discount. This completely defines the Fisher problem (Yaari, 1965).

Yaari, (1965) continues to derive functions based on a case of “no loved dependents” and realises that the subjective rate of discount  $-\frac{\dot{\alpha}}{\alpha}$  in fact becomes  $\pi_t(t) - \frac{\dot{\alpha}}{\alpha}$ , where  $\pi$  is the probability density function with respect to time.

Yaari (1965) continues to elucidate various cases of insurance consumption as given in Table 2. For the sake of brevity, and since we are only most interested in providing the information pertaining to the most relevant situation, only case D\* will be discussed.

In the case where an individual has assets and liabilities in the fashion of actuarial and regular notes, it is needed to characterise the best savings strategy ( $\dot{S}^*$ ), the best consumption strategy ( $c^*$ ) and the best “portfolio mix” between the normal and actuarial notes.

Yaari (1965), finds these to be:

$$\dot{S}^* = - \left\{ j(t) + \frac{\dot{\beta}(t)}{\beta(t)} \right\} \frac{\varphi'[c^*(t)]}{\varphi''[c^*(t)]}$$

**Equation 4**

And the exact same equation as Equation 3:

$$\dot{c}^* = - \left\{ j(t) + \frac{\dot{\alpha}(t)}{\alpha(t)} \right\} \frac{g'[c^*(t)]}{g''[c^*(t)]}$$

**Equation 5**

Where  $\varphi$  is a non-decreasing concave function, that is real and defined on the overall real number line and the whole utility of a consumption strategy,  $c$ , for a lifetime of span  $T$ , is provided by  $V(c) + \varphi[S(T)]$ . It follows that  $\varphi$  is a true penalty function, subject to the conditions:

$\varphi(x) = 0$  for  $x \geq 0$  and

$\varphi(x) < 0$  for  $x < 0$ ,

Equation 4 also depends on  $\beta$  which is the subjective weighting function for bequests. Fisher, (1930) states that the ambiguity of life, raises the rate of preference for current incomes above future incomes. Individuals with dependents may lead to a reduction in impatience. It is therefore important to note that  $\beta$  is a complication introduced since individuals are likely to value a given bequest according to its size and the time at which it is made. It may be noted that  $\beta$  is expected to be a hump-shaped curve since bequests are relatively more important in an individual’s middle years.

**Table 2: The various consumer situations described by Yaari, (1965)**

	<b>Fisher (1930) utility function with wealth constraint</b>	<b>Marshall (1920) utility function with no constraint</b>
<b>Insurance unavailable</b>	Case A	Case B
<b>Insurance available</b>	Case C	<b>Case D*</b>

It is also imperative to note some subjective differences of life insurance in the South African context, since these are important to consider when comparing other national contexts and insurance products. In South Africa, life insurance is regulated by the Long-Term Insurance Act, 1998 (Act No. 52 of 1998). It is different to short-term insurance which is regulated by the Short-Term Insurance Act, 1998 (Act No. 53 of 1998). The life-insurance products written on a long-term licence are very different to the insurance products written on a short-term licence.

Long-term insurance tends to be open ended, usually terminating in the event of the death of the client. Premiums under long-term insurance typically do not vary greatly as a function of time. It is however worth noting that premiums may be actuarially unfair in order to compensate for the risk associated with covering a life (Barseghyan et al., 2013). Funerals are a key life-cycle cost for low-income South African households. As such funeral insurance is a major form of life insurance in South Africa (Roth, 2000). It is for this reason why funeral insurance products are a key component for South African insurance company's product portfolios including the company's dataset under review in this research.

### **2.3 Factors of Insurance Take-up**

This paper defines the take-up of insurance as the purchasing of insurance and is analogous to a demand for insurance. The demand for insurance is primarily a response (a hedge) against risk. It is therefore a form of risk aversion behaviour. Since insurance ownership is a hedge against risk, it is intuitive that it could be used as a hedge against financial vulnerability. Bernheim, Forni, Gokhale, and Kotlikoff (2003) found no significant correlations between financial vulnerability and life insurance, but this was later contrasted by Lin and Grace (2007) who reported a positive relationship between the amount of life insurance consumed and levels of financial vulnerability.

## 2.4 Factors of Insurance Lapses

Hou, Sun and Webb (2015), report that low-income and low-wealth individuals are more likely to lapse their insurance policies. They continue to state that the penalties of lapsing are substantial, since individuals who lapse are correspondingly more likely to truly need the benefits of their insurance policies due to the lack of financial capacity to deal with life events. Moreover, having insurance cover and lapsing it, is actually counterproductive and goes against the financial intermediary purposes of insurance in the first place. Individuals who lapse not only relinquish the policy benefits, but adopt a tactic of drawing down their wealth in paying for insurance premiums while they still retain their coverage. It is for exactly this reason why it is important to understand which customers are at risk of lapsing their insurance products, and also to design insurance products that facilitate the circumvention of lapse altogether. This is the one side of the argument, but the affordability of insurance is also an important aspect preventing lapses. Eling and Kiesenbauer, (2014) reported that the tendency to lapse increases with a decrease in income, since more low income customers are less likely to consistently afford their premiums.

Lapses are typically defined as the termination of an insurance policy without any payout value, while the term “surrender” is used when a value is paid to the policyholder on termination of the policy. The term “lapses” is used to refer collectively to both lapses and surrender in this paper and this approach is consistent with Eling and Kiesenbauer (2014), Gatzert, Hoermann and Schmeiser (2009), and Kuo, Tsai and Chen (2003).

Lapses are especially important to life insurance managers, since they also affect an insurer’s profitability and liquidity. Lapses are similarly relevant for insurance regulators, since large scale lapse events may distress aggregate financial stability. The risk of lapses considerably affects an insurance company’s solvency capital requirements (Christiansen & Niemeyer 2014). However, Russell, Fier, Carson & Dumm, (2013) argue that a lapse may have a significantly adverse result on the policyholder’s wealth, since the surrender value of the insurance policy may be low compared to the value of an in-force life policy. This is normally what happens, especially with young policies and many insurance policies may even have no surrender value at all.

Lapse rates are naturally an important consideration when determining consumer appetite for insurance products, since the factors causing policyholders to lapse their policies, would similarly dissuade potential customers from taking out the policies in the

first place (Hou et al. 2015). This essentially confirmed earlier research as lapses are also important to consider, since they indicate a mind-set shift in the policyholder to relinquish insurance cover (Kuo, et al., 2003, Eling & Kiesenbauer, 2014). This makes studying lapse rates a useful tool for determining insurance propensity.

Eling and Kiesenbauer (2014) studied a large dataset to determine which features of the contracts influenced lapses. They found that product characteristics (product type and contract age) as well as policyholder characteristics (age and gender) were important drivers of lapse rates. While this is important to understand for the purposes of value and risk-based management practices, it is limited for the purposes of designing tailored products that would better serve customers' needs.

It has been well understood for a long time (Outreville, 1990) that there is a moderately close correlation between characteristics of the agent, factors related to the outside environment, characteristics of the insured, and characteristics of the product to lapse rates. This is confirmed by more current sources, such as Barsotti, Milhaud and Salhi (2016) who state that most activities of insurance companies, such as product design, reserving, pricing, capital allocation, asset and liability management, and risk management are influenced by policy holders' behaviours, especially lapse behaviours.

Traditionally, the analysis of lapse root causes has been the "interest rate" and "emergency fund" hypotheses (Kiesenbauer, 2012). The emergency fund hypothesis supposes that financial distress or lack of liquidity, causes policyholders to surrender their policies in order access the value of the policies. While the interest rate hypotheses supposes that lapse rates increase when there are conditions of increasing market interest rates or returns (external rates of return) and negatively by increasing internal rates of return (when surplus is returned to policyholders as an added insurance benefit) (Kiesenbauer, 2012). The question that this study argues is; which customer financial parameters (income, savings, and debt) provide insights into lapse behaviours? The findings may prove insightful to assist insurance companies to design products that are effective at reducing lapses.

Kuo et al., (2003) considered US insurance data from annual statements filed by various life insurance companies, government agencies, and trade associations. They considered annual voluntary termination rates for life insurance policies which were in force in the time period from 1951 to 1998. The ratio of the number of lapsed policies to the average number of policies, in force, is known as the voluntary termination rate. Their study spanned a period of time in which there were highly volatile interest rates. They found that the effects of interest rate fluctuations are only slightly significant as a

shorter-term reaction and this is consistent with emergency fund hypothesis. Furthermore, Kuo et al., (2003) showed findings built on, and were consistent with, Outreville's (1990) early work on the emergency fund hypothesis. However, in contrast to Outreville's (1990) work, Kuo et al., (2003) did additionally find a long-term relationship between lapse rate, interest rate, and unemployment rate which has a (statistically) significant faculty in explaining the longer-term behaviour of policyholder lapse rates and their work was later corroborated by Eling and Kiesenbauer (2014).

Kiesenbauer (2012) found that consumer confidence, yield of current interest rates, and GDP development were the most relevant economic indicators affecting lapse rates. The interest rate and emergency fund hypotheses do not hold for traditional, that is, not unit-linked, life insurance products. Both hypotheses, however, are supported when other business (representing almost exclusively unit-linked products) is considered.

From a regulatory standpoint as well as a drive to treat customers fairly, there has been consistent pressure on insurance companies to practice good product development in order to reduce the number of lapses for a very long time (Johne, 1993). There are a number of factors related to product development that influence lapses, namely pricing, arrears management, and premium collection methodology. However, from a consumer perspective, Fier and Liebenberg, (2013), reported that a far more important factor affecting lapses was customer debt and financial stress. They reported that greater debt was positively associated with increased lapses.

In order to price an insurance product effectively an insurance company must take into account operational aspects, such as acquisition costs and customer value based considerations such as expected claims experience, but the cost of offering a product is also highly dependent on the effective duration and convexity measures (Cummins & Santomero, 2012). Cummins and Santomero (2012) argue that convexity numbers are more sensitive to lapse assumptions (used to calculate product pricing) than duration numbers. They also state that a misspecification of interest rate sensitivity of lapses and other estimates, may lead to sizeable miscalculations in the effective duration estimates and may in turn create an even larger error in convexity estimates. Furthermore, they state that many insurance companies believe there is not sufficient reliable data on which to stipulate the relation of lapses to interest rate movements. This shortage of confidence that insurance companies have in this central parameter, again feeds back into greater lack of confidence in convexity estimates. According to Cummins and Santomero (2012), static lapse assumptions and exposure to interest rate risk are crucially important parameters for life insurance companies to consider



when pricing products (a fundamental of product development). This paper argues that this may be a central cause for insurance premiums to be actuarially unbalanced. This should be addressed in order to offer products that provide better value.

These economic influences in turn have distinct implications for further research on how consumer level financial behaviours (levels of income, debt, and savings) can be linked to lapses. In the same way that financial metrics such as income, debt, and savings can be linked to the propensity of consumers to take up insurance, so too can lapse rates be negatively linked to the same metrics. Finding correlations between the aforementioned metrics and lapse rates, may be insightful to insurance companies in designing insurance products that succeeds at maintaining consumer appetite for the insurance products in the long run.

## **2.5 Factors of Insurance Cancellations**

Cancellations (and to some degree lapses and purchasing) of insurance products are the outcome of a change in some indicator level that has implications at the individual level (Guillén, Nielsen, Scheike, & Pérez-Marín 2012). However, Valdez et al. (2014) argue that cancellations also have important implications at the company level. Therefore, being able to mitigate against cancellations, is important since cancellations may have negative consequences for the individual as well as the insurance company.

When a customer informs an insurance company of their policy cancellation, it denotes one of two things. Either the reason for the agreement is no longer applicable (this is referred to policy termination) or a rival insurer will take over the original agreement (this is known as policy cancellation) (Guillén, et al., (2012). For the purposes of this research, these two concepts are referred to interchangeably, since they both lead to the same outcome: the insurance policy is stopped with the insurance company concerned.

### **2.5.1 Policy Termination**

Valdez et al. (2014) published a study on life insurance policy termination and survivorship. They claim that policyholders who terminate their policies are believed to have better mortality risks than those who continue with their insurance. It is for this reason why policy termination is a particularly relevant factor to consider in insurance propensity studies, since it is desirable to have a favourable mix of mortality selection. They claim that individuals who terminate their insurance, are typically able to search for insurance cover elsewhere at potentially more desirable premium rates. Those customers who remain insured will subsequently have a worsened mortality selection

than originally estimated, thereby producing greater than expected early claims. These claims are further substantiated by the published monograph by the reinsurer; Munich RE (Donnelly, 2011).

### **2.5.2 Policy Cancellation**

Policy cancellation may be in the form of brand switching and since an individual may well be in possession of several agreements with a single insurer (to cover multiple risks such as life, property, health or liability); choices made about one of the products held by an individual may be affected by events that occurred with different insurance products, all held by the same individual, with the same insurer (Kamakura, Kossar, & Wedel, 2004).

Guillén, et al., (2012) disagrees with the above and reports that are other factors that may affect cancellations. These are primarily the way in which claims are processed. In the event of a claim, the financial compensation, as well as the manner in which insurance firms handle customers, will affect the customer lifetime duration. Assistance in filing a claim, the efficiency with which claims are reported, handled and processed may all affect customer retention. Schlesinger and Schulenburg (1993) found that among German insurance customers who switched from one insurance company to another; 52.5% of the claims made with the previous insurance companies took longer than three weeks for pay out. The consequence is that cancellations may often occur as a result of poor service delivery. These trends were also later confirmed by Kumar and Srivastava (2013). This argument may be an interesting starting point for future research, as it is unknown which of the above two arguments is stronger in the South African insurance context.

To create an even more complex scenario, Darooneh, (2007) reports on the experience rating mechanisms that insurance companies use which create a price increase for the insurance premium in the 12 months following a claim. This is often seen as an unfair penalty by the customer. Darooneh, (2007) contradict this switching theory by reporting that in some cases (where a poor claim experience is had), the insured customer may have to pay a higher than expected premium when changing to a new insurance provider. This may lead them to not cancelling with the existing insurance company delivering the poor service levels.

Guillén, et al., (2012) analysed insurance customer loyalty. They argued that the effect of claims and claims handling on customer lifetime duration and cancellations, is a difficult hypotheses to formulate. However, they propose that it is reasonable to assume that the effect may change over time, particularly when considering the effects

of delayed claim compensation payment by an insurer. Guillen et al., (2012) comment that insurance is automatically renewed each year in most European countries. It is typically levied from the policy holder's account with some warning. Renewal of the contract does not need to be demanded, however, if the customer wishes to end their insurance cover, then the insurer must be notified one month before the policy expires, lest it is renewed automatically. This is in contrast to how insurance contracts are handled in the South African context.

In South Africa, insurance companies are obliged to send an "annual statement" each year to their customers, typically on the anniversary of their policy commencement (Long-Term Insurance Act of 1998). They also do not have to renew their cover each year. However, unlike in a European context; many life insurance companies are allowing customers to cancel their cover at any time. In this study, data is used from a life insurance provider who allows customers to cancel their policies without the need to notify the company in advance, i.e. cancellations are immediate. This is an important difference to take note of, since most of the literature is therefore limited to applicability in a South African context.

Many Insurance providers strategically target newly acquired customers for cross-selling. Cross-selling is the sale of additional products that are different from those bought previously (Schmitz, Lee and Lilien, 2014). Once a first policy is transferred to a new insurer, the company will aim to transfer as many other insurance contracts. When notifying an insurer that the policy will be cancelled, insurers (on behalf of the customer) or customers themselves, typically announce the cancellation at the last minute, thereby leaving very little time to for the original insurer to retain the customer. Remaining insurance policies are therefore at risk of being moved to the new insurer after the first contract has been moved from one company to another. Guillén et al. (2012) continue to state that the expectation is that the influence is more significant directly following the first cancellation as compared to later, thereby introducing a time-changing effect. They report that the types of insurance policies that may be successfully retained after the first cancellation, could influence the success of retaining a client. For instance; keeping the household insurance in place after a policy cancellation has been seen to contribute to a greater than expected period with the original company after the first cancellation (Brockett, Cooper, Golden, Rousseau & Wang, 2005). Schmitz et al., (2014) however, argues that cross-selling in complex scenarios reduces the cross-selling performance. This concept can be linked back to tailoring initial cross-selling efforts because of the work done by Brockett et al. (2005). They provide further justification that some insurance products are seated more

strongly in the consumer's mind and therefore, it is probably more efficient to begin campaigns with these products and follow them with other products that do not have the same amplifying effect, in order to optimise the cross-selling efforts. It can be debated, however, whether stand-alone product design or cross-selling has the strongest effect on insurance consumption and the inverse, namely cancellations.

Kunreuther (2015) contested the above arguments. They report that one reason that customers cancel their insurance policies, is that insurance is often viewed as an investment rather than a financial means of protection. Individuals often purchase cover post experiencing loss as a result of a disaster. Yet at the same time, perceive that their premiums would be wasted if they do not claim over the few years after purchasing the insurance. For these customers, the probability of a disaster is so low that they ignore the consequences of a disaster and presume that insurance is not needed. Kunreuther (2015) goes on to explain that a normative model of choice (for example "expected utility theory") means that risk-averse consumers will place a higher value in insurance, since relative to their wealth, it guards them against large losses. Customers should, in theory, celebrate if they have not sustained any losses rather than cancelling their insurance because no claim has been made over the past few years. Most insurers face the challenge of informing their policyholders that the optimal return for any policy is no return at all.

This research paper, which is focused on finding relationships between individual-level financial behaviours by customers and insurance propensity, does not expect to be able to deduct any interesting findings pertaining to brand switching. Brand switching is important to understand, but it is outside of the scope of this study since brand loyalty is a large subject, worthy of its own focus. This can however form part of future studies with a more rigorous analysis, where switching to other insurers may be identified. In fact, future research may use a cancellation event as a trigger to analyse the behaviour of a specific client and monitor any further cancellations of other insurance products, as well as a change in premium paid to the new insurance company. This may provide further empirical evidence to substantiate the work done by Brockett et al (2005) and Guillén et al. (2012).

## **2.6 Risk Aversion**

*"Risk aversion is the primary reason for the existence of insurance markets" (Cohen & Einav, 2007)*

In his 1964 seminal work in *Econometrica*; John Pratt proved that a consumer has a larger risk aversion as compared to another consumer, if and only if, they are globally

more risk averse. By this he meant that; for all risks, the sum of money which they would exchange for the risk is less than for the other decision maker. Therefore, the expected monetary value minus cash equivalent (risk premium) is always larger, and they will always spend more on insurance. Another seminal paper on consumer behaviour towards insurance, is the work done by Yaari (1965) as discussed in 2.2 above (Pratt (1964) and Yaari (1965) are both still relevant foundations for researchers and textbooks alike despite their age according to Levy, (2015)). When assessing the relative appetite of insurance products, one assumes that consumers, who meaningfully plan for their financial future, consider the uncertainty of their length of life and risk of illness, accident or disability (Yaari, 1965). He showed how to solve the problem for determining a consumer's optimal consumption plan ( $c^*(t); t \geq 0$ ) until their time of death ( $t \leq D$ ) (for definitions of the variables please see section 2.2). He further derived the Euler-Lagrange equation for the most favourable trajectory of wealth and the function for its related consumption. His influential work (Shi et al. 2015) on lifetime uncertainty in a lifecycle model substantiated the assertion that lifetime uncertainty increases consumption and is analogous to behaviour under increased discount rates. This is a small but significant difference to the argument provided by Pratt (1964).

Barseghyan et al., (2013) used insurance decisions to approximate standard risk aversion; they state that households are generally averse to risk. Consumers therefore necessitate a premium to invest, and so they buy insurance at actuarially imbalanced rates. Barseghyan et al. (2013) explain that the general expected utility model ascribes risk aversion to a concave utility function, which is defined over different states of wealth with a declining marginal utility for increasing wealth. It is true that several empirical investigations on the risk preferences of households accept the expected utility and approximate a "standard" risk aversion as reported by Cohen and Einav, (2007).

Based on the above-mentioned literature, it may be confirmed that general risk aversion and insurance consumption are highly related. This has been understood for a long time in the context of developed nations. The link that this investigation aims to establish, is the relationship between risk-averse behaviours (such as having large savings and low amounts of debt) and the propensity for insurance consumption in the unique South African situation as expressed previously. The argument exists that there may be elements of the South African insurance market that is different to the status quo of the literature.

### **2.6.1 Risk Aversion Behaviour**

Individual attributes, as well as psychological dispositions or behaviours that profile widespread financial and investment practices, are a growing field of study according to Outreville (2014). He also argues that the study and understanding of the financial behaviour of policyholders, is an imperative issue in insurance. The consequence of not having sufficient understanding of how customer financial behaviour is linked to insurance consumption, is certainly a missed opportunity for improved customer value. If a company has a better understanding of its customer, then they will be better equipped to design products that serve their customers well and in so doing, ultimately be more competitive and profitable (Storbacka, 1997 and later Tomczyk, Doligalski & Zaborek, 2016). The consequences of understanding the behaviour of individuals facing uncertainty are valid for insurance and for most financial services sectors (Outreville, 2014). Excitingly, it has been argued that studying insurance poses an especially promising field for empirical research on contracts (Chiappori & Salanié, 2003).

Chiappoli and Salanié (2000) emphasise that there is a strong positive correlation between risk and demand for insurance, even though they are considered automobile insurance and not life insurance. They concluded that if various individuals have various levels of risk aversion, and assuming that individuals who are more risk adverse are expected to diminish the exposure and to purchase insurance, it would propose a negative relationship between insurance and accident rate of occurrence. Outreville, (2014) claims this form of correlation is a required circumstance for adverse selection, while the absence of a correlation is adequate to eliminate significant adverse selection behaviour.

Zimmer, et al., (2009) clearly demonstrated that the individual's attentiveness to their risk of default on their actual insurance premiums, has a marked influence on their insurance consumption behaviour. Zimmer et al., (2009) were able to experimentally test consumers' response to the risk of insurance default. They report that insurance with a risk of default is particularly unappealing to most consumers, since defaulting results in a partial or complete loss of invested premiums for the customer. Zimmer et al., (2009) report that a large portion of their study refused to accept any default risk; while another portion asked for a large drop (discount) in insurance premiums. A principle finding was that their experiment proved robust when testing different reasons of default, including insolvency. This provides a meaningful argument to test and understand a customer's appetite for insurance products with different levels of default risk.

The concept put forward by Zimmer et al., (2009); that customers are willing to pay different amounts for different insurance contracts based on the amounts of insurance default risk, highlights the fact that customers are sensitive to details of insurance policy features. This research paper similarly argues that there are linkages between the individual-level financial behaviour of customers and their insurance choices. Just as Zimmer et al., (2009) argue that a given customer with a certain psychological disposition towards default risk should have their product tailored (in this case with regard to price) according to their willingness to pay for such a product, so does this research paper argue that insurance companies must understand their customer responses to insurance products based on other metrics. In the case of this research, these are specifically the financial behaviours of customers (income, savings and debt).

### **2.6.2 The Risk-Aversion Function**

According to Simon and Blume (1994) the Arrow-Pratt degree of relative risk aversion is defined as

$$R(c) = cA(c)$$

$$= \frac{-cu''(c)}{u'(c)}$$

**Equation 6**

where  $A(c)$  is the absolute degree of risk aversion and where a greater curvature of the utility function  $u(c)$  denotes a greater degree of risk aversion.

The function as described by Pratt (1964) and Arrow (1965) begins from the notion of risk premium (the difference between the anticipated economic value of an uncertain income when price is static) and the certain economic value that the consumer would trade for the uncertain income. In the literature described by Deschamps (1973) there are two definitions of risk aversion that may be used as alternatives for one another.

Deschamps (1973) described the direct utility function as  $U(X)$  and the indirect utility function as  $V(y; P)$ . The functions are homogeneous of degree zero in  $(y, P)$ , where  $y$  is income and  $P$  is price.

Absolute risk aversion is defined as double the ordinary risk premium per infinitesimal unit of variance for risks. It measures the premium which a consumer is willing to pay for insurance cover to protect themselves against small risks. In some cases, (where the consumer is risk-loving) the absolute risk aversion may be negative.

By considering the indirect utility function, the absolute risk aversion function is

$$-\frac{V_{yy}}{V_y}$$

**Equation 4**

and depends only on  $y$  (which is income) and  $P$  (which is price).

The proportional (or relative) risk aversion is:

$$-y \frac{V_{yy}}{V_y}$$

**Equation 5**

This may be understood as the local measure of risk aversion as a proportion of consumer income. Considering the cardinal utility function, it may also be understood as the elasticity of marginal utility of income.

Deschamps (1973) goes on to describe absolute risk aversion functions and relative risk aversion functions as functions which remain unchanged when specified linear transformations is applied with respect to the utility function, but not to other transformations of the indirect utility function.

If  $W = F(V)$ , then

$$\begin{aligned} -\frac{W_{yy}}{W_y} &= \frac{V_{yy}}{V_y} - \frac{F''}{F'} V_y \\ &= -\frac{V_{yy}}{V_y} \end{aligned}$$

**Equation 6**

if and only if  $F'' = 0$ .

In the same way Deschamps states that:

$$-y \frac{W_{yy}}{W_y} = -y \frac{V_{yy}}{V_y} - \frac{F''}{F'} y V_y$$

**Equation 7**

And therefore, the absolute risk aversion function and relative risk aversion function are determined if (and only if) the utility function is cardinally defined.

Deschamps further stated that all hypotheses on risk aversion should satisfy at least some homogeneity conditions. Therefore, as  $V(y; P)$  is homogeneous of degree zero in income and price, the absolute risk aversion function is homogenous of degree minus one while the relative risk aversion functions are homogeneous of degrees zero.



Using normalized prices and income, the absolute risk aversion function is:

$$-\frac{V_{y^0y^0}^0}{V_{y^0}^0} = -p_1 \frac{V_{yy}}{V_y}$$

Equation 8

while the relative risk aversion function is:

$$-y^0 \frac{V_{y^0y^0}^0}{V_{y^0}^0} = -y \frac{V_{yy}}{V_y}$$

Equation 9

Deschamps notes that it is particularly interesting to notice that regardless of using normalised or non-normalised variables (therefore regardless of which measure of value or exchange is used), the relative risk aversion function remains the same, but the absolute risk aversion function does not behave this way.

When considering only normalized variables, the indirect utility function  $V^0(y^0; P^0)$  is not homogeneous. Therefore, any hypothesis about relative or absolute risk aversion does not need to comply with conditions of homogeneity.

Schroyen, (2013) showed that these measures of risk aversion are useful to understand how attitudes towards personal risk may change when experiencing income shocks or becoming unemployed. Due to the high levels of unemployment, some consumers may be using life insurance (such as retrenchment cover) to protect themselves from these risks Fields and Kanbur, (2007).

### **2.6.3 Relationships Between Functions of Demand and Risk Aversion**

The demand functions (and in particular – functions of insurance demand) are ordinal utility functions, but risk aversion functions are cardinal utility functions. It is thus impossible to establish a one to one relation between demand and risk aversion functions.

If one begins by considering the risk aversion function, the indirect utility function can be derived under the correct conditions of integration. Dechamps further establishes that all linear transformations of the indirect utility function correspond to the initial risk aversion function and thus the indirect utility function is cardinally defined.

On the other hand, when considering the demand functions, it is also possible to obtain the indirect utility function when under the correct integration conditions. As opposed to the risk aversion function, the demand function however is ordinal, and this leads to a group of risk aversion functions linked by equations Equation 6 and Equation 7 above.

Given the demand functions, we obtain the absolute risk aversion function  $-\frac{V_{yy}}{V_y}$  and the related equations seen in Equation 6 above.

Early studies on relative risk aversion have seen the publication of polarising results. Friend and Blume (1975) showed evidence that individuals invest comparatively larger proportions of their capital in risky investments with increasing wealth. Stiglitz, (1969) however, showed through his work, that relative risk aversion will increase with increasing wealth. There is still no agreement between economic scholars and this is still currently a contentious source of empirical studies, according to Oureville, (2014). It may thus be concluded that the consumer responses to how they manage their finances as a result of risk aversion, is still only partially understood. This further substantiates the reason for this investigation on the relationship between insurance purchases (an established form of risk aversion) and income and savings which are complementary measures of wealth.

Meyer and Meyer (2005) found that differences in the way in which a risky decision is measured, considerably changes the measure of the relative risk in the individual. In fact, relative risk aversion seems to depend highly on the way in which wealth is measured. In this study; income is a measure of the actual inflow of money into a customer's bank account, while savings is a measure of the cash savings and other cash investments that an individual is in possession of at any time. Therefore these may be quite accurate measures of wealth in comparison to what has already been studied. The relative sample of the study also affects this outcome as shown by Haushofer and Fehr, (2014). They found decreasing relative risk aversion with higher income and an increase in relative risk aversion with lower income. This investigation should, therefore, add to the body of knowledge on this matter, since it offers another approach of measuring both risk aversion and a very robust measure of income.

#### **2.6.4 Prudence and Temperance**

While there is a great deal of literature on the subject of relative risk aversion, there has been less work published in determining higher-order risk preferences of temperance and prudence which are also central concepts in economic decision making, according to Ebert and Wiesen (2014). Prudence and temperance are known as third-order and fourth order risk-aversion respectively. Both these concepts affect individual's behaviour towards risk and are therefore important to consider (albeit briefly) for the purposes of this research.

According to Kimball (1990), the concept of prudence and its inferences have been deployed in evaluating demand for savings from as early as 1968 by Leland and 1970 by Sandmo. Both authors showed that, within the expected utility setting, a risky income in the future does not necessarily equate to an increase in savings except if the consumer is prudent.

The term “temperance” was also coined by Kimball (1992). It is a concept implying that the initiation of any inescapable risk, would cause a customer to decrease their exposure to other risks even if they are statistically independent. According to Franke, Schlesinger, and Stapleton (2011), higher-order risk preferences (such as prudence and temperance) perform a vital part in choices where there is background risk.

Ebert and Wiesen (2014) showed that the aggregate traits of risk aversion, prudence, and temperance, correlate at the individual level. Interestingly, in considering gender, they found that women demand greater compensation for aggregated risk aversion, prudence, and temperance. This research does not consider gender as it is beyond the scope of study (this paper only considers financial factors as opposed to the already well understood demographic variables) but it is nonetheless important to bear in mind, since individual-level financial behaviours of consumers may similarly be affected by prudence and temperance.

Ebert and Wiesen (2014) also observed a substantially lower compensation for second-order risk than for the downside risk compensation. Unlike the frequently used utility functions, prospect theory (which was first described by Kahneman & Tversky, 1979) can accommodate these findings. Ebert and Wiesen (2014) find that both risk “lovers” and risk “averters” are prudent and thus the demand for risk reimbursements are not meaningfully lower for risk “lovers” or higher for risk “averters”. This is consistent with the theoretical arguments from Crainich, Eeckhoudt, and Trannoy, (2013) of prudence being a common behavioural trait, most likely shared by risk “lovers” and “averters” alike. Their findings make a strong argument that higher-order risk preferences are indeed relevant. The implications of this study, especially the manner in which individuals treat their finances (especially savings and debt as a ratio of income) has some implications for research on prudence and temperance. It is therefore suggested that future studies can consider the higher-order risk aversion behaviours by expanding on the work done in this research. Fei and Schlesinger (2008) found the concepts of prudence and temperance useful in determining choices on insurance demand.

Ebert and Wiesen (2014) further report that the empirical literature written on higher-order risk preferences is scarce. To address this gap, they assess the *intensity* of customers' risk preferences instead of only noting the *direction* of the risk preference. The scope of this research paper is limited to the direction of insurance choices only. In future studies, however, the number of insurance products purchased, or the amount of cover taken (level of cover), may also be used as a proxy for *intensity*.

## 2.7 Education

Dragos, (2014) reports that education is a demographic factor that is anticipated to have a positive influence on the propensity for the up-take of insurance. In the context of this research paper; Fernandes, Lynch & Netemeyer, (2014) reported that there is some relationship between education, financial literacy, and the downstream financial behaviours of individuals. They however, argue that the effects of financial education declines over time, with even substantial education interventions with numerous hours of tuition, having insignificant influences on financial behaviour 20 months from the time of education. This study again argues that directly relating individual-level financial customer behaviours to insurance take-up, lapse, and cancellations, is a good approach to customer understanding. The literature on the relationship between education and insurance demand will be detailed below.

Even though the association between risk aversion and level of education is not clear (Outreville, 2014), the effect of education on insurance consumption has been established to some degree (Dohmen, 2011; Dragos, 2014). Kimani, Ettarh, Warren and Bellows (2014) found a positive correlation between multiple factors such as employment status, marital status, media consumption, education achievement (secondary education or higher) and household wealth, and a greater consumption of health insurance among women in Kenya. Similarly, Kirigia, Sambo, Nganda, Mwabu, Chatora and Mwase (2005) determined that South African women who had completed high school or earned higher incomes, were more likely to consume health insurance.

However, Outreville, (2015), concluded that it is impossible to postulate that there is a positive association between education and degree of risk aversion. He further concluded that for emerging and developing countries, and especially in the insurance sector, there is a need for further empirical and theoretical research in this area of insurance. This is because firstly, there is a general scarcity of data related to the issue of education and insurance, and therefore the empirical analyses available has been constructed on an inadequate sample of countries (predominantly developed countries). Secondly, Outreville, (2015) proposes an extension to this concept in order

to analyse the wider concept of human capital development. He affirms that the results of such studies contain valuable implications for further macroeconomic research on the demand for life insurance or other financial products. This provides further justification for this research in the area of insurance demand in South Africa.

Frees and Sun, (2010) as well as Burnett and Palmer, (1984) state that education often provides individuals with the ability to manage risk more effectively and find a positive correlation between levels of education and the demand for life insurance. Szpiro, (1985) and Luciano, Outreville, and Rossi, (2016) however, found that there exists a negative correlation among levels of education and risk aversion. They find that that greater level of education precedes lower levels of risk aversion. Thereby leading to increased risk taking by well-educated individuals, again leading to the impression that the literature on the subject is divided.

Early studies postulate that increasing education is negatively related to the demand for life insurance (Duker, 1969; Anderson & Nevin, 1975), while later studies show that education doesn't seem to be a significant driver for life insurance consumption (Feyen, et al., 2011), further adding to the inconsistent nature of the literature. Zietz (2003) reported that a variety of research papers shows that education yields conflicting results as a model for life insurance propensity. It is ambiguous as to why, but education has been found to be largely insignificant for determining life insurance demand by a substantial share of literature. However, Outreville, (1996) stressed that persons with high levels of education were more cognisant of their risks and the relative consequence of poor risk management. But even so, it was impossible for him to empirically verify a significant relationship between education and life insurance. Beck and Webb (2003) showed that increasing the dependency period, schooling seems to have no strong influence on life insurance consumption demand. Individuals with higher education generally have higher incomes and tend to purchase life insurance. The findings seem somewhat confusing since there is a strong correlation between schooling and income (Dragos, 2014).

Hau (2000) found that it is indistinct whether education positively or negatively affects life insurance demand, while Dragos (2014) further proposed that education is a poor proxy for the aptitude of an individual to comprehend the intricacy of insurance products, since familiarity of insurance products may not be introduced properly at school level.

The primary reason why education seemingly affects insurance consumption, may be explained by the connection between relative risk aversion and insurance consumption.

The literature is inconsistent when it comes to the effect of education on relative risk aversion (Outreville, 2014). Hersch (1996) as well as Jianakopulos and Bernasek (1998) find a positive correlation between education and relative risk aversion while Riley and Chow (1992), Bellante and Green (2004), Harrison et al. (2007), Lin (2009) and Kapteyn and Teppa (2011) find a negative correlation between education and relative risk aversion.

Barsky et al. (1997) and Halek and Eisenhauer (2001) suggest that the relationship between education and relative risk aversion is most-likely nonlinear and that the number of years of education affects the positive or negative nature (as well as significance) of the relationship. More interestingly, financial risk aversion or risk-taking, may be a far more pertinent concept to study, since the access to financial knowledge rather than general education probably has a greater effect on the relationship (Bayer et al. 2009).

Bayer et al. (2009), Van Rooij *et al.* (2011) and Clark et al. (2012) all found that savings activity is significantly greater when employers offer financial education, such as retirement seminars. Not only did they find that financial education lead to relative risk aversion activity (savings), but the effects are amplified for employees that are paid less as compared to employees that are paid more. This is particularly interesting for this research, since looking at actual savings as well as another risk aversion behaviour namely; insurance consumption. Savings is the result of good financial education and planning, but culturally, South Africans have a very low domestic savings rate which propagates into low-growth (Aron & Muellbauer, 2000; Aron & Muellbauer, 2013).

This paper will not specifically look at levels of education as a predictor for propensity of insurance consumption, since this research paper is primarily concerned with the financial behaviours of customers and insurance propensity as discussed above. By providing a brief description of the work done by other authors, the reader may appreciate that differing levels of education does influence insurance consumption to some extent as evidenced by the above-mentioned literature, albeit that the results are of often ambiguous. The reader should also bear in mind that the concepts of education (as well as other demographic concepts) are in turn related to the individual-level financial behaviours of consumers (Fernandes et al., 2014).

## **2.8 Financial Sophistication**

Mulholland et al. (2016) report that that financial sophistication is an important issue in life insurance demand. They state that increasing financial sophistication leads to

increasing insurance consumption. To this end, income and savings and debt, are discussed in an effort to indicate what the literature argues on each concept.

### **2.8.1 Income**

Research into the demand of life insurance as a function of income, has been performed for quite some time (Yaari, 1965; Fischer, 1973; Campbell, 1980). These researchers all found that households systematically use life insurance as a protective risk-based measure against the loss of income flows (from labour) over the life cycle. Browne and Kim, (1993) report that many models interpret life insurance as a way of reducing uncertainty in income. This is of course in relation to the risk of possible death or disability of the principal income earner. The work was later confirmed by Akotey, Osei, and Gemegah, (2011) who showed that low income individuals demand insurance cover but cannot always afford it.

Lange et al. (2017) recently found that income is one of the most important drivers of the demand for supplemental health insurance. Frees and Sun, (2010) found that the influence of income on life insurance demand is analogous to the effect of wealth as a determinant of life insurance. By this, they mean that low income individuals may not always be able to afford life insurance and therefore cannot exhibit risk aversion (by purchasing life insurance). They used regular before tax salary and wage data as an amount of income. Similarly, Mulholland et al. (2016) found that households who take up cash value life insurance, are on average more financially sophisticated and have greater income. Therefore signifying that life insurance is used as a tax shield rather simply as protection against losses in income (human capital). This is somewhat contradictory to what Lin and Grace (2007) reported. They reported that an individual with decreasing risk aversion will consume lower amounts of insurance when at higher income levels. They however also argued that income levels which are higher may create increased risk for a household and in this way raise the demand for life insurance products. Other earlier research, such as Burnett and Palmer (1984) show confirmation of this second theory.

The findings by Lin and Grace (2007) and Mulholland et al. (2016) have an implication for insurance companies wishing to offer optimal insurance products for their customers, especially those customers who are price sensitive. It is important to note that the literature on this topic is quite old (Browne & Kim 1993 and Outreville, 1996). Additionally, only one of these papers considered life insurance in developing countries; namely Outreville (1996).

Outreville's (1996) eminent empirical study on the relationship between life insurance and income did not include South Africa as part of the research, further justifying the need for investigation. It is therefore unknown whether or not South African consumers conform to the same pattern as found in the above-mentioned research. South African trends in income distribution have been shown to be unique, due to the high poverty and inequality levels, the latter being among the highest in the world (Leibbrandt et al., 2010). The need for investigating the link between income and life insurance in the South African context, is further complicated by the nature of income distribution in South Africa and the complication of consumption patterns that have been linked to race (Kaus, 2013; Aron & Muellbauer, 2013).

### **2.8.2 Savings**

One core precept of this research will be to determine the consistency of Headen and Lee's (1974) insight that insurance and financial assets (savings) have a tendency to mutually proliferate with the South African context. Their work is reinforced by some sources, highlighting the substitutability or complementarity of savings and life insurance (Chen et al., 2006; Huang & Milevsky, 2008). However, Peter, (2017) makes a strong counter argument to this. He argues that heterogeneity is not clearly seen in the empirical data. He argues that prudent individuals are less likely to purchase insurance, and imprudent individuals are more likely to purchase insurance. He continues to argue that consumers with greater levels of savings are less likely to consume insurance. Therefore, this paper seeks to determine which literature is more relevant in the South African context, since culturally South Africans have a very low domestic savings rate (Aron & Muellbauer, 2013). To further contrast the literature, Shi et al. (2015) reported that both current and future household income, have a curvilinear impact on the demand for life insurance.

Risk aversion and savings have been shown to be well associated with one another (Ebert & Wiesen, 2014). As discussed previously, early studies by both Leland (1968) and Sandmo (1970) showed that, within the expected utility setting, a risky income in the future does not necessarily equate to an increase in savings unless the individual is prudent.

Peter, (2017) conversely argued that self-protection (insurance) is an expensive method in the reduction of risk as compared to savings. He showed that prudence is negatively related with the optimal savings for the purposes of protection against risk. However, he agrees with the effect of interest rates in conjunction with the interest rate hypothesis, where the individual seeks optimal self-protection in accordance with the



interest rate for savings and investments, and therefore seeks to increase the chances of having a less risky situation by choosing the option with the highest return. However, he continues to argue that the results show that insurance consumption centres on whether the individual also uses savings to improve their intertemporal consumption utility. In contrast to much of the literature, they find fundamentally different when an individual's portfolio comprises of insurance and saving or of insurance only. They also argue that it is irrelevant to the decision maker whether the insurance expenditures are upfront.

In 2008, Fei and Schlesinger used the impact of higher order risk aversion behaviours (prudence and temperance) to analyse insurance demand. It therefore seems a natural progression of closing the knowledge gap to consider the effect of savings on insurance demand in this study.

### **2.8.2.1 Background on Risk Aversion and Savings**

In the following section, the seminal work by Somerville (2004) is summarised as background to the relationship between risk aversion and savings.

Insurance is affected by risk and has an intertemporal aspect (as is the case with decisions concerning risk) (Somerville, 2004; Sliwinski, Michalski, & Roszkiewicz, 2013). Much of the work published regarding the economics of insurance is placed within a framework of only a single period, as is the case of the landmark work of Arrow (1963), Mossin (1968), and Rothschild and Stiglitz (1976). This single period approach is supportive of some significant findings, but an individual's administration of their exposure to risk will normally include savings behaviours as well (apart from only insurance) according to Somerville (2004). It is therefore important that an intertemporal model is required to discover the position of insurance cover in the individual's life-cycle consumption savings plan. This intertemporal aspect is also vital for the exploration into the impact of secular changes in the loss probability. Somerville (2004) extended the model for optimal consumption and saving by adapting the maximum principle to deal with the question of risk and insurance in continuous time. In a completely risk free environment the path of optimal consumption may be described by the differential:

$$\frac{\dot{c}}{c(t)} = \frac{r - \delta}{R(t)}$$

**Equation 10**

where  $R$  is the elasticity of marginal utility and can be described as the coefficient for relative risk aversion,  $\delta$  is the constant rate of preference of time as used for the

discounting utility, and  $r$  is the riskless interest rate received or paid. Convention (and convenience) however, makes it preferable to deal with absolute risk aversion here. Somerville (2004) further explained that when introducing risk but keeping insurance at zero, Equation 10 becomes:

$$\dot{c} = \frac{r - \delta}{A(t) + \ell \dot{\pi}}$$

**Equation 11**

where  $\pi$  is probability and  $\ell$  is loss. He states that this holds only if it is assumed that the absolute risk aversion,  $A$ , declines everywhere ( $A' < 0$ , and where:  $A = -\frac{u''}{u'}$ ). The added term is proportional to  $\pi$ , where  $\dot{\pi}(t)$  is the change in probability of loss as a function of time. When savings or even debt is justified by changes in income or interest-rates, Somerville (2004) describes that the optimal path involves precautionary debt or savings according to  $\dot{\pi}$  being greater than or less than zero respectively. It must be noted that it is the change in probability ( $\dot{\pi}$ ) that has this effect. In this context it means that when the change in probability is zero then the actual value of  $\pi$  does not have an effect on optimal choice.

When Somerville (2004) introduces insurance to the model for optimal consumption and saving, the optimal paths are linked to both cases mentioned above, i.e. to the riskless case (Equation 10) and the case where risk is present but without insurance (Equation 11). If the insurance is completely actuarially fair (an unreal scenario), then the chosen path is similar to the riskless case above but the difference must of course include the premium for the insurance. When this occurs, consumption  $c(t)$  does not change over time ( $t$ ) and the time path of the consumption function is independent of the chain of states and subsequently, there is a demand for insurance. Somerville (2004) continues to explain that the optimality conditions do not significantly depend on the change of the probability with time ( $\pi(t)$ ). A change in probability affects the magnitude of the premium through the effect on the premium rate. In this way, it causes a very similar effect as a change in income, affecting  $c(0)$  as well as borrowing or saving, but does not influence  $\dot{c}$  directly.

In a more realistic scenario, Somerville (2004) explains that if an actuarially unbalanced premium is considered, the demand for insurance is always sub-optimal at each moment. With actuarially unbalanced premiums, it can be expected that, overall, the optimum cover  $v(t)$  differs over time regardless of whether  $\pi$  is stable. The development of saving, consumption, and cover level is determined by interest rates (i.e., on  $r - \delta$ ), as is the case of no risk, and also on  $\dot{\pi}$ , as is the case of having no

insurance (the risky case). Similarly, as with no-insurance; if  $r = \delta$ , then precautionary borrowing or saving occurs according to  $\dot{\pi} < 0$  (and  $\dot{v} > 0$ ), or  $\dot{\pi} > 0$  (and  $\dot{v} < 0$ ) respectively, since  $\dot{\pi}$  is substituted straight into the equations for  $\dot{c}_1$ , and  $\dot{c}_2$  (if the insurance premium contains a risk premium). Furthermore, if  $r = \delta$  and the premium loading contains no risk premium then insurance steadies the optimum consumption function. If the insurance contains risk premium, then insurance will only partially stabilise the optimum level of state-two consumption.

If the situation exists where an actuarially unbalanced premium is present (as is the case for most insurance premiums),  $r \neq \delta$  and  $\dot{\pi} = 0$ , then the pathways in each state is delineated by the equations similar to the riskless case. If the case of declining absolute risk aversion is considered,  $|c_1|$  or  $|c_2|$ , so that  $\dot{v} < 0$  or  $\dot{v} > 0$  according to  $r > \delta$  or  $r < \delta$  respectively, and the size of  $\dot{v}$  is dependent on the dissimilarity in risk acceptance among the two conditions.

### **2.8.3 Debt**

Lin and Grace (2007) said that; for a household to have good risk management principles, the household should have protection against catastrophic losses. They continue to state that life insurance is a form of ensuring that debt, mortgages, and other obligations are settled on the death of an insured household member. Most importantly; they state that it is ambiguous whether there is a relationship between life insurance and debt. This provides the argument that this concept should be researched further.

To assist product development principles; this research aims to study which individuals (those with high or low levels of debt) hold life insurance. Previously, Frees and Sun (2010), modelled severity and frequency of demand for life insurance by building on Lin and Grace's (2007) work. Frees and Sun (2010) state that life insurance is oftentimes able to protect consumers against the financial burden of debt, but as a corollary to having this debt, it may also cause the life insurance to be unaffordable when under financial burden. They agree that the relationship between the demand for life insurance and debt is unclear and identify this as a research gap.

In the interim between the above call for further research, there has been some work done on the relationship between debt aversion and insurance take up by van Winssen, Kleef and van de Ven (2015). They write that debt aversion predicts that consumers do not like paying for healthcare after they have received the healthcare. Van Winssen et al., (2015) reported that customers who are overly debt averse, are also more risk averse and therefore demand full coverage offered by their insurance

policies. They also state that consumers choose flat-rate pricing schemes above payment decoupling. In the context of developing better products, this fact makes the concept of debt aversion germane for the choice to offer voluntary deductibles or not.

Van Winssen et al., (201) state that customers who are overly debt averse, may opt out of product features such as voluntary deductibles, since having a voluntary deductible policy in place will mean that part of their healthcare is paid for after their consumption thereof. These debt-averse customers may well prefer a flat rate pricing product feature where premiums are paid in advance. This may prove a strong relationship between debt aversion and risk aversion.

Therefore, if this investigation finds linkages between high or low levels of debt, and increased or decreased propensity to insurance purchases, it may have significant implications for designing insurance products with voluntary deductibles or not.

Frees and Sun (2010) conclude from their research that the amount of debt owed has a positive effect on an individual to have life insurance. With increasing debt, the likelihood of relatively inexpensive “term-life insurance” increases (term-life insurance is life insurance of a fixed duration). In an attempt to determine who purchases life insurance and how much they buy, Frees and Sun (2010) go on to suggest that there exists a negative relationship between term and whole-life insurance products. This is important to bear in mind when studying linkages between debt and life insurance, since it has the implication that may be a negative relationship between debt and whole-life insurance. The negative relationship they found is indicative of the frequency at which individuals take out life insurance as a response to debt. They also found a positive correlation between the amount of insurance purchased between term and whole-life insurance products. This assists the explanation of the severity of the consumption. They continue to clarify that the variegated nature of this effect may suggest that whole and term-life insurance products may in fact be forms of substitutes. On the other hand, they found a positive correlation in the amount of products purchased by individuals who own both types of insurance. To summarise; the substitution of whole and term-life insurance products clarifies the question on what frequency these products are purchased, while they are in fact complementary in determining the amount of insurance taken. It seems that it is therefore important to differentiate between which insurance products an individual holds, when studying debt linkages to insurance consumption.

## 2.9 Business Analytics

A core challenge of this research is to manage the large volume of data generated by the bank under study. Li, Moshirian, Nguyen and Wee (2007) found distinct connections between simple parameters, such as income and the demand for life insurance products. Alhassan and Biekpe (2016) discovered that in the African context, demographic factors are well suited to explain life insurance consumption. Through their sample of 31 African countries between 1996 and 2010, they examined the consequence of demographic and financial variables on life insurance consumption. They used the ratio of life insurance premiums to GDP as a proxy in their study. They found that demographic variables perform a crucial role in steering life insurance consumption in the African context. They found the macro factors such as health expenditure, dependency ratio, and education, as the noteworthy demographic determinants of life insurance. More interestingly for this study, they found financial development to be the foremost driver of life insurance consumption. Interestingly, Alhassan and Biekpe (2016) found that an increase in per capita income was negatively linked to life insurance and that an increase in dependency ratio also leads to a reduction in life insurance. This latter finding is also consistent with findings by Li, et al. (2007).

In 1965, Yaari postulated that “uncertainty of life” was the primary driving factor in the consumption of life insurance. Similarly, Karni and Zilcha (1985) found that behavioural risk aversion drove this consumption. Outreville (1996) examined the consumption of life insurance in 45 developing countries. He considered health, education, financial development, agricultural status and local market competition on the consumption of life insurance. Disposable income, local market competition and financial development were three factors most highly correlated to life insurance usage. More recently, Outreville (2014) showed linkages between demographic factors and the micro-economic demand for life insurance. From a conceptual point of view, this has particular implications for this research, since the literature has quite clearly pointed out strong links between the afore-mentioned macro-economic factors. However it has failed to address the links between personal savings, debt, and income, and the propensity for a particular consumer to take up insurance. This therefore provides exciting grounds for this research, since it considers a more granular look into the individuals handling of their finances and how this may influence their likelihood of purchasing and keeping insurance.

Elango and Jones (2011) examined the drivers of insurance demand in emerging economies between 1998 and 2008. They concluded that interest rates, gross national

income (GNI) per capita, and merchandise trade have a positive correlation on the degree with which life insurance is consumed in the general population. The authors were able to explain life insurance demand more closely with demographic factors than with economic and institutional factors. Even though the correlation between risk aversion and wealth is well understood (Outreville, 2014) and there is also evidence to suggest a positive correlation between risk aversion and insurance demand (Lewis, 1989). The work done on the direct relationship between individual-level income, savings, debt, and insurance demand is quite old and needs to be updated. This paper therefore aims to determine correlations between individual-level income and insurance demand directly, as there is a gap between Lewis (1989) and Outreville's (2014) work. By having a better understanding of the links between income and insurance demand, this paper aims to provide some guidance towards consumer insurance appetite as a function of income. Thereby providing some reference for developing products that can be suitably priced and will be more successful at attracting customers who are willing to pay for insurance products.

A study done by Feyen, et al. (2011) as cited by Alhassan and Biekpe (2015) showed that factors such as religion, demographic structures, population size, and density served as indicators of life insurance consumption. These findings leave a distinct opportunity for further research to ascertain influences, based on individual behaviour, for predicting insurance usage in specific environments. This shortfall that this paper will address is discussed in the below section.

## **2.10 Summary**

The literature reviewed in the above section serves as evidence that there is a marked research gap in terms of understanding the attributes of the African market demand for insurance products. Alhassan and Biekpe (2015) researched life insurance demand in Africa, but only considered the macro level. However they call for a micro-level analysis of life insurance consumption. Therefore, this paper argues that there is a need for linking internal firm data, such as spending patterns, saving habits, propensity for debt and income as predictors for insurance demand and claims experience.

Alhassan and Biekpe, (2015), state that demographic factors are better than financial factors in explaining the demand for life insurance. They also state that future studies could investigate the influence of dissimilar income levels (wealth) on the consumption of life insurance in African countries. A different methodological approach could be taken to their study, in that a non-linear econometric approach may be done to investigate the non-linear consequence of income effects on the consumption of life

insurance. Finally, they also suggest a micro-level analysis of life insurance consumption.

It appears that not much work has been published on individual's banking behaviour and the consumption of life insurance products. Additionally, there is also a lack of research linking individual banking behaviours (income, savings and debt) to risk aversion behaviours in the context of emerging markets.

Lin and Grace (2007) used a survey of consumer finances to investigate the life-cycle demand for life insurance. They tested for an aversion to income volatility through the purchase of life insurance. They also developed a "financial vulnerability index" as a measure of control for the amount of risk inherent in the household. What they discovered was in contradiction to previous research by Bernheim, Carman, Gokhale and Kotikoff (2001), that there is a relationship between financial vulnerability and the amount of life, or total life insurance purchased. They also found that comparatively older consumers of insurance cover utilised lower amounts of life insurance to defend their level of financial vulnerability, compared to their younger counterparts. They also found that their financial vulnerability index proved to be a significant variable in explaining household life insurance demand. These (somewhat contradictory) findings substantiate (to some extent) the need for more research into the financial determinants of insurance consumption, despite the efforts of Alhassan and Biekpe's (2015) work. This paper will also expand on Outreville's (2014) recommendations to look closely at the connection between propensity of debt and insurance consumption, since it has been shown that the amount of debt a household owes is positively correlated with their decision to hold life insurance. The more debt a household owes, the more likely that it is to have life insurance.

While there is a considerable amount of literature (much of which is on demographic and socio-economic variables (Outreville, 2014)) and information accessible on determinants of the demand for life insurance, there are multiple areas that necessitate additional consideration (Ziets, 2003). In view of the pace of change in the economic environment, demographic factors and technology, several of the findings from past studies may likely be deemed outdated. Moreover, considering the inattentiveness accorded to the fluctuating demand for various innovative products in the past, marketing and product development issues must be addressed more urgently than ever.

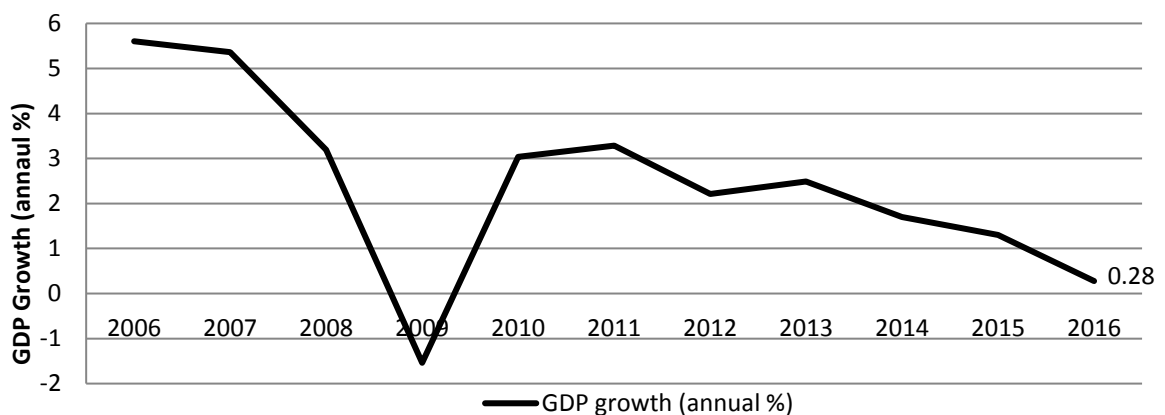
In their extensively cited and eminent research; Browne and Kim (1993) specifically called upon future research to extend their analysis to include a study of demand of the

growth and maturation of insurance markets. The call for more research was answered by Chang et al., (2014). Their paper also called for an extension of the investigation of insurance consumption in a broader sample of countries (which this research aims to satisfy). They continue to state that this extension of their study could possibly lead to a better understanding of insurance demand. Interestingly for this investigation; is that their research specifically looked at the links between insurance, and factors leading to differences in the demand for life insurance across different countries. Their research specifically identifies a void that will hopefully be filled by the findings of this investigation.

Another motivation for this particular research paper, is that insurance is often ignored in the academic literature (Chang et al. 2014); as opposed to the amount of research done on other areas of the financial sector, such as stock markets and banking, which attract copious consideration. The insurance industry has been described as the financial sector’s risk management service (Chang, et al., 2014). Further accentuating the need for the research.

This field of study is important in the South African context, since insurance has been shown to be linked to promote economic growth in some countries. Chang, et al., 2014 showed that insurance activities promote economic growth for Japan, the United Kingdom, Switzerland, France, and the Netherlands. Through Granger causes they showed that economic growth causes activities for different types of insurance in selected countries, such as the increase in life insurance in Italy and Canada, and non-life insurance in the United States of America. They were also able to show a two-way Granger causality between economic growth and life insurance in the USA.

**Figure 1: Graph showing South African GDP growth over the last 10 years. Source: World Bank national accounts data, and OECD National Accounts data files.**





## 3 Research Questions

### 3.1 introduction

To be able to meaningfully contribute to the existing knowledge base, the appropriate research methodology must be implemented (Kothari, 2004). To achieve this, the research must be well designed. The following section discusses the research questions and hypothesis of this paper. Banking data generally contains a great variety of information (Han et al. 2012). Therefore, it is important to carefully formulate the research questions so that the study does not become unmanageable in size and scope. Many banks offer a wide variety of investment, banking, and other services (such as credit, mortgage, and automobile loans). A unique characteristic of the banking data that this research has access to, is that it also contains insurance information as well as general banking information. By combining these distinct datasets, it makes it possible to answer the research questions posed below.

### 3.2 Core Research Question

The core research question follows on the research done by Mulholland et al., (2016). They report that households with the highest levels of general financial sophistication were at least twice as likely have life insurance. Furthermore, the chances of owning life insurance rose monotonically with increasing financial sophistication. The core research question seeks to determine if there are in fact linkages between financial behaviours of consumers and insurance demand:

*Is there a relationship between South African individual-level financial consumer behaviours, such as their levels of income, savings and, debt and their demand for life insurance?*

The findings would be able to confirm or dispute the suggestion that those households with high levels of financial sophistication are supporting their comparatively higher human financial capital with more sophisticated life insurance products (Mulholland et al., 2016).

### 3.3 Sub Research Questions

This paper provides some literature on the various financial determinants stated above. The sub research questions aim to partition the core research question into more manageable questions which can be expanded into hypotheses and analysed independently.

### **3.3.1 Research Question 1 - Income:**

Yaari, (1965), reported that households methodically consume life insurance as a protective risk-based measure against the loss of income flows (derived from labour) over their life cycle. Browne and Kim, (1993) report that many models portrayed in the literature interpret life insurance as a way of reducing uncertainty in income. This is of course in relation to the risk of possible death or disability of the principal income earner. Lange, et al., (2017) found that income is one of the most central drivers for the demand for insurance. Frees and Sun, (2010) found that the influence of income on life insurance demand is analogous to the effect of wealth as a determinant of life insurance. Outreville, (2014) reports that there is a limited number of investigations that have been conducted in developing countries and that the results are mixed. To address the comparative lack of research done in this field in emerging economies this paper seeks to find the answer to the following research question:

*Is there a relationship between the demand for life insurance and the amount of income an individual has?*

### **3.3.2 Research Question 2 - Savings:**

Since Headen and Lee's (1974) reported that insurance and financial assets have a tendency to mutually proliferate, this research seeks to determine if this concept is outdated or still valid. The validity of Headen and Lee's (1974) paper should also be tested in the South African context. Sommerville (2004) showed that a consumer's management of their exposure to risk typically includes both savings behaviours and insurance.

*Is there a relationship between the demand for life insurance and the amount of savings an individual has?*

### **3.3.3 Research Question 3 - Debt:**

Frees and Sun (2010) reported that the amount of debt owed has a positive effect on consumers to firstly purchase and then keep life insurance. They also state that with increasing debt, the likelihood of consuming relatively inexpensive forms of life insurance increases. This is of particular interest in the South African population who are already attracted to low-cost life insurance products such as funeral insurance (Roth, 2000).

Since Van Winssen, et al., (2016) reported on the links between debt aversion and insurance take up, this research seeks to determine if the results are reproducible and if the linkages are valid in the South African context.

*Is there a relationship between the demand for life insurance and the relative amount of debt an individual has?*

### 3.4 Hypotheses

**Table 3: Potential research hypotheses for the South African life insurance context**

Question	$H_0$	$H_1$
1	There is no significant relationship between the demand for life insurance and income.	There is a significant relationship between demand for life insurance and income.
2	There is no significant relationship between the demand for life insurance and savings	There is a significant relationship between demand for life insurance and savings
3	There is no significant relationship between the demand for life insurance and debt.	There is a significant relationship between the demand for life insurance and debt.

### 3.5 Summary

The table below shows how the research questions will be answered. In Chapter 4, a discussion of the research methodology will be provided.

**Table 4: The set-up of the research questions in relation to financial behaviours of customers and insurance demand**

		Measures of insurance demand		
		Take-up	Lapse	Cancel
Measures of financial behaviour of customers	Income	Research Question 1		
	Savings	Research Question 2		
	Debt	Research Question 3		

## 4 Research Methodology

### 4.1 Introduction

The purpose of this chapter is to detail the methodology used in this research. Banks have access to a large amounts of data on consumer transactions, savings levels, debt levels, income and general financial behaviour making it convenient to use this data to answer the research questions that this paper is concerned with. Secondary customer banking and insurance data was used to answer the research questions.

Finally, the methodology outlined in this research paper may be extended to other studies since it can be argued that for a firm to understand their data to usefully implementing findings from analysing it may lead to better decisions about which products to offer customers (Shanks, et al., 2010). This is the foundation for implementing business analytics – it has the ability to improve organisational performance and escalate competitive advantages.

Brynjolfsson et al., (2011) report that data driven decision making process are indeed associated with higher productivity and market value; this paper argues that without understanding the relative attributes and details (such as the relationships between individual level financial behaviours and insurance consumption) the benefits may be lost. The benefits may never be fully realised because “big data hubris” is the frequent assumption that big data is an alternate for good quality traditional data collection and analysis (Lazer, Kennedy, King, & Vespignani, 2014). Shanks et al., (2010) also warns that oftentimes firms speculate on the relative contribution of certain parameters to performance. This paper therefore aims to use a systematic approach, discussed in the this chapter, to attain useful information which would assist strategic decisions in the insurance marketplace.

#### 4.1.1 Definitions

The following table of definitions should be useful in defining the data concepts.

**Table 5: Key data and methodology definitions used**

Term	Definition	Sources
Claim	A claim denotes that the insurer must deliver the services which the individuals have been paying for. These services may extend beyond simple economic	Guillén, et al., 2012

compensation.

Debt	Total debt includes the following: Credit card, line of credit various mortgages, loans (consumer, education, motor vehicle, home improvement and land contract) as well as other debt and loans.	Frees and Sun, 2010
Financial vulnerability index	The volatility of a household's living standard if subjected to the death of a household member who earns an income. It is determined by the income through non-labour and labour, the probability of death of member, the household consumption to income ratio, and age effect on the needs of future consumption.	Lin and Grace, 2007
Lapse Rates	The termination of an insurance policy without pay out but for the purposes of this paper includes the definition of a surrender which is a term used when a value is paid to the policyholder on termination of the policy	Eling and Kiesenbauer (2014), Kuo et al., (2003)
Policy	Refers to the contract that an insurance company will pay the client for losses caused by covered hazards	Guillén, et al., (2012)
Unit-Linked	Unit-Linked insurance plans provide policyholders with investment (typically equity and debt schemes) and traditional insurance (mortality charges) under a single plan.	Li and Szimayer (2010)
Whole life insurance	Includes insurance and investment mechanisms. Insurance component compensates an already determined value	Frees and Sun, 2010

while investment component accumulates value over time as a savings component. The policy is applicable to whole of life which further differentiates it from a unit-linked plan.

## 4.2 Choice of Methodology

The aim of this paper was to model the insurance choices of individuals with respect to their financial behaviours. Variables were derived from their banking financial data. The study is based on a cross sectional analysis of a dataset of banking customers, who also have insurance with the bank (bancassurance). The methodology uses a binomial logistic regression model to investigate the linkages between each of the independent variables (see Table 6) and dependent variables (see Table 7)

According to Saunders, Lewis and Thornhill (2009) a study may be segmented into the “layers” of the research process. This paper sought to achieve answers to the research questions by the choices in methodology as shown in shown in Table 4; each “layer” will be discussed in detail below.

**Table 4: Research Methodology Choices**

Philosophy	Approach	Strategy	Choice	Time Horizon	Techniques & procedures
Pragmatic	Deduction	Archival research	Mono-Method	Cross-section	Data Collection & analysis

To answer the research questions the study used secondary data from the bank (where banking information was accompanied by corresponding insurance information for the same individuals). The assumption was made that it would be fitting to utilise this data given the research questions of this study. According to Saunders and Lewis (2012), choosing a research philosophy is guided by what is practical. This means that the research questions dictate that variables such as those provided in Table 6 and Table 7 were required at an individual level. The most appropriate place to gain access

to these variables was the secondary banking and insurance data that is readily available.

This is an appropriate methodology

My concern is that a pragmatic approach generally inculcates both quants and qual. You need to also defend – state why this is an appropriate methodology.

### **4.3 Population**

This research endeavours to establish insurance penetration among banking customers in South Africa. Therefore the population consisted of customers who were stable and active banking customers. For the lapse and cancel model these customers were also in possession of insurance products at the time of the observation point. For the take-up model the customers did not have life insurance at the time of the observation point.

### **4.4 Unit of analysis**

The unit of analysis in this study was the individual consumer of banking and insurance products. The individual customer banking and insurance data was averaged for the outcome period to determine the impact on insurance demand on an individual level. Consequently, the unit of analysis for the lapse and cancel model was the banking customers who had insurance, while the unit of analysis for the take-up model was those customers who did not have insurance at the time of the observation window. For the purposes of this research, the data was subdivided into two groups:

1. Consumers of banking products who had life insurance products during the observation period (the lapse and cancel model)
2. Consumers of banking products who did not have life insurance products during the observation period (the lapse and cancel model)

### **4.5 Sampling Method**

Sampling is the natural selection for the progressive fine-tuning of an abridged dataset, according to Han et al. (2012). A systematic random sampling method was applied to select customers with active bank accounts according to section 4.5.1 (Saunders & Lewis, 2012). The scope of the sample was limited to individual customers. No group schemes were included in the sample.

### **4.5.1 Individual Customers**

All individual customers with bank accounts and insurance relationships with the bank were included in the sample. Individual customers were identified from the historical customer database using the following high-level rules:

- Customer with significant credit turnover with the bank were selected
- Date of birth was available
- Insurance behaviour during the observation period for the lapse and cancellation models.

The above rules excluded any individual customers that opened bank accounts but did not ever take up insurance. The analysis was therefore only on customers who did take out insurance products at some point, but then specifically looked at the individual-level financial behaviours of these customers. This approach is useful for determining the financial indicator (savings, income or debt) that would be useful for indicating risk aversion behaviour.

The full transactional customer base (apart from specific exclusions described in ) was used. The variable selection process was only run on insurance customers who had the above mentioned behaviours due to time and system constraints. In future, the study could be extended to include customers who never take up any insurance products to determine if there is a difference between customers who are not disposed to the particular risk aversion behaviour but this is beyond the scope of this analysis.

### **4.5.2 Sample Size**

The sample is all the individuals who bank with the one of South Africa's four large banks. This data was made accessible to satisfy the research goals.

An optimum sample fulfils the requirements of efficiency, representativeness, reliability and flexibility (Kothari, 2004). The entire dataset was therefore not analysed. Han et al., (2012) reported that the reduced dataset may be further refined by increasing the sample size. This research was granted access to sufficient computational ability to analyse the an extensive dataset which was an important consideration for the logistic regression model.

The sample size for the three binary logistic regression models is shown below:



**Table 7: Table showing the relative sample sizes for the different logistic regression models**

<b>Model</b>	<b>Sample size</b>
Take-up	263 134
Lapse	115 671
Cancellation	154 086

#### **4.6 Data Gathering Process**

As discussed, the data was made available for the purposes of answering the research questions through the banking data warehousing systems. Statistical Analysis Software or “SAS” was used to analyse the data. The same logistic regressions were run in IBM® SPSS to ensure consistency of the results.

Payer-provider data, such as the data generated between banks and their customers, are an excellent source of big data according to Chen et al. (2012). This provides a convenient and good source of useful information for the purposes of this study. This research has access to one of the four large South African banks’ secondary data, which is also a provider of life insurance. The combination of these two services offered by one firm makes the information rich and particularly useful for answering the research questions. The research variables, their measures and source are shown in Table 7 below.

**Table 7: Data measurement and source**

<b>Variables</b>	<b>Measure</b>	<b>Source</b>
<b>Dependent Variables</b>		
Take-up	Clients who have taken insurance products. This is a measure of life insurance demand / market penetration and is to some extent synonymous to risk aversion.	Insurance
Lapse	Clients who already had taken-up insurance products who then fail to pay their premiums and therefore lapse. This is the inverse of life insurance demand and is to some extent	Insurance

	synonymous to risk loving.	
Cancellation	Clients who already had taken-up insurance products who then cancel their insurance. This is the inverse of life insurance demand and is to some extent synonymous to risk loving.	Insurance

### **Independent Variables**

Income	Income into transactional account	Bank
Savings	Amount of savings a customer had	Bank
Debt	Amount of debt a customer had	Bank

#### **4.6.1 Measurement Instrument**

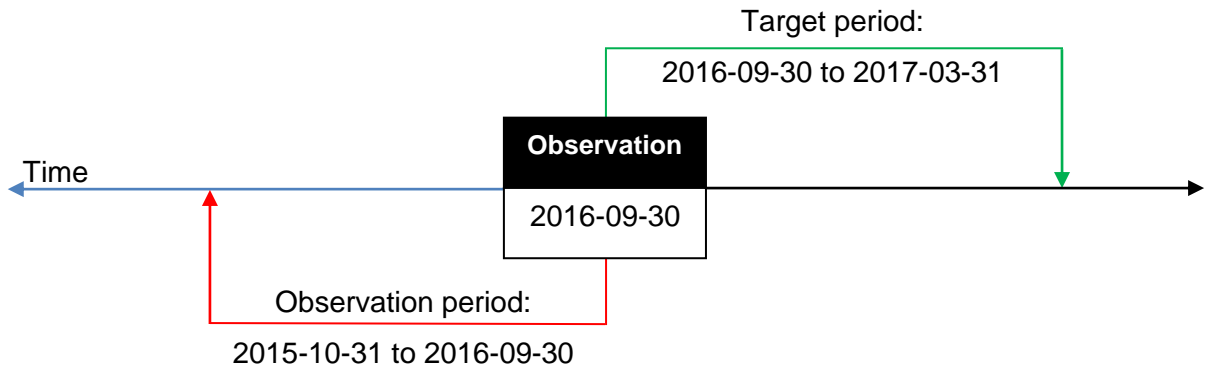
The measurement instrument for collecting the consumer information is through the various data capturing systems that are employed by the bank. When any banking customer transacts with their money, a historic record is kept of all transactions. The systems that collect the information are all legacy systems. Some of the systems specifically collect information when customers spend or withdraw money while others collect and store transactions of incoming monies. A separate system collects and stores customer's life insurance portfolios and consumption of life insurance products. All the collected data is stored in data warehouses which are accessible. The information may be anonymised quite easily thereby protecting individual rights.

#### **4.6.2 Time horizon**

A cross-sectional approach seems appropriate since this paper does not wish to establish the effect of an event or the implementation of some intervention. Saunders and Lewis (2012) explain; when there is no need to study "change and development" one does not need a longitudinal study. A Cross-sectional design takes a "snapshot" of a specific subject at a point in time.

The figure below illustrates the setup and development of data as well as total exposures for the target window.

**Figure 1: Time horizon of research**



### 4.6.3 Exclusions

The following table describes the exclusions made for the final model dataset. This is important to note for study reproducibility reasons.

**Table 9: Exclusions made to the research data**

Exclusion	Description and Motivation
Customers that passed away before or during the observation month.	If a customer passed away before or during the observation month, the customer was excluded from the dataset. These customers are removed from the model build so that only customers who were alive and active as at the observation month were included.
Customers who closed their accounts within the outcome period.	The final outcomes of these customers at the end of the outcome period as measured from the observation month are unknown. These customers were therefore removed from the model dataset.
Customers with a banking relationship of less than 12 months at the observation month.	Only customers with banking relationships of longer than 12 months preceding the observation month were included in the model build to ensure sufficient history and stability of banking data was available for each customer in the dataset.

Customers younger than 18. The model was only built for customers older than the ages of 18. The very low empirical rates for customers younger than 18 to have insurance products is the reason why these customers were excluded from the study.

---

## **4.7 Analysis Approach**

All research projects have either deductive or inductive approaches according to Tashakkori and Teddlie (2003). The five sequential stages in Saunders and Lewis (2012) were followed for the analysis approach. Firstly, the research questions were developed in response to established theory. Secondly it was determined how the questions were to be answered. Third the research searching for answers to the questions defined in the first stage through a binary logistic regression model. Fourth, the results were investigating in order to conclude if the theory was corroborated or refuted. Finally, the initial theory whereon the research questions were based was fortified.

### **4.7.1 Techniques**

According to Grunert and Weber (2009) banks need to collect and archive data on the spending and usage behaviours of their customers, especially for the purposes of establishing creditworthiness. Since the data was collected for a different purpose than for this study, it is secondary quantitative data. This data will need to undergo processing as it is in a “raw” format but some of the data may already be compiled where some summarising and selection has been executed. Using secondary data has may save time and provides access to a larger amount of data than what be collectable during the time of this study (Saunders & Lewis, 2012). Banking data is typically high quality and can be relatively easily “mined” according to Han et al, (2012).

Kjosevski (2012) used a panel data analysis with a fixed effects estimator (within estimator) for the coefficients in a study’s regression model for South-Eastern and Central European countries between 1998 and 2010. This was used to ascertain dependent variables as predictors for life insurance penetration. Some of these included social factors such as health expenditure, level of education as well as economic indicators such as GDP per capita and inflation.

### 4.7.2 Target Response Variable

The first step in deciding on the modelling approach was to decide on the response variables (i.e. the outcome that needs to be modelled). There were three target response variables modelled for in this study namely; take-up, lapse and cancellation.

A 6-month research window was used to determine the outcome. (i.e. 1 if the customer lapsed, cancelled or purchased insurance within 6 months of the observation point; 0 if the customer did not lapse, cancel or purchase insurance by the end of the outcome period) was set as the response variable given the following:

- The aim of the binary logistic regression model was to assess, and rank customers in terms of the probability of take-up, lapse and cancellation within an outcome period (and not the timing of the response within the outcome period). A binary response variable ties in with this aim.
- At the time of the research and model build, a 6-month outcome period (target window) was the longest available bearing in mind that a 12-month observation period preceding the outcome period was also incorporated (see Figure 3).

It was also intended for the model to be consistent with credit modelling techniques. Modelling a binary response variable is consistent with modelling default rates in a credit probability of default models (Thomas, Edelman, & Crook, 2017).

### 4.7.3 Regression Variables

In this section a detailed description of the regression variables is provided. The information is summarised into the following two tables of independent and dependent variables.

**Table 10: A summary of the independent variables and their definitions used in the study**

Independent Variable	Definition
Average Credit Turnover	Average credit turnover is the amount that a consumer spends on average per month. The figure is averaged over the last 12 months and therefore provides an accurate indication of the consumption of an individual.
Average Cheque Balance	This is the average amount of money a customer maintained in their primary cheque accounts. Since the account balance is most often highest after a salary deposit and lowest directly

before a salary deposit the average over the month is taken.

Average Income Estimate	<p>This is the income amount as estimated by the bank. It is based on the amount of money that is deposited into the customer's bank accounts. It does not account for any income that the individual may accrue through other means such as cash income (that is never banked) or income that is banked with another banking service provider.</p> <p>The income field is inflated to allow for general salary inflation from the calibration of the model to more recent months (i.e. income estimates used in calibrating the model need to be real in today's terms). However, given that wide income groupings were used, income estimates were left unadjusted and future development will consider refining the income calculations.</p> <p>The customers' income fields were grouped into different gross annual income bands, based on the outcome variable that was being tested.</p>
Average Savings and Investments	<p>This the average savings amount that a customer has. It is calculated by summing all savings and investment products available through the bank. Having large savings may be considered a measure of risk aversion. Savings may also be considered to be a measure of temperance.</p>
Total Debt	<p>The total debt is a summation of all available debt products offered by the bank. It includes the all personal loans, credit card debt, home loan debt and debt owed on a cheque account (this is an overdraft facility). The measure of debt can be interpreted as a measure of financial sophistication and only customers with sufficient risk profiles will be granted access to debt products. On the other hand one may argue that having debt is also a risk "loving" behaviour.</p>
Ratio of Savings to Income	<p>This is the ratio of savings to income. A higher ratio indicates that the individual is saving and investing a greater proportion of their income. This ratio is to some degree effective at negating the large differences in income, since even modest savings in</p>

---

	relation to modest income can lead to a high saving to income ratio. The variable is a measure of temperance.
Ratio of Debt to Income	The ratio of debt to income is calculated by dividing the Total Debt by the Average Income variables. It is similar to the Total Debt variable but enables a fair comparison of individuals with different incomes. A high ratio may be considered to be a sign that the individual is risk “loving”. A low ratio may also mean that the individual has a relatively high income or it may mean that they do not have access to many debt products with the bank.
Ratio of Credit Turnover to Income	The variable is calculated by dividing the average credit turnover by the income estimate variable. It is a measure of the portion of income an individual wishes to spend. A high ratio is an indication of high relative consumption behaviours while a low ratio means that the individual may be prudent.
Ratio of Average Cheque Account Balance to Income	The variable is calculated by dividing the average cheque account balance by the income estimate variable. It is a measure of the amount of money an individual keeps in their account in relation to the amount of income that they earn. A high ratio is an indication of aversion behaviour while a low ratio means that the individual may be more risk “loving”.
Ratio of Debt to Average Cheque Account Balance	The ratio of debt to average account balance is calculated by dividing the Total Debt by the Average Cheque Account Balance variable. It is similar to the Total Debt variable but enables a fair comparison of individuals with different account balances. A high ratio may be considered to be a sign that the individual is risk “loving”. A low ratio may also mean that the individual has a relatively high income or it may mean that they do not have access to many debt products with the bank.

---

**Table 11: A summary of the dependent variables and their definitions used in the study. Note that the information on target variables for life insurance products were included as binary variables in the model. This provides an overview of which target variables were specified for the research**

---

Dependent	Definition
-----------	------------

---

---

Variable	
Take-up	<p data-bbox="496 277 1406 360">If the customer purchased insurance within the research window, then the binary value one was allocated to the event.</p> <p data-bbox="496 398 1406 479">If the customer never purchased insurance within the research window, then the binary value zero was allocated to the event.</p>
Lapse	<p data-bbox="496 530 1406 613">If the customer lapsed their insurance within the research window, then the binary value one was allocated to the event.</p> <p data-bbox="496 651 1406 777">If the customer never lapsed their insurance within the research window (i.e. kept their insurance intact), then the binary value zero was allocated to the event.</p>
Cancellation	<p data-bbox="496 828 1406 911">If the customer cancelled their insurance within the research window, then the binary value one was allocated to the event.</p> <p data-bbox="496 949 1406 1075">If the customer never cancelled their insurance within the research window (i.e. kept their insurance intact), then the binary value zero was allocated to the event.</p>

---



#### 4.7.4 Logistic Regression

Given the binary response variables (within the 6-month outcome period) a logistic regression model (a type of generalised linear model) was the logical choice for this research. This modelling approach assumes that the distribution function of the indicator (take-up, lapse or cancel) is the Bernoulli distribution, that is:

$$Probability(Indicator\ for\ target\ variable = k) = \begin{cases} 1 - q & \text{for } k = 0 \\ q & \text{for } k = 1 \end{cases}$$

The expected value of the indicators is then  $q$ , i.e. the probability of take-up, lapse or cancel within the outcome period.

The logistic regression model uses the logit link function to link the probability of the deceased outcome to a linear function as follows:

$$logit(q) = \ln\left(\frac{q}{1-q}\right) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n$$

- The  $x_i$  represents the explanatory variables and  $\beta_i$  represents the coefficients of the input variables.
- For a numerical explanatory variable, the relationship between the logit of the probability of take-up, lapse or cancel and the variable will be linear (i.e. monotonically increasing or decreasing).
- For a categorical explanatory variable, the coefficient of that variable will contain the coefficients for each category of that input variable.

The Box-Tidwell (1962) technique was used to evaluate the linearity of the continuous variables with regard to the dependent variable logit. Tabachnick and Fidell, (2014) recommended using a Bonferroni correction. It was therefore applied using all seven terms in the model. Statistical significance was only established when  $p < 0.00625$ .

A logistic regression model was firstly done using PROC LOGISTIC function in SAS (please see Appendix A for the SAS code used in the model). Secondly, to ensure that the tool used in the study did not influence the outcome, all the tests were done using IBM® SPSS. The results for both analyses were equal thereby negating any effect the analysis tool may have on the outcome of the model.

##### 4.7.4.1 Null and Alternate Hypothesis for the Logistic Regression

The null hypothesis for the overall test is that each predictor is equal to zero, meaning each of the predictors is insignificant. The alternate hypothesis for the overall test is that at least one predictor is significant. Typically, this test is done first to make sure that at least one predictor is significant for predicting a response. The table "Testing for

Global Null Hypotheses: Beta = 0” is a test that all Betas are equal to zero. The likelihood ratio has a  $p$  value of 0.0001 and therefore, even if an alpha value of 0.05 is chosen, the test still indicates that at least one of the betas are significant.

The Wald Chi-Square value is a calculation may be shown by the following formula:

$$\left(\frac{\text{estimate value}}{\text{standard error}}\right)^2$$

The  $p$  – *value* will always be right tailed for the Chi Squared distribution and with one degree of freedom.

When it comes to the “Analysis of Maximum Likelihood”, or the specific tests on the predictor variables in the model, the null hypothesis is again that the beta is equal to zero. The alternate hypothesis it is that beta is not equal to zero. This test considers the marginal contribution for each predictor. The test provides the contribution of the predictor, given that the rest of the model stays the same, in other words; is the predictor significant, given that all other parameters are already in the model. If the  $p$  – *value* of this test is less than alpha (again chosen at 0.05) then the null hypothesis is rejected and we accept that the predictor is significant.

When considering whether or not to use the reduced model over the full model the “Model Fit Statistics” table is useful. The Intercept and Covariates value is deducted from the “Intercept Only” value. This provides a Chi-Squared statistic. Since there are seven more predictors in the full model over the reduced model it can be concluded that there are seven degrees of freedom.

#### **4.7.5 Background on Statistical Methods**

A logistic regression is the most popular means to model binary response data (Hilbe, 2011). A binomial logistic regression endeavours to predict the probability that an observation, based on one or more independent variables, satisfies either categories of a dichotomous dependent variable. The independent variables may be categorical or continuous (Laerd Statistics, 2015). Since this research aims to determine if income, savings and debt can be used as a prediction for insurance behaviours (take-up, lapse or cancel), a binomial logistic regression will be the primary means of analysing the data.

When the data response is binary, it usually is in the shape of a 1 or a 0, with 1 denoting a positive response and 0 indicating a negative response. The values that 1 and 0 may mean will vary according to the purpose of the study (Hilbe, 2011). For example, in this study of the odds of cancelling insurance; 1 has the value of a

successful cancel, and 0 of not-cancelled policy, while in the case of take-up; 1 has the value of a successful take-up, and 0 of not-taken-up. In the two instances the value 1 indicates different customer dispositions towards insurance. What is important, however, is that 1 indicated the primary subject of interest for which the binary response research was intended.

In the International Encyclopaedia of Statistical Science, Hilbe (2011) explains that, using a normal linear regression to modelling a binary response variables would have introduced considerable bias to the parameter estimates. A normal linear model assumes a Gaussian distributed response and error term, so that the variance,  $\sigma^2$ , is constant for all data observations, and that the data observations are independent of one another. However, when a binary variable is modelled using normal linear regression, the assumption of the Gaussian distributed response is violated. It is for this reason (that the normal regression model is based Gaussian probability distribution function) that a binary response model is used in this study. A binary response model is derived from a Bernoulli distribution; a subset of the binomial probability distribution function where the binomial denominator holds the value of 1. The Bernoulli probability distribution function is expressed as:

$$f(y_i; \pi_i) = \pi_i^{y_i} (1 - \pi_i)^{1-y_i}$$

Equation 12

#### **4.7.6 Logistic Regression Model Build-up**

A binary logistic regression models for each dependent variable (three models in total) were developed by in a stepwise manner by introducing new variables into the model and removing insignificant variables from previous runs.

This methodology was followed to “build-up” the model in the correct sequence, and to maintain track of the performance measures at each step of the logistic regression model.

The complete PROC LOGIC output from SAS is displayed in 11.2.

### **4.8 Summary**

In this chapter the methodology for answering the three research questions is provided. It offered an framework of the research methodology employed in evaluate the data so that statistical interpretations could be made relating to the research objectives. The study population and the sampling technique were described. The data gathering

process from the bank was also briefly discussed. The motives why the binary logistic regression was chosen was also provided. Chapter 5, which follows, presents the results of the data analysis.

## 5 Results

### 5.1 Introduction

The following section details the results of the binomial logistic regression for the take-up, lapse and cancellations respectively which are proxies for insurance behaviours (and thus risk aversion). Each section is accompanied by a graphical representation of the data. The three research questions under investigation were:

1. Is there a relationship between the demand for life insurance products and the amount of income an individual has?
2. Is there a relationship between the demand for life insurance products and the amount of savings an individual has?
3. Is there a relationship between the demand for life insurance products and the amount of debt an individual has?

This chapter includes the in depth analysis of the customer data with the use of a binomial logistic regression. The analysis begins with the results of the take-up model and highlights the most relevant results for the various predictor variables. Secondly the results of the lapse model are provided, followed lastly by the cancel model.

Section 5.2 delivers a short description of the data. In section 5.3 the reliability and validity of the data is discussed. After this the data transformations are discussed in section 5.4. Finally, in sections 5.5 to 5.7 the logistic regression analysis results are provided. At the end of each of the sections 5.5 to 5.7 a graphical representation of the results is displayed to facilitate a better understanding of the data.

### 5.2 Description of the Sample Data

A systematic random sampling method was employed to select the data. The sample omitted customers according to the exclusions described in section 4.14.1. The scope of the sample was limited appropriately for each analysis. Only customers with active life insurance products were selected for the lapse and cancel regression models while only customers without life insurance were selected for the take-up model.

There were 263 134 customers in the take-up sample, 115 671 customers in the lapse sample and 154 086 customers in the cancel sample.

### 5.3 Reliability and Validity of the Data

The data only included those customers who were acceptable for the study according to the exceptions provided in section 4.14. the 12 month observation period was designed so that only customers with valid and active banking and insurance relationships were used in the study. If the customer did not abide to the requirements then they were omitted from the sample. Detail of this process is visible in the SAS code provided in section 10.1.

### 5.4 Data Transformations

The customer data is typically available as a “snapshot” in time. The data was collected for 36 monthly snapshots. Therefore, one customer occupied 36 rows of data. The customer data had to be transformed from the 36 monthly snapshots into single rows of aggregated information for the customers. The variables therefore had to be averaged according to the outcome period of six months.

The data had to undergo multiple transformations to make it suitable for a binomial logistic regression analysis. To create the binary response variable the change in policy status had to be transformed. This was done as in the following way:

**Table 11: Data transformation into binary response variables**

<b>Model</b>	<b>Policy Status</b>	<b>Binary Response</b>
<b>Take-up</b>	Blank	0
	In Force	1
<b>Lapse</b>	In Force	0
	Lapsed	1
<b>Cancel</b>	In Force	0
	Cancelled	1

## 5.5 Take-up Model

Firstly, a binomial logistic regression was executed on the sample of banking customers who at the observation point had no life insurance. The logistic regression model was used to determine the effects of the following independent variables on the probability that individuals took-up (purchased) life insurance. Table 12 below shows both the significant and insignificant predictor variables for the logistic regression model on take-up (based on  $p < 0.05$ ):.

**Table 12: Significant and insignificant predictor variables for lapse**

Significant predictor variables	Insignificant predictor variables
Debt (total debt across a range of debt products)	Average income
Savings (average savings and investments balance)	Ratio of savings to income
Average credit turnover	Ratio of debt to income
Average cheque account balance	
Ratio of credit turnover to income	
Ratio of average cheque account balance to income	
Ratio of debt to average cheque account balance	

The assessment showed that seven of the ten continuous independent variables were found to be linearly related to the logit of the dependent variable. All the studentized residuals were retained in the model. The logistic regression model was statistically significant with  $\chi^2(7) = 7974.192$ ,  $p < .0001$ . The pseudo  $R^2$  (Nagelkerke) value was 6.5% which explained the variance in lapses. The model correctly classified 63.1% of cases. The positive predictive value of the model was 63% and negative predictive value was 36.8%.

Of the ten predictor variables, only seven of were statistically significant as shown in Table 12. According to the logistic regression an increase in the monthly average credit turnover, average cheque account balance and, average savings all decreased the likelihood of take-up. On the other hand increasing the total debt, ratio of credit turnover to income ratio of average cheque account balance to income and ratio of total debt to average cheque account balance all led to an increase in take-up.

**Table 13: Logistic Regression Predicting Likelihood of Take-up based on Average Credit Turnover, Average Cheque Account Balance, Savings, Total Debt, Ratio of Credit Turnover to Income, Ratio of Average Cheque Account Balance to Income and Ratio of Total Debt to Average Cheque Account Balance.**

	<i>B</i>	<i>SE</i>	Wald	<i>p</i>	Odds Ratio	95% CI for Odds Ratio	
						Lower	Upper
Average credit turnover	-0.00005	1.07E-06	2226.84	<.0001	1	1.000	1.000
Average cheque account balance	-8.05E-6	1.21E-06	44.6461	<.0001	1	1.000	1.000
Average savings	-4.29E-7	1.89E-07	5.13	0.0235	1	1.000	1.000
Total debt	0.000011	3.52E-07	973.237	<.0001	1	1.000	1.000
Ratio of credit turnover to income	0.5054	0.1024	24.3643	<.0001	1.658	1.356	2.026
Ratio of average cheque account balance to income	0.0468	0.0138	11.4627	0.0007	1.048	1.02	1.077
Ratio of total debt to average cheque account balance	0.00717	0.00069	107.605	<.0001	1.007	1.006	1.009
Intercept	-0.1270	0.0119	113.396	<.0001	0.877		

Note that the degrees of freedom for each variable was one. The degrees of freedom correspond to the predictor variables, each of which requires one degree of freedom. This describes the Chi-Square distribution to test if the regression coefficients are 0, if the other predictor variables are already in the model.

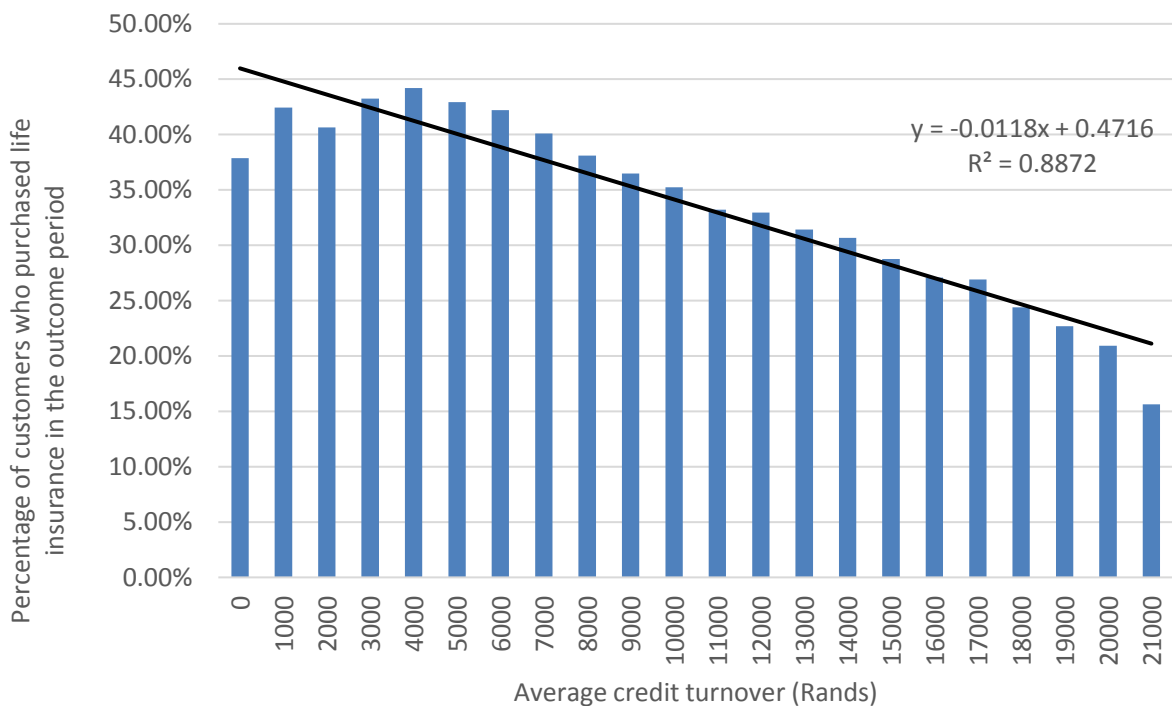


### 5.5.1 Graphical Representation of Take-up Predictor Variables

In the following section a graphical representation of the data is shown. An emphasis has been placed on those variables that were found to be significant on the take-up dependent variable in the logistic regression model.

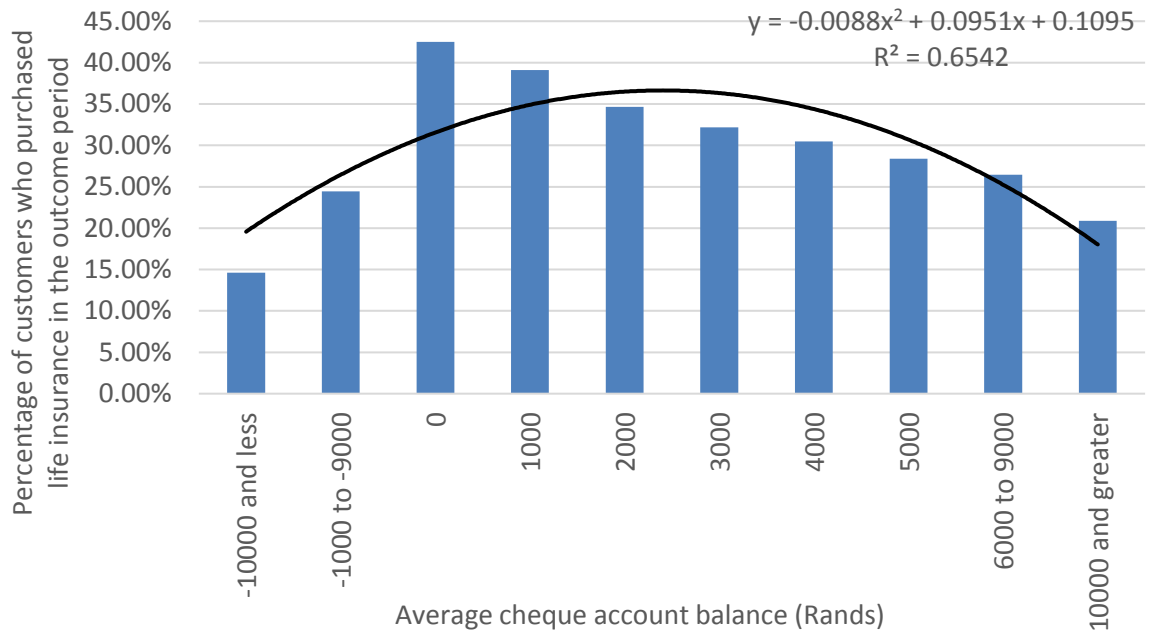
In the figure below the percentage of take-up of life insurance increased with increasing credit turnover but only until about R4 000 monthly spend. In the customers who spend more than R4 000 monthly, there was a decrease in life insurance purchased with increasing credit turnover. A linear trendline was fit to the data with a  $R^2$  of 0.8872. The maximum take-up was approximately 45% with customers spending R4 000 monthly, and a minimum take-up was approximately 16% in those customers who spend R21 000 monthly.

**Figure 4: The effect of average credit turnover on take-up**



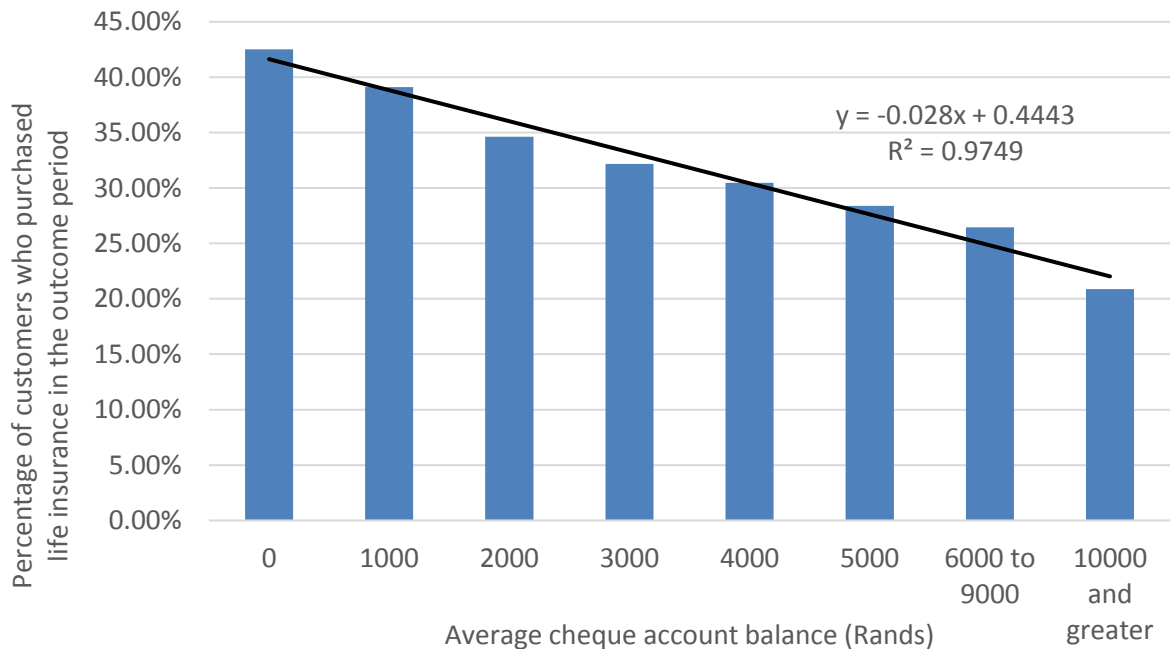
In the figure below the percentage of take-up of life insurance increased slightly with decreasing negative cheque account balances. After this, in customers who had a positive cheque account balance, there was a decrease in life insurance purchased with increasing cheque account balance. A second order polynomial trendline was fit to the data with a  $R^2$  of 0.6542. The maximum take-up was approximately 42% with customers who on average had almost no money in their cheque accounts, and a minimum take-up was approximately 15% in those customers who had an account balance of more than R10 000.

**Figure 5: The effect of average cheque balance on take-up, including customers with a negative account balance**



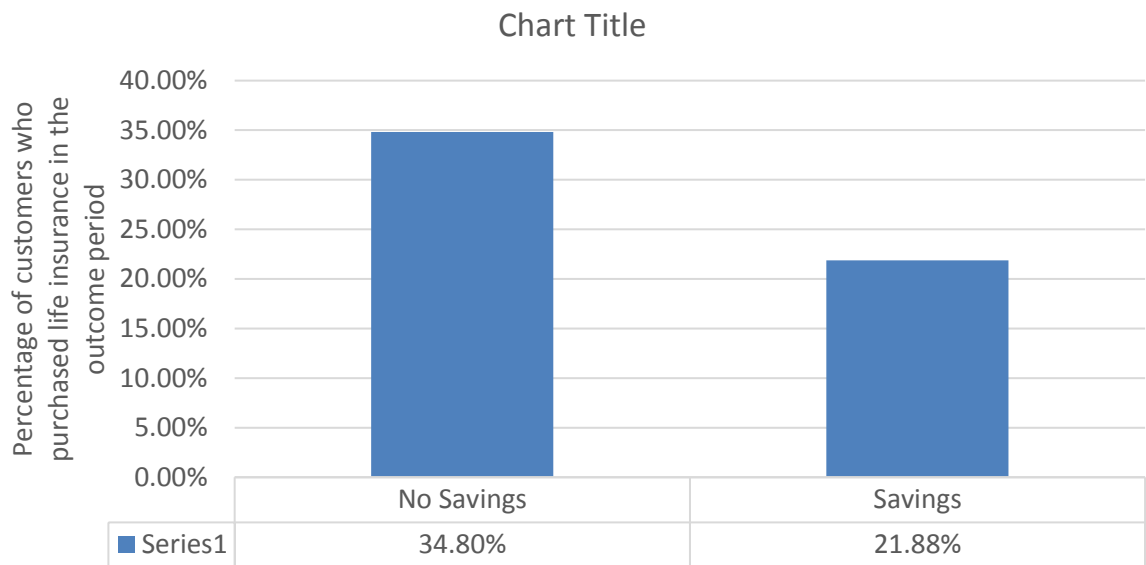
In the figure below the effect of average cheque account balance is reproduced with those customers who had a negative account balance being omitted. A clear linear trendline provides a  $R^2$  of 0.9749.

**Figure 6: The effect of average cheque balance on take-up, excluding customers with a negative account balance.**



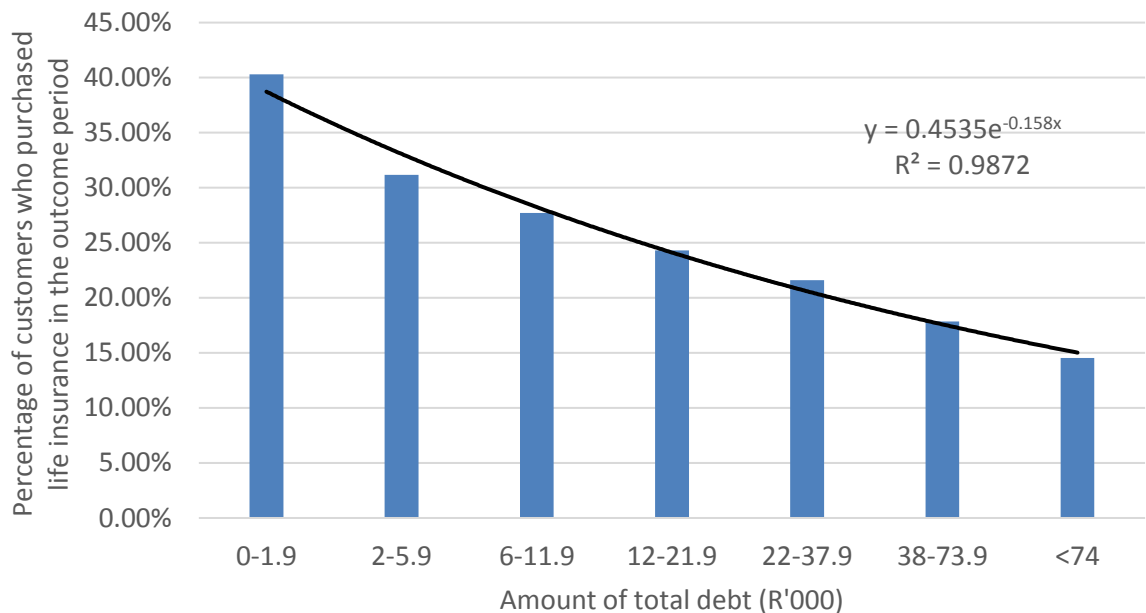
The figure below shows the effect of not having savings on the likelihood of taking up life insurance. Customers with no savings were more likely to take-up life insurance savings (34.8%) as compared to those customers who had savings (21.88%).

**Figure 7: The effect of savings on take-up**



The figure below shows the effect of total debt on the likelihood of taking life insurance. There was a decrease in life insurance take-up with an increase in the amount of debt. Approximately 40% of those customers who had between no debt and R1 900 of debt took out life insurance while only 15% of customers with more than R74 000 of debt were likely to take-up life insurance. An exponential trendline had an  $R^2$  of 0.9872.

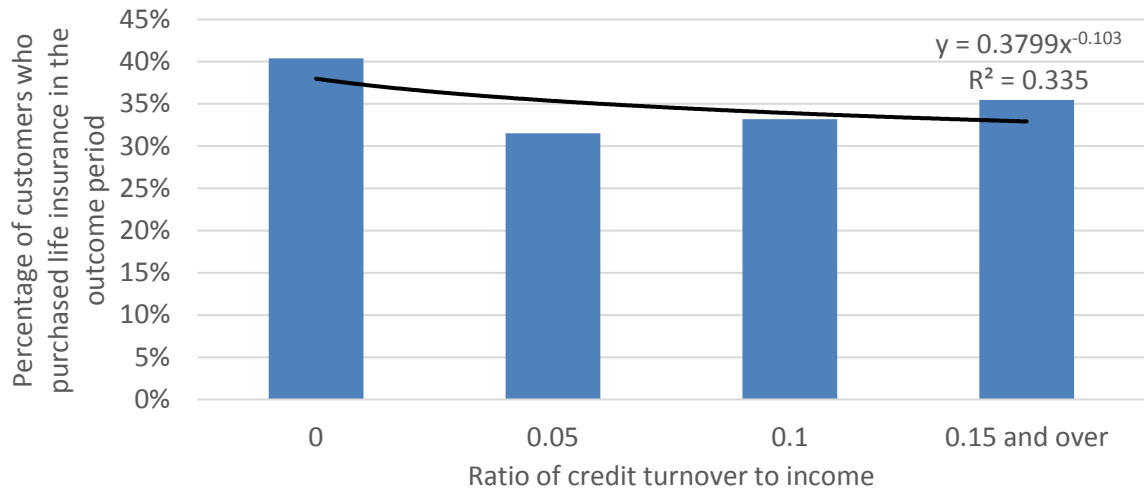
**Figure 8: The effect of amount of total debt on take-up**



The below figure shows the effect of the ratio of credit turnover to income. The trend was mixed with a ratio of 0.05 showing the lowest take-up of life insurance. With increasing credit turnover to income there was an increasing take-up of insurance. The

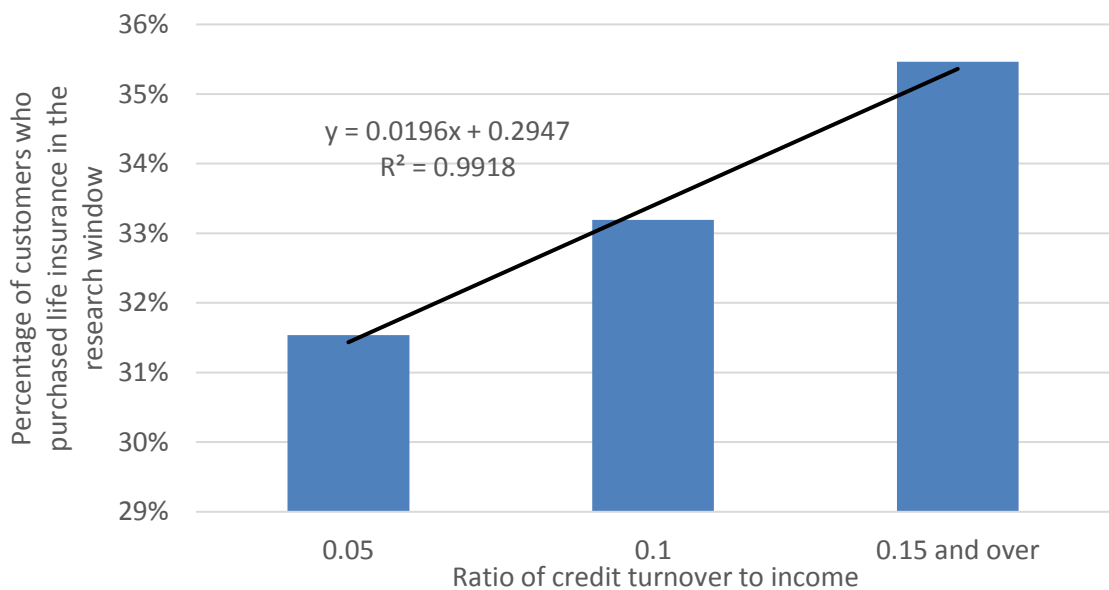
maximum take-up was 40% among those customers who had a ratio of credit turnover to income of close to zero. A power function trendline had a  $R^2$  of 0.335.

**Figure 9: The effect of the ratio of average credit turnover to income**



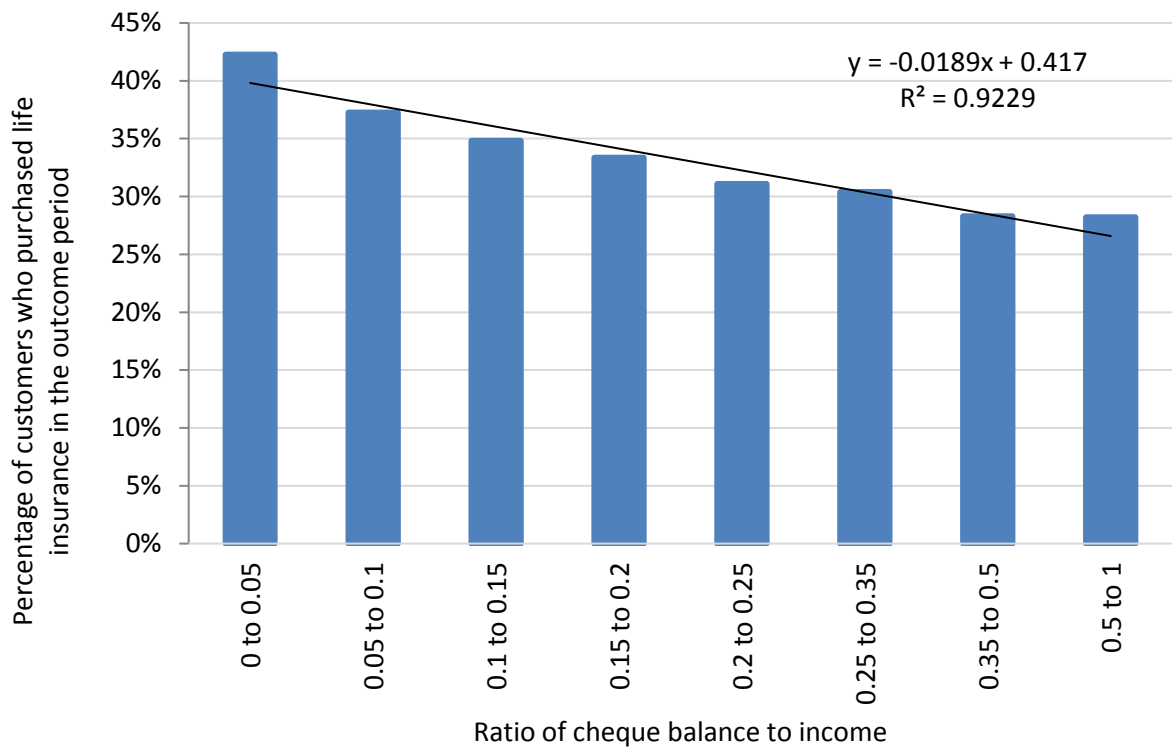
The figure below shows the effect of the ratio of credit turnover to income on take-up. There was an increase of take-up with an increase in the ratio of credit turnover to income. The minimum of 32% of take-up was seen in those customers who had a very low ratio of credit turnover to income, and a maximum take-up of 35% among those customers who spent the largest part of their income in a month. A linear trendline was fit to the data with a  $R^2$  of 0.9918. Those customers with very large income and very small spend were excluded from this figure.

**Figure 10: The effect of the ratio of average credit turnover to income on take-up**



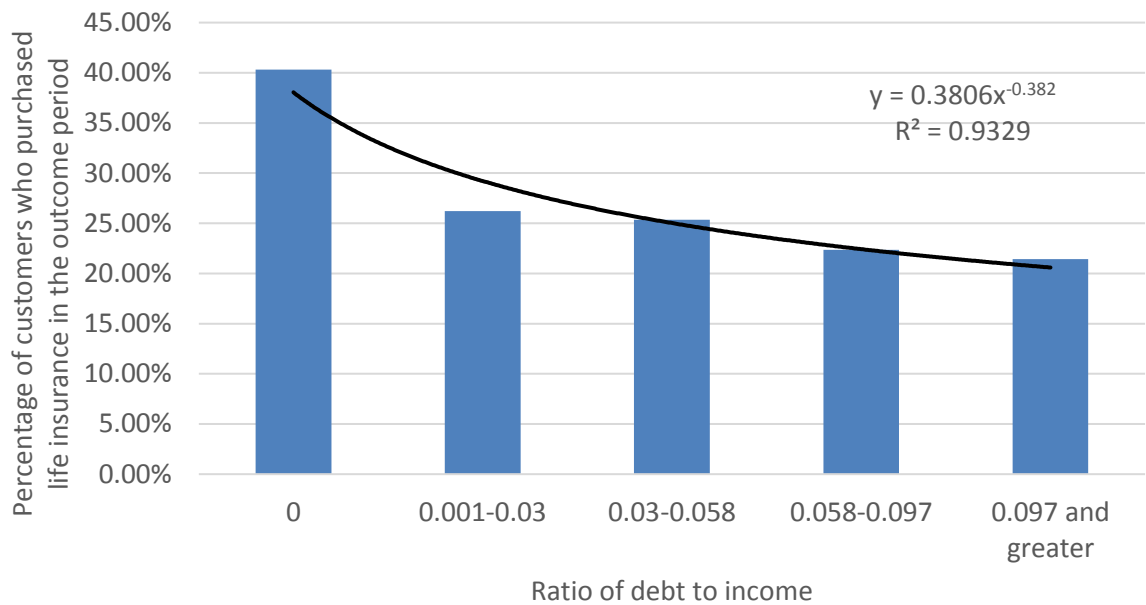
The following figure shows the ratio of cheque balance to income on take-up. There was a clear decrease in insurance take-up with increasing ratio of cheque balance to income. Even a linear trendline had a corresponding  $R^2$  of 0.9229. The highest take-up (42%) was among those customers who had a very low ratio of credit turnover to income as compared to the lowest take-up of 28% among customers with the highest ratios of account balance to income.

**Figure 11: The effect of the ratio of average cheque account balance to income on take-up**



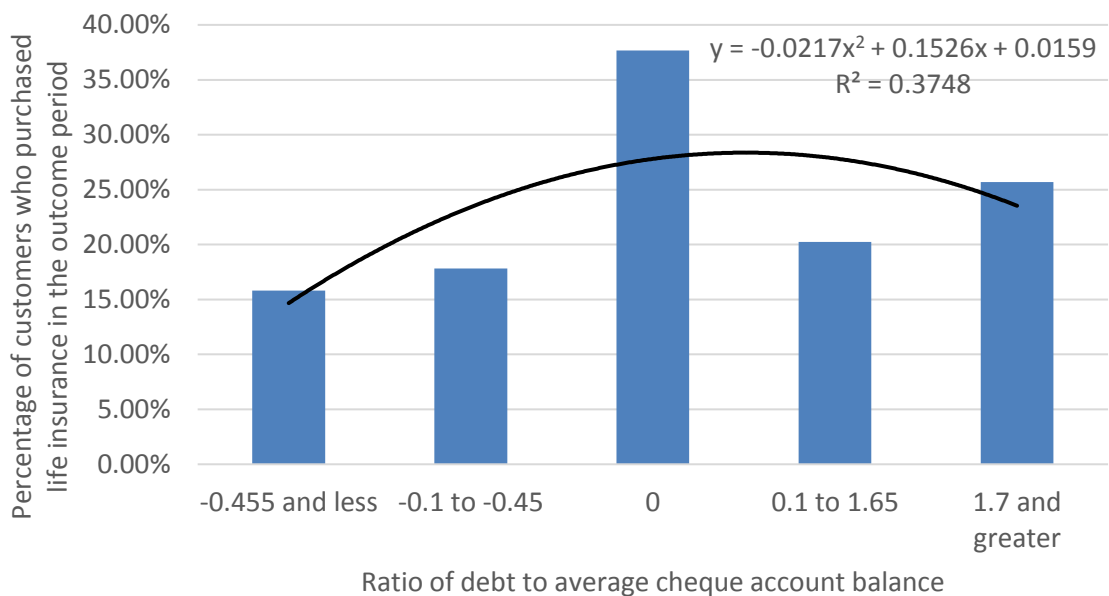
The following figure shows the effect of the ratio of total debt to income on take-up. With an increasing ratio of debt to income there is a decreasing take-up of life insurance. The maximum take-up (40.33%) was among customers who had a ratio of debt to income of close to zero. The minimum take-up (21.44%) was among customers who had the largest ratio of debt to income. A power function trendline had a  $R^2$  of 0.9329.

**Figure 12: The effect of the ratio of total debt to income on take-up**



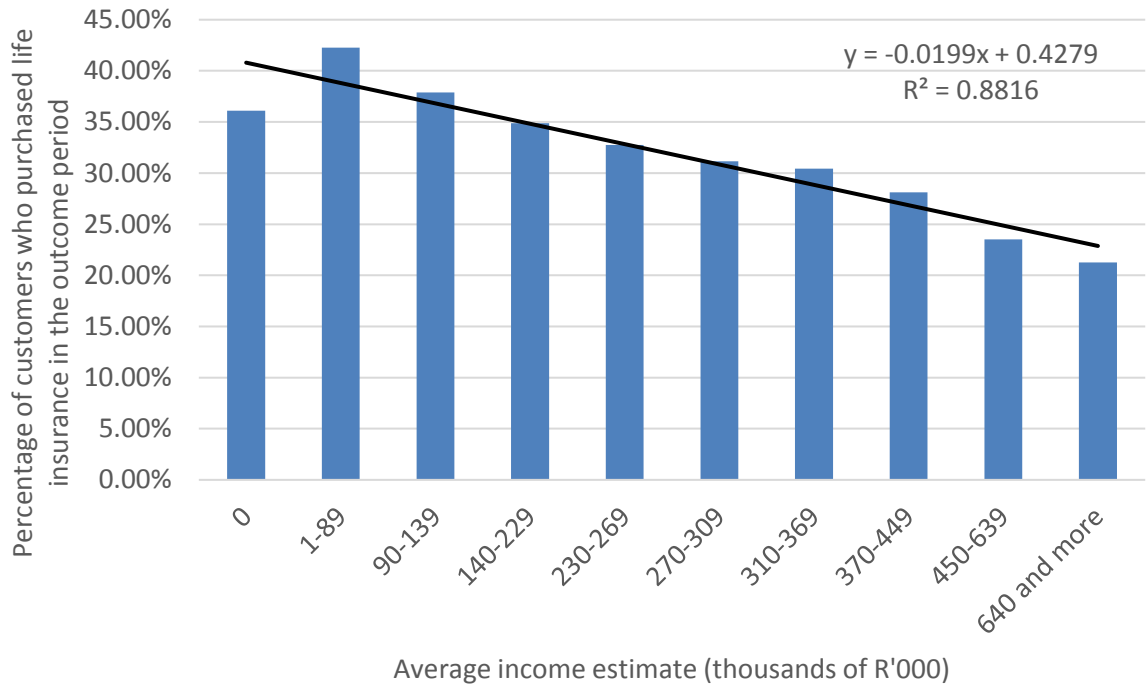
The final figure shows the effect of the ratio of debt to cheque account balance. Customers who had a ratio of debt to account balance had the highest take-up (37.68%). The best fit trendline was a second order polynomial trendline but it had a  $R^2$  of only 0.3748.

**Figure 13: The effect of the ratio of total debt to average cheque account balance on take-up**



The below figure shows the influence of income on take-up. The variable was not significant in the model but it shows that there is a trend between increasing income and a decrease in insurance consumption as evidenced by the linear trendline with a  $R^2$  value of 0.8816.

**Figure 14: The effect of income on take-up**



## 5.6 Lapse Model

Secondly, another separate binomial logistic regression was executed on a separate sample of banking customers who also had an active life insurance product at the time of the observation point. The logistic regression model was used to determine the effects of the following independent variables on the probability that individuals lapsed their life insurance (failed to pay their insurance premiums). Table 13 below shows both the significant and insignificant predictor variables for the logistic regression model on lapse (based on  $p < 0.05$ ):

**Table 14: Significant and insignificant predictor variables for lapse**

Significant predictor variables	Insignificant predictor variables
Debt (total debt across a range of debt products)	Average income
Savings (average savings and investments balance)	Ratio of savings to income
Average credit turnover	Ratio of debt to income
Average cheque account balance	
Ratio of credit turnover to income	
Ratio of average cheque account balance to income	
Ratio of debt to average cheque account balance	

The assessment showed that seven of the ten continuous independent variables were found to be linearly related to the logit of the dependent variable. All the studentized residuals were retained in the model. The logistic regression model was statistically significant with  $\chi^2(7) = 2254.132$ ,  $p < .0001$ . The pseudo  $R^2$  (Nagelkerke) value was 4.9% which explained the variance in lapses. The model correctly classified 62.5% of cases. The positive predictive value of the model was 62.4% and negative predictive value was 37.4%.

Of the ten predictor variables, only seven of were statistically significant as shown in Table 9. According to the logistic regression an increase in the monthly average credit turnover, average cheque account balance, average savings, and, ratio of cheque balance to income all decreased the likelihood of lapsing. On the other hand increasing the total debt, ratio of credit turnover to income and ratio of debt to average balance all led to an increase in lapses.



**Table 15: Logistic Regression Predicting Likelihood of Lapse based on Average Credit Turnover, Average Cheque Account Balance, Average Savings and Investments Balance, Total Debt, Ratio of Credit Turnover to Income, Ratio of Average Cheque Account Balance to Income and Ratio of Total Debt to Average Cheque Account Balance.**

	<i>B</i>	<i>SE</i>	Wald	<i>p</i>	Odds Ratio	95% CI for Odds Ratio	
						Lower	Upper
Average credit turnover	-0.00004	2.076E-6	456.96	<.0001	1.000	1.000	1.000
Average cheque account balance	-0.00005	3.155E-6	239.58	<.0001	1.000	1.000	1.000
Average savings	-4.44E-6	1.532E-6	8.39	0.0038	1.000	1.000	1.000
Total debt	0.000016	9.633E-7	290.84	<.0001	1.000	1.000	1.000
Ratio of credit turnover to income	1.5418	0.1702	82.02	<.0001	4.673	3.347	6.524
Ratio of average cheque account balance to income	-0.1715	0.0262	42.84	<.0001	0.842	0.800	0.887
Ratio of total debt to average cheque account balance	0.0262	0.00151	300.99	<.0001	1.027	1.023	1.030
Intercept	-1.1359	0.0210	2931.8126	<.0001	0.318		

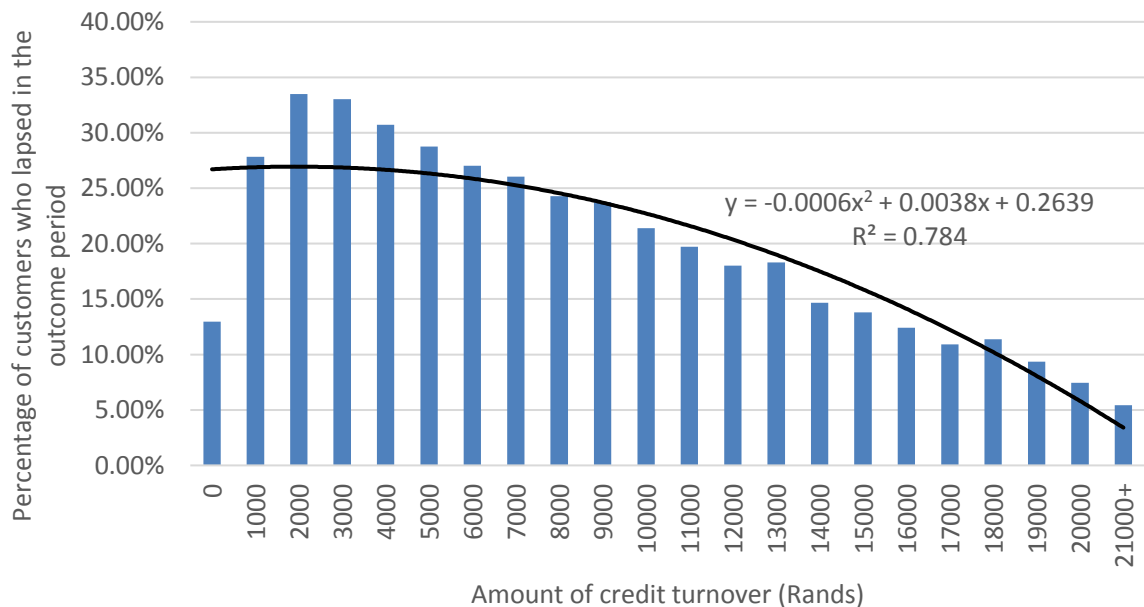
*Note that the degrees of freedom for all variables was one.*

### 5.6.1 Graphical Representation of Lapse Predictor Variables

In the following section a graphical representation of the data is shown. An emphasis has been placed on those variables that were found to be significant on the lapse dependent variable in the logistic regression model.

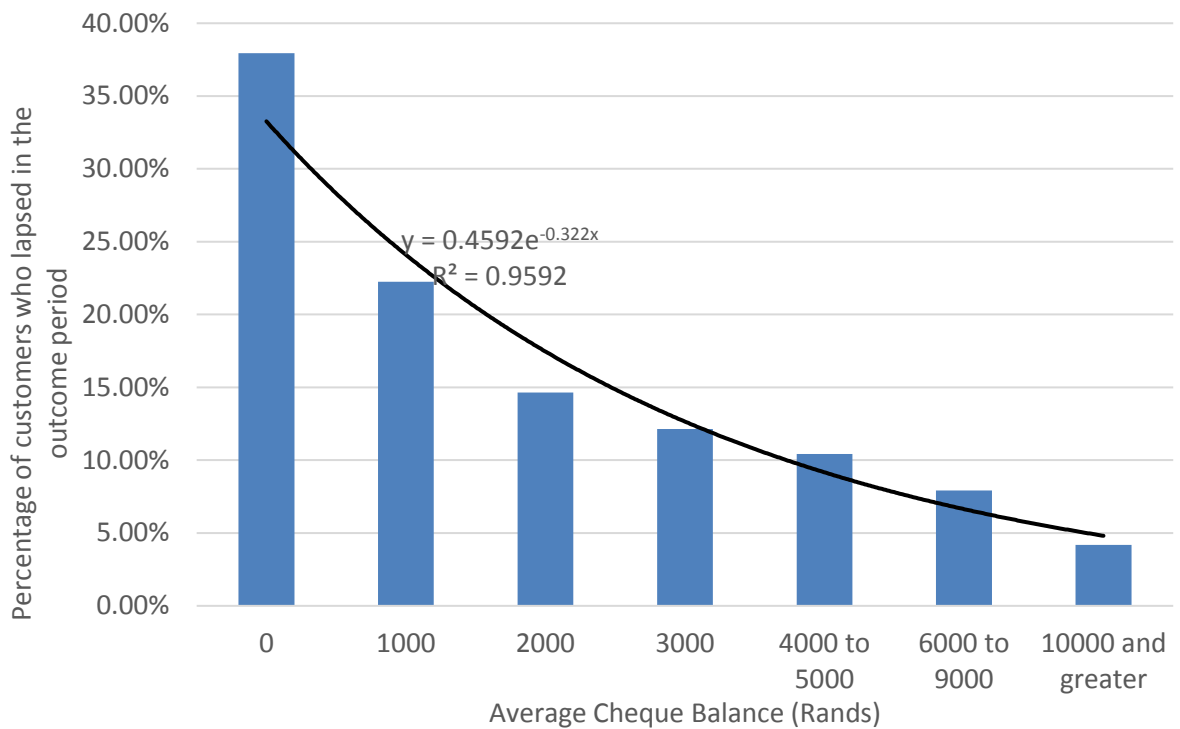
In the figure below the percentage of lapse of life insurance increased slightly with increasing credit turnover but only until about R2 000 monthly spend. In the customers who spend more than R2 000 monthly, there was a decrease in life insurance purchased with increasing credit turnover. A second order polynomial trendline was fit to the data with a  $R^2$  of 0.784. The maximum lapse was 33.48% with customers spending R2 000 monthly, and a minimum lapse was approximately 5.43% in those customers who spend R21 000 monthly or more.

**Figure 15: the effect of credit turnover on lapse**



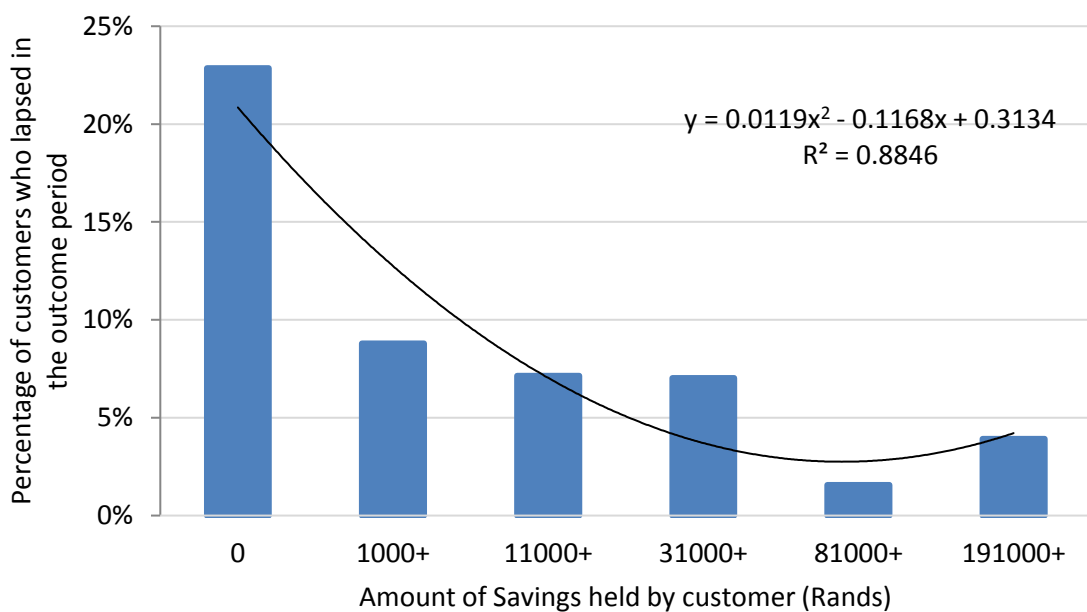
In the figure below customers who had a positive cheque account balance, there was a decrease in life insurance purchased with increasing cheque account balance. An exponential function trendline was fit to the data with a  $R^2$  of 0.9592. The maximum lapse was approximately 37.96% with customers who on average had almost no money in their cheque accounts, and a minimum lapse was approximately 4.18% in those customers who have R10 000 or more in their cheque account.

**Figure 16: The effect of average cheque balance on lapse**



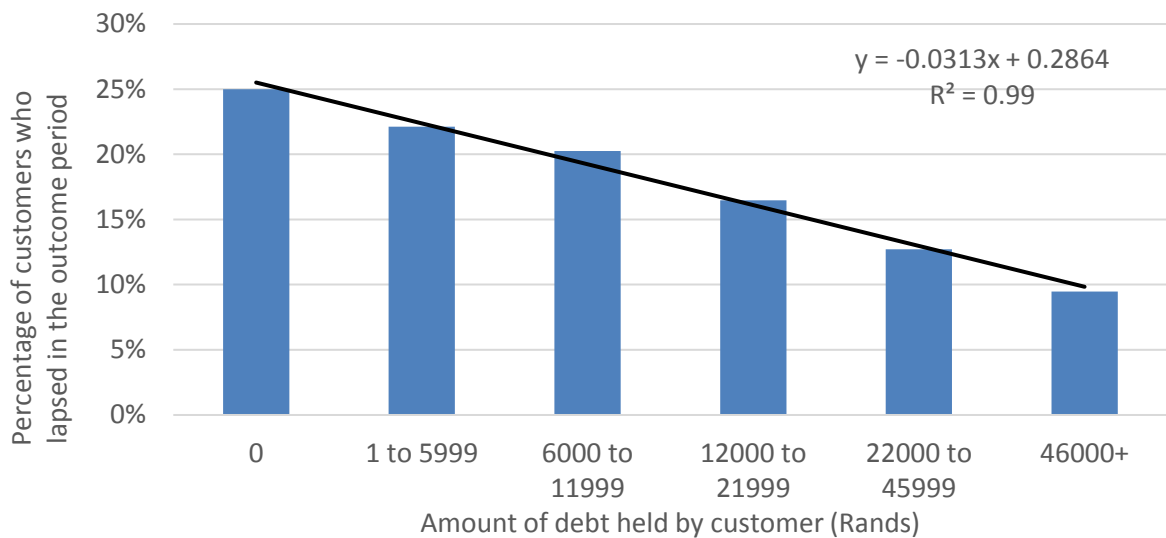
The figure below shows the effect of savings on the likelihood of lapsing life insurance. Customers with no savings were more likely to lapse life insurance savings (23%) as compared to those customers who large amounts of savings (between 2 and 4% of customers with R81 000 and greater, and R191 000 and greater, respectively).

**Figure 17: The effect of savings on lapse**



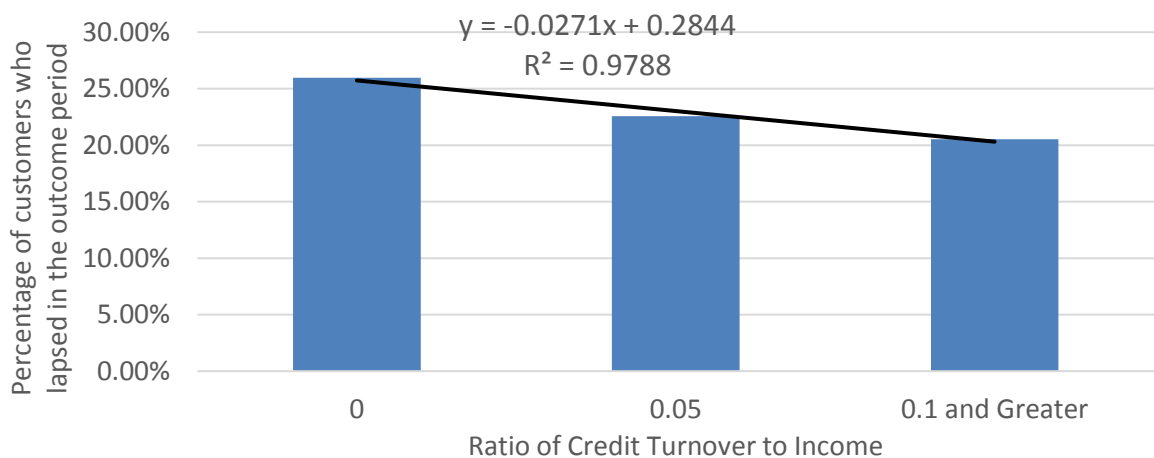
The figure below shows the effect of total debt on the likelihood of lapsing insurance. There was a decrease in life insurance lapse with an increase in the amount of debt. Approximately 25% of those customers who had no debt lapsed their life insurance while only 9% of customers with more than R46 000 of debt were likely to lapse life insurance. A simple linear trendline had an  $R^2$  of 0.99.

**Figure 18: The effect of total debt on lapse**



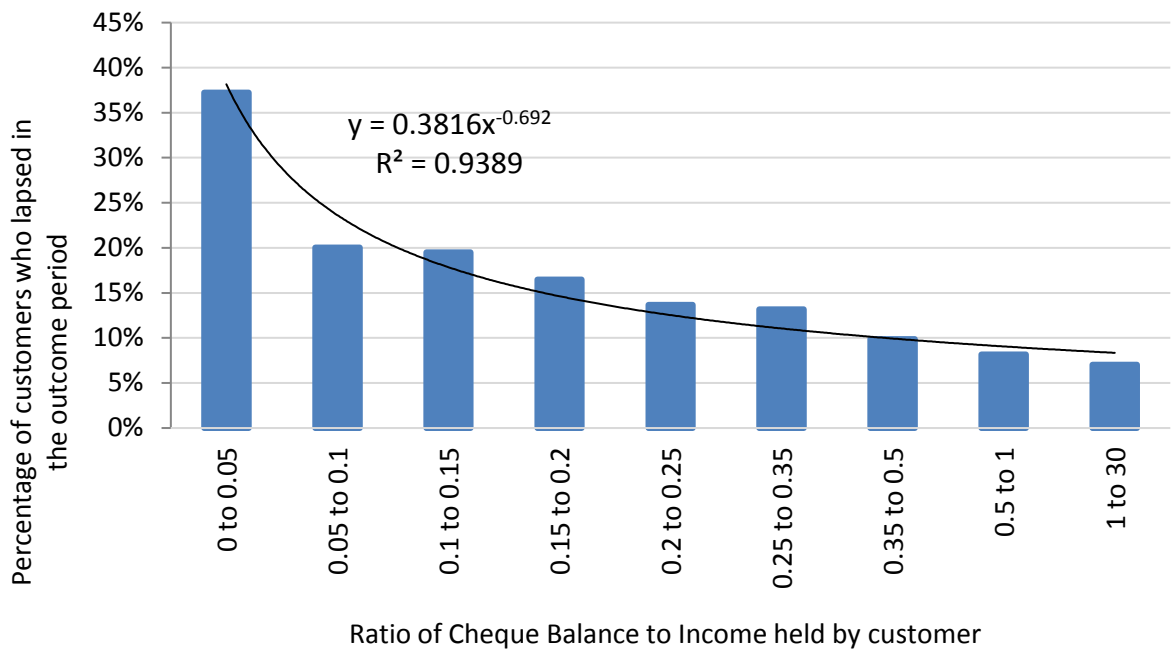
The below figure shows the effect of the ratio of credit turnover to income on lapsing. The trend a decrease in lapse rate with increasing ratio of credit turnover to income. The maximum lapse was 25.96% among those customers who had a ratio of credit turnover to income of close to zero. The minimum lapse was 20.54% among customers who had a ratio of credit turnover to income of 0.1 and more. A linear trendline had a  $R^2$  of 0.9788.

**Figure 19: The effect of the ratio of credit turnover to income on lapse**



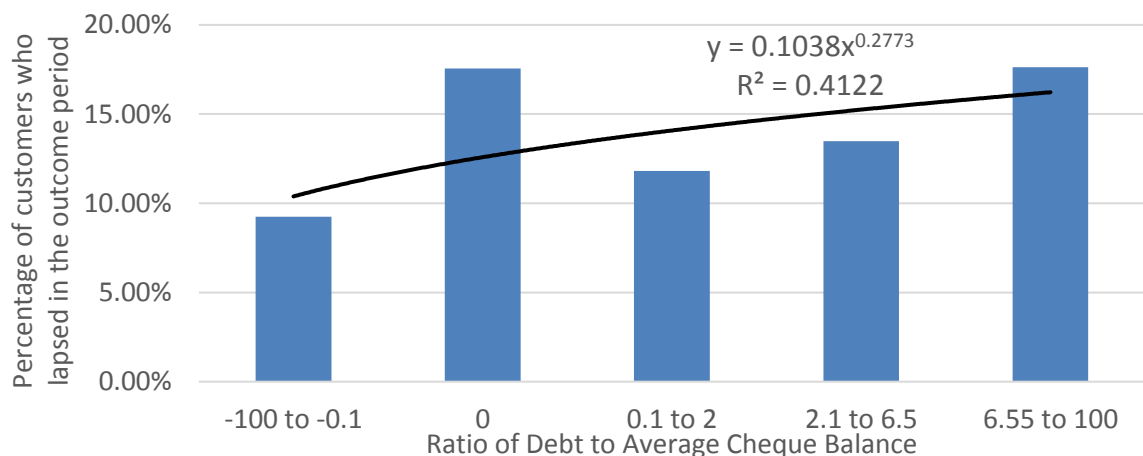
The following figure shows the ratio of cheque balance to income on lapse. There was a clear decrease in insurance lapse with increasing ratio of cheque balance to income. A power function trendline had a corresponding  $R^2$  of 0.9389. The highest lapse (37%) was among those customers who had a very low ratio of credit turnover to income as compared to the lowest lapse of 7% among customers with the highest ratios of account balance to income.

**Figure 20: The effect of the ratio of cheque balance to income on lapse**



The final figure shows the effect of the ratio of debt to cheque account balance on lapse. Customers who had a ratio of debt to account balance of close to zero had the highest lapse (17.55%). The best fit trendline was a power function trendline but it had a  $R^2$  of only 0.4122.

**Figure 21: The effect of total debt to average cheque account balance on lapse**



## 5.7 Cancel Model

The final binomial logistic regression was executed on a sample of banking customers who also had an active life insurance product at the time of the observation point. The logistic regression model was used to determine the effects of the following independent variables on the probability that individuals cancelled their life insurance; Table 16 below shows both the significant and insignificant predictor variables for the logistic regression model on cancellations (based on  $p < 0.05$ ):

**Table 16: Significant and insignificant predictor variables for cancellations.**

Significant predictor variable	Insignificant predictor variables
Average credit turnover	Average income
	Average cheque account balance
	Debt
	Savings
	Ratio of average cheque account balance to income
	Ratio of debt to average cheque account balance
	Ratio of savings to income
	Ratio of debt to income

The assessment showed that one of the ten continuous independent variables were found to be linearly related to the logit of the dependent variable. All the studentized residuals were retained in the model. The logistic regression model was statistically significant with  $\chi^2(1) = 138.857$ ,  $p < .0001$ . The pseudo  $R^2$  (Nagelkerke) value was 0.7% which explained the variance in cancellations. The model correctly classified 54.9% of cases. The positive predictive value of the model was 52.2% and negative predictive value was 42.4%.

Of the ten predictor variables, only one of was statistically significant as shown in Table 16. An increase in the average credit turnover lead to a decrease in cancellations.

**Table 17: Logistic Regression Predicting Likelihood of Cancellation based on Average Credit Turnover.**

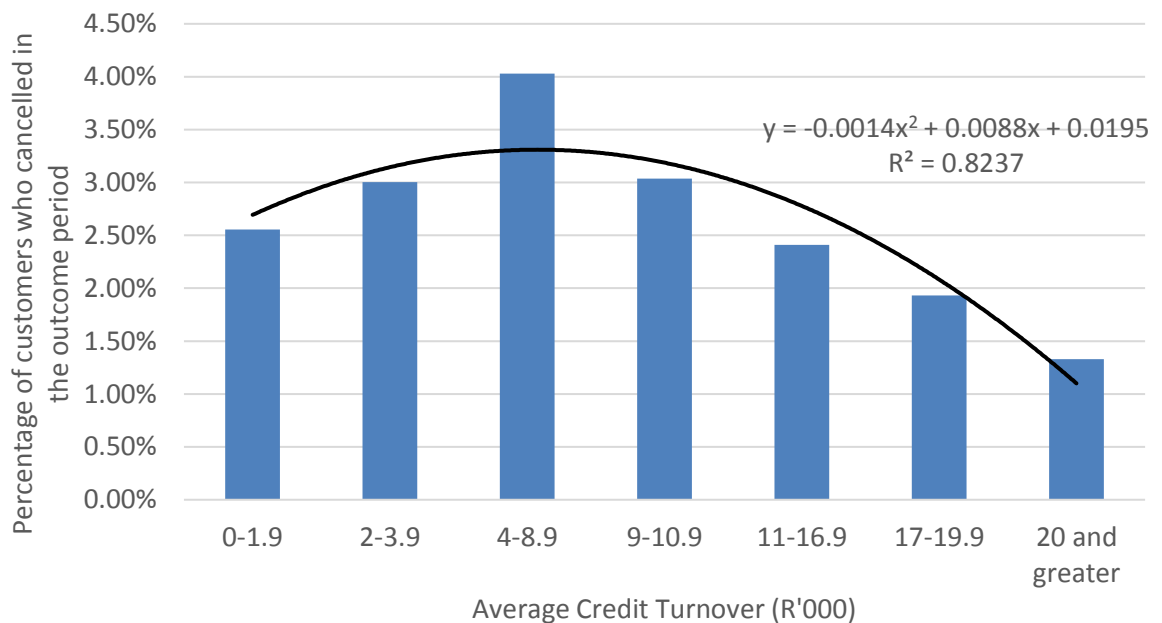
	<i>B</i>	<i>SE</i>	Wald	<i>p</i>	Odds Ratio	95% CI for Odds Ratio	
						Lower	Upper
Average credit turnover	-0.00003	3E-6	456.96	<.0001	1.000	1.000	1.000
Intercept	-1.1359	0.0210	12137.2859	<.0001	0.041		

*Note that the degrees of freedom for all variables was one.*

### 5.7.1 Graphical Representation of the Cancellation Predictor Variables

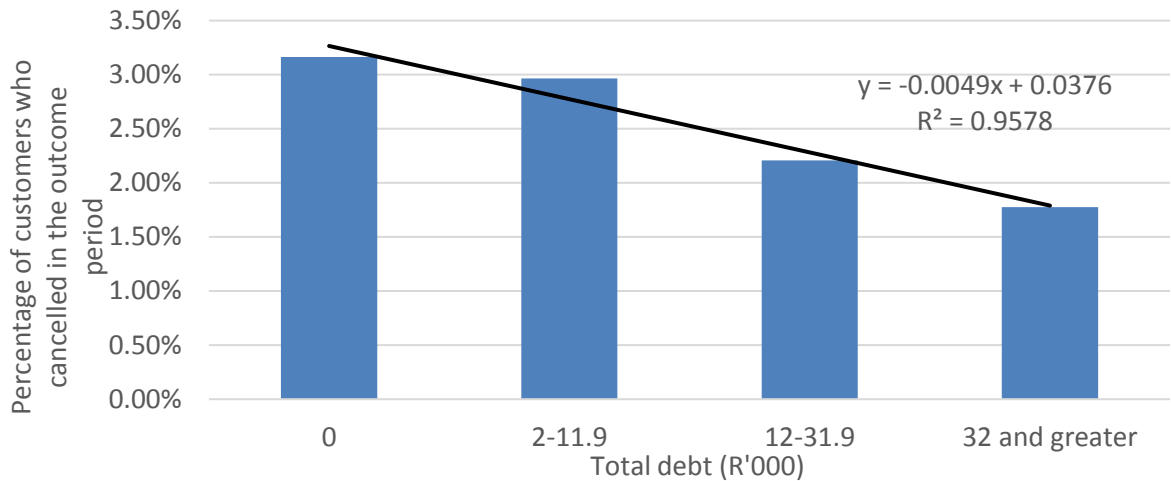
In the figure below the percentage of cancellations of life insurance increased with increasing credit turnover but only until about the bucket of customers spending between R4 000 and R8 900 monthly. In the customers who spend more than R9 000 monthly, there was a decrease in life insurance purchased with increasing credit turnover. A second order polynomial trendline was fit to the data with a  $R^2$  of 0.834. The maximum cancellation was 4.03% with customers spending between R4 000 and R8 900 monthly, and a minimum lapse was approximately 1.33% in those customers who spend R20 000 monthly or more.

**Figure 22: The effect of average credit turnover on cancellations, this was the only significant variable in the logistic regression model**



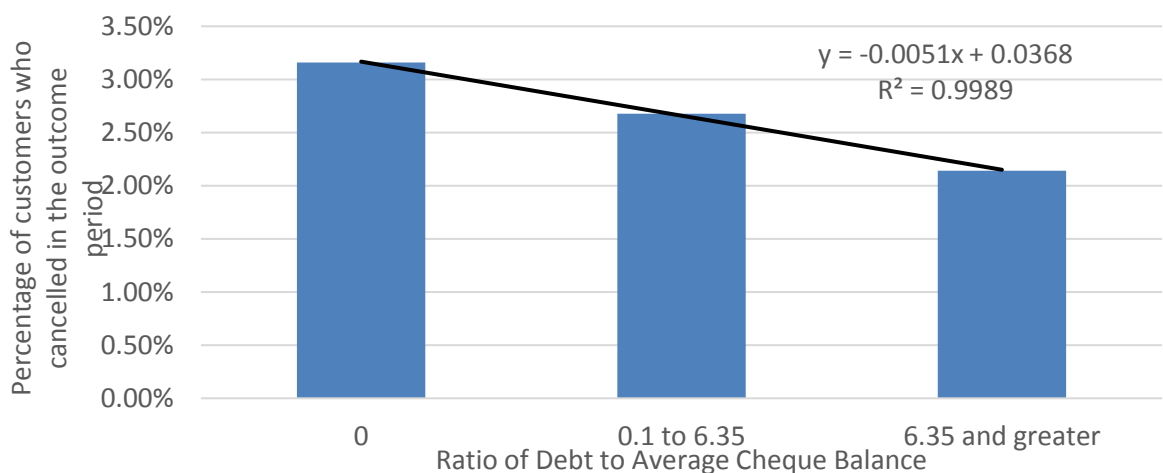
The figure below shows the effect of total debt on the likelihood of lapsing insurance. This was an insignificant predictor in the logistic regression model. There was a decrease in life insurance lapse with an increase in the amount of debt. Approximately 3.16% of those customers who had no debt cancelled their life insurance while only 1.78% of customers with more than R46 000 of debt were likely to cancel their life insurance. A simple linear trendline had an  $R^2$  of 0.958.

**Figure 23: The effect of total debt on cancellations**



The following figure shows the effect of the ratio of debt to cheque account balance on cancellations. Customers who had a ratio of debt to account balance of close to zero had the highest cancellations (3.16%). The best fit trendline was a linear trendline with a  $R^2$  of 0.9989. This predictor variable was also insignificant in the logistic regression model but it is included in the results section to illustrate the small variance in cancellations.

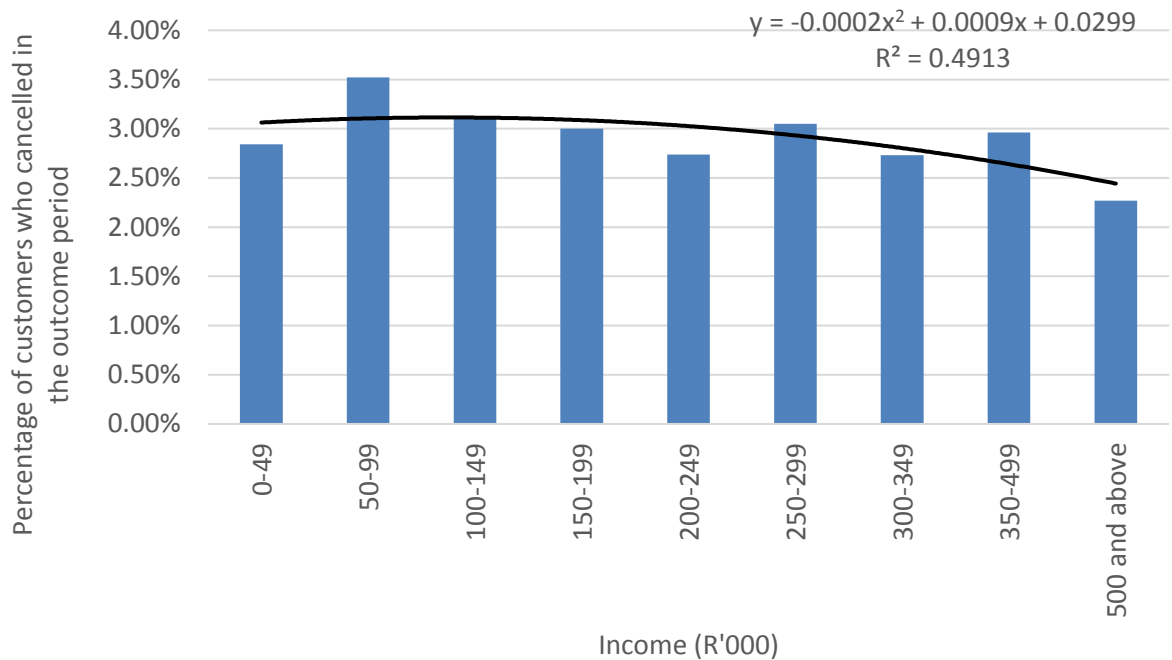
**Figure 24: The effect of the ratio of debt to average cheque balance on cancellations**





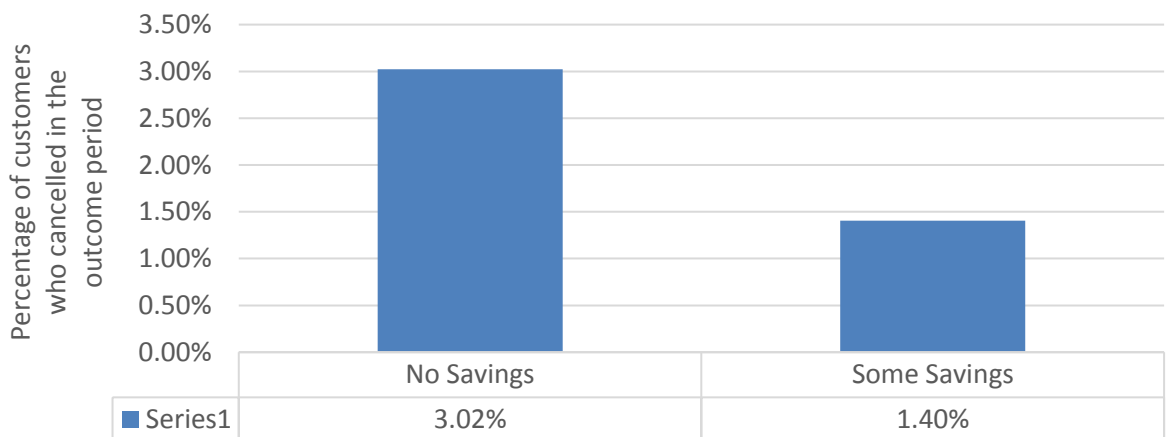
The following figure shows the effect of income on cancellations. This was an insignificant predictor in the binomial regression model. The results are again mixed but there is a slight tendency for cancellations to decrease with increasing income. The maximum cancellations occurred among customers earning between R100 000 and R149 000 while the minimum cancellations occurred among customers earning R500 000 and more annually.

**Figure 25: The effect of average income on cancellations**



The figure below shows the effect of not having savings on the likelihood of cancelling life insurance. Customers with no savings were more likely to cancel life insurance savings (3.02%) as compared to those customers who had savings (1.4%).

**Figure 26: The effect of savings on cancellations**



## 6 Discussion of Results

### 6.1 Introduction

In the following chapter, the results provided in the previous chapter will be discussed. The analysis of the customer records proves insightful in the context of the research questions and literature reviewed in Chapter 2 above. The discussion that this chapter provides is furthermore insightful to consider for the purposes of designing insurance products that are appealing to a given customer, or is effective in avoiding lapses and cancellations. The discussion focuses on the results given by the binary logistic regression analysis as well as the graphical results shown in Chapter 5. The discussion also draws comparisons to the findings presented in the literature from previous research. To summarise, the research questions that were investigated are:

1. Is there a relationship between the demand for life insurance products and the amount of income an individual has?
2. Is there a relationship between the demand for life insurance products and the amount of savings an individual has?
3. Is there a relationship between the demand for life insurance products and the amount of debt an individual has?

The discussion of each of the research questions (also provided in Chapter 3) is presented in detail below. The discussion of the research questions take place in the order provided above. For each research question, each of the three models is discussed with reference to the applicable logistic regression models in order of take-up, lapse, and lastly cancel. Concluding remarks are provided at the end of each section on the concepts relating to the research questions. Finally, all concerns that arose from the discussion are summarised in Section 6.5

### 6.2 Discussion of Research Question 1: The influence of income on the demand for life insurance products

In the following section, the results of the various logistic regression models will be discussed in relation to the first research question. The first research question is concerned with how income may affect the propensity to take-up, lapse, or cancel life insurance. Since life insurance is a form of protection against risks, this is risk aversion behaviour. We begin by discussing the propensity for insurance consumption, followed by the propensity to lapse and finally, discuss the propensity to cancel.

### **6.2.1 Discussion of Income in the Take-up Model**

The binomial logistic regression model did not find the average income estimate to be a significant predictor variable for take-up. The model did however find that an increase in the ratio of credit turnover to income, significantly lead to an increase in take-up. This may indicate that customers who had high incomes in relation to their monthly commitments were less likely to take-up insurance. There are important possible implications for this finding, namely that customers who had greater financial means to cover their monthly expenses, likely had the capacity to deal with risks and uncertainty more sufficiently. Low income individuals often do not have financial recourse in times of uncertainty. This is also in line with literature explaining that less wealthy individuals are less able to deal with negative income shocks (Fier & Liebenberg, 2013). This variable had the highest  $\beta$  which means that it is the strongest predictor for influencing take-up. The results are aligned to the findings of Lin and Grace (2007) who reported that an individual with decreasing risk aversion will consume lower amounts of insurance when at higher income levels.

The model also found that there was a significant increase in take-up with increasing ratios of average account balance to income. This variable may also be interpreted as a measure of relevant affluence. By increasing the denominator of the variable (increasing income), the ratio is lowered and there is a decrease in take-up. Therefore, relatively higher income customers may be taking-up less insurance. Increasing values of the ratio is a sign of customers who keep larger portions of their income in their cheque accounts, while customers who are earning large sums are likely investing the surplus elsewhere (such as in savings or investment accounts). This theory is corroborated by Sliwinski et al. (2013) and Lin et al., (2017). They state that some individuals with larger incomes are more likely to move their funds into more financially sophisticated investments.

The model predicted a small decrease in take-up with increasing average credit turnover. This is consistent with the above two findings. Since customers who spent greater amounts were found to take-insurance less. Customers with lower amounts of disposable income are therefore less likely to purchase the life insurance products offered by the insurer. Figure 4 shows the downward trend.

The average balance in the customer's cheque account, also had a small but significant effect on the take-up of insurance. With an increasing amount of money in a cheque account, there was a decrease in take-up of insurance products (see Figure 6).

This finding is also consistent with the other results as less wealthy customers were more likely to purchase insurance.

There was an interesting finding that can be seen when comparing Figure 5 and Figure 6. This is that customers who had access to an overdraft facility, were less likely to purchase insurance. Individuals who are granted access to overdraft facilities are usually quite financially sophisticated and credit worthy. We therefore see a decrease in the likelihood of insurance consumption which is in contradiction with the literature on financial sophistication (Mulholland & Finke, 2014; Lin et al., 2017). It is important to address the issue of context now. In much of the literature the life insurance products are suitable to individuals with high financial literacy (Fernandes et al., 2014). These individuals are concerned with estate planning and bequests (Lin et al., 2017). However, it must be stressed that in the South African context, some insurers actively aim to offer simple to understand and inexpensive insurance products, which are suitable for low income individuals. This paper therefore concludes that context is an important factor to consider when reporting on trends and relationships concerning insurance. These results therefore may help to raise more questions for future research.

### **6.2.2 Discussion of Income in the Lapse Model**

The results of the logistic regression model for predicting lapses, showed that there may be a number of income related individual-level financial behaviours that may be significantly correlated to lapses. The findings are consistent with the findings of Hou, et al., (2015), who reported that low-income and low-wealth individuals may be more likely to lapse their insurance policies. The model is significant in expounding some finer details of which measures may be linked to the increase in lapses.

The ratio of credit turnover to income had the strongest predictive value in the lapse model. The variable is a measure of the amount a customer spends in a month compared to their income. Increasing ratios may mean that customers are spending larger fractions of their total income. This could indicate that they are more financially stressed. Increasing the ratio led to an increase in lapses (see Figure 19). Typically, low-income and low-wealth individuals could be more financially stressed and it therefore is corroborated by the analysis that they were more likely to lapse their policies. This is also consistent with Kiesenbauer, (2012) who reported that one of the root causes of lapses, is the “emergency fund” hypotheses, where customers are using the money that they would have used in paying for their insurance premiums on more

urgent matters. More affluent individuals typically have other means for paying for emergencies.

An increase in the ratio of average account balance to income may have resulted in a decrease in lapses, but the result seems quite mixed (also see Figure 20). Customers who keep a large amount of money in their cheque accounts are seemingly more risk averse. Increasing values of the ratio is perhaps a sign of customers who maintain larger sums in their cheque accounts in portion to their income. While customers who are earning large sums are likely investing the surplus elsewhere (such as in savings or investment accounts) in accordance with Clark et al., (2012). The model shows a decrease in likelihood of lapsing with increasing ratios. This variable may be useful in identifying those customers who are financially stressed and are lapsing, but due to the mixed results, no definitive conclusions should be made from this finding.

The model identified the average credit turnover as a significant predictor variable to lapsing. It must be said that the effect was, however, very small as can be seen from the  $\beta$  value. Nonetheless, the model found that with increasing credit turnover, there could have been a likelihood of increasing lapses. This is somewhat contradictory to the findings of the take-up model. As can be seen from Figure 4, there may be an increase in lapse with increasing credit turnover, directly followed by a decrease in lapses. Since most customers in the sample are low income customers, the logistic regression likely identifies this to be the dominant trend, due to the larger total of responses. However, from Figure 4 it is apparent that, at higher turnover amounts, an increasing value may lead to lower lapses. The findings are therefore that the lowest income customers are less likely to lapse their income, primarily because they probably value their life insurance greatly. The tendency to lapse increases up to a point, after which it again decreases with increasing income as expected, since more affluent customers are more likely to afford their premiums (Eling & Kiesenbauer, 2014).

The result for the average cheque balance predictor variable is that with increasing money in their accounts, there may be a lower tendency for the customer to lapse. Therefore, the conclusion is that maintaining higher amounts of money on hand, which is in itself a form of risk aversion behaviour, may be correlated to lower lapses, which is also a form of risk aversion. The increase in cheque account balance is also a measure of income and as expected more affluent customers are less likely to miss premium payments, which are in accordance with Fier and Liebenberg, 2013.

Interestingly, the model also did not include income as a significant variable in predicting lapse. This is may be due to comparatively homogenous consumption across the sample with income as an absolute value.

### **6.2.3 Discussion of Income in the Cancel Model**

The logistic regression model concerned with cancellations only found one significant predictor value, namely the average credit turnover. The relationship was negative, indicating that with increasing monthly spend, customers were less likely to cancel their insurance. The result is important since cancellations are a highly unwanted behaviour in insurance, both from a corporate and individual level (Christiansen & Niemeyer 2014). Losing insurance cover means that all the premiums paid are effectively lost according to Guillén et al., (2012). The sense is that low income customers (customers who spend the least amount of money) are more likely to cancel their policies (probably due to affordability). This means that they both lose the premiums paid in the past as well as the benefits and protection that the insurance affords (Guillén et al., 2012).

The result may be reconciled with Kunreuther (2015) who reported that customers cancel their insurance policies, since they view their insurance as an investment rather than financial protection. The result may also be due to financially unsophisticated individuals, purchasing cover as a result of experiencing a loss. These customers change their disposition towards their insurance and do not value their original purchase decision. The result is that they rather cancel their insurance. Considering the other side of the income spectrum; there may be a situation where more affluent customers exhibit this behaviour less often. This may be evidenced in the result presented in Figure 22.

### **6.2.4 Summary of Income**

The results between the variables are consistent with one another and all point in the same direction. Even though none of the regression models identified income as a significant predictor variable, the other predictor variables used as a proxy for understanding wealth were significant. It can be seen from Figure 14 that there is a definite downward trend in insurance consumption with increasing income. Since low income households often have greater levels of uncertainty of income, the findings are also consistent with the findings by Browne and Kim (1993) who interpreted life insurance as a way of reducing uncertainty in income.

Low income consumers are more likely to take-up insurance, but at the same time, they are also more likely to lapse their insurance which is a regrettable issue that should be addressed. Product developers should look at features that are effective in

preventing their customers from lapsing. Such features include “premium holidays” as discussed by Gruber and McKnight (2016).

Affluent consumers likely feel that there are more efficient ways to cover their risks. Additionally, with increasing financial sophistication, there may be a tendency for these customers to look at new products which may seem like better investments than the traditionally actuarially unfair insurance premiums (Barseghyan et al., 2013).

The findings may be somewhat surprising, given the literature on the subject. However South African trends in income distribution have been shown to be unique. This is due to the high poverty and inequality levels with South Africa having one of the highest Gini coefficients in the world (Leibbrandt, et al., 2010).

### **6.3 Discussion of Research Question 2: The influence of savings on the demand for insurance products**

In the following section the results of the various logistic regression models will be discussed in relation to the second research question. The second research question is concerned with how savings affects the propensity to take-up, lapse, or cancel life insurance. This section begins by discussing the propensity for insurance consumption, followed by the propensity to lapse, and finally, discuss the propensity to cancel.

#### **6.3.1 Discussion of Savings in the Take-up Model**

Two predictor variables related to savings were shown to be significant in the take-up logistic regression model. These are average savings balance and average cheque balance.

The model showed that increased savings had a small but significant negative effect on the propensity to purchase insurance. This is in contrast with the findings by Headen and Lee (1974), but is in alignment with the findings of Peter, (2017). By referring to Figure 7, one may see that there is nearly a 14% decrease in take-up among those customers who had savings, as compared to the customers who had no savings. Referring to this result and the regression model it may be concluded that there is therefore, a negative correlation between savings and insurance demand.

The average cheque balance predictor variable has already been used in the discussion of income above, but is also relevant for the discussion on savings behaviours. Since there is a negative effect of average cheque balance predictor variable in the logistic regression model, the conclusion is that there is a decrease in the likelihood for individuals to take up insurance with increasing savings.

### **6.3.2 Discussion of Savings in the Lapse Model**

Two predictor variables related to savings were shown to be significant in the lapse logistic regression model. These are average savings balance and average cheque balance. Increased savings is shown to have a negative effect on the likelihood of lapsing as may be expected. This is probably in part because the likelihood of access to life insurance savings decreases with age, while the relative accumulation of savings typically increases (Eling & Kiesenbauer, 2014). What is also highly likely is that consumers with savings are in possession with emergency funds to cope with other financial stresses. Customers with savings are more able to protect themselves from the risk of lapsing.

Average cheque balance, as used above, is also relevant for the discussion on savings as the two are very similar. The logistic regression shows that there is a small but significant negative effect of cheque balance on the likelihood of lapsing. This is consistent with the findings of savings and is intuitively sensible. The same reasons for a decrease in lapse rates due to increased savings are valid for the cheque account predictor variable. Intuitively, there was the sense that the cheque account balance would have a stronger effect in preventing lapses, since many individuals use the funds in the account to pay for their insurance premiums. When there are low funds in the account, there will be a greater chance of lapsing. While if there are sufficient funds in the account, there is a lower chance of the premium not being paid. Clearly there is some effect of the risk aversion behaviour of individuals to ensure that there are funds in the account. This may be the case of individuals with the financial capacity to store funds in their cheque accounts only, and as such, is not always a clear prediction of risk aversion as it is also due to financial means.

The results largely agree with the work by Hou, et al., (2015), who reported that low-income and low-wealth individuals have an increased likelihood to lapse their insurance policies.

In affirmation of the literature, the results show an increase in lapses among the absolute poorest customers (see Figure 17). There was, however, a small increase in those customers with very large savings. However this was among a relatively small sample. This is perhaps due to these customers keeping most of their funds in savings products instead of the account from which insurance premiums are collected.

### **6.3.3 Discussion of Savings in the Cancel Model**

According to the results, there was no significant effect of any savings related predictor variables retained in the cancellation logistic regression model. This is may be due to



the data being insufficient in expounding any results since there are so few cancellations even with the large sample size.

#### **6.3.4 Summary of Savings**

The findings point to the differences between behaviours in individuals: When individuals have savings, they are likely empowered to cope with risky situations without the need of separate insurance. Savings is likely seen as substitute to insurance, rather than a complimentary as argued by Chen et al., (2006) and Huang and Milevsky, (2008). In this way the findings may contradict their assertions. However, the results may indicate support for the findings by Peter, (2017) who reports that consumers with savings and who are prudent, may not be inclined to take-up insurance.

The findings also show that a lack of savings exposes individuals to a risk of lapsing. Lapses may have a further undesirable effect on an individual's wealth, since the surrender value of the insurance policy may be low compared to the value of the in-force policy (Russell et al., (2013).

The sample used in the analysis showed that there were in fact very few individuals with savings. This may have implications in a macro-economic context and needs to be researched further, since Aron and Muellnauer, (2000) reported that savings and investments have an effect on economic growth. Outreville, (2015) further reported that insurance is also essential for economic growth. As discussed, the bulk of the products sold were relatively low cost insurance products (in comparison to competitor offerings).

### **6.4 Discussion of Research Question 3: The influence of debt on the demand for insurance products**

In the following section, the results of the various logistic regression models will be discussed in relation to the third research question. The third research question is concerned with how debt affects the propensity to take-up, lapse, or cancel life insurance. This section starts by discussing the propensity for insurance consumption, followed by the propensity to lapse and finally discuss the propensity to cancel.

#### **6.4.1 Discussion of Debt in the Take-up Model**

The results of the logistic regression model for predicting take-up showed that there are two debt related individual-level financial behaviours that are significantly correlated to take-up. If we consider the fact that low-income consumers typically do not have

access to debt products (as they do not meet the credit-worthiness requirements by the bank), then the findings are also consistent with Hou, et al., (2015) who reported that low-income and low-wealth individuals are more likely to take-up insurance policies.

The regression model also found a significant positive relationship between the ratio of debt to account balance predictor variable and the likelihood of purchasing insurance. This was the strongest predictor variable related to debt in the model. With increasing ratios of debt to account balance, individuals are more financially vulnerable and the result is therefore aligned with the findings of Van Winssen et al., (2016) who found that customers who are overly debt averse, are also more risk averse and demand full coverage offered by their insurance policies.

The total amount of debt was found to have a very small positive but significant effect on the likelihood of purchasing insurance. This is in contrast to Lin and Grace (2007) who reported that it is ambiguous whether there is a relationship between life insurance and debt.

It is interesting to note that the model did not find the relationship between the ratio of debt to income and take-up to be significant.

#### **6.4.2 Discussion of Debt in the Lapse Model**

The results of the logistic regression model for predicting lapses showed that there are a number of debt related individual-level financial behaviours that are significantly correlated to lapses. The findings are consistent with the findings of Fier and Liebenberg, (2013) who found that among individuals with larger levels of debt, and who underwent negative income shocks, the likelihood of willingly lapsing increased.

The logistic regression found a small but significant positive relationship in the likelihood of lapsing with increasing debt. This finding is consistent with the reports by Frees and Sun (2010) that the amount of debt owed has a positive effect on an individual to have whole life insurance. The results are however contrasted with the findings of Frees and Sun (2010) that increasing debt increased the likelihood of relatively inexpensive term-life insurance. This is however intuitively sensible since they found a negative relationship between term and whole-life insurance. It is therefore important to note that the specific types of life insurance products contained in the sample may influence the results.

The regression model also found a significant positive relationship between the ratio of debt to account balance predictor variable and the likelihood of lapsing. With increasing ratios of debt to account balance, individuals are more financially vulnerable and the

result is therefore aligned with the findings of Lin and Grace (2007) who found an increasing likelihood of lapsing insurance with increasing financial vulnerability.

### **6.4.3 Discussion of Debt in the Cancel Model**

The results of the logistic regression model for predicting cancellations showed that there was no debt related individual-level financial behaviours that were significantly correlated to cancellations. As discussed, this may be due to the data being insufficient in expounding any results since there are so few cancellations even with the relatively large sample size.

### **6.4.4 Summary of Debt**

High levels of debt may be an indication of financial sophistication. Customers with high levels of debt are usually not in part of the low-income group. This is because the debt products include items such as home loans, personal loans, and overdraft facilities. Debt may therefore be interpreted as a measure of financial sophistication and ultimately wealth. Bernheim, et. al., (2003) found no significant correlations between financial vulnerability and life insurance, but this was later contrasted by Lin and Grace (2007) who reported a positive relationship between the amount of life insurance consumed and levels of financial vulnerability.

The ratio of debt to account balance is a measure of how much debt a customer has in relation to their account balance. High ratios are indicative that a consumer may have relatively low account balances and may exhibit financial vulnerability. The logistic regression model found a positive relationship between the ratio of debt to account balance and take-up, indicating that individuals who are financial vulnerable may seek protection from their risks.

## **6.5 Concerns**

The type of insurance products offered by the insurance company probably had an influence on the results of the analysis. The products from the insurer under study are known to be quite low cost compared to their competitors. This is important to consider when interpreting the results as it is a business strategy to be a low-cost competitor. The insurance products are therefore seemingly attractive to lower income consumers and may be inappropriate for more affluent consumers.

The data on savings was dominated by customer profiles that did not have any form of savings. This was particularly striking and goes some way in demonstrating the problem discussed by Aron and Muellbauer in 2000 and again in 2013. There were

very few individuals in the sample with savings. Approximately 99% of the sample did not have any savings products with the bank. This is a clearly a pressing matter since saving is an important means of self protection (Peter,2017).

## 7 Conclusion

### 7.1 Introduction

The following chapter will summarise the most significant findings from Chapter 6. It will also provide some implications for management especially with regard to insurance product design. The concluding chapter will also elaborate on the limitations of this research and provide some recommendations for future research.

This research analysed the relationships between individual-level financial behaviours of consumers on the take-up, lapse, and cancellations behaviours in a sample of South African life insurance market. The individual-level financial behaviours, such as levels of income, savings and debt, are considered. Three binary logistic regression models were used to analyse the individual customer profiles with regard to propensity to take-up, lapse and cancel insurance.

### 7.2 Principal Findings

The results of the analysis show that with increasing income, there is a negative relationship to purchase insurance. This means that there may be an increase in risk aversion (interpreted by insurance consumption behaviour) with decreasing income. All the significant predictor variables related to income were consistent in the direction of influence. The amount of monthly spend (credit turnover), average account balance, ratio of credit turnover to income, and ratio of account balance to income, were all consistent with one another and significant in predicting the purchasing of life insurance behaviour. None of the regression models identified average income itself as a significant predictor variable. This was interesting since there is a clear relationship between take-up of insurance and the other predictor variable which are used as complementary measures of income. The conclusion is therefore that the predictor variables used as a proxy for understanding wealth were significant in their relationship to take-up while an absolute measure of income was not significant.

This is downward trend in insurance consumption with increasing income is consistent with at least some of the literature (Lin & Grace, 2007). They reported that individuals with decreasing risk aversion will consume lower amounts of insurance at higher income levels.

There are, however, contradictory arguments that exist in the literature. For instance: Lange et al., (2017) reported that increasing income is an important driver for the demand for insurance. Frees and Sun, (2010) reported that low income individuals may

not always be able to afford life insurance and therefore exhibit lower levels of insurance consumption. Similarly, Mulholland, et al., (2016) reported that with increasing financial sophistication, there is an increase in insurance demand.

This proves that context is extremely important when considering the results. The sample data analysed in this research was from an insurer who is associated with having low cost insurance products, suitable for the majority of the South African population. It makes intuitive sense therefore to observe that insurance consumption is highest among the lower income customers.

Since low income households often have greater levels of uncertainty of income, the findings are also consistent with Browne and Kim (1993) who interpreted life insurance as a way of reducing uncertainty in income, and later by Akotey et al., (2011) who showed that low income individuals demand insurance cover, but cannot always afford it. This is important for the context of emerging economies.

According to the results presented in chapter 6, low income consumers are more likely to take-up life insurance. At the same time, however, these same customers are also more likely to lapse their insurance. This finding is consistent with Hou et al., (2015). This is highly a regrettable issue that should urgently be addressed. Features such as “premium holidays” as discussed by Gruber and McKnight (2016) may be effective in protecting low income consumers against lapsing.

Affluent consumers likely feel that there are more efficient ways to cover their risks. Additionally, with increasing financial sophistication, there may be a tendency for these customers to look at new products which would seem like better investments than the traditionally actuarially unfair insurance premiums.

Individuals who are granted access by the bank to financially sophisticated products (overdraft facilities and home loans) are typically credit worthy and more affluent. In the results, there may be a decrease in the likelihood of insurance consumption among these customers, which is in contradiction with the literature on financial sophistication and income (Mulholland & Finke, 2014; Lin et al., 2017). It is important to address the issue of context now. In the academic literature, many available life insurance products are suitable to individuals with high financial literacy (Fernandes et al., 2014). These individuals are concerned with estate planning and bequests (Lin et al., 2017). However, in the South African context, some insurers actively aim to offer simple to understand and inexpensive insurance products, which are suitable for low income individuals. This paper therefore concludes that context is an important factor to

consider when reporting on trends and relationships concerning insurance. These results therefore may help to raise more questions for future research.

The findings of the relationship between income and insurance consumption may be surprising given that some of the literature is contradictory to the analysis, but this shows that the South African context is clearly unique. This is substantiated by the trends in South African income distribution, which have been shown to be unique due to high poverty and inequality levels (Leibbrandt, et al., 2010).

With regard to savings, the results indicate that there are significant differences between behaviours in individuals. The findings show that there is a relationship between individuals who do not have savings to take up life insurance. When individuals have savings, they are likely empowered to cope with income shocks and other risks without the need of separate life insurance (Nolte & Schneider, 2017). Savings is likely seen as substitute to insurance rather than a complimentary as argued by Chen et al., (2006) and Huang and Milevsky, (2008). In this way, the findings may contradict their assertions. However, the results indicate support for the findings by Peter, (2017) who reports that consumers with savings and are prudent may not be inclined to take-up insurance.

The findings also show that a lack of savings may expose individuals to an increased risk of lapsing. Lapses may have a further undesirable effect on an individual's wealth, since the surrender value of the insurance policy may be low compared to the value of the in-force policy (Russell et al., 2013).

It is important to note that the sample used in the analysis, showed that there was in fact very few individuals with savings. This may have implications in a macro-economic context and needs to be researched further. Aron and Muellnauer, (2000) as well Outreville, (2015) reported that savings and investments have an essential positive effect on economic growth.

In the context of this research, high levels of debt may be an indication of financial sophistication. Customers with high levels of debt are usually not part of the low-income group. This is because the debt products include items such as home loans, personal loans and overdraft facilities which banks typically only grant to more affluent customers based on their risk profiles. Debt may therefore be interpreted as a measure of financial sophistication and ultimately wealth. Bernheim, et al., (2003) found no significant correlations between financial vulnerability and life insurance, but this was later contrasted by Lin and Grace (2007) who reported a positive relationship between the amount of life insurance consumed and levels of financial vulnerability.

The ratio of how much debt a customer has in relation to their account balance, was another significant predictor variable in the take-up and lapse model. High ratios are indicative that an individual may have relatively low account balances and may exhibit financial vulnerability. The logistic regression model found a positive relationship between the ratio of debt to account balance and take-up, indicating that individuals who are financial vulnerable may seek protection from their risks. This is risk aversion behaviour and is consistent with Outreville (2014).

The logistic regression model that was developed to predict take-up of insurance was 63.1% accurate in discriminating the correct response. The lapse model was slightly less accurate with a 62.5% accuracy of discriminating the correct response while the cancellation model which only had one significant predictor variable was 54.9% accurate at predicting a response.



**Table 18: Summary of the direction of the relationships between individual level financial behaviours of individuals regarding insurance demand.**

Increasing predictor variable	Increasing independent variable		
	Take-up	Lapse	Cancellation
Average credit turnover	–	–	–
Average cheque balance	–	–	NS
Average income	NS	NS	NS
Average savings	–	–	NS
Total debt	+	+	NS
Ratio of savings to income	NS	NS	NS
Ratio of credit turnover to income	+	+	NS
Ratio of debt to income	+	+	NS
Ratio of average balance to income	+	–	NS
Ratio of debt to average balance	+	+	NS

*NS denotes not significant*

### **7.3 Implications for Management**

By definition, insurance is intended to act a financial intermediary aimed at providing individuals with the financial means to weather undesirable events (Brockett et al. 2005). Life insurance is meant to be a facilitator for financial capacity or income replacement, according to Outreville (1996). McNeil, Frey, and Embrechts (2015) showed that it is possible to calculate the capital charges for different kinds of risks. Since risk aversion is the principal motivation for the existence of insurance products (Cohen & Einav, 2007) it is important to understand how certain individual level financial behaviours are linked to the propensity of insurance consumption.

We examine the hypothesis of connections between the way in which individuals treat their finances, specifically, income, savings, and debt, and their insurance choices which provides an indication of their risk preferences.

Low income consumers are also more likely to lapse their insurance which is a regrettable issue that should be addressed. Product developers should look at features that are effective in preventing their customers from lapsing. Such features include “premium holidays” as discussed by Gruber and McKnight (2016).

It is clear that due to the uncertainty associated with claims experiences in insurance products that businesses often avoid risk by offering less value to the customer through higher pricing. This is known as actuarially imbalanced rates Barseghyan et al., (2013). This paper aims to promote understanding of the market and thereby facilitates the ability of firms to offer better value to their customers. This also provides a competitive advantage to firms as well as improved products to consumers.

Chang et al., (2014) reported that insurance activity promotes economic growth. Life insurance policies are effective tools to ensure that mobilised capital is made accessible for financial agents, owing to intermediation activities of the financial markets. This makes insurance a somewhat unconventional but real resource for internal fund mobilisation in an economy such as South Africa. This in turn has the ability to reduce dependence on other sources (Alhassan & Biepke, 2015).

### **7.4 Limitations of the Research**

Only one bank was approached for data. This research therefore did not accumulate data from any other bank. This means that the interpretation of the results should be done with caution, since many customers will have bank accounts with multiple banks.

Their financial profile with the bank may not be an accurate reflection of their true financial behaviour.

The results of the research are of course a generalisation, based on aggregate behaviours. This may be useful to managers in their formulation of business strategy but does not necessarily reflect the truth for specific individual.

The results are specific to the data used in the study. As such, the findings are most applicable to the firm whose data was researched. Other organisations may use the findings to improve their products but since their specific products, customers and operational structure is significantly different from the company whose data was used in the study, the findings may have limited use. It would be better for other organisations to employ a similar study, by following the same methodology.

As mentioned above, the evidence presented by this research report is based on an empirical examination on the relationship between the consumption of life insurance and individual-level financial factors and thus any evidence gained from this study will only point to associations between the variables, and not to the real nature of causal links between variables.

## **7.5 Suggestions for Future Research**

In the following section a number of suggestions are proposed that may assist with future research.

### ***7.5.1 Incorporating Transactional Data***

Apart from the various calls for future research already mentioned, this study further recommends for future research may be expanded to find linkages between transactional banking data and the consumption of insurance products.

### ***7.5.2 Extend Observation Period***

A dataset which is valid over a longer period will enable the extension of the observation period (i.e. tracking valid research subjects or customer profiles, over a longer period). It is recommended that the extension is for another 12 months making the total observation period 24 months.

### ***7.5.3 More Variables***

Future research may look at more transactional variables and combinations of variables can be tested and included in the model (for example, the proportional spend on insurance and proportional number of insurance purchases).

#### **7.5.4 *Extend Outcome Period***

As in 8.1 a dataset which is valid over a longer period will enable the outcome period to be extended (i.e. tracking insurance consumption, lapse or cancellation over a longer period). It is recommended that the outcome period is extended for another six months, making the entire observation period 12 months.

#### **7.5.5 *Post-model Adjustments***

To enhance the usability of the model in a business context, it may be useful to develop and apply some post-model adjustments, for example excluding certain customers based on their demographic profile (as they might be deemed as likely to lapse, cancel or purchase insurance in the long-term while the model might predict them to be unlikely to lapse, cancel or purchase insurance in the short-term).

#### **7.5.6 *Brand Switching***

The cancel logistic regression model was relatively insufficient at accurately predicting cancellations. This research however identified that a study on brand switching may allow for a better understanding of the nature of cancellations. Brand switching is important to understand since brand loyalty is a large and important subject, worthy of its own focus. A more rigorous analysis may be used to identify switching to other insurers. In fact, future research may use a cancellation event as a trigger to analyse the behaviour of a specific client and monitor any further cancellations of other insurance products as well as a change in premium paid to the new insurance company. This may provide further empirical evidence to substantiate the work done by Brockett et al., (2005) and Guillén et al., (2012).

## 8 References

- Alhassan, A. L., & Biekpe, N. (2016). Determinants of life insurance consumption in Africa. *Research in International Business and Finance*, 37, 17–27.  
<http://doi.org/10.1016/j.ribaf.2015.10.016>
- Agarwal, B. S., & Qian, W. (2014). Consumption and Debt Response to Unanticipated Income Shocks: Evidence from a Natural Experiment in Singapore. *American Economic Review*, 104(12), 4205–4230.
- Akotey, O. J., Osei, K. A., & Gemegah, A. (2011). The demand for micro insurance in Ghana. *The Journal of Risk Finance*, 12(3), 182–194.  
<http://doi.org/10.1108/15265941111136932>
- Anderson, D., & Nevin, J. (1975). Determinants of Young Marrieds' Life Insurance Purchasing Behavior: An Empirical Investigation. *The Journal of Risk and Insurance*, 42(3), 375–387. <http://www.jstor.org/stable/251694>
- Aron, J., & Muellbauer, J. (2000). Personal and corporate saving in South Africa. *The World Bank Economic Review*, 14(3), 509–544.  
<http://doi.org/10.1093/wber/14.3.509>
- Aron, J., & Muellbauer, J. (2013). Wealth, credit conditions, and consumption: Evidence from South Africa. *Review of Income and Wealth*, 59(SUPPL1).  
<http://doi.org/10.1111/roiw.12033>
- Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review*. <http://doi.org/10.1215/03616878-26-5-851>
- Arrow, K. J. (1965). The theory of risk aversion. In *Aspects of the Theory of Risk Bearing* (pp. 90–109).
- Athreya, K. B. (2008). Default, insurance, and debt over the life-cycle. *Journal of Monetary Economics*, 55(4), 752–774.  
<http://doi.org/10.1016/j.jmoneco.2008.05.006>
- Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The quarterly journal of economics*, 116(1), 261–292.
- Barney, J. B., & Arikan, A. M. (2001). *The resource-based view: origins and implications. The Blackwell handbook of strategic management.*  
<http://doi.org/10.1177/014920639101700107>

- Barseghyan, L., Molinari, F., O'Donoghue, T., & Teitelbaum, J. C. (2013). The nature of risk preferences: Evidence from insurance choices. *American Economic Review*, 103(6), 2499–2529. <http://doi.org/10.1257/aer.103.6.2499>
- Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 112(2), 537-579.
- Barsotti, F., Milhaud, X., & Salhi, Y. (2016). Lapse risk in life insurance: Correlation and contagion effects among policyholders' behaviors. *Insurance: Mathematics and Economics*, 71, 317–331. <https://doi.org/10.1016/j.insmatheco.2016.09.008>
- Bayer, P. J., Bernheim, B. D., & Scholz, J. K. (2009). The effects of financial education in the workplace: Evidence from a survey of employers. *Economic Inquiry*, 47(4), 605-624.
- Beck, T. (2003). Economic, Demographic, and Institutional Determinants of Life Insurance Consumption across Countries. *The World Bank Economic Review*, 17(1), 51–88. <http://doi.org/10.1093/wber/lhg011>
- Beer, M., Voelpel, S. C., Leibold, M., & Tekie, E. B. (2005). Strategic Management as Organizational Learning : Developing Fit and Alignment through a Disciplined Process. *Long Range Planning*, 38, 445–465. <https://doi.org/10.1016/j.lrp.2005.04.008>
- Bellante, D., & Green, C. A. (2004). Relative risk aversion among the elderly. *Review of Financial Economics*. <http://doi.org/10.1016/j.rfe.2003.09.010>
- Benartzi, S., & Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American economic review*, 79-98.
- Bernheim, D. B. (1991). How Strong Are Bequest Motives? Evidence Based on Estimates of the Demand for Life Insurance and Annuities. *Journal of Political Economy*, 99(5), 899. <https://doi.org/10.1086/261783>
- Bernheim, B. D., Carman, K. G., Gokhale, J., & Kotlikoff, L. J. (2001). *The mismatch between life insurance holdings and financial vulnerabilities: evidence from the Survey of Consumer Finances* (No. w8544). National Bureau of Economic Research.
- Bernheim, B. D., Forni, L., Gokhale, J., & Kotlikoff, L. J. (2003). The mismatch between life insurance holdings and financial vulnerabilities: Evidence from the health and

- retirement study. *American Economic Review*, 93(1), 354–365.  
<http://doi.org/10.1257/000282803321455340>
- Bernoulli, D. (1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica*, 22(1), 23–36.
- Box, G. E. P., & Tidwell, P. W. (1962). Transformation of the independent variables. *Technometrics*, 4, 531-550.
- Brockett, P. L., Cooper, W. W., Golden, L. L., Rousseau, J. J., & Wang, Y. (2005). Approach to Efficiency of Marketing Distribution Systems and Organizational Structure of Insurance Companies. *The Journal of Risk and Insurance*, 72(3), 393–412.
- Browne, M. J., & Kim, K., (1993). An International Analysis of Life Insurance Demand, *Journal of Risk and Insurance*, 60: 616-634.
- Bryman, A. & Bell, E. (2011). *Research Methodology: Business and Management Contexts*. Cape Town: Oxford University Press.
- Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in Numbers: How does data-driven decision-making affect firm performance? *ICIS 2011 Proceedings*, 18.  
<http://doi.org/10.2139/ssrn.1819486>
- Burnett, J. J., & Palmer, B. A. (1984). Examining life insurance ownership through demographic and psychographic characteristics. *Journal of risk and insurance*, 453-467.
- Campbell, R. A. (1980). The demand for life insurance: An application of the economics of uncertainty. *The Journal of Finance*, 35(5), 1155-1172.
- Chang, T., Lee, C.-C., & Chang, C.-H. (2014). Does insurance activity promote economic growth? Further evidence based on bootstrap panel Granger causality test. *The European Journal of Finance*, 20(12), 1187–1210.  
<http://doi.org/10.1080/1351847X.2012.757555>
- Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal Strat. Mgmt. J*, 37, 1597–1614.  
<https://doi.org/10.1002/smj.2407>
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data To Big Impact. *Mis Quarterly*, 36(4), 1165–1188.  
<http://doi.org/10.1145/2463676.2463712>

- Chiappori, P.-A., & Salanie, B. (2000). Testing for Asymmetric Information in Insurance Markets. *Journal of Political Economy*, 108(1), 56–78.  
<http://doi.org/10.1086/262111>
- Chiappori, P. a., & Salanié, B. (2003). Testing Contract Theory: A Survey of Some Recent Work. *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress.*, 115–149.  
<http://doi.org/10.1017/CBO9780511610240.005>
- Christiansen, M. C., & Niemeyer, A. (2014). Fundamental Definition of the Solvency Capital Requirement in Solvency II. *ASTIN Bulletin: The Journal of the IAA*, 44(3), 501–533. <http://doi.org/10.1017/asb.2014.10>
- Chui, A. C. W., & Kwok, C. C. Y. (2008). National culture and life insurance consumption. *Journal of International Business Studies*, 39(1), 88–101.  
<http://doi.org/10.1057/palgrave.jibs.8400316>
- Clark, R. L., Morrill, M. S., & Allen, S. G. (2012). The role of financial literacy in determining retirement plans. *Economic Inquiry*, 50(4), 851–866.  
<https://doi.org/10.1111/j.1465-7295.2011.00390.x>
- Cohen, A., & Einav, L. (2007). Estimating Risk Preference from Deductible Choice. *American Economic Review*, 97(1994), 745–788.  
<http://doi.org/10.1257/aer.97.3.745>
- Cohen, A., & Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77(1), 39–84. <http://doi.org/10.1111/j.1539-6975.2009.01337.x>
- Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, 70, 379–390.  
<http://doi.org/10.1016/j.jbusres.2016.08.011>
- Crainich, D., Eeckhoudt, L., & Trannoy, A. (2013). Even (mixed) risk lovers are prudent. *American Economic Review*. <http://doi.org/10.1257/aer.103.4.1529>
- Cummins, D., & Santomero, A. (Eds.). (2012). *Changes in the life insurance industry: Efficiency, technology and risk management* (Vol. 11). Springer Science & Business Media.
- Darooneh, A. H. (2007). Insurance pricing in small size markets. *Physica A: Statistical Mechanics and its Applications*, 380, 411-417.



- DeSarbo, W. S., Jedidi, K., & Sinha, I. (2001). Customer value analysis in a heterogeneous market. *Strategic Management Journal*, 22(9), 845–857.  
<http://doi.org/10.1002/smj.191>
- Deschamps, B. Y. R. (1973). Risk Aversion and Demand Functions. *Econometrica*, 41(3), 455–465. Retrieved from <http://www.jstor.org/stable/1913370>
- Detragiache, E., Tressel, T., & Gupta, P. (2008). Foreign Banks in Poor Countries : Theory. *The Journal of Finance*, LXIII(5), 2123–2161.
- Dionne, G. (2013). *Handbook of insurance: Second edition. Handbook of Insurance: Second Edition*. <http://doi.org/10.1007/978-1-4614-0155-1>
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G.G. (2011) Individual risk attitudes: measurement, determinants and behavioral consequences. *Journal of the European Economic Association* 9(3): 522–550.
- Donnelly, P., 2011. How Life Insurers Combat Anti-Selection, Monograph published by Munich RE Retrieved September 19, 2017, from, Available at <http://automation.munichre.com>.
- Duker, J. M. (1969). Expenditures for Life Insurance among Working-Wife Families. *The Journal of Risk and Insurance*, 36(5), 525–533.
- Ebert, S., & Wiesen, D. (2014). Joint measurement of risk aversion, prudence, and temperance. *Journal of Risk and Uncertainty*, 48(3), 231–252.  
<http://doi.org/10.1007/s11166-014-9193-0>
- Elango, B., & Jones, J. (2011). Drivers of insurance demand in emerging markets. *Journal of Service Science Research*, 3(2), 185–204. article.  
<http://doi.org/10.1007/s12927-011-0008-4>
- Eling, M., & Kiesenbauer, D. (2014). What Policy Features Determine Life Insurance Lapse? An Analysis of the German Market. *Journal of Risk and Insurance*, 81(2), 241–269. <http://doi.org/10.1111/j.1539-6975.2012.01504.x>
- Emamgholipour, S., Arab, M., & Mohajerzadeh, Z. (2017). Life insurance demand: Middle East and North Africa. *International Journal of Social Economics*, 44(4), 521–529. <https://doi.org/10.1108/IJSE-04-2015-0106>
- Enz, R. (2000). The S-Curve Relation Between Per-Capita Income and Insurance Penetration. *The Geneva Papers on Risk and Insurance*, 25(July), 396–406.  
[http://doi.org/10.1016/S0267-3649\(00\)88914-1](http://doi.org/10.1016/S0267-3649(00)88914-1)

- Ernst, H., Hoyer, W. D., & Rübstaamen, C. (2010). Sales, marketing, and research-and-development cooperation across new product development stages: implications for success. *Journal of Marketing*, 74(5), 80-92.
- Fei, W., & Schlesinger, H. (2008). Precautionary insurance demand with state-dependent background risk. *Journal of Risk and Insurance*, 75, 1–16.
- Fields, G. S., & Kanbur, R. (2007). Minimum wages and poverty with income-sharing. *Journal of Economic Inequality*, 5(2), 135–147. <http://doi.org/10.1007/s10888-006-9037-5>
- Ferber, R., & Lee, L. C. (1980). Acquisition and accumulation of life insurance in early married life. *Journal of Risk and Insurance*, 713-734.
- Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science*, 60(8), 1861–1883. <https://doi.org/10.1287/mnsc.2013.1849>
- Feyen, E., Lester, R. R., & Rocha, R. D. R. (2011). What drives the development of the insurance sector? An empirical analysis based on a panel of developed and developing countries.
- Fischer, S. (1973). A life cycle model of life insurance purchases. *International Economic Review*, 132-152.
- Franke, G., Schlesinger, H., & Stapleton, R. C. (2011). Risk taking with additive and multiplicative background risks. *Journal of Economic Theory*, 146(4), 1547–1568. <http://doi.org/10.1016/j.jet.2011.03.008>
- Frees, E. W., & Sun, Y. (2010). Household life insurance demand: A multivariate two-part model. *North American Actuarial Journal*, 14(3), 338–354. Retrieved from <https://www.scopus.com/inward/record.uri?eid=2-s2.0-79952018986&partnerID=40&md5=61f53914b1461f7774c1aa4f38ec3194>
- Friend, I., & Blume, M. E. (1975). The demand for risky assets. *The American Economic Review*, 65(5), 900-922.
- Gatzert, N., Hoermann, G., & Schmeiser, H. (2009). The impact of the secondary market on life insurers' surrender profits. *Journal of Risk and Insurance*, 76(4), 887–908. <http://doi.org/10.1111/j.1539-6975.2009.01320.x>
- Greene, M. R. (1963). Attitudes toward risk and a theory of insurance consumption. *The Journal of Insurance*, 30(2), 165-182

- Gruber, J., & McKnight, R. (2016). Controlling health care costs through limited network insurance plans: Evidence from Massachusetts state employees. *American Economic Journal: Economic Policy*, 8(2), 219–250.  
<http://doi.org/10.1257/pol.20140335>
- Grunert, J., & Weber, M. (2009). Recovery rates of commercial lending: Empirical evidence for German companies. *Journal of Banking & Finance*, 33(3), 505–513.  
<http://doi.org/10.1016/j.jbankfin.2008.09.002>
- Guillén, M., Nielsen, J. P., Scheike, T. H., & Pérez-Marín, A. M. (2012). Time-varying effects in the analysis of customer loyalty: A case study in insurance. *Expert Systems with Applications*, 39(3), 3551–3558.  
<http://doi.org/10.1016/j.eswa.2011.09.045>
- Hakansson, N. H. (1969). Optimal investment and consumption strategies under risk, an uncertain lifetime, and insurance. *International Economic Review*, 10(3), 443-466.
- Halek, M., & Eisenhauer, J. G. (2001). Demography of Risk Aversion. *Source: The Journal of Risk and Insurance* *The Journal of Risk and Insurance*, 68(1), 1–24.  
<http://doi.org/10.2307/2678130>
- Hammond, J. D., Houston, D. B., & Melander, E. R. (1967). Determinants of household life insurance premium expenditures: An empirical investigation. *Journal of Risk and Insurance*, 397-408.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. San Francisco, CA, Ltd: Morgan Kaufmann. <http://doi.org/10.1016/B978-0-12-381479-1.00001-0>
- Harrison, G. W., Lau, M. I., & Rutström, E. E. (2007). Estimating risk attitudes in Denmark: A field experiment. *Scandinavian Journal of Economics*, 109(2), 341–368. <http://doi.org/10.1111/j.1467-9442.2007.00496.x>
- Hau, A. (2000). Liquidity, Estate Liquidation, Charitable Motives, and Life Insurance Demand by Retired Singles. *The Journal of Risk and Insurance*, 67(1), 123–141.
- Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. *Science*, 344(6186), 862-867.
- Headen, R. S., & Lee, J. F. (1974). Life insurance demand and household portfolio behavior. *Journal of Risk and Insurance*, 685-698.

- Herath, T., & Rao, H. R. (2009). Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness. *Decision Support Systems*, 47(2), 154–165. <https://doi.org/10.1016/j.dss.2009.02.005>
- Hersch, J. (1996). Smoking, seat belts, and other risky consumer decisions: differences by gender and race. *Managerial and Decision Economics*, 17(5), 471–481. <http://doi.org/10.2307/2487821>
- Hilbe, J. M. (2011). Logistic regression. In *International Encyclopedia of Statistical Science* (pp. 755-758). Springer Berlin Heidelberg.
- Hill, T. (2000). Operations strategy. In Hill, T. (ed.), *Operations management: Strategic context and managerial analysis* (pp. 25-52). Palgrave: Hampshire.
- Hofstede, G. (1983). The cultural relativity of organizational practices and theories. *Journal of international business studies*, 14(2), 75-89.
- Hofstede, G. (1995). Insurance as a product of national values. *The Geneva Papers on Risk and Insurance Issues and Practice*, 20(4), 423-429.
- Hou, W., Sun, W., & Webb, A. (2015). Why do People Lapse their Long-Term Care Insurance? *Center for Retirement Research*. Boston College, (15).
- Janson, N. (2004). Managing financial risk. *Australian Journal of Pharmacy*, 85(1011), 440–441. <http://doi.org/10.1002/jcaf>
- Jehle, G. A., & Reny, P. J. (2011). *Advanced Microeconomic Theory* (Third Edition). Harlow: Pearson Education Limited.
- Jianakoplos, N. A., & Bernasek, A. (1998). Are Women More Risk Averse. *Economic Inquiry*, 36(4), 62–630. <http://doi.org/10.1111/j.1465-7295.1998.tb01740.x>
- Johne, A. (1993). Insurance product development: managing the changes. *International Journal of Bank Marketing*, 11(3), 5–14. <http://doi.org/10.1108/02652329310027693>
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–292.
- Kamakura, W. A., Kossar, B. S., & Wedel, M. (2004). Identifying Innovators for the Cross-Selling of New Products. *Management Science*, 50(8), 1120–1133. <http://doi.org/10.1287/mnsc.1040.0256>
- Karni, E., & Zilcha, I. (1985). Uncertain lifetime, risk aversion and life insurance. *Scandinavian Actuarial Journal*, 1985(2), 109-123.

- Kapteyn, A., & Teppa, F. (2011). Subjective measures of risk aversion, fixed costs, and portfolio choice. *Journal of Economic Psychology*, 32(4), 564–580.  
<http://doi.org/10.1016/j.joep.2011.04.002>
- Kaus, W. (2013). Conspicuous consumption and “race”: Evidence from South Africa. *Journal of Development Economics*, 100(1), 63–73.  
<http://doi.org/10.1016/j.jdeveco.2012.07.004>
- Kimani, J. K., Ettarh, R., Warren, C., & Bellows, B. (2014). Determinants of health insurance ownership among women in Kenya: evidence from the 2008-09 Kenya demographic and health survey. *International Journal for Equity in Health*, 13(1), 1–8. <https://doi.org/10.1186/1475-9276-13-27>
- Kimball, M. S. (1990). Precautionary Saving in the Small and in the Large. *Econometrica*, 58(1), 53–73.
- Kimball, M. S. (1992). Precautionary Motives for Holding Assets. *The New Palgrave Dictionary of Money and Finance*. <http://doi.org/10.3386/w3586>
- Kiesenbauer, D. (2012). Main Determinants of Lapse in the German Life Insurance Industry. *North American Actuarial Journal*, 16(1), 52–73.  
<http://doi.org/10.1080/10920277.2012.10590632>
- Kirigia, J. M., Sambo, L. G., Nganda, B., Mwabu, G. M., Chatora, R., & Mwase, T. (2005). Determinants of Health Insurance Ownership among South African Women. *BMC Health Services Research*, 5(1), 1472–6963.  
<https://doi.org/10.1186/1472-6963-5-17>
- Kjosevski, J. (2012). The Determinants of Life Insurance Demand in Central and Southeastern Europe. *International Journal of Economics and Finance*, 4(3), 237–247. <http://doi.org/10.5539/ijef.v4n3p237>
- Kogan, N., & Wallach, M. A. (1964). *Risk taking: A study in cognition and personality*. *The American Journal of Psychology* (Vol. 78). <http://doi.org/10.2307/1420596>
- Kothari, C. R. (2004). *Research Methodology: Methods & Techniques*. New Age International (P) Ltd.
- Kumar, R. K., & Srivastava, M. (2013). The antecedents of customer loyalty: An empirical investigation in life insurance context. *Journal of Competitiveness*, 5(2).
- Kunreuther, H. (2015). The Role of Insurance in Reducing Losses from Extreme Events: The Need for Public–Private Partnerships. *The Geneva Papers on Risk*

- and Insurance - Issues and Practice*, 40(4), 741–762.  
<http://doi.org/10.1057/gpp.2015.14>
- Kuo, W., Tsai, C., & Chen, W.-K. (2003). an Empirical Study on the Lapse Rate: the Cointegration Approach. *The Journal of Risk and Insurance*, 70(3), 489–508.  
<http://doi.org/10.1111/1539-6975.t01-1-00061>
- Laerd Statistics (2015). Binomial logistic regression using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from <https://statistics.laerd.com/>
- Lange, R., Schiller, J., & Steinorth, P. (2017). Demand and Selection Effects in Supplemental Health Insurance in Germany. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 42(1), 5–30. <http://doi.org/10.1057/s41288-016-0023-2>
- Lavalle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big Data, Analytics and the Path from Insights to Value. *MIT Sloan Management Review*, 52(2), 21–32.
- Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343(6176), 1203–1205.  
<https://doi.org/10.1126/science.1248506>
- Leibbrandt, M., Woolard, I., Finn, A., & Argent, J. (2010). *Trends in South African Income Distribution and Poverty Since the Fall of Apartheid. OECD Social, Employment and Migration Working Papers*.  
<https://doi.org/10.1787/5kmms0t7p1ms-en>
- Leland, H. E. (1968). Saving and uncertainty: The precautionary demand for saving. *The Quarterly Journal of Economics*, 82(3), 465-473.
- Levy, H. (2015). *Stochastic Dominance: Investment Decision Making under Uncertainty*. book, Springer International Publishing. Retrieved from [https://books.google.co.za/books?id=VL\\_YCgAAQBAJ](https://books.google.co.za/books?id=VL_YCgAAQBAJ)
- Lewis, F. D. (1989). Dependents and the Demand for Life Insurance. *American Economic Review*, 79, 452–467. <http://doi.org/10.1126/science.151.3712.867-a>
- Li, D., Moshirian, F., Nguyen, P., & Wee, T. (2007). The demand for life insurance in OECD countries. *Journal of Risk and Insurance*, 74(3), 637–652.  
<http://doi.org/10.1111/j.1539-6975.2007.00228.x>

- Li, J., & Szimayer, A. (2010). The effect of policyholders' rationality on unit-linked life insurance contracts with surrender guarantees. *Bonn Econ Discussion Papers*, 22, 63. <http://doi.org/10.1007/s10273-011-1262-2>
- Liebenberg, A. P., Carson, J. M., & Dumm, R. E. (2012). A Dynamic Analysis of the Demand for Life Insurance. *Journal of Risk and Insurance*, 79(3), 619–644. <http://doi.org/10.1111/j.1539-6975.2011.01454.x>
- Lin, Y., & Grace, M. F. (2007). Household life cycle protection: Life insurance holdings, financial vulnerability, and portfolio implications. *Journal of Risk and Insurance*, 74(1), 141–173. <http://doi.org/10.1111/j.1539-6975.2007.00205.x>
- Luciano, E., Outreville, J. F., & Rossi, M. (2016). Life Insurance Ownership by Italian Households: A Gender-Based Differences Analysis. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 41(3), 468-490.
- MacCrimmon, K. R. and D. A. Wehrung (1986) Taking Risks: The Management of Uncertainty, *The Free Press*.
- Makadok, R. (1999). Interfirm differences in scale economies and the evolution of market shares. *Strategic Management Journal*, 20, 935–952. [http://doi.org/10.1002/\(sici\)1097-0266\(199910\)20:10<935::aid-smj56>3.0.co;2-g](http://doi.org/10.1002/(sici)1097-0266(199910)20:10<935::aid-smj56>3.0.co;2-g)
- Mantis, G., & Farmer, R. N. (1968). Demand for Life Insurance. *Journal of Risk and Insurance*, 35(2), 247–256. <https://doi.org/10.2307/250834>
- Mayers, D., & Smith Jr, C. W. (1983). The Interdependence of Individual Portfolio Decisions and the Demand for Insurance. *The Journal of Political Economy*, 91(2), 304. <http://doi.org/10.1086/261145>
- McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. book, Princeton University Press. Retrieved from <https://books.google.co.za/books?id=SfJnBgAAQBAJ>
- Meyer, D., & Meyer, J. (2005). Relative Risk Aversion: What Do We Know? *Journal of Risk and Uncertainty*, 31(3), 243–262. <http://doi.org/10.1007/s11166-005-5102-x>
- Millo, G., & Carmeci, G. (2011). Non-life insurance consumption in Italy: a sub-regional panel data analysis. *Journal of Geographical Systems*, 13(3), 273-298.
- Montibeller, G., & von Winterfeldt, D. (2015). Cognitive and Motivational Biases in Decision and Risk Analysis. *Risk Analysis*, 35(7), 1230–1251. <https://doi.org/10.1111/risa.12360>

- Mossin, J. (1968). Aspects of Rational Insurance Purchasing. *Journal of Political Economy*. <http://doi.org/10.1086/259427>
- Mulholland, B., Finke, M., & Huston, S. (2016). *Understanding the Shift in Demand for Cash Value Life Insurance*. *Risk Management and Insurance Review* (Vol. 19). <http://doi.org/10.1111/rmir.12031>
- Nakata, H., & Sawada, Y. (2007). Demand for non-life insurance: A cross-country analysis. *CIRJE Working Paper F-461*.
- Newman, K. L., & Nollen, S. D. (1996). Culture and congruence: The fit between management practices and national culture. *Journal of international business studies*, 27(4), 753-779.
- Nolte, S., & Schneider, J. C. (2017). Don't lapse into temptation: a behavioral explanation for policy surrender. *Journal of Banking and Finance*, 79, 12–27. <http://doi.org/10.1016/j.jbankfin.2017.02.011>
- Oldenboom, N., & Abratt, R. (2000). Success and failure factors in developing new banking and insurance services in South Africa. *International Journal of Bank Marketing*, 18(5), 233–245. <https://doi.org/10.1108/02652320010356762>
- Outreville, J. F. (1990). Whole-life insurance lapse rates and the emergency fund hypothesis. *Insurance Mathematics and Economics*, 9(4), 249–255. [http://doi.org/10.1016/0167-6687\(90\)90002-U](http://doi.org/10.1016/0167-6687(90)90002-U)
- Outreville, J. F. (1996). Life Insurance in Developing Countries. *The Journal of Risk and Insurance*, 63(2), 263–278.
- Outreville, J. F. (2013). The relationship between insurance and economic development: 85 empirical papers for a review of the literature. *Risk Management and Insurance Review*, 16(1), 71–122. <http://doi.org/10.1111/j.1540-6296.2012.01219.x>
- Outreville, J. F. (2014). Risk Aversion, Risk Behavior, and Demand for Insurance: A Survey. *Journal of Insurance Issues*, 37(2), 158–186. <http://doi.org/10.2139/ssrn.2363877>
- Outreville, J. F. (2015). The relationship between relative risk aversion and the level of education: A survey and implications for the demand for life insurance. *Journal of Economic Surveys*, 29(1), 97–111. <http://doi.org/10.1111/joes.12050>



- Peppers, D., Rogers, M., & Kotler, P. (2016). *Managing Customer Experience and Relationships: A Strategic Framework*. book, Wiley. Retrieved from <https://books.google.co.za/books?id=8lVeDQAAQBAJ>
- Peter, R. (2017). Optimal self-protection in two periods: On the role of endogenous saving. *Journal of Economic Behavior and Organization*, 137, 19–36. <http://doi.org/10.1016/j.jebo.2017.01.017>
- Peteraf, M. A., & Barney, J. B. (2003). Unraveling the resource-based tangle. *Managerial and Decision Economics*. <http://doi.org/10.1002/mde.1126>
- Pratt J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, 32(1), 122–136. <http://doi.org/10.2307/1913738>
- Riley, W. B., & Chow, K. V. (1992). Asset Allocation and Individual Risk Aversion. *Financial Analysts Journal*, 48(6), 32–37. <http://doi.org/10.2469/faj.v48.n6.32>
- Roth, J. (2000). *Informal micro-finance schemes: The case of funeral insurance in South Africa*. International Labour Office, Geneva.
- Rothschild, M., & Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics*, 90(4), 629–649. <http://doi.org/10.2307/1885326>
- Russell, D. T., Fier, S. G., Carson, J. M., & Dumm, R. E. (2013). An Empirical Analysis of Life Insurance Policy Surrender Activity. *Journal of Insurance Issues*, 36(1), 35–57.
- Salunke, S., Weerawardena, J., & McColl-Kennedy, J. R. (2011). Towards a model of dynamic capabilities in innovation-based competitive strategy: Insights from project-oriented service firms. *Industrial Marketing Management*, 40(8), 1251–1263. <http://doi.org/10.1016/j.indmarman.2011.10.009>
- Sandmo, A. (1970). The effect of uncertainty on saving decisions. *The Review of Economic Studies*, 37(3), 353-360.
- Saunders, M., & Lewis, P. (2012). *Doing Research in Business and Management*. Pearson: Edinburgh Gate
- Saunders, M., Lewis, P. & Thornhill, A. (2009). *Research Methods for Business Students*, 5<sup>th</sup> Edition, Harlow: FT Prentice Hall.
- Schlesinger, H. (1981). The optimal level of deductibility in insurance contracts. *Journal of risk and insurance*, 465-481.

- Schlesinger, H., & von der Schulenburg, J. M. G. (1993). Consumer information and decisions to switch insurers. *Journal of Risk and Insurance*, 591-615.
- Schmitz, C., Lee, Y. C., & Lilien, G. L. (2014). Cross-selling performance in complex selling contexts: an examination of supervisory-and compensation-based controls. *Journal of Marketing*, 78(3), 1-19.
- Shanks, G., Sharma, R., Seddon, P., & Reynolds, P. (2010). The Impact of Strategy and Maturity on Business Analytics and Firm Performance: A Review and Research Agenda. *ACIS 2010 Proceedings*, 51.
- Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, 23(4), 433–441. <http://doi.org/10.1057/ejis.2014.17>
- Shi, X., Wang, H. J., & Xing, C. (2015). The role of life insurance in an emerging economy: Human capital protection, assets allocation and social interaction. *Journal of Banking and Finance*, 50, 19–33. <http://doi.org/10.1016/j.jbankfin.2014.08.028>
- Siegel, F. W., & Hoban, J. P. (1991). Measuring risk aversion: Allocation, leverage, and accumulation. *Journal of Financial Research*, 14(1), 27-35.
- Skiba, F., & Herstatt, C. (2009). Users as sources for radical service innovations: opportunities from collaboration with service lead users. *International Journal of Services Technology and Management*, 12(3), 317-337.
- Slack, N., & Lewis, M. (2015). *Operations Strategy*. Pearson Education Limited. <http://doi.org/10.1002/9781118785317.weom100160>
- Sliwinski, A., Michalski, T., & Roszkiewicz, M. (2013). Demand for Life Insurance—An Empirical Analysis in the Case of Poland. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 38(1), 62–87. <http://doi.org/10.1057/gpp.2012.21>
- Siegel, F. W., & Hoban, J. P. J. (1982). Relative Risk Aversion Revisited. *The Review of Economics and Statistics*.
- Simon, C. P., & Blume, L. (1994). *Mathematics for Economists*. W.W. Norton & Company, New York, London. <http://doi.org/10.1007/b97511>
- Slovic, P. (2016). *The Perception of Risk*. book, Taylor & Francis. Retrieved from <https://books.google.co.za/books?id=BLYyDQAAQBAJ>

- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263–286. <http://doi.org/10.1016/j.jbusres.2016.08.001>
- Smidts, A. (1997). The relationship between risk attitude and strength of preference: A test of intrinsic risk attitude. *Management Science*, 43(3), 357–370. Retrieved from [http://www.nd.edu/~nmark/gradinter\\_files/Mark-Wu-EJ.pdf](http://www.nd.edu/~nmark/gradinter_files/Mark-Wu-EJ.pdf)
- Somerville, R. A. (2004). Insurance, Consumption, and Saving: A Dynamic Analysis in Continuous Time. *The American Economic Review*, 94(4), 1130–1140. Retrieved from <http://www.jstor.org/stable/3592808>
- Stark, J. (2015). Product lifecycle management. In *Product Lifecycle Management (Volume 1)* (pp. 1-29). Springer International Publishing.
- Stiglitz, J. E. (1969). The effects of income, wealth, and capital gains taxation on risk-taking. *The Quarterly Journal of Economics*, 83(2), 263-283.
- Storbacka, K. (1997). Segmentation based on customer profitability—retrospective analysis of retail bank customer bases. *Journal of Marketing Management*, 13(5), 479-492.
- Szpiro, G. G. (1985). Optimal insurance coverage. *The Journal of Risk and Insurance*, 52(4), 704-710.
- Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th ed.). Essex, UK: Pearson.
- Tashakkori, A., & Teddlie, C. (2003). *Handbook of Mixed Methods in Social & Behavioral Research*. book, SAGE Publications. Retrieved from <https://books.google.co.za/books?id=F8BFOM8DCKoC>
- Thomas, L. C., Edelman, D. B., & Crook, J. N. (2017). *Credit Scoring and Its Applications*. Society for Industrial and Applied Mathematics. Retrieved from <https://books.google.co.za/books?id=GMWcWuBDJZUC>
- Tomczyk, P., Doligalski, T., & Zaborek, P. (2016). Does customer analysis affect firm performance? Quantitative evidence from the Polish insurance market. *Journal of Business Research*. <http://doi.org/http://dx.doi.org/10.1016/j.jbusres.2016.03.026>
- Urban, G.L., & Hauser, J.R. (1993) Design and marketing of new products. *Prentice-Hall, Englewood Cliff N.J.*

- Valdez, E. A., Vadiveloo, J., & Dias, U. (2014). Life insurance policy termination and survivorship. *Insurance: Mathematics and Economics*, 58, 138–149.  
<https://doi.org/10.1016/j.insmatheco.2014.06.011>
- Van Der Boor, P., Oliveira, P., & Veloso, F. (2014). Users as innovators in developing countries: The global sources of innovation and diffusion in mobile banking services. *Research Policy*, 43(9), 1594–1607.  
<https://doi.org/10.1016/j.respol.2014.05.003>
- Van Kleef, R. C., Van De Ven, W. P. M. M., & Van Vliet, R. C. J. A. (2006). A voluntary deductible in social health insurance with risk equalization: “Community-rated or risk-rated premium rebate?” *Journal of Risk and Insurance*, 73(3), 529–550.  
<http://doi.org/10.1111/j.1539-6975.2006.00187.x>
- Van Rooij, M. C. J., Lusardi, A., & Alessie, R. J. M. (2011). Financial literacy and retirement planning in the Netherlands. *Journal of Economic Psychology*, 32(4), 593–608. <https://doi.org/10.1016/j.joep.2011.02.004>
- Van Winssen, K. P. M., van Kleef, R. C., & van de Ven, W. P. M. M. (2016). Potential determinants of deductible uptake in health insurance: How to increase uptake in The Netherlands? *European Journal of Health Economics*, 17(9), 1059–1072.  
<http://doi.org/10.1007/s10198-015-0745-2>
- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356–365.  
<http://doi.org/10.1016/j.jbusres.2016.08.009>
- Ward, D., & Zurbrugg, R. (2000). Does insurance promote economic growth? Evidence from OECD countries. *Journal of Risk and Insurance*, 489-506.
- World Bank national accounts data, and OECD National Accounts data files, GDP growth (annual %). (n.d.). Retrieved September 26, 2017, from <https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=ZA>
- Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, 26(1), 97–107.  
<http://doi.org/10.1109/TKDE.2013.109>
- Yaari, M. E. (1964). On the consumer's lifetime allocation process. *International Economic Review*, 5(3), 304-317.

- Yaari, M. E. (1965). Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *Review of Economic Studies*, 32(2), 137.
- Yilma, Z., Van Kempen, L., & De Hoop, T. (2012). A perverse “net” effect? Health insurance and ex-ante moral hazard in Ghana. *Social Science and Medicine*, 75(1), 138–147. <https://doi.org/10.1016/j.socscimed.2012.02.035>
- Yousefi, A. (2016). The relationship between service quality with customer satisfaction and words of mouth. *International Business Management*, 10(5), 6542–6549.
- Zietz, E. N. (2003). An Examination of the Demand for Life Insurance. *Risk Management and Insurance Review*, 6(2), 159–191. <http://doi.org/10.1046/J.1098-1616.2003.030.x>
- Zimmer, A., Schade, C., & Gründl, H. (2009). Is default risk acceptable when purchasing insurance? Experimental evidence for different probability representations, reasons for default, and framings. *Journal of Economic Psychology*, 30(1), 11–23. <http://doi.org/10.1016/j.joep.2008.09.001>
- Zou, X., & Scholer, A. A. (2016). Motivational Affordance and Risk-Taking Across Decision Domains. *Personality and Social Psychology Bulletin*, 42(3), 275–289. <https://doi.org/10.1177/0146167215626706>

## 9 Appendix

## 10 Appendix

### 10.1 Appendix A

**Table 18: Consistency Matrix**

<b>Propositions/ Questions/ Hypotheses</b>	<b>Literature Review</b>	<b>Data Collection Tool</b>	<b>Analysis</b>
1. There is a significant correlation between demand for life insurance products and income.	Chen et al. (2012); Yaari, (1965); Kothari, 2004; Han et al. 2012; Li et al. 2007; Alhassan & Biekpe, 2016	Banking and Insurance legacy systems (secondary data)	Binary logistic regression analysis and modelling
2. There is a significant correlation between demand for life insurance products and savings as a percentage of income.	Chen et al. (2012); Yaari, (1965); Kothari, 2004; Han et al. 2012	Banking and Insurance legacy systems (secondary data)	Binary logistic regression analysis and modelling
3. There is a significant correlation between the demand for life insurance products and the amount of debt a customer has.	Chen et al. (2012); Yaari, (1965); Kothari, 2004; Han et al. 2012	Banking and Insurance legacy systems (secondary data)	Binary logistic regression analysis and modelling

## 10.2 Appendix B

In this section the SAS code used to analyse the data is provided for purposes of guiding future research efforts.

```
libname new '/grid/nfsshare/department';

*No insurance data;
data sales_new;
set new.sales_data;
    where sold_date between '01SEP2017'd and '30SEP2017'd;
run;

proc sql;
create table exclusion as
select a.*
from new.debt_all_ as a
where policyholder_acc_no not in (select distinct
policyholder_acc_no from sales_new);
quit;

*Credit turn over first 12 months;
proc sql;
create table snapshot_check as
select *
from (
    select policyholder_acc_no, sum(case when snapshot_date ne .
then 1 else 0 end ) as no_turnovers
from exclusion
group by policyholder_acc_no
) as a
where no_turnovers<=36

;
quit;

proc sql;
create table selection as
select *, case when MTH_END_BAL_CHEQ_SAV < 0 then
MTH_END_BAL_CHEQ_SAV else 0 end as debt_cheque
from exclusion as a
inner join snapshot_check as b on a.policyholder_acc_no =
b.policyholder_acc_no
;
quit;

data selection;
set selection;
if personal_loan_balance = . then personal_loan_balance=0;
if home_loan_balance = . then home_loan_balance=0;
```

```

if credit_card_balance= . then credit_card_balance=0;
total_debt = sum(personal_loan_balance*-1,credit_card_balance*-
1,debt_cheque);
total_debt_HL = sum(personal_loan_balance*-
1,credit_card_balance*-1,home_loan_balance*-1,debt_cheque);
run;

/*****
***** LAPSED *****/
*****/

data lapse_base;
set selection;
    where snapshot_date = '30SEP2016'd and
PolicyStatusDescription in ('In Force','Active');
run;

proc sql;
create table behavioural_info_L12M as
select * from (
select policyholder_acc_no
,sum(case when MTH_CR_TURNOVER_CHEQ_SAV>0 then 1 else 0 end ) as
no_turnovers
,round(avg(case when MTH_CR_TURNOVER_CHEQ_SAV > 20000 then 21000
else MTH_CR_TURNOVER_CHEQ_SAV end),1000) as avg_cheq_CR_TO
,round(avg(case when AVE_BAL_CHEQ_SAV > 20000 then 21000 when
AVE_BAL_CHEQ_SAV<=-20000 then -21000 else AVE_BAL_CHEQ_SAV
end),1000) as avg_cheq_BAL
,round(avg(income_estimate),10000) as avg_income_est
,case when round(avg(MTH_END_BAL_SNI),1000) = . then 0 else
round(avg(MTH_END_BAL_SNI),1000) end as avg_SNI_Bal
,round(avg(case when total_debt <=-100000 then -100000 else
total_debt end),2000) as total_debt
,round(avg(case when total_debt_HL <=-100000 then -100000 else
total_debt_HL end),2000) as total_debt_HL
from selection
where snapshot_date between '31OCT2015'd and '30SEP2016'd
and policyholder_acc_no in (select policyholder_acc_no from
lapse_base)
group by policyholder_acc_no)
as a
where no_turnovers between 10 and 12
;
quit;
/**/
/*proc univariate data=behavioural_info_L12M;var
avg_cheq_CR_TO;quit;*/
/*proc univariate data=behavioural_info_L12M;var
avg_cheq_BAL;quit;*/
/*proc univariate data=behavioural_info_L12M;var
avg_income_est;quit;*/

```



```

/*proc univariate data=behavioural_info_L12M;var
avg_SNI_Bal;quit;*/
/*proc univariate data=behavioural_info_L12M;var
total_debt;quit;*/

proc sql;
create table target_info_F6M_Lapse as
select policyholder_acc_no, max(target_lapse) as target_lapse
from (
select policyholder_acc_no, case when
upcase(PolicyStatusDescription) in ('LAPSED', 'AUTO LAPSE') then
1 else 0 end as Target_Lapse
from selection
where snapshot_date between '01OCT2016'd and '31MAR2017'd and
policyholder_acc_no in
(select policyholder_acc_no from behavioural_info_L12M)
) as a
group by policyholder_acc_no
;
quit;

proc sql;
create table lapse_all as
select
a.policyholder_acc_no
,a.cust_age
,b.*
,c.*
,round((avg_sni_bal/avg_income_est),0.05) as ratio_saving_income
,round(((c.total_debt*-1)/avg_income_est),0.05) as
ratio_debt_income
,round((avg_cheq_CR_TO/avg_income_est),0.05) as
ratio_CRTO_income
,round((avg_cheq_BAL/(avg_income_est/12)),0.05) as
ratio_avgbal_income
,round(((c.total_debt*-1)/avg_cheq_BAL),0.05) as
ratio_debt_avgbal
from lapse_base as a
left join target_info_F6M_Lapse as b on a.policyholder_acc_no =
b.policyholder_acc_no
left join behavioural_info_L12M as c on a.policyholder_acc_no =
c.policyholder_acc_no
where a.policyholder_acc_no in
(select policyholder_acc_no from behavioural_info_L12M)
;
quit;

proc logistic data=lapse_all;
model target_lapse(event = '1') = avg_cheq_CR_TO avg_cheq_BAL
avg_income_est avg_SNI_Bal total_debt ratio_saving_income
ratio_debt_income ratio_CRTO_income ratio_avgbal_income
ratio_debt_avgbal;

```

```
quit;

/*****
/***** CANCELLED *****/
/*****/

data cancel_base;
set selection;
    where snapshot_date = '30SEP2016'd and
PolicyStatusDescription in ('In Force', 'Active');
run;

proc sql;
create table behavioural_info_L12M_canc as
select * from (
select policyholder_acc_no
, sum(case when MTH_CR_TURNOVER_CHEQ_SAV > 0 then 1 else 0 end ) as
no_turnovers
, round(avg(case when MTH_CR_TURNOVER_CHEQ_SAV > 20000 then 21000
else MTH_CR_TURNOVER_CHEQ_SAV end), 1000) as avg_cheq_CR_TO
, round(avg(case when AVE_BAL_CHEQ_SAV > 20000 then 21000 when
AVE_BAL_CHEQ_SAV <= -20000 then -21000 else AVE_BAL_CHEQ_SAV
end), 1000) as avg_cheq_BAL
, round(avg(income_estimate), 10000) as avg_income_est
, case when round(avg(MTH_END_BAL_SNI), 1000) = . then 0 else
round(avg(MTH_END_BAL_SNI), 1000) end as avg_SNI_Bal
, round(avg(case when total_debt <= -100000 then -100000 else
total_debt end), 2000) as total_debt
, round(avg(case when total_debt_HL <= -100000 then -100000 else
total_debt_HL end), 2000) as total_debt_HL
from selection
where snapshot_date between '31OCT2015'd and '30SEP2016'd
and policyholder_acc_no in (select policyholder_acc_no from
cancel_base)
group by policyholder_acc_no)
as a
where no_turnovers between 6 and 12
;
quit;
/**/
/*proc univariate data=behavioural_info_L12M; var
avg_cheq_CR_TO; quit; */
/*proc univariate data=behavioural_info_L12M; var
avg_cheq_BAL; quit; */
/*proc univariate data=behavioural_info_L12M; var
avg_income_est; quit; */
/*proc univariate data=behavioural_info_L12M; var
avg_SNI_Bal; quit; */
/*proc univariate data=behavioural_info_L12M; var
total_debt; quit; */
```

```

proc sql;
create table target_info_F6M_cancel as
select policyholder_acc_no, max(target_cancel) as target_cancel
from (
select policyholder_acc_no, case when
upcase(PolicyStatusDescription) like '%CANCEL%' then 1 else 0
end as Target_cancel
from selection
where snapshot_date between '01OCT2016'd and '31MAR2017'd and
policyholder_acc_no in
(select policyholder_acc_no from behavioural_info_L12M_canc)
) as a
group by policyholder_acc_no
;
quit;

proc sql;
create table cancel_all as
select
a.policyholder_acc_no
,a.cust_age
,b.*
,c.*
,round((avg_sni_bal/avg_income_est),0.05) as ratio_saving_income
,round(((c.total_debt*-1)/avg_income_est),0.05) as
ratio_debt_income
,round((avg_cheq_CR_TO/avg_income_est),0.05) as
ratio_CRTO_income
,round((avg_cheq_BAL/(avg_income_est/12)),0.05) as
ratio_avgbal_income
,round(((c.total_debt*-1)/avg_cheq_BAL),0.05) as
ratio_debt_avgbal
from cancel_base as a
left join target_info_F6M_cancel as b on a.policyholder_acc_no =
b.policyholder_acc_no
left join behavioural_info_L12M_canc as c on
a.policyholder_acc_no = c.policyholder_acc_no
where a.policyholder_acc_no in
(select policyholder_acc_no from behavioural_info_L12M_canc)
;
quit;

proc logistic data=cancel_all;
model target_cancel(event ='1') = avg_cheq_CR_TO avg_cheq_BAL
avg_income_est avg_SNI_Bal total_debt ratio_saving_income
ratio_debt_income ratio_CRTO_income ratio_avgbal_income
ratio_debt_avgbal;
quit;

```

```

/*****
/***** TAKE UP *****/
/*****

data Takeup_base;
set selection;
    where snapshot_date = '30SEP2016'd and PolicyNumber='';
run;

proc sql;
create table behavioural_info_L12M_TU as
select * from (
select policyholder acc no
, sum(case when MTH_CR_TURNOVER_CHEQ_SAV>0 then 1 else 0 end ) as
no_turnovers
, round(avg(case when MTH_CR_TURNOVER_CHEQ_SAV > 20000 then 21000
else MTH_CR_TURNOVER_CHEQ_SAV end),1000) as avg_cheq_CR_TO
, round(avg(case when AVE_BAL_CHEQ_SAV > 20000 then 21000 when
AVE_BAL_CHEQ_SAV<=-20000 then -21000 else AVE_BAL_CHEQ_SAV
end),1000) as avg_cheq_BAL
, round(avg(income_estimate),10000) as avg_income_est
, case when round(avg(MTH_END_BAL_SNI),1000) = . then 0 else
round(avg(MTH_END_BAL_SNI),1000) end as avg_SNI_Bal
, round(avg(case when total_debt <=-100000 then -100000 else
total_debt end),2000) as total_debt
, round(avg(case when total_debt_HL <=-100000 then -100000 else
total_debt_HL end),2000) as total_debt_HL
from selection
where snapshot_date between '31OCT2015'd and '30SEP2016'd
and policyholder_acc_no in (select policyholder_acc_no from
Takeup_base)
group by policyholder_acc_no)
as a
where no_turnovers between 10 and 12
;
quit;
/**/
/*proc univariate data=behavioural_info_L12M;var
avg_cheq_CR_TO;quit;*/
/*proc univariate data=behavioural_info_L12M;var
avg_cheq_BAL;quit;*/
/*proc univariate data=behavioural_info_L12M;var
avg_income_est;quit;*/
/*proc univariate data=behavioural_info_L12M;var
avg_SNI_Bal;quit;*/
/*proc univariate data=behavioural_info_L12M;var
total_debt;quit;*/

proc sql;
create table target info F6M Takeup as

```

```

select policyholder_acc_no,PolicyStatusDescription,
max(target_Takeup) as target_Takeup
from (
select policyholder_acc no,PolicyStatusDescription,case when
sold_date ne . then 1 else 0 end as Target_Takeup
from selection
where snapshot_date between '01OCT2016'd and '31MAR2017'd and
policyholder_acc_no in
(select policyholder_acc_no from behavioural_info_L12M_TU)
) as a
group by policyholder_acc_no,PolicyStatusDescription
;
quit;

proc sql;
create table Takeup_all as
select
a.policyholder_acc_no
,a.cust_age
,b.*
,c.*
,round((avg_sni_bal/avg_income_est),0.05) as ratio_saving_income
,round(((c.total_debt*-1)/avg_income_est),0.05) as
ratio_debt_income
,round((avg_cheq_CR_TO/avg_income_est),0.05) as
ratio_CRTO_income
,round((avg_cheq_BAL/(avg_income_est/12)),0.05) as
ratio_avgbal_income
,round(((c.total_debt*-1)/avg_cheq_BAL),0.05) as
ratio_debt_avgbal
from Takeup_base as a
left join target_info_F6M_Takeup as b on a.policyholder_acc_no =
b.policyholder_acc_no
left join behavioural_info_L12M_TU as c on a.policyholder_acc_no
= c.policyholder_acc_no
where a.policyholder_acc_no in
(select policyholder_acc_no from behavioural_info_L12M_TU)
;
quit;

proc logistic data=Takeup_all;
model target_Takeup(event='1') = avg_cheq_CR_TO avg_cheq_BAL
avg_income_est avg_SNI_Bal total_debt ratio_saving_income
ratio_debt_income ratio_CRTO_income ratio_avgbal_income
ratio_debt_avgbal;
quit;

```

### 10.3 Appendix C

## The LOGISTIC Procedure

Model Information	
Data Set	WORK.TAKEUP_ALL
Response Variable	target_Takeup
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	263134
Number of Observations Used	171184

Response Profile		
Ordered Value	target_Takeup	Total Frequency
1	0	118846
2	1	52338

Probability modeled is target\_Takeup=1.

Note: 91950 observations were deleted due to missing values for the response or explanatory variables.

## Stepwise Selection Procedure

Step 0. Intercept entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 210779.07

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
7508.0073	9	<.0001

Step 1. Effect avg\_cheq\_CR\_TO entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	210781.07	204182.25
SC	210791.12	204202.35
-2 Log L	210779.07	204178.25

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6600.8131	1	<.0001
Score	6529.7515	1	<.0001
Wald	6349.2166	1	<.0001

### The LOGISTIC Procedure

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
1268.1308	8	<.0001

Note: No effects for the model in Step 1 are removed.

#### Step 2. Effect total\_debt entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	210781.07	203007.92
SC	210791.12	203038.07
-2 Log L	210779.07	203001.92

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7777.1488	2	<.0001
Score	7367.7877	2	<.0001
Wald	6974.3483	2	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
197.4287	7	<.0001

Note: No effects for the model in Step 2 are removed.

#### Step 3. Effect ratio\_debt\_avgbal entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	210781.07	202922.56
SC	210791.12	202962.76
-2 Log L	210779.07	202914.56

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7864.5036	3	<.0001
Score	7406.0250	3	<.0001
Wald	6967.0844	3	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
110.6816	6	<.0001

### The LOGISTIC Procedure

#### Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

#### Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	210781.07	202870.04
SC	210791.12	202920.29
-2 Log L	210779.07	202860.04

#### Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7919.0292	4	<.0001
Score	7447.1269	4	<.0001
Wald	7003.4845	4	<.0001

#### Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
54.6816	5	<.0001

Note: No effects for the model in Step 4 are removed.

#### Step 5. Effect avg\_cheq\_BAL entered:

#### Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

#### Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	210781.07	202835.56
SC	210791.12	202895.86
-2 Log L	210779.07	202823.56

#### Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7955.5082	5	<.0001
Score	7469.3943	5	<.0001
Wald	7016.4911	5	<.0001

#### Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
18.5517	4	0.0010

Note: No effects for the model in Step 5 are removed.

#### Step 6. Effect ratio\_avgbal\_income entered:

#### Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.



### The LOGISTIC Procedure

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	210781.07	202825.64
SC	210791.12	202895.99
-2 Log L	210779.07	202811.64

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7967.4246	6	<.0001
Score	7487.4630	6	<.0001
Wald	7032.4369	6	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
6.7917	3	0.0788

Note: No effects for the model in Step 6 are removed.

#### Step 7. Effect avg\_SNI\_Bal entered:

Model Convergence Status	
Convergence criterion (GCONV=1E-8) satisfied.	

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	210781.07	202820.87
SC	210791.12	202901.28
-2 Log L	210779.07	202804.87

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7974.1918	7	<.0001
Score	7491.3741	7	<.0001
Wald	7034.8864	7	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
1.1827	2	0.5536

Note: No effects for the model in Step 7 are removed.

Note: No (additional) effects met the 0.05 significance level for entry into the model.

Summary of Stepwise Selection							
Step	Effect		DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
	Entered	Removed					
1	avg_cheq_CR_TO		1	1	6529.7515		<.0001
2	total_debt		1	2	1098.1890		<.0001
3	ratio_debt_avgbal		1	3	88.2476		<.0001

## The LOGISTIC Procedure

Summary of Stepwise Selection							
Step	Effect		DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
	Entered	Removed					
4	ratio_CRTO_income		1	4	56.1201		<.0001
5	avg_cheq_BAL		1	5	36.3747		<.0001
6	ratio_avgbal_income		1	6	11.9192		0.0006
7	avg_SNI_Bal		1	7	5.2274		0.0222

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.1270	0.0119	113.3964	<.0001
avg_cheq_CR_TO	1	-0.00005	1.067E-6	2226.8364	<.0001
avg_cheq_BAL	1	-8.05E-6	1.205E-6	44.6461	<.0001
avg_SNI_Bal	1	-4.29E-7	1.894E-7	5.1300	0.0235
total_debt	1	0.000011	3.524E-7	973.2365	<.0001
ratio_CRTO_income	1	0.5054	0.1024	24.3643	<.0001
ratio_avgbal_income	1	0.0468	0.0138	11.4627	0.0007
ratio_debt_avgbal	1	0.00717	0.000692	107.6048	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
		avg_cheq_CR_TO	1.000
avg_cheq_BAL	1.000	1.000	1.000
avg_SNI_Bal	1.000	1.000	1.000
total_debt	1.000	1.000	1.000
ratio_CRTO_income	1.658	1.356	2.026
ratio_avgbal_income	1.048	1.020	1.077
ratio_debt_avgbal	1.007	1.006	1.009

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	63.0	Somers' D	0.262
Percent Discordant	36.8	Gamma	0.263
Percent Tied	0.2	Tau-a	0.111
Pairs	6220161948	c	0.631

### The LOGISTIC Procedure

Model Information	
Data Set	WORK.LAPSE_ALL
Response Variable	target_lapse
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	115671
Number of Observations Used	80043

Response Profile		
Ordered Value	target_lapse	Total Frequency
1	0	67159
2	1	12884

Probability modeled is target\_lapse=1.

Note: 35628 observations were deleted due to missing values for the response or explanatory variables.

### Stepwise Selection Procedure

Step 0. Intercept entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 70640.205

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
1988.7133	9	<.0001

Step 1. Effect avg\_cheq\_CR\_TO entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	70642.205	69360.692
SC	70651.496	69379.272
-2 Log L	70640.205	69356.692

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1283.5138	1	<.0001
Score	1244.7131	1	<.0001
Wald	1217.4618	1	<.0001

### The LOGISTIC Procedure

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
860.9194	8	<.0001

Note: No effects for the model in Step 1 are removed.

#### Step 2. Effect avg\_cheq\_BAL entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	70642.205	69052.717
SC	70651.496	69080.588
-2 Log L	70640.205	69046.717

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1593.4886	2	<.0001
Score	1501.2440	2	<.0001
Wald	1460.3042	2	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
563.8172	7	<.0001

Note: No effects for the model in Step 2 are removed.

#### Step 3. Effect total\_debt entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	70642.205	68838.182
SC	70651.496	68875.344
-2 Log L	70640.205	68830.182

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1810.0232	3	<.0001
Score	1635.8946	3	<.0001
Wald	1571.4416	3	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
459.3791	6	<.0001

### The LOGISTIC Procedure

#### Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

#### Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	70642.205	68508.698
SC	70651.496	68555.150
-2 Log L	70640.205	68498.698

#### Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2141.5071	4	<.0001
Score	1824.7297	4	<.0001
Wald	1730.6197	4	<.0001

#### Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
106.9141	5	<.0001

Note: No effects for the model in Step 4 are removed.

#### Step 5. Effect ratio\_CRTO\_income entered:

#### Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

#### Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	70642.205	68457.953
SC	70651.496	68513.695
-2 Log L	70640.205	68445.953

#### Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2194.2528	5	<.0001
Score	1876.6248	5	<.0001
Wald	1779.6400	5	<.0001

#### Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
47.9536	4	<.0001

Note: No effects for the model in Step 5 are removed.

#### Step 6. Effect ratio\_avgbal\_income entered:

#### Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

### The LOGISTIC Procedure

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	70642.205	68416.015
SC	70651.496	68481.047
-2 Log L	70640.205	68402.015

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2238.1909	6	<.0001
Score	1983.6421	6	<.0001
Wald	1870.1375	6	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
7.3546	3	0.0614

Note: No effects for the model in Step 6 are removed.

#### Step 7. Effect avg\_SNI\_Bal entered:

Model Convergence Status	
Convergence criterion (GCONV=1E-8) satisfied.	

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	70642.205	68402.073
SC	70651.496	68476.396
-2 Log L	70640.205	68386.073

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2254.1323	7	<.0001
Score	1986.2456	7	<.0001
Wald	1866.3089	7	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
0.3257	2	0.8497

Note: No effects for the model in Step 7 are removed.

Note: No (additional) effects met the 0.05 significance level for entry into the model.

Summary of Stepwise Selection							
Step	Effect		DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
	Entered	Removed					
1	avg_cheq_CR_TO		1	1	1244.7131		<.0001
2	avg_cheq_BAL		1	2	320.1162		<.0001
3	total_debt		1	3	197.4075		<.0001

**The LOGISTIC Procedure**

Summary of Stepwise Selection							
Step	Effect		DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
	Entered	Removed					
4	ratio_debt_avgbal		1	4	351.2000		<.0001
5	ratio_CRTO_income		1	5	56.7138		<.0001
6	ratio_avgbal_income		1	6	40.2495		<.0001
7	avg_SNI_Bal		1	7	7.1346		0.0076

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.1359	0.0210	2931.8126	<.0001
avg_cheq_CR_TO	1	-0.00004	2.076E-6	456.9666	<.0001
avg_cheq_BAL	1	-0.00005	3.155E-6	239.5763	<.0001
avg_SNI_Bal	1	-4.44E-6	1.532E-6	8.3884	0.0038
total_debt	1	0.000016	9.633E-7	290.8389	<.0001
ratio_CRTO_income	1	1.5418	0.1702	82.0203	<.0001
ratio_avgbal_income	1	-0.1715	0.0262	42.8410	<.0001
ratio_debt_avgbal	1	0.0262	0.00151	300.9902	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
avg_cheq_CR_TO	1.000	1.000	1.000
avg_cheq_BAL	1.000	1.000	1.000
avg_SNI_Bal	1.000	1.000	1.000
total_debt	1.000	1.000	1.000
ratio_CRTO_income	4.673	3.347	6.524
ratio_avgbal_income	0.842	0.800	0.887
ratio_debt_avgbal	1.027	1.023	1.030

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	62.4	Somers' D	0.250
Percent Discordant	37.4	Gamma	0.251
Percent Tied	0.2	Tau-a	0.068
Pairs	865276556	c	0.625

### The LOGISTIC Procedure

Model Information	
Data Set	WORK.CANCEL_ALL
Response Variable	target_cancel
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	154086
Number of Observations Used	111635

Response Profile		
Ordered Value	target_cancel	Total Frequency
1	0	108317
2	1	3318

Probability modeled is target\_cancel=1.

Note: 42451 observations were deleted due to missing values for the response or explanatory variables.

### Stepwise Selection Procedure

Step 0. Intercept entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 29867.720

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
198.0455	9	<.0001

Step 1. Effect avg\_cheq\_CR\_TO entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	29869.720	29732.863
SC	29879.343	29752.109
-2 Log L	29867.720	29728.863

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	138.8573	1	<.0001
Score	133.8747	1	<.0001
Wald	132.6260	1	<.0001



### The LOGISTIC Procedure

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
65.5333	8	<.0001

Note: No effects for the model in Step 1 are removed.

#### Step 2. Effect ratio\_saving\_income entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	29869.720	29728.456
SC	29879.343	29757.325
-2 Log L	29867.720	29722.456

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	145.2645	2	<.0001
Score	160.8469	2	<.0001
Wald	134.5103	2	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
38.6342	7	<.0001

#### Step 3. Effect ratio\_saving\_income is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	29869.720	29732.863
SC	29879.343	29752.109
-2 Log L	29867.720	29728.863

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	138.8573	1	<.0001
Score	133.8747	1	<.0001
Wald	132.6260	1	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
65.5333	8	<.0001

Note: No effects for the model in Step 3 are removed.

**The LOGISTIC Procedure**
**Summary of Stepwise Selection**

Step	Effect		DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
	Entered	Removed					
1	avg_cheq_CR_TO		1	1	133.8747		<.0001
2	ratio_saving_income		1	2	24.7100		<.0001
3		ratio_saving_income	1	1		2.1283	0.1446

**Analysis of Maximum Likelihood Estimates**

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.1996	0.0290	12137.2859	<.0001
avg_cheq_CR_TO	1	-0.00003	3E-6	132.6260	<.0001

**Odds Ratio Estimates**

Effect	Point Estimate	95% Wald Confidence Limits	
		Lower	Upper
avg_cheq_CR_TO	1.000	1.000	1.000

**Association of Predicted Probabilities and Observed Responses**

Percent Concordant	52.2	Somers' D	0.099
Percent Discordant	42.4	Gamma	0.104
Percent Tied	5.4	Tau-a	0.006
Pairs	359395806	c	0.549