KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

Monetary Valuation of Microinsurance against Fire Risk: Mixed Logit Approach

By

Emmanuel Asare-Bediako

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN INDUSTRIAL MATHEMATICS

October 9, 2014

Declaration

I hereby declare that this submission is my own work towards the award of the MSc degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of the university, except where due acknowledgement had been made in the text.

Emmanuel Asare-Bediako	....	....
Student	Signature	Date

Certified by:

Dr. G.A. Okyere	....	....
Supervisor	Signature	Date

Certified by:

Prof. S. k. Amponsah	....	....
Head of Department	Signature	Date

Dedication

I dedicate this work to all students in Ghana. I encourage then to contribute their quota, however small it may be to move Ghana forward.

Abstract

The non-existence of separate insurance policy for people in the retail sector in Ghana at this moment may be due to supply side problems, such as correlated risks, uncertainty of risks, adverse selection and moral hazard, or because of a lack of demand for insurance coverage. This thesis uses mixed logit estimation methods to examine the effects of fire risk on the demand of microinsurance. The study took place at Kumasi Central Market. The Hierarchical Bayesian was used to estimate the parameters of the mixed logit model. The distribution of the parameters was normal distribution. It was found that positive relationship exist between utility and high damage cost, high fire risk and high insurance coverage, however, utility decline with insurance premium. Also, estimation results support that offering fire insurance may be profitable in the current situation. Stake holders in the insurance industry are therefore recommended to ensure the introduction of separate insurance for the retail sector of the economy of Ghana.

Acknowledgements

May the name of Almighty God be praised forever and over; it is His bountiful endowment of grace through Christ Jesus that kept me through. I would like to acknowledge the support and guidance received from my supervisor, Dr. G. A. Okyere. I am grateful to Pro. S. K. Amponsah, head of Mathematics department for his encouragement. I would also want to thank all traders at Kumasi Central Market, for their assistance. I am very grateful to all members of Mathematics department for their support and encouragement. My special thanks go to my wife Faustina Asare-Bediako and children, Nyameahoden Kwame Asare-Bediako, Nyameadom Ama Asare-Bediako and Christable Ama Asare-Bediako, for their patience and encouragement throughout my work.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background of the Thesis

One recurrent issue that Ghanaian markets countenance is the problem of market fires. Faulty electrical wiring has been claimed as the causes of the fire Adofo (2013). This is expected to have significant consequences for the insurance sector. This thesis is about using mixed logit model to estimate willingness to pay (WTP) microinsurance premium.

Insurance for that matter, microinsurance is an arrangement that offers financial protection to individuals and groups with a low-income against specific risks in exchange for premium payments (Tan et al., 2012). Microinsurance is run in accordance with generally accepted insurance practices but is designed to meet the needs of those who are otherwise unable to access mainstream insurance. For microinsurance, the premium is low. Microinsurance is targeted at those with low-incomes (in particular, those living on between approximately $1 and $4 per day).

Mixed logit models are the statistical tool applied in analysis of discrete choices data and it has assumed vitality in health economics (Hall et al., 2006; King et al., 2007; Paterson et al., 2008), environmental economics (Train, 2003; Brownstone and Train, 1999; Train and Wilson, 2007) and marketing (Revelt and Train, 1998; Hensher et al., 2003). Mixed logit model makes it probable to report heterogeneity in choices which are unrelated to observed features. It can be used to approximate discrete choice random utility model (McFadden and Train, 2000). Mixed logit model is a model which must always follow a particular distribution, for instance normal distribution, exponential distribution etc.. The parameters of this distribution are estimated by using either the Hierarchical Bayesian estimation methods or Maximum Likelihood estimation .

WTP is defined as the maximum amount that an individual is willing to bid for a public goods or services while remaining on her indifference curve that is without losing any utility. WTP is the unavoidable economic value that equates the utility with and without the non-market goods. WTP measures are considered useful for a number of reasons. Firstly, they can directly inform policy makers by proving information about how much people value some goods or services and can influence the pricing of these goods or services (Hensher et al., 2003; Hole and Kolstad, 2010). Secondly, WTP measures can be essential inputs in economic appraisals such as cost benefit analyses (Oliver et al., 2002; Negrín et al., 2008; Hole and Kolstad, 2010) Lastly, WTP measures can be a convenient tool to make relative comparisons and rankings of the desirability of goods and services (Hole and Kolstad, 2010). Mixed logit has been used to estimate willingness to pay (WTP) by many authors (McFadden and Train, 2000; Train, 2003; Regier et al., 2009), nevertheless, little it has been seen in the insurance industry.

The WTP for insurance premium is expected to go up due to growth in the probability of distressing weather-related damage, unpredictable fire outbreak and increasing rampant arm robbery, but it is uncertain how large this growth will be. Indeed, insight into the pressure of these challenges, WTP for microinsurance is necessary so that insurers can evaluate the future profitability of offering coverage in opposition to damage caused by disasters and crime. This is very apt given that climate change, electrical fault and crime are to be expected to continue in the coming years due to rapid projected growth of emissions (Pielke et al., 2007; Botzen and van den Bergh, 2008), growth of human population and crime along streets and market places.

Kumasi central market has experienced number of fire outbreaks. Last year alone, there were three fire outbreaks that were reported at the market. According to Kwakye (2013), the deputy Minister of Information, it is estimated that 4500 traders had substantial properties loss to fire outbreaks last year.This study looks at how mixed logit model can be use to predict insurance premium using data from second hand clothes sellers in Ghana. The recent devastating fire outbreaks at the markets places may have motivated this study.

1.2 Problem Statement

Decision to buy microinsurance policy sometimes largely depends on unexpected circumstances. Rick taking has come to stay as far as disasters are part of the unexpected circumstances. Many developed countries have laws about risk taken. However, in the third world countries these laws just exist in books. This may cause people not to buy insurance policies for the following reasons: non-enforcement of insurance laws; untimely payment of benefit when you are due; high levels of premium charge; and lack of education on insurance policies and premium charge.

Most insurance companies in Ghana find it difficult to calculate the income of people in the retail sector. As a result there is no separate insurance policy for the retail sector. The nonexistence of separate insurance policy for people in the retail sector in Ghana at this moment may be due to supply side problems, such as correlated risks, uncertainty of risks, adverse selection and moral hazard, or because of a lack of demand for insurance coverage. de Vries (1998) suggested that problems with adverse selection may be severe in the case of offering disaster insurance, because only individuals who live in unprotected areas with high disaster risks would demand insurance. Examining how WTP relates to actual risk derived from the retail industry characteristics will provide insight into potential problems with adverse selection.

The government can grant partly compensation of damage caused by large-scale fire outbreaks, as is also the situation in several other countries (Crichton, 2008). Decisions about granting relief and its extent are a political decision. As a consequence, other sectors of the economy may expect that the government will compensate future fire damage unconditional on the risk they take. This may lessen the desirability of private insurance, which is often referred to as crowding out (Harrington, 2000).

1.3 Objectives of the Thesis

Microinsurance in Ghana is expected to grow due to increasing incidents of fire disasters at the market places. The thesis is aimed at:

1. To examine whether demand is the main obstruction to the establishment microinsurance market for the retail industry by estimating the level of WTP relative to the expected value of loss.

2. To analyse the effect of the current institutional setting, characterized by availability of government compensation of fire victims, on microinsurance industry.

1.4 Scope of the Study Area

Ashanti region is one of the ten regions in Ghana, which is the highest populated region in Ghana. Ashanti region is centrally located in the middle belt of Ghana, it lies between longitudes $0.15W$ and $2.25W$ and latitude $5.50N$ and $7.46N$. The region shares boundaries with four of the ten regions, Brong Ahafo to the north, Eastern region to the east, Central region to the south and Western region to the south west. The population of the region is concentrated in a few districts, Kumasi metropolis alone accounts for nearly one-third of the regions population.

The high level of urbanization in the region is due mainly to the high level of concentration of the population in the Kumasi metropolis.

The region occupies a land area of $24,389$ square kilometres representing 10.2% of the total land area of Ghana in which Kumasi alone is 250 square kilometers. It is the third largest region after Northern and Brong Ahafo regions. The region has a population density of 148.1 persons per square kilometre and Kumasi has a population of about 1.5 million people. The people of the region are into farming, mining and trading. Tradition is held very high in the region and blends well with modernity. Residential land use in Kumasi forms about 60 % of the total land use in the metropolitan area and they are categorised into three zones namely; the low income, middle, and the high income zone. The metropolitan area has one teaching hospital, 9 hospitals and many private hospitals and clinics, two public universities and six private universities. The metropolitan area also has very big market that is Kumasi central market. Most middle income people work at the market. The market over the years have experienced fire outbreaks. This has left many people unemployed since they did not either insured their goods or their stores.

1.5 Methodology

The study is about using discrete choice experiment to estimate WTP of insurance premium. WTP for insurance is elicited by means of a choice experiment. The choice experiment values insurance with different levels of coverage in situations with varying fire probabilities and damages caused by fire at the market places. The distribution of the study shall follow normal distribution and lognormal. Bayesian estimation methods will be employed to determine the parameters of the distribution. Mixed logit model will be used to estimate the insurance premium. Convenience sampling method will be used to select six hundred second clothes sellers. The study area is the Kumasi central market. Structured questionnaires will be employed to collect data. The software employed to solve the problem is sawtooth and STATA version 12.

1.6 Justification

The study is of much importance and relevance for number of reasons; the chief amongst them is to partially fulfil the requirement for the award of degree in industrial mathematics. The study will also estimate WTP for microinsurance premium in Ghana under the current institutional setting. This is of practical importance for insurers and the government in evaluating whether demand for insurance in the retail industry will be sufficiently high to make a private market feasible. The role of expectations about government compensation of disaster damage will be analysed by comparing WTP with and without relief of fire damage by the government. This can aid the government in assessing what circumstances need to be created to stimulate, or at least not impede, the emergence of a market for insurance in the retail industry in the wake of rampant fire outbreaks.

Furthermore, the challenges of socio-economic developments on the demand for microinsurance will be assessed. This is consummated by obtaining microinsurance demand under different scenarios of increased fire outbreak probabilities due to electrical faults and varying levels of expected fire outbreak damage. In addition, offer functions will be estimated to discover factors behind WTP using as explanatory variables perceptions of microinsurance premium based on estimates of individual risk aversion, actual insurance purchase behaviour, and socio-economic characteristics.

Lastly, the study will also serve as a source of reference material to students, government, private and school libraries. And also serve as a research paper for further research work.

1.7 Organisation of the Thesis

This thesis has five chapters. Chapter One contains the background to the study, statement of the problem, objectives of the study, Scope of the study area, Methodology and Justification. Chapter Two reviews the relevant literature on mixed logit models and WTP.

Chapter Three focuses on the methodology of the study. It describes the study design, the target population, the sampling procedure, sample size, the research instruments used, data and sources, data processing and analysis, Mixed logit model methods and the ethical issues arising from the research. Chapter Four concerns itself with the data analysis, the presentation and discussion of results while Chapter Five provides the summary, conclusions and recommendations of the study.

Chapter 2

Literature Review

2.1 Introduction

This chapter reviews the various works that are relevant to the study. The topics reviewed include the Insurance Premium Estimation, Discrete Choice Experiment, Mixed Logit Approach, Terminology Mixed Logit Approach and WTP. Other important topics reviewed are Estimation of mixed logit model, Unobserved Heterogeneity and Distributional Assumptions.

2.2 Microinsurance

Microinsurance is an arrangement that offers financial protection to individuals and groups with a low-income against specific risks in exchange for premium payments (Tan et al., 2012). The International Association of Insurance Supervisors ($IAIS$) also defines microinsurance as "protection of low income people against specific perils in exchange for regular premium payments proportionate to the likelihood and cost of the risk involved." Microinsurance is run in accordance with generally accepted insurance practices but is designed to meet the needs of those who are otherwise unable to access mainstream insurance. Microinsurance is based on the same principles as most other forms of insurance, with a number of key differences:transaction sizes are smaller and premiums are lower; the target audience consists of those with low incomes, traditional jobs/businesses and a limited knowledge of insurance and the benefits that it affords; Products offered are simpler, often providing just one type of cover rather than bundles; more flexibility is possible in both product design and premium payments. Due to the irregular income stream of the clients, microinsurance arrangements may allow payments to be made at irregular times and in uneven amounts; the claims process is faster and less complicated, involving substantially less documentation. As a result, microinsurance claims are usually resolved quicker than most insurance claims, which must go through a more stringent assessment process; Less regulation is imposed and enforced, though some countries are beginning to regulate the microinsurance industry in a more structured manner; and Underwriting is a simpler process (i.e. there are fewer terms and conditions and exclusions) as relatively small sums are involved.

Microinsurance is aimed at a broad range of clients including individuals, households and whole communities in both rural and urban areas. Microinsurance is targeted at those with low-incomes (in particular, those living on between approximately \$1 and \$4 per day). Without a way to manage risk they can be hit hard by sudden events for which they have not planned (including crop failures, death of relatives, ongoing health issues, fire outbreaks). Many of those who use microinsurance work informally or in unpredictable sectors such as trading and, therefore, do not have a regular flow of income (Tan et al., 2012). Access to insurance at an affordable cost to the poor is now seen in many countries as a condition necessary for poverty reduction and social harmony in the financial landscape (Matul et al., 2010). Access to microinsurance is generally very limited. Compared to some other regions, in Africa, there is still a dramatic lack of microinsurance available to the low-income market. According to the Landscaping Study of the Microinsurance Centre, only 3.5 million people in Africa use microinsurance to secure risks. In most sub-Saharan countries less than 2 % of the poor and vulnerable have access to microinsurance Matul et al. (2010). For this study, microinsurance is limited to people with limited knowledge of insurance and the benefits that it affords, have minimum capital around GHS 1000 and have monthly income after tax around GHS 700.

2.3 Microinsurance Premium Estimation

The premium or price of microinsurance is the monetary value for which two parties agree to exchange risk and "certainty" (Laeven and Goovaerts, 2011). This sometimes based on a lot of factors. Some of the factor are age of the people buying insurance, the kind of policy, risk involve in the policy and income of the people buying insurance. According to Laeven and Goovaerts (2011), there are two commonly encountered situations in which the price of insurance is subjected to: when an individual bearing an insurable risk, buys insurance from an insurer at an agreed periodic premium; and when insurance portfolios (that is, a collection of insurance contracts) are treated in financial industry. However, the intention of this study is to predict the insurance premium using data from market survey.

Luan (2001) examined an insurance or risk premium calculation method called the mean-value-distortion pricing principle in the general framework of anticipated utility theory. He concluded that essential properties such as non-negative loading, non-excessive loading, scale and translation invariant, stop-loss order preservation, and sub-additivity are preserved in the analysis of the pricing principle. This fact suggests consideration of microinsurance problems in a larger theoretical frame. This method cannot be used to estimate optimal microinsurance premium. Since mean-value-distortion pricing principle does not give optimal microinsurance premium, to solve this problem, mixed logit model will be use.

Wuthrich (2003) estimated value at risk for sums of dependent random variables. He specify the dependent strength and the marginal behaviour and concluded that the dependence structure has a rather large impact on joint extreme value calculations. Botzen and van den Bergh (2008) also estimated risk premiums for flood insurance demand in the Netherlands. They indicated that rising flood probabilities from 1 in 1250 up to 1 in 550 cause WTP to rise more than expected value of the loss. This study will look at how fire probabilities will affect WTP. Fire risk has increase in the country for the pass decade. Many trader have loss their capitals due to market fires. It is estimated that in every fire outbreak that occurs at the market places over 500 traders get their goods destroy. With choice modelling and mixed logit estimation methods Botzen and van den Bergh (2008) estimated the dependence of WTP on risk aversion and socioeconomic characteristics. Their results indicate that opportunities for a private flood insurance market exist. Therefore, the influence of socio-economic development on WTP for disaster insurance needs to be analysed in addition to potential effects of fire to arrive at reliable estimates for future demand. Botzen and van den Bergh (2008) approach in calculating insurance premium will be considered in this study. Mixed logit model will be employ to estimate willingness to pay insurance premium using data from market related discrete choice experiment.

2.4 Discrete Choice Experiment

The discrete choice experiment (DCE) method has become an established stated preference approach for many economic, marketing, health and educational surveys. It has also been used for valuing insurance benefits. This methodology, since the work by Adamowicz (1994), has gain strong theoretical background. It is consistent with both Lancasterian microeconomic approach to utility derivation (Lancaster, 1966) and is behaviourally stuck in random utility theory McFadden (1974).

Campbell (2006) reported the findings from a discrete choice experiment study designed to estimate the economic benefits associated with rural landscape improvements in Ireland. Using a mixed logit model, the panel nature of the dataset is exploited to retrieve willingness to pay values for every individual in the sample. This departs from customary approaches in which the willingness to pay estimates are normally expressed as measures of central tendency of a priori distribution. Gu et al. (2007) use discrete choice experiments ($DCEs$) to estimate health state utility values. They compared DCE and standard methods such as Time Trade-off (TTO). It was established that the two methods were fundamentally conceptually different and have different interpretations in policy evaluation. However, it was pointed out that $DCEs$ estimate the " average valuation " for a health state better. A basic decision when designing $DCEs$ is whether to use labelled or unlabelled choice tasks. The labelled and the unlabelled discrete choice experiments have been applied by many researchers (Blamey et al., 2000; McClure et al., 2004; Shen and Saijo, 2009; Doherty et al., 2011).

Blamey et al. (2000) gave an advantage of why it is good to assign labels. It enables responses better reflect the emotional context in which preferences are ultimately revealed. This was supported by McClure et al. (2004) and Shen and Saijo (2009). Labelling alternatives enables factors to be captured more accurately. An alternative label is different from other attributes because it is independent from the quantifiable characteristics of the alternative, and thus can invoke different emotions from respondents (Czajkowski and Hanley, 2009). Indeed, within the context of fire disaster, using labels to represent the different types of policies has particular advantages. it gives the respondents predisposition toward choosing particular types of insurance because it could invoke memories of past fond experiences (Blamey et al., 2000). Adversely, labelling alternatives may result in the labels having a significantly superior impact on how respondents reach their choice outcomes than may be expected when designing $DCEs$. Doherty et al. (2011) use discrete choice experiment data aimed at eliciting the demand for recreational walking trails on farmland in Ireland. It was used to explore whether some respondents reach their choices solely on the basis of the alternative's label. To investigate this type of processing strategy, the paper exploits a discrete mixtures approach which encompasses random parameters for the attribute. It was evidence that respondents employ different processing strategies for different alternatives and differences in processing emerge between rural and urban based respondents. Results highlight that model fit and policy conclusions are sensitive to assumptions related to processing strategies among respondents.

Blamey et al. (2000) and Bekker-Grob et al. (2010) advocate that respondents have a higher tendency to ignore attributes when labelled alternatives are included in the choice experiment. The study provided comprehensively analysis to probabilistically agree on each alternative, great proportion of respondents made their choices based on its label only. For this study, unlabelled discrete choice experiment will be use. This is because respondents employ different processing strategies for different alternatives.

2.5 Mixed Logit Approach

Mixed logit is a highly flexible model that can approximate any random utility model McFadden and Train (2000). The mixed logit model is an interesting, very flexible and useful modelling alternative, allowing to model and estimate correlation and heteroscedasticity (Munizaga, 2000). In this context, it can become a real competitor to Probit, usually considered as the only or principal way to make more flexible the modelling of discrete choices. Nevertheless it is quite important to have a clear assessment of its properties and to justify properly any specific structure over the basis of theoretical considerations prior to the estimation of the parameters. The covariance matrix associated to mixed logit depends on the specification given to the additional error terms and it can be as general as desired. In this sense, it offers a more flexible structure than other models, in particular it has the capacity of recognising correlated alternatives and taste variations expressed through random parameters. In the work of Munizaga (2000), two numerical applications were presented, one based on simulation experiments and another one with real data, both in the context of similar alternatives. The latter motivates a nesting error structure. It is shown analytically that the Nested Mixed Logit is not equivalent to Nested Logit at least considering its covariance structure. However, for the reported correlation level, if mixed logit is not adjusted to obtain a homoscedastic covariance matrix, then the predicted market shares for both do not present severe differences. So, it is understood that these models could approximate a situation like the one presented here. The fact that the nested structure for mixed logit commits homoscedasticity when defining correlation, could be seen as a problem, if you want to compare it with Nested Logit, or as an advantage from the point of view that it allows to represent another covariance structure, that includes correlation and heteroscedasticity (Munizaga, 2000).

Mixed logit obviates the three restrictions of standard logit by permitting for random taste variation, unrestricted substitution patterns, and correlation in unobserved factors over time. It is not restricted to normal distributions. Its derivation is straightforward, and simulation of its choice probabilities is computationally simple. The mixed logit model has been known for many years but has only become fully applicable since the advent of simulation (Train, 2003; Yannis and Antoniou, 2007).

Algers et al. (1998) use mixed logit specifications to allow parameters to vary in the population when estimating the value of time for long-distance car travel. Their main conclusion was that the estimated value of time is very sensitive to how the model is specified. They found that the ratio of coefficients in a mixed logit specification differ significantly from the ones in a traditional logit specification. This is contrary to the results obtained by Brownstone and Train (1999) and Train (1998). Whether the ratios will differ or not depends on the model and the data generating process at hand.

Gu et al. (2007) propose two types of models, one using preference space and the other using Quality-Adjusted Life Year ($QALY$) space. The estimation approach was based on the mixed logit. Their results demonstrated that the preferred $QALY$ space mode provides lower estimates of the utility values than the conditional logit, with the divergence increasing with worsening health states. Brownstone et al. (2000) compare multinomial logit and mixed logit models using data from California household on stated preference for auto mobiles. It was concluded that mixed logit models provide improved fits over logit that are highly significant, and show large heterogeneity in respondents' preferences for alternative fuel vehicles. Taste heterogeneity is another important aspect of discrete choice modelling. Train and Weeks (2005) show that WTP estimates can be estimated directly in a mixed logit model by re-formulating the modelling such a way that the estimated parameters represent the parameters of the WTP distribution rather than the parameters of the usual coefficients.

Campbell (2006) use mixed logit model to identify the determinants of willingness to pay for rural landscape improvements. It was asserted that mixed logit add more considerably validity and explanatory power to welfare estimates. Mixed Logit models are designed to capture observed as well as unobserved heterogeneity, there is no reason to expect these tastes for different things to be independent. King et al. (2007) also analyse patients' preferences for managing asthma using mixed logit models with random intercepts. They find that the mixed logit models fit the data better than a standard logit and that a substantial amount of heterogeneity is unaccounted for observable characteristics.

Hole (2008) examines patients' preferences for the attributes of a general practitioner appointment using mixed and latent class logit models. Significant preference heterogeneity is found for all attributes including cost and the mixed and latent class logit modes fit the data considerably better than the standard logit model. Negrín et al. (2008) apply mixed logit models to analyse the willingness to pay for alternative policies for patients with Alzheimer's disease and they discovered that there is significant heterogeneity in the preferences for all the attributes including cost. The researchers report WTP measures calculated at the means of the coefficient distributions.

Paterson et al. (2008) study smokers' preferences for increased efficacy and other attributes of smoking cessation therapies. Using a mixed logit model, they estimate the willingness to pay for different treatments among groups of smokers. They find evidence of substantial preference heterogeneity and demonstrate that allowing for heterogeneity both improves the fit of the model and enhances our understanding of the smokers' preferences. Regier et al. (2009) analyse preferences regarding genetic testing for developmental delay using mixed logit models. They concluded that WTP measures are derived from the assumptions affect the WTP estimates. Ozdemir et al. (2009) analyse how "cheap talk" affects estimates of the willingness to pay for health care using a mixed logit model estimated in WTP space. They conclude that being exposed to "cheap talk" has an impact on the estimated willingness to pay.

This study therefore looks at how mixed logit model can be used to estimate insurance premium using data from market related survey. This study will emulate McFadden and Train (2000) and Train (2003). The framework of random utility maximisation is deep-rooted to model such choices but there are still many issues that deserve attention. This thesis investigates how taste correlation should be incorporated into applied mixed logit estimation.

The challenges mixed logit models face is from the quality of the data (Sennhauser, 2010). Mixed logit certainly demands better quality data than MNL (Train, 1998) since it offers an extended framework within which to capture a greater amount of true behavioural variability in choice making. Mixed logit align itself much more with reality where every individual has their own inter-related systematic and random components for each alternative in their perceptual choice set(s). Although there is a level of irreducible variability in everyone, its does have some basis in the fact that individuals do not do the same thing all the time for a variety of reasons that analysts cannot fully observe or explain (and probably neither can the individuals themselves) Michaud et al. (2012). The mixed logit probability according to McFadden and Train (1997); Train (1998) and Brownstone and Train (1999) can be derived from utility-maximizing behaviour in several ways that are formally equivalent but provide different interpretations. The most uncomplicated derivation, and most widely used in recent applications (Ben-Akiva and Lerman, 1985; Bhat, 1998; Train, 1998), is based on random coefficients.

The derivation can also be based on the error component. The error components emphasize the fact that the unobserved portion of utility consists of a number of components and that these components can be specified to provide realistic substitution patterns rather than to represent random parameters Brownstone and Train (1999). This term encompasses any interpretation that is consistent with the functional form. This study will base on random coefficient parameters estimation.

2.6 WTP

Direct questions regarding willingness to pay are cognitively tricky to answer directly and respondents may have incentives to answer strategically (Arrow et al., 1993; Carson et al., 2001; Hanley et al., 2002; Ryan, 2004). However, WTP measures can be derived from discrete choice models estimated using either revealed preference data or data from discrete choice experiments ($DCEs$). In these cases, the WTP for an alternative attribute can be calculated as the ratio of the attribute coefficient to the price coefficient (Train, 2003).

WTP is defined as the maximum amount that an individual is willing to bid for a public goods or services whiles remaining on her indifference curve that is without losing any utility Hole and Kolstad (2010). Boccaletti and Moro (2000) investigated consumers' willingness to pay (WTP) a price premium for two environmental attributes of a non-food agricultural product. They studied individual preferences for roses associated with an eco-label and a carbon footprint using an economic experiment combining discrete choice questions and real economic incentives involving real purchases of roses against cash. The data were analysed with a mixed logit model and reveal significant premium for both environmental attributes of the product.

WTP is the unavoidable economic value that equates the utility with and without the non-market goods. WTP measures are considered useful for a number of reasons. Firstly, they can directly inform policy makers by proving information about how much people value some goods or services and can influence the pricing of these goods or services (Hanley et al., 2002; Hole and Kolstad, 2010). Secondly, WTP measures can be essential inputs in economic appraisals such as cost benefit analyses (Oliver et al., 2002; Negrín et al., 2008; Hole and Kolstad, 2010). Lastly, WTP measures can be a convenient tool to make relative comparisons and rankings of the desirability of goods and services (Hole and Kolstad, 2010). Mixed logit has been used to estimate willingness to pay (WTP) by many authors (McFadden and Train, 2000; Train, 2003; Regier et al., 2009). Gu et al. (2007) did three separate investigations on taste correlation in willingness-to-pay estimation. The first contribution addresses how to incorporate taste correlation in the estimation of the value of travel time for public transport. The second contribution examines how different distributional assumptions are affected by the inclusion of taste correlation. The third contribution investigates the correlation patterns between willingness to pay measures for different public transport modes and how to capture them in the simplest possible way. A general feature of the three investigations is that they discovered scale heterogeneity. Since this induces correlation it is an important aspect of taste correlation to specify the scale correctly. It was concluded that scale heterogeneity may be partly explicated by background variables. Examining the three contributions on taste correlation there seems to be the general conclusion that considerable taste correlation is often present and that it sometimes has an effect on willingness to pay evaluation.

Eboli and Mazzulla (2008) examined passengers' willingness to pay (WTP) for improving the quality levels of a bus service. The objective of the study was to provide a tool for evaluating passenger willingness to pay by considering some qualitative service aspects, in addition to the calibration of behavioural models based on user choices. The WTP values were obtained as marginal rates of substitution between some service quality attributes and travel cost at constant utility.

While willingness-to-pay (WTP) measures derived from individual choice models provide an alternative assessment, antitrust law is, however, framed in terms of the likely price effects of mergers. In their paper "What does Willingness to pay reveal about hospital market power in merger cases?" Fournier and Gai (2006) examine the connection between health plan prices and WTP. They used merger cases in Florida and New York State to evaluate the accuracy of pre-merger predictions from patient-level choice models to assess mergers' effects on patients' aggregate WTP.

Fournier and Gai (2006) find that the method can provide reliable predictions of patients' post-merger willingness-to-pay, and thereby help inform the pre-merger investigation concerning likely price effects. Campbell (2006) in his paper "combining mixed logit models and random effects models to identify the determinants of willingness to pay for rural landscape improvements", departed from customary approaches in which the willingness to pay estimates are normally expressed as measures of central tendency of a priori distribution. He used random effects modes for panel data to identify the determinants of the individual-specific willingness to pay estimates. In comparison with the standard methods used to incorporate individual-specific variables into the analysis of discrete choice experiments, the analytical approach outlined add considerably more validity and explanatory power to welfare estimates. Paterson et al. (2008) also reported the WTP for the non-monetary attributes by calculating the median of the coefficient distributions. Hole (2008) examined patients' preferences for the attributes of a general practitioner appointment using mixed and latent class logit models. The WTP distributions are found to be right-skewed as the mean WTP is substantially higher than the median WTP.

Hole and Kolstad (2010) use different approaches in modelling the distribution of WTP by comparing using stated preference data on Tanzanian Clinical Officers' job choices and mixed logit models. The standard approach of specifying the distributions of the coefficients and deriving WTP as the ratio of two coefficients (estimation in preference space) is compared to specifying the distributions for WTP directly at the better than the corresponding models in WTP space although the difference between the best fitting models in the two estimation regimes is minimal. Moreover, the willingness to pay estimates derived from the preference space models turn out to be unrealistically high for many of the job attributes. The results suggest that sensitivity testing using a variety of model specifications, including estimation in WTP pace, is recommended when using mixed logit models to estimate willingness to pay distributions.

2.7 Estimation of Mixed Logit Model

There are many ways of estimating the mixed logit coefficient. However, the common ones are maximum simulated likelihood and Bayesian methods. Many researchers have tried to compare the two. Train and Weeks (2005) and Spanier and Maize (1991) employ stated preference data on the choice of cars with different fuel systems to compare the performance of models in WTP space to models in preference space. Both studies draw on hierarchical Bayes to estimate the mixed logit models and their results were parallel in that it was found that the models in preference space fit the data better than the models estimated in WTP space. Nevertheless, the models in WTP space were found to produce more pragmatic WTP measures. Scarpa et al. (2008) use revealed preference data on destination choices to estimate models in preference and WTP space using both maximum simulated likelihood and hierarchical Bayes. Scarpa et al. (2008) discovered that both models fit the data better and produces more realistic WTP estimates. They therefore conclude that there is no tradeoff between goodness of fit and reasonable WTP estimates.

Negrín et al. (2008) apply mixed logit models to analyse the willingness to pay for alternative policies for patients with Alzheimer's disease. All coefficients were specified to be normally distributed and both maximum simulated likelihood and hierarchical Bayes methods were used to estimate the modes. They found that there is significant heterogeneity in the preferences for all the attributes including cost. Regier et al. (2009) estimated mixed logit models using hierarchical Bayes and maximum simulated likelihood. WTP measures were derived from the coefficients in the estimated models and it was demonstrated that different distributional assumptions affect the WTP estimates. It was noted that when the cost parameter is assumed to be log-normally distributed some WTP estimates were found to be very high.

However, Bayesian procedures have been use by researchers such as (Haan et al., 2012; Ozaki et al., 2005; Train and Sonnier, 2003) etc.. Train (2001) asserted that the Bayesian procedures operate as effectively with log-normals as normals because the log-normal is simply a transformation of the normal that does not entail any other parameters. Train and Sonnier (2003) were convinced that the use of a joint normal distribution for part worths is computationally attractive, particularly with Bayesian procedures, and yet is unrealistic for any attribute whose part worth is logically bounded (e.g., is necessarily positive or cannot be unboundedly large).

Ozaki et al. (2005) focused on the modelling of agricultural yield data using hierarchical Bayesian models. They considered the temporal, spatial and spatio-temporal relationships pertinent to the prediction and pricing of insurance contracts based on regional crop yields. The methodology used in this article proposes improvements in the statistical and actuarial methods often applied to the calculation of insurance premium rates. These improvements are especially relevant to situations of limited data. These conditions are often encountered, in particular at the individual level.

Haan et al. (2012) applied Bayesian procedures as a numerical tool for the estimation of a female labour supply model based on a sample size which is typical for common household panels. They provided two important results for the practitioner: First, for a specification with a multivariate normal distribution for the unobserved heterogeneity, the Bayesian estimator yields almost identical results as a classical Maximum Simulated Likelihood (MSL) estimator. Second, they observed that imposing distributional assumptions which are consistent with economic theory, e.g. log-normally distributed consumption preferences, the Bayesian method performs well and provides reasonable estimates, while the

MSL estimator does not converge. These results indicate that Bayesian procedures can be a beneficial tool for the estimation of dynamic discrete choice models.

2.8 Bayesian Approach

Bayesian approach was introduced by (Allenby, 1997) for mixed logits with normally distributed coefficients. Train (2001) extended this procedure for mixed logit to non-normal distributions of coefficients, including lognormal, uniform and triangular distributions. According to Train (2003), Bayesian procedure do not require maximization of any function like classical procedures, since with some mixed logit models (example, lognormal distributions), maximization of the simulated likelihood function can be difficult numerically. Often the algorithm fails to converge for following reasons: The choice of starting values is often critical, with the algorithm converging from starting values that are close to the maximum but not from other starting values; and the issue of local versus global maxima complicates the maximization further, since convergence does not guarantee that the global maximum has been attained.

Also, desirable estimation properties such as consistency and efficiency, can be attained under more relaxed conditions with Bayesian methods than Classical ones. Maximum Simulated Likelihood (MSL) is consistent only if the number of draws used in simulation is considered to rise with sample size and efficiency is attained only if the number of draws rises faster than the square root of sample size. Nevertheless, the Bayesian estimators are consistent for a fixed number of draws used in simulation and are efficient if the number of draws rises at any rate with sample size.

To simulate relevant statistics that are defined over a distribution, the Bayesian procedures according to Train (2003), use an iterative process that converges, with a sufficient number of iterations, to draw from that distribution. This convergence is different from the convergence to a maximum that is needed for classical procedures and involves its own set of difficulties. It is difficult to determine whether convergence has actually been achieved since Bayesian procedures trade the difficulties of convergence to a maximum Train and Sonnier (2003).

Lastly, Bayesian procedures are far faster in some modes and distribution and are more straightforward from a programming perspective than classical methods Train (2003)

2.9 Unobserved Heterogeneity

Copious studies have documented that discrete choice models without unobserved heterogeneity require either very strong or often implausible assumptions or lead to biased estimates of central parameters. Han et al. (2001) extend a mixed logit model to hold the random heterogeneity across drivers and to handle the correlation between repeated choices.

Hess et al. (2004) employ mixed logit models that okay random taste heterogeneity for the computation of value-of-time. From Train and Wilson (2007), unobserved heterogeneity in discrete choice models can be complex and therefore it is often necessary to allow for a general specification with potential correlations of the different processes.

According to Haan et al. (2012), this is true for dynamic models which analyse the role of state dependence in the behaviour of agents. In Bayesian models it is necessary to disentangle true state dependence from individual specific effects van den Berg (2001); Dube et al. (2010); Prowse (2010). Bayesian procedures, such as Markov Chain Monte Carlo ($MCMC$) techniques serve as an imperative alternative for the estimation of non-linear models with unobserved heterogeneity.

Since the Bayesian $MCMC$ estimator does not involve maximization of a likelihood function, the numerical problems of classical procedures such as maximum simulated likelihood (MSL) do not arise.

2.10 Distributional Assumptions

According to Algers et al. (1998), the estimated parameters are very sensitive to how the model is specified. They found that it is significantly lower when the coefficients are assumed to be normally distributed in the population, as compared to the traditional case when they are treated as fixed. This helps to obtain optimum estimation. Train and Weeks (2005) stated that to models unobserved taste heterogeneity, distributional assumptions can be placed in two ways: by specifying the distributional of coefficients in the utility function and deriving the distribution of willingness to pay (WTP); and by specifying the distribution of WTP and deriving the distribution of coefficients. In general the two approaches are equivalent, in that any mutually compatible distributions for coefficients and WTP can be represented in either way. However, in practice, convenient distributions such as normal or log-normal, are usually specified, and these convenient distributions have different implications when placed on $WTP's$ than on coefficients (called models in preference space) with models using these distributions for WTP (called models in WTP space).

Train and Weeks (2005) find that the models in preference space fit the data better but provide less reasonable distributions of WTP than the models in WTP space. Our findings suggests that further work is needed to identify distributions that either fit better when applied in WTP space or imply more reasonable distributions of WTP when applied in preference space.

Train (2001)) points out that in the classical approach finding the maximum of the likelihood is considerably more complex with log-normal distributions. And, yet if a maximum is found, it may occur that the Hessian is singular at this point. According to Train and Sonnier (2003), mixed logit is specified with part worths that are transformations of normally distributed terms, where the transformation induces bounds; examples include censored normals and log-normals, and distributions which are bounded on both sides. The model retains the computational advantages of joint normals while providing greater flexibility for the distributions of correlated part worths.

Bounded distributions can and have been used in mixed logits estimated by both the classical and Bayesian procedures (e.g., Bhat (2000); Revelt and Train (1998); Train (1998); Revelt and Train (1999); Brownstone and Train (1999); Train (2001); Johnston et al. (2002); Boatwright et al. (2003)). However, each estimation procedure, while feasible with bounded distributions, entails numerical difficulties that are intrinsic to its form, as described and illustrated by Train (2001). In particular: Classical procedures handle triangular, truncated normal, and similarly bounded distributions easily while Bayesian procedures are relatively slow with these distributions. On the other hand, fully correlated part worths are difficult to handle in classical procedures due to the proliferation of parameters, while the Bayesian procedures accommodate these correlations readily.

Several researchers have executed log-normal distributions within mixed logit, though usually without allowing full correlation; see, e.g., Bhat (1998, 2000); Revelt and Train (1998); Train (1998); Johnston et al. (2002) examined censored normals and found that they provided more reasonable results and better fit than uncensored normals in his application.

Bayesian procedures operate effectively with normals because of the convenient posteriors that arise with normals. This study will build upon the observation in

Train (2001) that the Bayesian procedures operate as effectively with log-normals as normals because the log-normal is simply a transformation of the normal that does not entail any other parameters.

Chapter 3

Methodology

3.1 Introduction

In this chapter, we shall consider the data collection method for the study and mathematical formulations of mixed logit model.

3.2 The Experimental Design

The experimental design for this study was the discrete choice experiment (DCE). The choice experiment involves three circumstances in which insurance is available with as attributes the number of times fire outbreaks occur in a year, expected damage on goods, the percentage of insurance coverage, and the expected insurance premium. An " opt out " option is included for respondents who do not want insurance. The lowest insurance coverage level was chosen to be 75 percent in this experiment since this equals to the maximum allowed deductible in disaster insurance markets (Kunreuther et al., 2008). The expected damage of goods was chosen to be GHS 2000.00. This is the estimated capital per second hand clothes sellers at Kumasi central market, while the other two levels of the experiment (GHS 1000.00 and GHS 5000.00) can be regarded as minimum and maximum estimates. The respondents indicate whether they prefer to buy insurance and if yes which insurance policy they favour. The method used in this study tries to evaluate the factors that influence insurance decision. For instance, individuals choose whether to buy a certain degree of insurance coverage against a risk and damage for a certain premium.

As an advantage the choice experiment provides more information about the factors that influence demand for insurance. It allows for simultaneously examining effects of varying probabilities, expected damages, coverage levels, and premiums on choices for insurance. In addition, the experiment is closer to reality where respondents can choose between different insurance options without a need to state a maximum WTP amount, which may result in smaller biases Botzen and van den Bergh (2008). In total, 300 versions of the design have been generated to which respondents were purposively assigned. This means that many combinations of the levels of the attributes appear in the experiment. The generated design has been checked for strictly dominant choices, which were then excluded from the final design. Each respondent answered twelve random choices.

3.2.1 Pre-tests

The questionnaire was reviewed by experienced lecturers. After including their comments, pretest of the questionnaire were conducted using face-to-face interviews. Thirty second hand clothes sellers at Asafo Market in Kumasi were interviewed. Particular attention was paid to the intellectual capacity on fire risk and the choice experiment. Labelled experiment (one label per insurance type with label specific attributes) consisting of insurance options that cover fire damage on store content, store, both, and no insurance. The consequential choice experiment turned out to be exceedingly complex. Instead, an experiment with unlabelled alternatives that values an insurance covering both damage on store contents and store was used in the survey. This turned out to be easier for respondents. Monthly premiums were provided in the choice experiment and their levels were derived from answers to open-ended WTP questions.

3.2.2 The Structure of the Questionnaires

The questionnaire was opened with questions about whether respondents have experience fire outbreak, fire damage and have been evacuated because of fire threats. In addition, several questions address the perception of fire risks using both qualitative and quantitative answer categories. The answers to these risk perception questions are discussed in detail in Botzen and van den Bergh (2008). These questions familiarize the respondents with fire risks. The assessed perceptions may be important in decision making under risk and serve as explanatory variables. Moreover, questions are included about risk aversion and actual insurance purchases. The uncertainty about receiving relief from government on fire damage is mentioned.

The choice experiment values fire insurance with varying coverage levels in situations with number of times fire outbreaks occur in a year and expected damages. An unlabelled experiment is used where respondents choose between insurance "Situation A", "Situation B", "Situation C" or an " opt out". Respondents are instructed to choose the " opt out" in case they do not want insurance or find the insurance in both situations unattractive. The questionnaire concludes with the usual socio-demographic questions.

3.2.3 Administration of the survey questionnaire and sample characteristics

The survey was administered using Sawtooth *CBC* software. This computer based method which has the following advantage: follow-up questions can be automated; high quality graphics can be included; a large underlying design for the choice model can be applied; interviewer effects can be avoided; and a geographically spread sample can be obtained at relatively low costs. Respondents were selected from the second hand clothes sellers at Kumasi central market. The sample consists of random respondents that sell at "18 line" where there is at least one fire outbreak in a year. A total of 137 respondents filled out the questionnaire. The questionnaire was administered using sawtooth software.

3.3 Derivation of Choice Probabilities

There are different types of discrete choice models that has been use to estimate WTP. The most common among them is the probit. However, since the discovery of simulation technique mixed logit model has been use by researchers in different field to estimate WTP. This derivation is from Train (2003)'s work. Mixed logit model can be derived under a variety of different behavioural specifications. For utility maximisation, the respondent always choose the alternative that provides the greatest utility. Let U_{nj} be the utility where n is the respondent and J is the alternatives $j = 1, 2, ..., J$. The behavioural model is therefore: choose alternative i if and only if

$$U_{ni} > U_{nj} \forall j \neq i.$$

Some attributes of the alternatives are observed, label $x_{nj} \forall j$ and some attributes of the respondent, label S_n are also observed. We therefore specify a function that relates that all the observed factors to the respondent. The function is denoted by $V_{nj} = V(x_{nj,S_n}) \forall j$ and it is called representative utility.

Since there are aspects that cannot be observe, $V_{nj} \neq U_{nj}$. Utility is decomposed as

$$U_{nj} = V_{nj} + \varepsilon_{nj}$$

where ε_{nj} represent factors that affect utility but are not captured by V_{nj}. This decomposition is fully general, since ε_{nj} is simply defined as the difference between true utility U_{nj} and the utility that is captured in V_{nj}. Given its definition, the characteristics of ε_{nj}, such as its distribution, depend critically on the specification of V_{nj}. ε_{nj} is not defined for choice situation, rather, it is defined relative to our representation of that choice situation.

Since ε_{nj} is on unknown, it is treated as random. The joint density of the random vector

$$\varepsilon'_{nj} = \varepsilon_{n1}, ..., \varepsilon nj$$

is denoted by $f(\varepsilon_n)$. We can therefore make probabilistic statement about the respondent's choice. The probability that the respondent n chooses alternative i is

$$P_{ni} = Prob(U_{ni} > U_{nj} \forall j \neq i)$$

$$= Prob(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \forall j \neq i)$$

$$= Prob(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i)$$

$$P_{ni} = Prob(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i) \tag{3.1}$$

This probability is a cumulative distribution *ie* the probability that each $\varepsilon_{nj} - \varepsilon_{ni}$ is below the observe quantity $V_{ni} - V_{nj}$. For the density $f(\varepsilon_n)$, the cumulative probability can be rewritten as

$$P_{ni} = \int_{\varepsilon} Prob(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i) f(\varepsilon_n) d\varepsilon_n \tag{3.2}$$

where

$$I(.) = \begin{cases} true, & \text{for } (\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i) = 1 \\ otherwise, & \text{for } (\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i) = 0 \end{cases}$$

This is a multidimensional integral over the density of the unobserved portion of utility, $f(\varepsilon_n)$.

Different discrete choice models are obtained from different specifications of this density, that is, from different assumptions about the distribution of the unobserved portion of utility. The integral takes a closed form only for certain specifications of $f(.)$. Logit and nested logit have closed-form expressions for this integral. They are derived under the assumption that the unobserved portion of utility is distributed *iid* extreme value and a type of generalized extreme value, respectively. Probit is derived under the assumption that $f(.)$ is a multivariate normal, and mixed logit is based on the assumption that the unobserved portion of utility consists of a part that follows any distribution plus a part that is *iid* extreme value. With probit and mixed logit, the resulting integral does not have a closed form and is evaluated numerically through simulation.

3.4 Properties of Choice Models

Several aspects of the behavioural decision process affect the specification and estimation of any discrete choice model. The issues can be summarized easily in two statements: " only differences in utility matter" and "The scale of utility is arbitrary."

3.4.1 Differences in Utility Matter

The absolute level of utility is irrelevant to the model. If a constant is added to the utility of all alternatives, the alternative with the highest utility doesn't change. That is the respondent chooses the same alternative with $U_{nj} \forall j$ as with $U_{nj} + K \forall j$ for any constant K The choice probability is

$$P_{ni} = Prob(U_{ni} > U_{nj} \forall j \neq i)$$

$$= Prob(U_{ni} - U_{nj} > 0 \forall j \neq i)$$

, which depends only on the difference in utility not its absolute level. When utility is decomposed into the observed and unobserved parts as in the equation

$$P_{ni} = Prob(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i)$$

which also depends only on differences. This is an indication that parameters that capture differences across alternative can be estimated.

Alternative-specific Constant

$$V_{nj} = x'_{nj}\beta + K_j \forall j$$

where x_{nj} is a vector of variables that relate to alternatives j as faced by the respondent n, β are coefficient of these variables, and K_j is a constant that is specific to alternative j. The utility is linear in parameters with the constant. The alternatives-specific constant for an alternative captures the average effect on utility of all factors that are not included in the model. Thus they serve a similar function as the constant in regression model. When alternative-specific constants are included, the unobserved portion of utility ε_{nj}, has zero mean by construction. If ε_{nj} has a nonzero mean when the constants are not included, then adding the constants makes the remaining error have zero mean. Thus, if

$$U_{nj} = x'_{nj}\beta + \varepsilon^*_{nj} \forall j$$

with

$$E(\varepsilon_{nj})^* = K_j \neq 0$$

, then

$$U_{nj} = x_{nj}\beta + K_j + \varepsilon_{nj} \forall j$$

with $E(\varepsilon_{nj}) = 0$. It is reasonable therefore, to include a constant in V_{nj} for each alternative. However, since only differences in the alternative-specific constants are relevant, not their absolute levels. For this, it is good to set the overall level of these constants. Any model with the same difference in constants is equivalent. In terms of estimation, it is impossible to estimate the two constants themselves, since an infinite number of values of the two constants result in the same choice probabilities. To account for this fact, the absolute levels of the constants must be normalized. The standard procedure is to normalize one of the constants to zero. With j alternatives, at most $J - 1$ alternative-specific constants can enter the model, with one of the constants normalized to zero. It is irrelevant which constant is normalized to zero, the other constants are interpreted as being relative to whichever one is set to zero. It is possible to normalize to other values than zero, however, normalizing to zero is easier.

3.4.2 The scale of utility is arbitrary

The alternative with the highest utility doesn't change no matter how utility is scaled. The model $U^o_{nj} = V_{nj} + \varepsilon_{nj} \forall j$ is equivalent to $U^1_{nj} = \lambda V_{nj} + \lambda\varepsilon_{nj} \forall j$ for any $\lambda > 0$. To take account of this fact, the scale of utility must be normalized. The standard way to do that is to normalize the variance of the error terms. The scale of utility and variance of the error terms are definitionally linked, thus when the utility is multiplied by λ, the variance of the ε_{nj} alters by λ^2 *ie* $var(\lambda\varepsilon_{nj}) = \lambda^2 var(\varepsilon_{nj})$. Therefore, normalizing the variance of the error terms is equivalent to normalizing the scale of utility.

3.5 Mixed Logit Model

The logit model is the most commonly used choice model. However, it exhibit restrictive independence assumption because unobserved characteristics associated with alternatives in a choice situation may be similar. Moreover, unobserved factors that affect the choice in one choice situation may affect the choice in a subsequent choice situation, which induces dependence among choices over time. The more general mixed logit is highly flexible random utility model McFadden and Train (2000) and can overcome these problems by allowing for random taste variation, unrestricted substitution pattens and correlation in unobserved characteristics over different choice situations.

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

where L_{ni} is logit probability evaluated at parameters β

$$L_{ni} = \frac{\ell^{v_{ni}(\beta)}}{\Sigma_{j=1}^{j} \ell^{v_{nj}(\beta)}}$$

v_{ni} is the observed portion of the utility which depends on the parameters β. If utility is linear in β, then $V_{ni} = \beta' x_{ni}$

$$P_{ni} = \int (\frac{\ell^{\beta' x_{ni}}}{\Sigma_j \ell^{\beta' x_{nj}}}) f(\beta) d\beta$$

The mixed logit probability is a weighted average of the logit formula evaluated at different values of β with the weights given by the density $f(\beta)$. The density of β is specified to be normal with mean b and standard deviation s. The choice probability under the density becomes

$$P_{ni} = \int (\frac{\ell^{\beta' x_{ni}}}{\Sigma_j \ell^{\beta' x_{nj}}}) \phi(\beta/b, s) d\beta$$

$\phi(\beta/b, s)$ is the normal density with mean b and standard deviation s.

$$P_{ni} = \int \prod_{t=1}^{T} (\frac{\ell^{\beta' x_{nit}}}{\Sigma_j \ell^{\beta' x_{njt}}}) f(\beta) d\beta \tag{3.3}$$

The panel data structure is presented by the time subscript t and is explicitly modelled since respondents were asked to answer sequential choice card (so that $T = 4$).

Random Coefficients

$$U_{nj} = \beta_n' x_{nj} + \varepsilon_{nj} \tag{3.4}$$

where x_{nj} is observed variables that are related to the alternatives and respondent, β_n is a vector of coefficient of these variables for respondent n representing that respondents' taste and ε_{nj} is a random term which is *iid* extreme value. The coefficients vary over respondents in a population with density $f(\beta)$. This density is a function of parameters θ that represent the mean and covariance of the β's in the population. The respondent knows the value of his/her own β_n and ε_{nj}'s. He/she chooses alternative i if and only if $U_{ni} > U_{nj} \forall j \neq i$. The researcher observes x_{nj}'s not β_n or ε_{nj}'s. If β_n is observed, then the choice probability would be standard logit, since the ε_{nj}'s are *iid* extreme value. The probability condition on β_n is

$$L_{ni}(\beta_n) = \frac{\ell^{\beta_n' x_{ni}}}{\Sigma_j \ell^{\beta_n' x_{nj}}}$$

However, the β_n is unknown so one cannot condition on β. The unconditional choice probability is the integral of $L_{ni}(\beta_n)$ over all possible variables.

$$P_{ni} = \int L_{ni}(\beta_n) f(\beta) d(\beta) \tag{3.5}$$

Specify a distribution for the coefficients and estimate the parameters of the distribution.

3.6 Bayesian Concepts

Consider a model with parameters θ. The ideal about the parameters are represented by a probability distribution over all possible values that the parameters can take. These ideals are represented by a density on θ, called the prior distribution and denoted $k(\theta)$. Data are collected to improve ideals about the value of θ. Suppose a sample size of N is observed. Let y_n denote the observed choice (or choices) of respondents n and let the set of observed choices for the entire sample be labelled as $Y = y_1, ...y_n$. Based on this sample information, ideals about θ are updated. This is done by representing the updated ideals by new density on θ, labelled $K(\theta/Y)$ called the posterior distribution. This posterior distribution depends on Y

3.6.1 Bayes' Rule

let (y_n/θ) be the probability of outcome (y_n) for respondent n. This probability is the behavioural model that relates the explanatory variables and parameters to the outcome. The probability of observing the sample outcomes Y is $L(Y/\theta)=\Pi_{n=1}^{N} p(y_n/\theta)$

This is the Likelihood function of the observed choices which is a function of the parameters θ. By this rule of conditioning, ways are provided by Bayes' rule to improve on θ.

$$K(\theta/Y)L(Y) = L(y/\theta)k(\theta) \tag{3.6}$$

where $L(Y)$ is the marginal probability of Y, marginal over θ. Both side of equation 3.6 represent the joint probability of Y and θ with the conditioning in opposite directions. The left hand side is the probability of Y times the probability of θ given Y, while the right hand side is the probability of θ times the probability of Y given θ. This then become

$$K(\theta/Y) = \frac{L(Y/\theta)k(\theta)}{L(Y)} \tag{3.7}$$

This equation 3.7 is Bayes' rule applied to prior and posterior distributions. Bayesian statistics arises when the unconditional probability is the prior distribution and the conditional probability is the posterior distribution. The marginal probability of Y, $L(Y)$, is constant with respect to θ and, more specifically is the integral of the numerator of equation 3.7. Also, $L(Y)$ is simply the normalizing constant that assures that the posterior distribution integrates to 1, as required for any proper density. By this fact, posterior distribution is proportional to the prior distribution times the likelihood function.

$$K(\theta/Y) \propto L(Y/\theta)k(\theta) \tag{3.8}$$

The mean of the posterior distribution is

$$\bar{\theta} = \int \theta K(\theta/Y) d\theta \tag{3.9}$$

$\bar{\theta}$ is the value of θ that minimizes the expected cost of being wrong about θ, if the cost of error is quadratics in the size of the error. If say θ_o is use in decision when the true value is $\theta*$, the cost of being wrong is

$$C(\theta_o, \theta*) = (\theta_o - \theta*)' B(\theta_o - \theta*)$$

where B is a matrix of constants. The true value of θ is unknown, however, it is belief that it has it value in $K(\theta/Y)$. Therefore, the expected value of cost be wrong can be found by using the value of θ_o.

$$EC(\theta_o) = \int C(\theta_o, \theta) K(\theta/Y) d\theta$$

$$= \int (\theta_o - \theta)' B(\theta_o - \theta) K(\theta/Y) d\theta$$

The value of θ_O that minimizes this expected cost is determined by differentiating $EC(\theta_o)$, then set it to zero and solve for θ_o

$$\frac{\partial EC(\theta_o)}{\partial \theta_o} = \int \frac{\partial[(\theta_o - \theta)' B(\theta_o - \theta)]}{\partial \theta_o} K(\theta/Y) d\theta$$

$$= \int 2(\theta_o - \theta)' B K(\theta/Y) d\theta$$

$$= 2\theta_o' B \int K(\theta/Y) d\theta - 2(\int \theta K(\theta/Y) d\theta)' B$$

$$= 2\theta_o' B - 2\bar{\theta}' B$$

$$2\theta_o' B - 2\bar{\theta}' B = 0$$

$$\theta_o' B = \bar{\theta} B$$

$$\theta_o = \bar{\theta}$$

The mean of the posterior, $\bar{\theta}$, is the value of θ that would optimally act upon, if the cost of being wrong about θ rises quadratically with the distance to the true θ. To calculate the mean of the posterior distribution, simulation procedures are required.
A simulated approximation of this integral is obtained by taking draws of θ from the posterior distribution and averaging the results. The simulated mean is

$$\check{\theta} = \frac{1}{R}\Sigma_{r=1}^{R}\theta^{r} \tag{3.10}$$

where θ^r is the rth draw from $K(\theta/Y)$. The standard deviation of the posterior, which serves as the standard error of the estimates, is simulated by taking the standard deviation of the R draws. The simulated mean of the posterior (SMP) are attained with more relaxed conditions on the number of draws.

3.6.2 Hierarchical Bayes for Mixed Logit

In all statistical analysis, there are three kinds of concepts: data, models and parameters Gelman et al. (1995). For this study, data are the choices that respondents make. Models are the assumptions about data, for example, the distribution for this study is normal distribution. The Parameters are numerical values in the models. The parameters for this study is the mean and the standard deviation.

The hierarchical Bayesian for mixed logit was developed by Allenby (1997). Sawtooth software is used to implement it.

Let the utility of respondent n obtains from alternative j in time period t be

$$U_{njt} = \beta_n' x_{njt} + \varepsilon_{njt}$$

ε_{njt} is iid extreme value and $\beta_n \sim N(b, W)$

$y_{nt} = i$ is the observed choice (alternative) $\iff U_{nit} > U_{njt} \forall j \neq i$

Prior: $k(b, W)$ where $k(b)$ is $N(b_0, s_0)$ with extremely large variance; $k(W)$ is $IW(d, I)$. The prior on W is inverted Wishart with d degrees of freedom and scale matrix I. Sample of N respondents is observed, the chosen alternatives in all time periods for respondent n is denoted $y_n = y_{n1}, ..., y_{nT}$, add the choices of the entire sample to get $Y = y1, ..., y_n$. The probability of respondent n's observed alternative, conditional on β is

$$L(y_n/\beta) = \Pi_t(\frac{\ell^{\beta' x_{nit}}}{\Sigma_j \ell^{\beta' x_{njt}}})$$

The probability not conditional on β is the integral of $L(y_n/\beta)$;

$$L(y_n/b, W) = \int L(y_{n/}\beta)\Phi(\beta/b, W)d(\beta)$$

where $\Phi(\beta/b, W)$ is the normal density with mean b and variance W. $L(y_n/b, W)$ is the mixed logit probability.The posterior distribution of b and W is

$$K(b, W/y_n) \propto \Pi_n(L(y_n/b, W)k(b, W) \tag{3.11}$$

where $k(b, W)$ is the prior on b and W. It is possible to draw directly from $K(b, W/y_n)$ with Metropolis Hasting (MH) algorithm. The disadvantage of MH algorithm is that, it is computationally very slow. It would be necessary to calculate the right-hand side of equation 3.6. The choice probability $L(y_n/b, W)$ is an integral without a closed form and must be approximated through simulation. Each iteration of the MH algorithm would require simulation of $L(y_n/b, W)$ for each n.

However, since the computation is time consuming draws can be taken from the posterior without the need to simulate the choice probabilities. Drawing from $K(b, W/y_n)$ is fast simple if each β_n is considered to be a parameter along with b

and W. The posterior for b, W and β_n $\forall n$

$$K(b, W, \beta_n \forall n / Y) \propto \Pi_n (L(y_n / \beta_n) \Phi(\beta_n / b, W) k(b, W)$$

Draws from this posterior are obtained through Gibbs sampling which is also called Monte Carlo Markov Chain ($MCMC$). A draw of each parameter is taken conditional on the other parameters *ie* draw $\beta_n \forall n$ conditional on values of b and W. The posterior for each respondent's β_n, conditional on their choices and the population mean and variance of β_n is

$$K(\beta_n / b, W) \propto (L(y_n / \beta_n) \Phi(\beta_n / b, W) \forall n$$

$K(b, W, \beta_n \forall n)$ is $N \sim (\bar{\beta}, W/N)$,, where $\bar{\beta} = \Sigma_n \beta_n / N$, $K(W/b, \beta_n \forall n)$ is $IW(d + N, \frac{dI + N\bar{S}}{d+N})$ where $\bar{S} = \Sigma_n (\beta_n - b)(\beta_n - b)' / N$

3.6.3 The Metropolis Hasting (MH) Algorithm

(a) Start with a value β_n^o (b) Draw K independent values from a standard normal density and stack the draws into a vector labelled η'. (c) Create a trial value of β_n' as

$$\bar{\beta'}_n = \beta_n^o + \rho L \eta'$$

, where ρ is a scalar which need to be specified, L is the choleski factor of W (d) Draw a standard uniform variable μ' (e) Calculate the ratio

$$F = \frac{L(y_n / \bar{\beta}') \phi(bar \beta_n' / b, W)}{L(y_n / \beta_n'^o) \phi(\beta_n'^o / b, W)}$$

(f) If $\mu' \leq F$, accept $\bar{\beta}_n'$ and let $\beta_n' = \bar{\beta}_n{}'$. If $\mu' > F$, reject $\bar{\beta}_n$ and let $\beta_n' = \beta_n^o$ (g) Repeat the process many times

3.6.4 Gibb Sampling

The t^{th} iteration of the Gibbs sampler consists of these steps 1. Draw b^t from $N \sim (\beta^{-t-1}, W^{t-1}/N)$, where β^{-t-1} is the mean of the β^{t-1}'s 2. Draw W_t from $IW(d+N, (dI+NS^{t-1})/(d+N)$ where

$$s^{t-1} = \Sigma_n(\beta_n^{t-1} - b^t)(\beta_n^{t-1} - b^{t\prime})/N$$

3. For each n, draw β_n^t using one iteration of the MH algorithm, starting from β^{t-1} and using the normal density $\phi(\beta_n/b^t, W^t)$. The process is repeated for many iterations. The resulting values converge to draws from the joint posterior of b, W and $\beta_n \forall n$. Once the converged draws from the posterior are obtained, the mean and standard deviation of the draws can be calculated to obtain estimates and standard errors of the parameters. The procedure provides information about β_n for each n. The iterations prior to convergence are called burn-in.

Chapter 4

Analysis

4.1 Introduction

This chapter discuses how data collected was used for the intended analysis based on the Hierarchical Bayesian discussed in chapter three. Respondents for the study were second hand clothes sellers who were selected from Kumasi Central Market. A total of 137 traders filled out the questionnaire which was administered using sawtooth software. Each of the respondents was presented with 12 hypothetical choice situations. The sawtooth software and Stata 12 were used for the analysis. See appendix C for example of the choice card.

4.2 Socio-economic Observations

The survey had more females (72 %) than males. The average age of the respondents was 43 years. The proportion of the respondents who were married was about 56 %. The proportion of respondents who had more that two children was 32%. 90 % of the respondents had at least basic school leaving certificate. Families with more children and with a high education value insurance more than smaller families and those with a low education (Botzen and van den Bergh, 2008). Most of the respondents (41 %) had goods which value GHS 4000.00. According to Botzen and van den Bergh (2008), people with a higher value of property are more likely to self-insure and demand less disaster insurance. The after-tax income of the traders was less than GHS 700.00. This suggest that the income levels of the country may be low. Low income individuals have less taste for insurance Slovic (2000). See Table 4.1 for the descriptive statistics.

Table 4.1: Descriptive Statistics

	N.obs.	Proportion	Response
Experience with fire and evacuation	137	0.53	Yes
Experience of fire damage	137	0.56	Yes
Knowledge on the causes of fire outbreaks	137	0.91	Yes
Illegal/faulty electrification causes higher fire risk	137	0.97	Yes
Fire outbreak is exogenous to human control	137	0.65	Disagree
Risk of suffering fire damage	137	0.40	High risk
Distance from fire threat	137	0.89	No
Lower fire risk than average residential area	137	0.65	No
Expected fire damage	137	0.52	GHS 1500.00
Zero expected return period	137	0.73	Yes
Insurance purchase index	137	0.78	No
Risk seeking index	137	0.45	Not risk averse
Government relief	137	0.67	Yes
Age	137	033	42.50
Sex	137	0.72	Female
Value of goods	137	0.86	GHS 2000
Marital Status	137	0.56	Married
Children	137	0.63	$\geqslant 2$
JHS graduate	137	0.90	Yes
Income	137	0.78	$\geqslant$ GHS 700

4.3 Examination of Fire Risk at Kumasi Central Market

The proportion of respondents who had experienced fire outbreaks and had been evacuated was 53 %. 56 % of the respondents had experienced fire damage. Individuals who have experienced a fire outbreaks and have been evacuated for a threat of fire outbreaks are more likely to demand insurance (Michel-Kerjan and Kousky, 2008). However, 58 % of the traders did not want to purchase insurance. This might due to their low levels of after-tax income. Most of the traders (91 %) have knowledge on the causes of fire outbreaks. 97 % of the traders believed that fire outbreaks at the markets is due to illegal/ faulty electrification.

Majority (65 %) of them disagreed that fire outbreaks at the markets is as a result of climate change or natural conditions. This is an indication that most of the fire outbreaks are human controlled. 40 % of the respondents reported that there was a high risk of fire outbreak one can suffer from at the market. Adverse selection could hamper the development of fire insurance market if only high risk individuals who live in unprotected areas are interested in purchasing insurance and insurers are unable to adequately distinguish low from high risk customers and charge the latter a higher premium. This further provides relevant insight into risk characteristics of individuals with low-probability of fire risk.According to Botzen and van den Bergh (2008), high-risk individuals place a larger value on fire insurance coverage than individuals who face a lower risk.

64 % of the traders perceived that there was high fire outbreak risk at the market compared to the average residential areas. The probability of choosing fire insurance is lower if individuals expect their fire risk to be lower than an average residential area (Slovic, 2000). The probability of choosing fire insurance is also lower if individuals expect that the return period of fire outbreak equal to zero (Slovic, 2000). Majority (73 %) of the traders expected that the return period of fire outbreak equals zero. This confirms Slovic (2000) assertion. The average fire damage by the traders was GHS 1500.00. Fire insurance is positively related to expected fire damage (Botzen and van den Bergh, 2008). This means that there is a possibility that demand for fire insurance by the traders could be high. This explains why 52 % of the traders took no insurance option.

Purchasing insurance may be a good indicator of risk aversion since they represent revealed preferences for financial protection Michel-Kerjan and Kousky (2008). 78 % of the respondents had not purchased one or more of the following insurance: health insurance, life, funeral, education, home, home contents, continuous travel, all-risk car, and disability insurance. This is an indication that individuals many not purchase fire insurance. A risk seeking index has been derived by asking individuals how well they correspond to a risk averse. Most risk seeking individuals are less likely to insure. 45 % of the respondents risk seeking index is not risk averse.

A variable was included about the availability of compensation of fire damage via the government to estimate differences in insurance demand. 65 % of the traders perceived that government compensate fire damage. This makes choosing insurance not attractive. This crowds out demand for private insurance. Perception of the risk of fire outbreaks are important determinant in the choice of fire insurance. See Table 4.1 and appendix A.

4.4 Estimation Results of the Choice Model for Fire Insurance

The discrete choice experiment was unlabelled, which indicates that there was no rationale to expect a general preference for one of the three situations with insurance shown to respondents. The three situations (A, B and C) were chosen 11 %, 18 % and 12% in that order. The " opt out " or no insurance was chosen 59 % of the time.

The concepts of Hierarchical Bayesian for mixed logit with normally distributed coefficients was introduced by Allenby (1997). He showed how the parameters of the model can be estimated without the need to calculate the choice probabilities. The Hierarchical Bayesian procedures was used because of the following reasons: first, it does not require maximization of the function. Second, desirable estimation properties, such as consistency and efficiency can be attained with Bayesian procedures (Train, 2001). For Bayesian estimators, fixed number of draws are used in simulation and are efficient if the number of draws rises at any rate with sample size (Train, 2001). The Bayesian procedures uses an iterative process. For convergence, there should be sufficient number of iterations.

For Hierarchical Bayesian procedures, the draws are correlated over iterations even after convergence has been achieved. This is because each iteration builds on the previous one. 100000 iterations of $MCMC$ were specified in total. 20000 for burn in and 80000 after convergence. To avoid correlation of draws, every 100th draws were retained and the rest discarded. A total of 1000 draws was retained. The mean and standard deviations of these draws constitutes the estimates.

The following utility specification was used for the model that includes only the attributes of the experiment.

$$U_{insurance} = \beta_1 * fire + \beta_2 * damage_1 + \beta_3 * damage_2 + \beta_4 * coverage + \beta_5 * price$$

and

$$U_{noinsurance} = \beta_6 * contant \tag{4.1}$$

The utility of having insurance is dependent on the expected number of fire outbreaks (fire), expected fire damage, insurance coverage and insurance premium (price). There is no a prior reason to expect that the attributes have a different effect on the utility in the generic situations A or B or C. The utility of the option without insurance is modelled with coefficient of constant term (none). The Table 4.2 below shows the estimation results of equation 4.1. The model was estimated under the assumption that the coefficients are independently normally distributed in the population. That is, $\beta_n \sim N(b, W)$ with population. The population parameters are the mean and standard deviation of each coefficient. Table 4.2 gives the simulated mean of the posterior of these parameters. The parameters of the attributes are the same for all situations because the experiment was unlabelled.

Table 4.2: Mixed Logit Model Estimated Parameters

		Coefficients
Premium (price)	**Mean (Std. error)**	-0.084(0.015)
	Std. dev.(Std. error)	0.049(0.003)
Damage of GHS 1000	**Mean (Std. error)**	-0.046(0.028)
	Std. dev. (Std. error)	0.032(0.021)
Damage of GHS 2000	**Mean (Std. error)**	0.067(0.002)
	Std dev. (Std. error)	0.047(0.022)
High insurance coverage (100 %)	**Mean (Std. error)**	0.029(0.013)
	Std. dev. (Std. error)	0.034(0.011)
Number of fire outbreak in a year	**Mean (Std. error)**	0.021(0.019)
	Std. dev. (Std. error)	0.030(0.012)
" Opt out"	**Mean (Std. error)**	8.510 (1.398)
	Std. dev. (Std. error)	6.274(3.866)

The damage and the coverage variables were coded using dummy. The third option of damage *ie* GHS 5000.00 and the second option of coverage *ie* 75 %, were excluded so that the coefficients β_2, β_3 and β_4 measure the effect relative to having insurance under damage of GHS 5000.00 and coverage of 75 %, in that order. Using dummies allows one to examine non-linear effects without restricting the functional form of this non-linearity.

The attributes of the choice experiment determine the utility of having insurance so that a positive coefficient indicates a positive relation between the attributes and the value placed on insurance. The utility of insurance increases with damage of GHS 2000.00, high insurance coverage and number of fire outbreaks in a year. It however, decreases with damage of GHS 1000.00. This is an indication that utility of insurance increases with high damage, high insurance coverage and number of fire outbreaks that occurs in a year. The percent certainty which measures goodness of fit of the model was 77.2 %. This is an indication that all the coefficients of the attributes were statistically significant.

All the parameters of the attributes except that of price and fire were specified to be random so as to capture unobserved individual heterogeneity in the response. The percent certainty also indicates that preference heterogeneity exists in the coefficient.

It is noted that the " opt out " variable determine the utility of having no insurance. The positive coefficient of the " opt out " indicates that a negative relation between the " opt out " and the value placed on fire insurance. See the variance covariance matrix below.

Table 4.3: Variance-covariance matrix

$$\begin{bmatrix} 0.779 & & & & \\ -0.081 & 0.994 & & & \\ 0.226 & -0.110 & 2.332 & & \\ 0.004 & 0.025 & 0.054 & 0.385 & \\ -0.492 & -.2.395 & -14.906 & -1.026 & 0361.766 \end{bmatrix}$$

4.5 Willingness to pay Estimates for Fire Insurance using Mixed Logit

Willingness to pay (WTP) measures for changes in the attribute values are commonly computed as the ratio of the coefficient value of the attribute of interest to the coefficient of the cost attribute. The price coefficient was held fixed, so that the distribution of WTP is simply the distribution of the attribute's coefficient. Table 4.4 shows the monthly WTP estimates for the damage of GHS 1000.00 and coverage attributes of the choice experiment. The ratios $-\frac{\beta_2}{\beta_5}$ and $-\frac{\beta_4}{\beta_5}$ give the estimates for damage and coverage for individuals respectively. The mean WTP

Table 4.4: WTP estimates

		Coefficients
Damage of GHS 1000.00	Mean	-0.55
	Std. dev.	1.87
High insurance premium	Mean	0.35
	Std. dev.	0.87

per damage of GHS 1000.00 is -0.55. This is an indication that a rise in expected damage would increase WTP by GHS 0.55. The WTP estimate for coverage indicates that increasing coverage by about 10 percent point increases WTP by about GHS 3.5 per month for individuals. Individual with high cost of damage and need total coverage of insurance is require to pay high premium.

Chapter 5

Conclusion

5.1 Introduction

This paper has examined the mixed logit model and how it is used to estimate insurance premium. The survey has been conducted among second hand clothes sellers in Kumasi Central Market which is vulnerable to fire outbreaks. The chapter gives conclusion and recommendations for further research work.

5.2 Conclusion

The results of this survey provided three main insights for the feasibility of introducing fire insurance: There is positive relationship between utility and high damage cost, high fire risk and high insurance coverage, however, utility decline with insurance premium; Estimation results support that offering fire insurance may be profitable in the current situation; and the effect of crowding out of demand by the availability of government compensation is shown to be considerable.

With respect to the first point, traders with high damage cost and trade at where fire risk is high want total insurance coverage, however, paying of equal insurance premium has been an obstacle. This may due to the economy of these individuals. secondly, it should be noted that a considerable proportion of the traders are willing to pay for fire insurance. Demand is expected to increase on the condition that government abolishes the current regulation of damage relief and refrains from compensating fire damage, according to the third point. However, there remains a large proportion of traders that are unwilling to insure even if the government regulation will be abolished. This could be overcome by making insurance coverage compulsory for all traders.

5.3 Recommendations

It is therefore recommended that, the model should be study by the stake holders in the insurance industry to ensure traders are captured to protect them from fire risk.

The estimation results suggest that offering fire insurance may be profitable in the current situation. The government and insurers should investigate the possibilities of introducing fire insurance.

An important lesson for policy-makers is that they can play a significant role in stimulating or at least not preventing the emergence of private insurance markets by refraining from ex-post damage compensations.

Further research is recommended on how to use mixed logit to solve problems in other sectors of the economy such as education, health, marketing and so on.

REFERENCES

Adamowicz, W. L. (1994). Habit formation and variety seeking in a discrete choice model of recreation demand. *Journal of Agricultural and Resources Economics*, 19(3):19–31.

Adofo, R. (2013). Insightful expert's view on the possible causes of market fire outbreaks in ghana. *ModernGhana*, 256:2–3.

Algers, S., Bergström, P., Dahlberg, M., and Dillén, J. L. (1998). Mixed logit estimation of the value of travel time. In *JEL Classification: C15, C25, R41*.

Allenby, G. (1997). An introduction to hierarchical bayesian modeling. In *Tutorial Notes, Advanced Research Techniques Forum*. American Marketing Association.

Arrow, K., Solow, R., Portney, P., Leamer, E., Radner, R., and Schumann, H. (1993). Report of the noaa panel of contingent valuation. *Federal Register*, 10:4601–4614.

Bekker-Grob, E. W. D., Hol, L., Donkers, B., van Dam, L., Habbema, J. D. F., van Leerdam, M. E., kuipers, E. J., Essink-Bot, M. L., and Steyerberg, E. W. (2010). L abeled versus unlabeled discrete choice experiment in health economics: an application to cloorectal cancer screening. *Value Health*, 13(2):315–323.

Ben-Akiva, M. E. and Lerman, S. R. (1985). *Dicrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, Mass.

Bhat, C. (2000). *Flexible model structures for discrete choice analysis. Handbook of transport modelling*.

Bhat, C. R. (1998). Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling. *Transportation Research Part A*, 32:495–507.

Blamey, R. K., Bennett, J. W., Louviere, J. J., Morrison, M. D., and Rolfe, J. C. (2000). A test of policy labels in environmental choice modelling studies. *Ecological Economics*, 32(2):269–286.

Boatwright, P., Borle, S., and Kadane, J. (2003). A model of the joint distribution of purchase quantity and timing. *Journal of the American Statistical Association*, 98:564-572.

Boccaletti, S. and Moro, D. (2000). Consumer willingness-to-pay for gm food products in italy. *AgBioForum*, 3(4):259–267.

Botzen, W. J. W. and van den Bergh, J. C. J. M. (2008). Insurance against climate change and flooding in the netherlands: Present, future and comparison with other countries. *Risk Analysis*, 28(2):413–426.

Brownstone, D., Bunch, D., and Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research B*, 34:315–338.

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

Campbell, D. (2006). Combining mixed logit models and random effects models to identify the determinants of willingness to pay for rural landscape improvements. In *Paper written for the Agricultural Economics Society 81st Annual Conference Reading*.

Carson, R. T., Flores, N. E., and Meade, N. S. (2001). Contingent valuation: controversies and evidence. *Environmental and Resource Economics*, 19:173–210.

Crichton, D. (2008). Role of insurance in reducing flood risk. *The Geneva Papers*, 33(1):117–132.

Czajkowski, M. and Hanley, N. (2009). Using labels to investigate scope effects in stated preference methods. *Environment and Resource Economics*, 44(4):521–535.

de Vries, E. J. (1998). Vergoeding van rampschade. *Nederlands Juristenblad*, 42:1908–915.

Doherty, E., Campbell, D., Hynes, S., and van Rensburg, T. (2011). The effect of using labelled alternatives in stated choice experiments: an exploration focusing on farmland walking trails in ireland. In *Paper proposal for the Second International Choice Modelling Conference Annual Conference*.

Dube, J. P., Hitsch, G. J., and Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *Rand Journal of Economics*, 41:417–445.

Eboli, L. and Mazzulla, G. (2008). Willingness-to-pay of public transport users for improvement in service quality. *European Transport*, 38:107–118.

Fournier, G. M. and Gai, Y. (2006). What does willingness-to-pay reveal about hospital market power in merger cases. In *JEL Classification: L40, I11, I18*.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (1995). *Bayesian Data Analysis*. Chapman and Hall, Suffolk.

Gu, Y., Norman, R., and Viney, R. (2007). Estimating health stat utility values from dicrete choice experiments- a qaly space model approach. In *Centre for Health Economics Research and Evaluation*, Sydney, Australia. University of Technology.

Haan, P., Kemptner, D., and Uhlendorff, A. (2012). Bayesian procedures as a numerical tool for the estimation of dynamic discrete choice models. In *Discussion Paper No. 6544*.

Hall, J., Fiebig, D. G., King, M. T., Hossain, I., and Louviere, J. (2006). What influences participation in genetic carrier testing? results from a discrete choice experiment. *Journal of Health Economics*, 25:520–537.

Han, B., Algers, S., and Engelson, L. (2001). Accommodating drivers' taste variation and repeated choice correlation in route choice modeling by using the mixed logit model. In *Proceedings of the 80th Annual Meeting of the Transportation Research Board*.

Hanley, N., Wright, R. E., and Koop, G. (2002). Modelling recreation demand using choice experiments: Climbing in scotland. *Environmental and Resource Economics*, 449-466:22.

Harrington, S. E. (2000). Rethinking disaster policy: Breaking the cycle of free disaster assistance, subsidized insurance, and risky behaviour. *Regulation*, 23(1):40–46.

Hensher, D. A., Greene, W. H., and Rose, J. M. (2003). Using classical inference methods to reveal individual-specific parameter estimates to avoid the potential complexities of wtp derived from population moments. In *Working Paper ITS-WP-03-18*. Institute of Transport Studies, the University of Sydney and Monash University.

Hess, S., Bierlaire, M., and Polak, J. (2004). Estimation of value of travel-time savings using mixed logit models. *Transportation Research Part A*, 39(3):221–236.

Hole, A. R. (2008). Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment. *Journal of Health Economics*, 27:1078–1094.

Hole, A. R. and Kolstad, J. R. (2010). Mixed logit estimation of willingness to pay distributions: A comparison of models in preference and wtp space using data from a health- related choice experiment. In *Working Papers in Economics*, number 3.

Johnston, R. J., Swallow, S. K., and Bauer, D. M. (2002). Spatial factors and stated preference values for public goods: considerations for rural land use. *Land Economics*, 78:481–500.

King, M. T., Hall, J., Lancsar, E., Fiebig, D., Hossain, I., Louviere, J., Reddel, H., and Jenkins, C. R. (2007). Patient preferences for managing asthma: results from a discrete choice experiment. *Health Economics*, 16:703–717.

Kunreuther, H. C., Michel-Kerjan, E. O., Doherty, N. A., Grace, M. F., Klein, R. W., and Pauly, M. V. (2008). Managing large-scale risks in a new era of catastrophes: Insuring, mitigating and financing recovery from natural disasters in the united states. Master's thesis, Whartson Risk Management and Desicion Processes Center, Georgia State University, and the Insurance Information Institute, Philadelphia.

Laeven, R. J. A. and Goovaerts, J. M. (2011). *Premium Calculation and Insurance Pricing*. R J A Laeven and J M Goovaerts, 3 edition.

Lancaster, K. J. (1966). A new approach to consumer theory,. *The Journal of Political Economy*, 74(2):132–157.

Luan, C. (2001). Insurance premuim calculations with anticipated utility theory. *Astin Bullentin*, 31(1):23–35.

Matul, M., McCord, M., Phily, C., and Harms, J. (2010). The landscape of microinsurance in africa. In *Geneva: International Labour Organization*.

McClure, S. M., Li, J., Tomlin, D., Cypert, K., Montague, L. M., and Montague, P. (2004). Neural correlates of behavioural preference for culturally familiar drinks. *Neuron*, 44:379–387.

McFadden, D. and Train, K. (1997). Mixed multinomial logit models for discrete response, working paper. In *Department of Economics*, Berkeley. University of California.

McFadden, D. and Train, K. (2000). Mixed mnl for discrete response. *Journal of Applied Econometrics*, 15:447–470.

McFadden, D. L. (1974). *Conditional logit analysis of qualitative choice behavior, Frontiers in Econometrics.* Academic Press, New York.

Michaud, C., Llerena, D., and Joly, I. (2012). Willingness to pay for environmental attributes of non food agricultural products: A real choice experiment. In *Working Paper GAEL; 2012-01.*

Michel-Kerjan, E. and Kousky, C. (2008). Come rain or shine: Evidance on flood insurance purchases in florida. Woeking paper, The Wharton School of University of Pennsylvania and The Kennedy School of Government Harvard University.

Munizaga, A. M. (2000). Evaluation of mixed logit as a practical modelling alternative. Universidad De Chile, Casilla 228-3, Santiago, Chile.

Negrín, M. A., Pinilla, J., and León, C. J. (2008). Willingness to pay for alternative policies for patients with alzheimer's disease. *Health Economics, Policy and Law*, 3:257–275.

Oliver, A., Healey, A., and Donaldson, C. (2002). Choosing the method to match the perspective: economic assessment and its implications for health-services efficiency. *Lancet*, 359:1771–1774.

Ozaki, A. V., Ghosh, S. K., Goodwin, K. B., and Shirota, R. (2005). Spatio-temporal modeling of agricultural yield data with an application to pricing crop insurance contracts. In *Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting*. Providence, Rhode Island.

Ozdemir, S., Johnson, F. R., and Hauber, A. B. (2009). Hypothetical bias, cheap talk, and stated willingness to pay for health care. *Journal of Health Economics In press*, 39:234–245.

Paterson, R. W., Boyle, K. J. L. K. J., Parmeter, C. F., Neumann, J. E., and Civita, P. D. (2008). Heterogeneity in preferences for smoking cessation. *Health Economics*, 17:1363–1377.

Pielke, R. A., Prins, G., Rayner, S., and Sarewitz, D. (2007). Lifting the taboo on adaptation. *Nature*, 445:597–598.

Prowse, V. (2010). Modeling employment dynamics with state dependence and unobserved heterogeneity. In *IZA Discussion-Paper, 4489*.

Regier, D. A., Ryan, M., and andC A Marra, E. P. (2009). Bayesian and classical estimation of mixed logit: An application to genetic testing. *Journal of Health Economics*, 28:598–610.

Revelt, D. and Train, K. (1999). Customer-specific taste parameters and mixed logit. In *Department of Economics*, Berkeley. University of California.

Revelt, D. and Train, K. E. (1998). Mixed logit with repeated choices: Households' choices of appliance effciency level. *The Review of Economics and Statistics*, 80:647–657.

Ryan, M. (2004). A comparison of stated preference methods for estimating monetary values. *Health Economics*, 13:291–296.

Scarpa, R., Thiene, M., and Train, K. (2008). Utility in wtp space: a tool to address confounding random scale effects in destination choice to the alps. *American Journal of Agricultural Economics*, 90(4):994–101.

Sennhauser, M. (2010). Why the linear utility function is a risky choice in discrete-choice experiments. In *Working Paper*, number 1014.

Shen, J. and Saijo, T. (2009). Does an energy efficient label alter consumers' purchasing decisions? a latent class approach based on a stated choice experiement in shanghai. *Journal of Environmental Management*, 90(1):3561–3573.

Slovic, P. (2000). *Perception of Risk*. Earthscan, London.

Spanier, J. and Maize, E. (1991). Quasi-random methods for estimating integrals using relatively small samples. *SIAM Review*, 36:18–44.

Tan, V., Allen, and LLP, O. (2012). Microinsurance. In *Advocates For International Development*.

Train, K. (1998). Recreation demand models with taste variation. *Land Economics*, 74(2):230-239.

Train, K. (2001). A comparison of hierarchical bayes and maximum simulated likelihood for mixed logit. In *Paper presented at the Asilomar Invitational Choice Symposium*, California.

Train, K. and Sonnier, G. (2003). Mixed logit with bounded distributions of partworths. In *Working Paper*, Berkeley. Department of Economics, University of California.

Train, K. and Weeks, M. (2005). Applications of simulation methods in environmental and resource economics. In *Working Paper*, Berkeley. Department of Economics, University of California.

Train, K. and Wilson, W. (2007). Estimation on stated-preference experiments constructed from revealed-preference choices. *Transportation Research B, forthcoming*, 1:28,29.

Train, K. E. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press.

van den Berg, G. J. (2001). *Duration Models: Specification, Identification, and Multiple Durations*, volume 5. Handbook of Econometrics, North-Holland, Amsterdam.

Wuthrich, M. V. (2003). Asymptotic value-at-risk estimates for sums of dependent random variables. *Astin Bullentin*, 33(1):75–92.

Yannis, G. and Antoniou, C. (2007). A mixed logit model for the sensitivity analysis of greek drivers' behaviour towards enforcement for road safety. *European Transport*, 37:62–77.

Appendix A

Table 5.1: Descriptive Statistics of the Explanatory Variables

	N.obs.	Mean	Std. Dev.
Experience with fire and evacuation	137	1.47	0.50
Experience of fire damage	137	1.44	0.49
Knowledge on the causes of fire outbreaks	137	1.09	0.28
Illegal/faulty electrification causes higher fire risk	137	1.03	0.17
Fire outbreak is exogenous to human control	137	2.33	1.25
Risk of suffering fire damage	137	4.00	1.47
Distance from fire threat	137	1.65	0.4
Lower fire risk than average residential area	137	1.65	0.48
Expected fire damage	137	GHS 1500.00	1.67
Zero expected return period	137	1.27	0.45
Insurance purchase index	137	1.72	0.45
Risk seeking index	137	3.28	1.03
Government relief	137	1.32	0.47
Age	137	42.5	1.63
Female	137	1.72	1.28
Value of goods	137	GHS 2000	0.41
Marital Status	137	1.69	0.59
Children	137	1.37	0.48
JHS graduate	137	1.09	0.29
Income	137	700	0.60

Appendix B

Questionnaire

I would be much honoured if you give me some few minutes to enable me ask you some few questions that would assist me to determine microinsurance premiums. It is purely academic exercise. Your candid response shall be deemed very confidential and useful for this study.

Q1. Have you ever experienced fire outbreak(s) and have been evacuated?

Yes

No

Q2. Have you ever experienced fire damage(s)?

Yes

No

Q3. Do you have knowledge on the cause(s) of fire outbreak?

Yes

No

Q4. Does illegal/faulty electrification causes higher risk of fire outbreaks

Yes

No

I don't know

Q5. Climate or natural conditions and not human control are causes of fire outbreaks

Strongly disagree

Disagree

Unknown

Agree

Strongly agree

Q6. Is there a fire outbreak risk that one can suffer from at this market?

No risk

Extremely low risk

Low risk

Unknown

High risk

Extremely high risk

Q7. Are you 2 kilometres away from fire threat?

Yes

No

I don't know

Q8. There is very low fire outbreak risk at the market compared to average residential areas?

Yes

No

I don't know

Q9. If your goods were to be destroyed by fire, what will be the total damage?

< GHS 1000.00

GHS 1000.00-2000

GHS 2000.00-3000.00

GHS 3000.00-4000.00

GHS 4000.00-5000.00

> GHS 5000.00

Q10. Do you expect zero return period fire outbreak?

Yes

No

Q11. What fire outbreak return period do you expect?

1-10 years

10-20 years

20-30 years

30-40 years

40-50 years

50-60 years

None

Q12. Have you purchased one or more of the following insurances: health insurance, life, funeral, education, home, home contents, all-risk car and disability insurance

Yes

No

Q13. Risk seeking index

Very risk averse

Risk averse

Unknown

Not risk averse

Not risk averse at all

Q14. Does the government compensate fire victims?

Yes

No

Close Window

Q15. Do you want insurance? Yes

No

Q16. Age (in years)

< 20

20-25

25-30

30-35

35-40

40-45

45-50

50-55

55-60

60-65

> 65

Q17. Sex

Female

Male

Q18. Value of your goods

$<$ GHS 1000.00

GHS 1000.00-3000.00

GHS 3000.00-5000.00

$>$ GHS 5000.00

Q19. Marital status

Single

Married

Widowed/divorced/separated

Q20. Number of children

0-2

> 2

Q21. Your level of education is at least JHS

Yes

No

Q22. What is your monthly income after tax?

$<$ GHS 700.00

GHS 700.00-GHS 1000.00

$>$ GHS 1000.00

Appendix C

Figure 5.1: Example of Choice Card