PREMIUM PRICING IN EMPLOYER BASED MEDICAL INSURANCE SCHEMES

SIELE DICKSON CHERUIYOT

A Research Project Submitted to the Graduate School in Partial Fulfillment for the Requirements of the award of Post-graduate Diploma in Actuarial Science in the School of Mathematics, University Of Nairobi.

UNIVERSITY OF NAIROBI

JULY 2013

DECLARATION AND RECOMMENDATION

Declaration

This research project is my original work and has not been submitted or presented for examination in any other institution.

Signature: ………………………………….. Date: ……………………………..

Siele Dickson Cheruiyot

I46/80836/2012

Recommendations

This research project has been submitted for examination with our approval as university of Nairobi supervisors.

Signature: ………………………………….. Date: ……………………………..

PROF. PATRICK G.O. WEKE

School of Mathematics

Faculty of science

University of Nairobi

COPY RIGHT

DEDICATION

In memory of my beloved and departed friend and colleague, Dr. Jesse Wachira Mwangi.

ACKNOWLEDGEMENT

First I would like to thank the Lord Almighty for bestowing upon me a good health, sane mind, strength and sufficient grace to see me through this study. Thank you and Glory be to You Lord!

Secondly, I pass my regards to my supervisors; Prof. Weke and Prof. Jamm Otieno for their support and technical advice and immeasurable value addition to the achievement of the goal of this study.

My sincere gratitude goes to all those who have made worthy contribution to this work. My colleagues too have so far been supportive and have provided guidance wherever needed. Thank you and God bless you.

Special thanks go to my Parents and Family for their encouragement through the taxing period of preparation of this work. May The Almighty God Bless us all immensely.

TABLE OF CONTENTS

ABSTRACT

In health care insurance, the main problem at task is “how much?” This would pose a rather complex situation in the case that we are projecting on future possible financial scenarios. With the rising cost of medical insurance, insurance providers argue that pooling of risk is a possible way of eventually reducing the cost of medical financing in the long-run. The goal of every organization that offers medical insurance to its employees is to find a precise and accurate estimate of premiums to be paid to the insurance company in the provision of essential medical services. This work takes into account several assumptions of risk including: homogeneity of risk, non-discriminatory employment and thus non-discriminatory premium allocation, minimum deductibles and co-insurance. This work thus comes up with a mathematical estimation procedure stipulating the theoretical premium amounts that is contributed to medical insurers to offset the rising financial costs based on past experience on claims. Results for predictions of premiums to be paid in current times are based on claim experiences. This work then performs comparative studies on future credibility premiums based on both the Buhlman’s and the Buhlmann Straub procedures.

Results show that the Buhlmann Straub procedure yields higher premium amounts. For all contracts, the individual premiums are higher than in the case of Buhlmann procedure. This may be due to weighting of claim amounts thus reducing variance components.

DEFINITION OF TERMS

Experience rating: This is the correct premium that is assigned to each individual risk. This depends exclusively on the (unknown) claims distribution of individual risk for the same period.

Risk premium: This is the correct premium to charge to an insured's risk level that is known. This is simply the expected value of the insured's aggregate claim amount in one period, given his or her risk level.

Collective premium (m): This is the pure premium charged when nothing is known about the insured's risk level. This is average value of all possible risk premiums.

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Health insurance is a protection provided to individuals who pay for health care expenses. Studies have shown that individual policies are 30% more costly than group insurance. Kenya's vision 2030 and its constitution advocates for universal health care for its population, irrespective of status of employment of its citizens. The main objective of this is to reduce the cost of medical financing that’s placing lots of financial burden to its citizens that eventually ends up consuming a large portion of the citizen's income. An insurance premium is an amount of money a person pays to an insurance company for an insurance policy. This payment could be regarded as transferring some or all of the risk (or cost) of loss or damage. The cost of an insurance premium needs to take into account the expected number of claims and the expected average claim size.

Insurance premiums may be set depending on risk factors, including:

- Individual’s medical history.
- Individual’s occupation- some occupations may be considered more risky and dangerous. This calls for rated premiums above the normal premiums.
- Also with some types of injury / disability insurance, if the person has what is called a “pre-existing condition” – an existing or previous illness that the insurer believes is likely to worsen or recur – there may be some insurers that will not cover this risk.

There are changing rules in the market with respect to medical insurance coverage. Short term rules indicate new requirements for benefits, such as providing coverage for adult children through age 26. There are also new minimum standards for loss ratios, and rebates that must be paid if those standards are not met. In 2030 the changes become more significant due to rising cost of medical insurance and numerous compensations. Many insurance providers will be willing to provide cover with the most effective and efficient services. Perhaps most importantly, companies will no longer be able to deny coverage based on poor health, consider health status when setting premiums, and will be limited in their ability to reflect age in their premiums.

Many institutions and parastatals in Kenya provide group medical insurance cover for its employees. This is an incentive for its worker population that enjoys eighty percent of free cover and only twenty percent out-of-pocket contribution toward offsetting medical cost on drugs whenever a member falls sick. This cost-sharing attitude relieves the employer of extraneous budgetary allocation towards medical cover. There are different levels of coverage ranging from individual to family coverage. The following are features of the institutional group medical scheme:

- All institutional employees are covered to a certain maximum amount irrespective of family status whether single or married. The age factor is not normally a consideration.
- An employee's medical coverage begins immediately on employment and ends with termination of service.
- All employees are covered irrespective of disability status.
- There is no discrimination of employment of individuals on health status. HIV patients enjoy the same rights of employment as well as coverage.
- The amount of coverage of hospital expenses is independent of salary cadre of an individual-assumes homogeneous population.
- Family coverage features include; employee and spousal full coverage as long as employment status is concerned, legally acquired dependents are covered up to 25 years of age but up to a certain percentage.

Institutions offering medical cover to its employees enjoy tax benefits up to a certain maximum. In Kenya companies and individuals who set aside a portion of their income on insurance enjoy a tax free status on the amount catered for the insurance. In addition to this, the institution best places itself at a statutory class of most searched employers (self-advertiser for employees). More-over, medical coverage to its employees is later reflected in the face of employee performance since this is a reflection of a healthy population. None-the less, growth in family income reflects the economic growth in households thus translating to contribution towards poverty eradication as envisaged in vision 2030. This is as a result of employees setting up a large junk of money that would otherwise have been used to cover medical expenses, in investment projects.

Health insurance may help employers to:

- Recruit high quality workers
- reduce staff turnover
- reduce the cost of absenteeism
- limit disability and workers compensation claims

Health premiums paid by employers are tax deductible as a business expense. Moreover,

Health insurance gives you access to price reductions that health insurance companies negotiate with healthcare providers.

Premium calculation is rather pluralistic in nature. Several aspects must be considered in premium calculation:

- Cost of paying benefits.
- Cost of administering the program - issuing policies, adjudicating claims, collecting premiums, filing annual statements, etc.

- Cost of marketing and distributing policies - this includes the commissions paid to agents and brokers.

- Companies also need to cover their cost of capital and maintain adequate financial reserves in case costs are higher than they expect.

Pricing of random claims has ever been one of the core subjects in both actuarial and financial mathematics and there exist various approaches for calculating (fair) prices. The actuarial way of pricing usually considers the classical premium calculation principles that consist of net premium and safety loading: Thus apart from the cost of paying benefits, there must be stipulated some loading factor usually a percentage of the discretized monthly/annual premium.

1.2 Statement of The Problem

With group health insurance, the focus is on the aggregate cost of the group. Except for the very smallest firms, health plans focus on the historical claim levels for the group, rather than on the health of specific employees. Institutions allocate a large portion of their income in paying medical expenditure for its employees. Many insurers normally make payments on behalf of institutions who in turn are responsible for payment of premiums. The question now here is "how much of a premium is to be paid projected from the past expenditure history on medical cover?" This is an ingredient to selection of insurance providers by concept of optimal pricing. The University seeks to find an insurance medical provider with a minimal premium cost but effective and satisfactory service provision to its members. This would largely lead to saving/investing of large amounts of money that would otherwise been used to cover medical costs. For all customers, the need for being charged 'the right price' is paramount.

This study makes use of credibility theory in pooling of risk with the assumption that premium pricing is independent of age, sex and health status-homogeneity of population.

1.3 Objectives

1.3.1 General Objective

The general objective of this study is to find the optimal price of premiums paid to an insurance company to cover for medical expenditure based on historical expenditure.

1.3.2 Specific Objectives

In line with the above objective, we aim to fulfill the following specific objectives:

(i) To model group expenditure claim cost using distribution based techniques to ascertain estimates of cost of expenditure.

(ii) To come up with a credible risk premium values of institutional claim experience to be used as an average premium regulator.

(iii) To project future financial cost of providing medical cover due to rising costs brought about by inflation and technological advancement.

1.4 Assumptions

(i) Health plans focus on the historical claim levels for the group, rather than on the health of specific employees.

(ii) Administration and marketing costs are a percentage of net annual premiums.

(iii) Premiums paid are a reflection of group insurance without adherence to age and sex of individuals.

(iv) Observations from a given entity/contract both within one year and from year to year are mutually independent.

(v) The contracts/hospitals are non-homogenous. The risk classes are also a cost dependent.

(vi) The cost of medical claims is inclusive of administration costs and includes the extra cost that may be brought about by the difference in age factor.

(vii) Mean square error will be minimized to determine the estimate.

1.5 Justification

The cost of healthcare in Kenya is soaring high each coming year due to several factors affecting the economy. These include the cost of inflation and technological advancement. Institutions hire the services of insurance companies in providing medical cover to its employees. A large number of insurance providers would want to take advantage of the lack of expatriates that determines the correct premiums. This has subsequently led to unscrupulous insurance companies taking advantage of clients by charging high cost of insurance. This substantially allows them make super normal profits. The aim of this study is to provide an optimal premium paying function obtained by experience rating- premium is based on the group's own experience - past year's claims are projected forward and used as the basis for this year's premiums. The obtained limits of premium payments would aid the employer in determining the 'fair' price that is charged by the health insurer.

The study will provide substantial knowledge to actuarists in determining gross premiums affected by fluctuation in incidental economic factors, an extension of credibility theory.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

The study of insurance dates back to the 19^{th} century during the famous era of London's fire. In the aftermath, London's residents made contributions to a group account in form of savings. Some insurance companies directly paid for some of Londons' city fire brigades. People who paid insurance companies to insure their homes were given a 'fire plate' showing the insurance company's logo. This fire plate was fixed near the front door of their house. If a house caught fire, the fire brigade would check if the house had a fire plate. If the house did not have a fire plate, or had the fire plate of another insurance company, the fire brigade would let the house burn down. People who wanted the fire brigade to help them if their home caught fire, would each put in a little money to help pay for the fire brigade to protect their house. The problem that was of concern was; residents were of different classes of wealth and the kind of lifestyle one lived was dependent upon the class (the population was heterogeneous in nature). The size of wealth one owned was evidenced by the kind of housing. The main issue now was 'How much one is expected to pay monthly that is proportional to the value of the property'. Since this era, actuarists have developed much theory to determine the premium amounts per period of insurance. Much of the theories developed include credibility theory-measure of predictive value attached to a particular class of data based on experience rating-and premium rating-process of determining premium estimates of expected values of future costs per unit time of exposure for group of risks. .

2.1 Credibility theory

Credibility theory is a set of techniques of calculating insurance premiums for short-term nonlife insurance contracts. This technique makes use of:

- Historical data related to the actual risk.
- Data from other related but relevant sources commonly referred to as collateral data.

The credibility premium as derived by Waters (1987), in the special note, is of the form:

$$\hat{m} = z\bar{x} + (1 - z)\,\mu$$

Where $\hat{m}$ is the premium, z - the weight or credibility factor usually between zero and 1. The credibility factor here is an increasing function for large values of *n*. The mean parameter $\bar{x}$ is the observed mean claim amounts per unit risk exposed for individual contract/risk itself. μ is the parametric estimate of the proposed data in the case than an assumption of the underlying distribution is made. For a series of risks, μ is the corresponding portfolio (set of risks) mean. The following are features of the credibility formula:

- It is a linear combination of estimates to a pure premium policy based on observed data from the risk itself and the other based on projected risks.
- The credibility factor Z shows the degree of reliability of the observed risk data in the sense that high values of Z implies high reliability.
- The credibility factor is a dependent function of the number of claims. This implies that the higher the claim numbers, the larger the credibility factor.
- The value of Z is between zero and one, i.e., $0 \le Z \le 1$.

2.1.1 Credibility theory development

The origin of credibility dates back to early 1900's. It was originally developed by North American actuarists over a large period of time. The earliest of the most practical solution to premium calculation was developed by Mowbray (1914). This came to be called the American credibility theory. This work is sometimes referred to as 'limited credibility theory' or 'the Fixed effect credibility'. In this work, the assumption was that the annual claims $X_1, X_2, ..., X_N$ are independent and identically distributed random variables from a probabilistic model with means $m(\alpha)$ and variance $s^2(\alpha)$. The assumption is that the collateral data follows a normal distribution distributed as $m \sim N\left(m(\alpha), {s^2(\alpha)}/{\sqrt{n}}\right)$. This theory asserts that the size of claims *n* should be large.

This theory received a lot of criticisms from many researchers including Whitney (1918). He proposed that claims are random in nature and thus the assumption of fixed effects model was invalid. Moreover, the American credibility theory was faced with the problem of partial credibility. It was difficult to determine the value of the credibility factor. With regards to this after the World War II revolution, Whitney's random effect model came into place.

Nelder and Verall derived credibility functions by generalized linear model approach and consequently included the random effects model. This has provided a wide platform of actuarial application including premium rating and reserving. Despite the wide research findings, it was observed that the limited credibility theory was unable to solve the problem of credibility. This was partly attributed to very poor or undeveloped statistical background.

A major breakthrough was made in the year 1967 and 1970 when Bulhmann derived the credibility premium formula in a distribution free-way such that there was no assumption of prior distribution of claims. Several assumptions of using the credibility premium formula in this paper were clarified in this work in 1971 (see Bulhmann 1971). This major breakthrough has seen much of the research tilting to the development of Bayesian estimation techniques by Jewell (1974, 1975), Hachmeister (1975), Devylder (1976, 1986), Zehnith (1977) and, Gooverts and Hoogstad (1987). Jewell (1974) showed that for exponential family distributions, the best linear approximation to Bayesian estimate is obtained using quadratic loss functions. Hachmeister (1975) extended the Bulhmann Straub (B-S) model for a class of business by use of matrix methods. Much of this work is summarized by Noberg (1979). The vast literature developed thus far is referred to as 'The European Credibility theory', 'Greatest accuracy credibility theory' or 'The empirical Bayes credibility theory'. In the preceding sections, we study each of the credibility theories in detail.

CHAPTER THREE

METHODOLOGY

3.1 Credibility Theory

The credibility premium is a linear function of the form: $\hat{m} = z\bar{x} + (1-z)\mu$. In this case, Z is the amount of credibility that is assigned to a certain data set originating from past experience data. The main problem of actuarists is how much information/observations are required for one to attain 100% credibility. This leads us to determining conditions necessary to attain full credibility and partial credibility. In most practical situations, full credibility is a rare phenomenon. The fixed effects model of Mowbray (1914) deduced a criterion for determining the sample size required for partial credibility as of the form: $n \geq \frac{Z^2_{1-\frac{\varepsilon}{2}} \hat{S}^2}{k^2 \hat{m}^2}$.

This approach came under so much criticism due to its fixed effects. This led to the adoption of Whitney's random effects model that mainly focused on the estimation of the credibility function. This opened a wide area of research where experience rating problems were seen to be a matter of estimating the random variables $\hat{m}$ from observed mean of information, μ, of the individual data sets. The main aim was to minimizing the Mean square error (MSE)

The optimal estimator, μ, is obtained by conditional approach $\hat{m} = E(m/x)$. The most important computational functions include:

$$E(X) = E\left[E(X/Y)\right] \text{ and}$$

$$Var(X) = E\left[Var(X/Y)\right] + Var\left[E(X/Y)\right]$$

Thus the MSE is thus obtained as follows:

$$\rho(\hat{m}) = E\left[Var\left(m(\varphi)/Y\right)\right]$$

$$= Var(m) - Var(\hat{m})$$

$$= E\left[m(\varphi) - E(m/X)\right]^2 + E\left[E(m/X) - m(X)\right]^2$$

Thus $\rho(\hat{m}) = E\left[m(\varphi) - \hat{m}(X)\right]^2$.

The above derivation of the MSE mostly gave restrictions on distribution functions. This form of MSE was then modified to avoid much restriction on distribution function which eventually gave rise to a a linear credibility function of the form $\tilde{m}(X) = a + b\hat{m}(X)$

The linear Bayes estimator of the form:

$$\bar{m} = E(m(\varphi)) + \frac{Cov[m, \hat{m}]}{Var(\hat{m})}(\hat{m} - E(\hat{m})).$$

The Linear Bayes risk is thus given by the function

$$\bar{\rho} = Var(\hat{m}) - \frac{Cov^2[m, \hat{m}]}{Var(\hat{m})}$$

The linear Bayes risk approaches zero with increasing amounts of data. The sufficient conditions that must hold include:

(i) $E(\hat{m} - E(\hat{m}))^2 \to 0,$

(ii) $E\left[Var(\hat{m}/\varphi)\right] \to 0$ and

(iii) $E\left[\left(\hat{m}/\varphi\right)\right]=m\left(\varphi\right)$

For these conditions in place,

$$E\left(\hat{m}\right)=E\left(m\right)$$

$$Cov\left[m,\hat{m}\right]=Var(m)$$

$$Var\left(\hat{m}\right)=Var(m)+E\left[Var\left(m\left(\varphi\right)/Y\right)\right],$$

The credibility function Z is thus given as

$$Z=\frac{Var\left(m\left(\varphi\right)\right)}{Var(m)+E\left[Var\left(m\left(\varphi\right)/Y\right)\right]}$$

Various models have been suggested for calculation for the credibility premiums in the vast literature of empirical Bayes credibility. The model assumptions are that the aggregate claims are independent and identically distributed in nature. In most life situations, this is not normally the case since to analyze for risk, we need different variables that are not necessarily dependent on one another. We relax this assumption of independence and we assert that the aggregate claims are not necessarily identically distributed.

3.2 Empirical Bayes credibility

Denote a given random data set of aggregate claims for successive years for a particular class of risk, say $Y_1, Y_2, ..., Y_n$ for successive years. Let $P_1, P_2, ..., P_n$ be a corresponding sequence of known constants, in this case the number of policies issued in a year. Let $X_1, X_2, ..., X_n$ be a sequence of random variables such that: $X_j = Y_j / P_j$

The assumption on the distribution of the random variable X_j is dependent on a fixed parameter ϕ and is denoted by $U(\phi)$.

Assumptions:

i. $X_j / \phi, \quad j = 0,1,2,...$ are independent and not necessarily identically distributed.

ii. is not dependent on J.

iii. $P_j Var\left[X_j / \phi\right]$ is not dependent on J.

Intuitively, Y_j s are the aggregate claims from different amounts of business. X_j s are the standardized Y_j s obtained by reducing effects of business levels (by smoothing).

We define

$$m(\phi) = E\left[X_j / \phi\right] \text{ and}$$

$$S^2(\phi) = P_j Var\left[X_j / \phi\right]$$

3.3 Derivation of the credibility premium

3.3.1 Buhlmann,s credibility

By Bulhamann’s approach, the credibility premium is a linear function of observed values X_j which gives the best approximation to $E\left[m(\phi)/X\right]$.

The observed values are linear and are of the form

$$a+\sum_{j=1}^{n}a_jX_j \text{ , j=1,2,3,….n}$$

We seek to find the constants a_j that minimize the mean square error. i.e, $E\left\{E\left[m(\phi)/X\right]-\left(a+\sum_{j=1}^{n}a_jX_j\right)\right\}^2$.

We solve this by differentiating the function above with respect to a_j s.

This leads to a new set of equations that can be solved iteratively:

$$E\left[m(\phi)-a_0-\sum_{j=1}^{n}a_jX_j\right]=0$$

$$E\left[X_km(\phi)-a_0X_k-\sum_{j=1}^{n}a_jX_kX_j\right]=0, \quad k=1,2,..,n \text{ .}$$

Equation (i) and (ii) are reducible to the forms:

$$a_0=\left\{1-\sum_{j=1}^{n}a_j\right\}E\left(m(\phi)\right)\text{. And}$$

$$E\left[m^2(\phi)-a_0E\left(m(\phi)\right)-a_k\,{E\left[S^2(\phi)\right]}\big/{P_j}-\sum_{j=1}^{n}a_jE\left(m^2(\phi)\right)\right]=0$$

Substituting (iii) into (iv), we obtain:

$$P_k Var\left[m(\phi)\right]\left\{1-\sum_{j=1}^{n} a_j\right\} = a_k E\left[S^2(\phi)\right]$$

By summing up the equation (v) for diferent policy values we note that

$$\sum_{j=1}^{n} a_j = \sum_{j=1}^{n} P_j \Bigg/ \left\{\sum_{j=1}^{n} P_j + E\left[S^2(\phi)\right] \big/ Var\left[m(\phi)\right]\right\}$$

Using (vi) in (v) and (iii) we obtain the results of a_0 and a_k as

$$a_0 = E\left(m(\phi)\right)\left[E\left[S^2(\phi)\right]\big/Var\left[m(\phi)\right]\right]\Bigg/\left[\sum_{j=1}^{n} P_j + E\left[S^2(\phi)\right]\big/Var\left[m(\phi)\right]\right]$$

$$a_k = P_k \Bigg/ \left\{\sum_{j=1}^{n} P_j + E\left[S^2(\phi)\right]\big/Var\left[m(\phi)\right]\right\}$$

We thus substitute equations (vii) and (viii) in ther linear form of credibility premium to obtain

The pure premium per unit volume of risk. This equation is of the form;

$$E\left(m(\phi)\right)\left[E\left[S^2(\phi)\right]\big/Var\left[m(\phi)\right]\right] + \sum_{j=1}^{n} Y_j \Bigg/ \left[\sum_{j=1}^{n} P_j + E\left[S^2(\phi)\right]\big/Var\left[m(\phi)\right]\right]$$

Taking $\bar{X} = \sum_{j=1}^{n} P_j X_j \Big/ \sum_{j=1}^{n} P_j$

$$Z = \sum_{j=1}^{n} P_j \Bigg/ \left[\sum_{j=1}^{n} P_j + E\left[S^2(\phi)\right]\big/Var\left[m(\phi)\right]\right],$$

we obtain the linear form of credibility premium $\hat{m} = z\bar{x} + (1-z)\mu$.

3.3.2 Parameter Estimation

In this section, we estimate parameters contained in the credibility premium formula from a suitable data set. The parameter estimates are proved to be unbiased estimators.

Suppose we define a single risk from a class of N risks. Let $Y_{i1}, Y_{i2}, ..., Y_{in}$ denote the aggregate claims in successive years from the risk with $P_{i1}, P_{i2}, ..., P_{in}$ being the corresponding risk volumes of known constant value. Define $X_{ij} = {Y_{ij}}/{P_{ij}}$

The assumptions under this model are that:

- $X_{i1}/\phi_i, X_{i2}/\phi_i, ..., X_{in}/\phi_i$ are independent but not necessarily identically distributed
- $\phi_1, \phi_2, ..., \phi_n$ are independent and identically distributed.
- We possess the same number of observed risks.

We assume that

$m(\phi_i) = E\left[X_{ij}/\phi_j\right]$ and

$S^2(\phi_i)/P_{ij} = Var\left[X_{ij}/\phi_j\right]$

Because of the assumption of identical distribution of ϕ_i's, the distribution of $m(\phi_i)$ and $S^2(\phi_i)$ are the same for all i's.

We adopt the following notations:

$$\bar{P}_i = \sum_{j=1}^{n} P_{ij}$$

$$\bar{P} = \sum_{j=1}^{N} \bar{P}_i$$

$$P^* = \left[\sum_{j=1}^{N} \bar{P}_i \left(1 - \bar{P}_i / \bar{P}\right) \right] \Big/ (Nn - 1)$$

$$\bar{X}_i = \sum_{j=1}^{n} P_{ij} X_{ij} \Big/ \bar{P} \quad \text{and}$$

$$\bar{X} = \sum_{j=1}^{N} \bar{P}_i \bar{X}_i \Big/ \bar{P} = \sum_{i=1}^{N} \sum_{j=1}^{n} P_{ij} X_{ij} \Big/ \bar{P}$$

Note that the unbiased estimators were of the form:

$$E\left[m(\phi)\right] = \bar{X}$$

$$E\left[S^2(\phi)\right] = \sum_{j=1}^{n} P_{ij} \left(X_{ij} - \bar{X}\right)^2 \Big/ N \sum_{i=1}^{n} (n-1)$$

$$Var\left[m(\phi)\right] = P^{*-1} \left\{ \frac{\sum_{i=1}^{N} \sum_{j=1}^{n} P_{ij} \left(X_{ij} - \bar{X}\right)^2}{(Nn-1)} - \sum_{j=1}^{n} P_{ij} \left(X_{ij} - \bar{X}\right)^2 \Big/ N \sum_{i=1}^{n} (n-1) \right\}$$

3.4 Bulhmann-Straub Model

This is a generalization of the classical credibility premium of Bulhmann(1969) and has been used to rate reinsurance treaties where much of this has been applied in auto and reinsurance sectors.

This consists of a portfolio of N insureds each characterized by an unobservable random risk parameter ϕ_i and . Let X_{it} be a list of available observations such as the average claim amount or claim loss ratio. for $t = 1, 2, ..., T_i$ and $i = 1, 2, ..., N$. The number of periods of experience depends on the insured. To each X_{it} , a weight w_{it} is assigned. The weights can be valid measures such as no of claims in one year or the premium volumes.

Model assumptions

The insureds' vectors $\left(X_{i1,...,}X_{iT_i}, \phi_i\right)$, $i = 1, 2, ..., N$ are mutually independent.

The risk parameters ϕ_i are independent and identically distributed.

The variables X_{it} have finite variance

For $i = 1, 2, ..., N$

$$E\left(X_{it} / \phi_i\right) = \mu\left(\phi_i\right) \quad \text{(i)}$$

$$\text{cov}\left(X_{it}, X_{it} / \phi_i\right) = \frac{\sigma^2\left(\phi_i\right)}{w_{it}} \quad \text{(ii)}$$

Equation (ii) reflects the non-correlation between the insured's claim experience across the years and the homogeneity in time.

Equation (i) shows that the risk premiums $\mu\left(\phi_i\right)$ is time invariant.

3.4.1 Parameter Estimation of the B-S credibility premium

The structural parameters are as follows.

$$m = E\left(X_{it}/\phi_i\right) = \mu\left(\phi_i\right) \qquad s^2 = E\left(\sigma^2\left(\phi_i\right)\right)$$

$$a = \operatorname{var}\left(\mu\left(\phi_i\right)\right) \qquad w_{i\bullet} = \sum_{t=1}^{T_i} w_{it}$$

$$w_{\bullet\bullet} = \sum_{i=1}^{I}\sum_{t=1}^{T_i} w_{it} \qquad X_{i\bullet} = \sum_{t=1}^{T_i} \frac{w_{it}}{w_{i\bullet}} X_{it}$$

$$X_{\bullet\bullet} = \sum_{i=1}^{I}\sum_{t=1}^{T_i} \frac{w_{it}}{w_{\bullet\bullet}} X_{it} \qquad z_{\bullet} = \sum_{i=1}^{I} z_i$$

and $X^{z}_{\bullet\bullet} = \sum_{i=1}^{I} \frac{z_i}{z_{\bullet}} \sum_{t=1}^{T_i} \frac{w_{it}}{w_{\bullet\bullet}} X_{it}$.

The credibility premium P_i is found by minimizing the mean square error. This is estimated as

$P_i = z_i X_{i\bullet} + \left(1 - z_i\right) m$ where,

$$z_i = \frac{w_{i\bullet}}{w_{i\bullet} + k} \text{ and } k = \frac{s^2}{a}.$$

s^2 is a measure of the stability of portfolio claim experience (homogeneity within the insureds). The lower this value, the larger the credibility factor. a is a measure of variation of various individual risk premiums and denotes homogeneity between the insureds. An increase in this leads to an increase in the credibility factor, Gowell, (1998).

The estimates of the structural parameters are:

$m = X^{z}_{\bullet\bullet}$ which is the pseudo estimator, a function of unknown parameters s^2 and a.

$s^2 = \frac{1}{N-I}\sum_{i=1}^{I}\sum_{t=1}^{T_i} w_{it}\left(X_{it} - X_{i\bullet}\right)^2$, the unbiased estimator of s^2.

$\hat{a} = \frac{w_{\bullet\bullet}}{w_{\bullet\bullet}^2 - \sum_i w_{i\bullet}^2}\left(\sum_{i=1}^{I} w_{i\bullet}\left(X_{i\bullet} - X_{\bullet\bullet}\right)^2 - (I-1)s^2\right)$, the estimator obtained by ANOVA which is sometimes negative.

$\hat{a} = \frac{1}{I-1}\sum_{i=1}^{I} z_i\left(X_{i\bullet} - X_{\bullet\bullet}^{z}\right)^2$, the Bichsel-Straub Estimator which is always positive.

3.5 Model Application and Methodology

A family of distributions for the number of claims N can be generated by assuming that the Poisson parameter is random variable with pdf $f(\lambda)$ with $\lambda > 0$. The conditional distribution of N is also a Poisson with parameter λ. When the variance of the number of claims exceeds its mean, the Poisson distribution is not appropriate. Rather, a negative binomial distribution is used, Bowers, et al (1997).

For a collective risk model, we assume a random process that generates claims for a portfolio of policies. Each of the claim amounts X_i, then $S = X_1 + X_2 + ... + X_N$ represents the aggregate claims for the portfolio for the period under study. The random variables $X_1, X_2, ...X_N$ also measure the severity of the claims. For this reason of stability, we simulate claim numbers from a Poisson distribution where the mean and variance components are equal. i.e , $E\left(\mu(\theta_i)\right) = Var\left(\mu(\theta_i)\right) = \theta_i$.

For many insurance claims, the claim amount random variable is only positive and its distribution is usually skewed to the right. These properties resemble the properties to the gamma distribution. In this study we adopt the above two essential properties to perform simulation of claim amounts. The distribution of claim amounts may not be of a simple form, but the convolution of claim amounts may yield a compound Poisson distribution. We may opt to choose a discrete claim distribution and calculate the required convolutions numerically. This may be a new line of investigation for credibility premium calculations.

3.5.1 Simulation

In this study, we simulate five contracts depicting five different insureds/contracts, in our case, the contracts are the Hospitals. These contracts (hospitals) each involves observed data for a period of five years with each claim size and the respective number of claims in that year being obtained.

The simulation procedure first begins with the generation of weights, W_{it}, from a Uniform distribution such that on (a, b), $0 \leq a < b$. In our case we have chosen to simulate ten variables from uniform (500, 1000) by the function, unif(10,(500,1000)). These weights may be the total number of claims in the respective year of interest or any chosen function of the claim amount, say the square root of the total claim amount in that year.

This is then followed by generation of the risk levels from a gamma distribution function using the R function, rgamma (3,2). The risk levels are also functions of weights generated as above.

The aggregate claims N_{it} are for different contracts and/or insureds is obtained from a *poisson* distribution using the function rpois (weights * contracts), where each claim is generated from a gamma distribution of parameters α and β such that each claim amount is also gamma distributed.

The total amounts made for claims, S_{it} is the sum of all N_{it} for all insureds or contracts.

Finally, we obtain the claim ratios from dividing the total claim amount S_{it} by the weights obtained above. This is represented as $X_{it} = \frac{S_{it}}{W_{it}}$.

In order to fit the credibility model to the above data as obtained from simulation, the procedure requires that a package *actuar* be loaded such that the linear modeling parameters be installed.

We thus extract the data and process it in a simplified form as shown in the appendix.

In the analysis, we attempt to make projections of credibility predictions based on both the Buhlman's credibility and the Buhlman Straub credibility approaches. We tabulate results for the weights that are obtained, the various variance components and predict future claim experience by experience rating.

Where possible, we attempt to provide graphical comparison for the same procedures above.

3.5.2 Analysis

In this study, we analyze tabulated information that gives projections for the credibility premiums. This involves computation of between and within portfolio variances with the view of finding credibility premiums by linear estimation from the credibility formula. In the process, individual means are thus tabulated with respect to each of the above procedures.

The results are obtained by simulation procedure using the R.13.0 package.

We find the unbiased estimators for the mean and variance functions for both the Buhlmann and Buhlmann-straub procedures.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.0 Presentation and discussion of results

The following data was generated by the simulation approach as described in chapter three above. The tables have been modified as they would appear in R script output. The result below is one of the outputs for a single simulation, rather there are many outcomes expected since we are simulating from continuous distributions. We use the data set below for our analysis. The discussion presented is similar for all types of scenarios that can be generated.

TABLE 4.1 Aggregate claim amounts.

Contract	Year 1	Year 2	Year 3	Year 4	Year 5	Totals
1	149908.7	362442.6	977109.3	1064314.7	417414.6	2971189.9
2	1438904.0	1027173.7	634067.0	623405.8	531180.4	4254730.9
3	1206991.0	526138.3	585919.8	415930.3	1896456.2	4631435.6
4	519796.7	932460.7	759577.3	2387072.1	818550.8	5417457.6
5	964479.7	1962366.8	1930013.6	999284.5	1149114.7	7005259.3

TABLE 4.2: Ratios

ratio.1	ratio.2	ratio.3	ratio.4	ratio.5	sum(ratios)
541	874	1427	1434	931	5207
1093	912	732	721	680	4138
1304	856	918	749	1622	5449
983	1336	1176	2100	1217	6812
1502	2129	2087	1562	1592	8872

TABLE 4.3: weights

weight.1	weight.2	weight.3	weight.4	weight.5	sum (weights)
277.096	414.694	684.7297	742.1999	448.3508	2567.0701
1316.47	1126.29	866.2118	864.6405	781.1477	4954.759
925.607	614.648	638.2569	555.3141	1169.209	3903.0337
528.786	697.95	645.8991	1136.701	672.5972	3681.933
642.13	921.732	924.7789	639.7468	721.8057	3850.1932

Fitting the Buhlman's credibility yields the following results for prediction for the sixth year claim experience: The structural parameters being the components of the table below.

Table 4.4 Buhlmann's Result

Between contract variance/covariance: 108981.8

Within contract variance: 118167.5

Collective premium: 1219.12

Contract	$\bar{X}_{i\bullet}$	credib Fac($Z_{i\bullet}$)	cred premium
1	1041.4	0.821789	1073.0717
2	827.6	0.821789	897.3732
3	1089.8	0.821789	1112.8462
4	1362.4	0.821789	1336.8659
5	1774.4	0.821789	1675.443

Fitting the Buhlman-Straub credibility for the sixth year claim experience yields the following linear prediction for the five contracts:

TABLE 4.5 - RESULTS

Collective premium: 1297.027

Between contract variance/covariance: 109431.8

Within contract variance: 91987995

Contract	indiv.mean	weight	cred.Factor	cred premium
1	1157.425	2567.07	0.753322	1191.8615
2	858.716	4954.759	0.8549534	922.2916
3	1186.625	3903.034	0.8227947	1206.1885
4	1471.362	3681.933	0.8141313	1438.959
5	1819.456	3850.193	0.8207985	1725.8364

Results show that the Buhlmann Straub procedure yields higher premium amounts. For all contracts, the individual premiums are higher than in the case of Buhlmann procedure. This may be due to weighting of claim amounts thus reducing variance components.

The individual contract means above were obtained by, $\bar{X}_i = \sum_{i=1}^{5} \frac{sum(ratio)}{5}$.

This gives rise to the group mean $\bar{X} = \sum_{i=1}^{5} \bar{X}_i$.

The unbiased estimator for the contract variances is the sample variance, 109431.8

We can readily observe that the credibility factor for the Buhlmann procedure is a constant for all the hospitals/contracts. In the case of the Buhlmann Straub procedure, the credibility premium varies with the associated weights.

The following table shows the numerical summary of the structural parameters for the Buhlmann straub procedure for credibility theory.

TABLE 4.6 SUMMARY DATA-BUHLLMANN PARAMETERS

$W_{i\bullet}$	$X_{i\bullet}$	$S_{i\bullet}$	$\bar{X}_{i\bullet}$	$Z_{i\bullet}$	**Cred. premium**
2567.0701	5207	2971189.9	1157.425	0.753322	1191.8615
4954.759	4138	4254730.9	858.716	0.8549534	922.2916
3903.0337	5449	4631436	1186.625	0.8227947	1206.1885
3681.933	6812	5417458	1471.362,	0.8141313	1438.959
3850.1932	8872	7005259	1819.456	0.8207985	1725.8364

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion

This study has focused mainly computation of credibility premiums by the Buhlmann and Bulhmann Straub credibility theory. These methods are rather linear approximation techniques as opposed to techniques that are normally parametric in nature. Most of the parametric simulations are distribution sensitive and much of this has had little or no application in real life. We have assessed several portfolios that describe the real life application as in the medical sector.

In the case of hospital claims, there are the risk levels being the outpatient and in-patient entities. Our procedure only looks at only one side of the scenario being either of the two. We can simulate the other scenario in the same way and come up with the same conclusions.

Real data is important in determining the physical financial scenario of a company. In the medical sector, real sector is difficult to obtain. Data from institutions such as the National Hospital Insurance Fund (NHIF) and one of the Universities was difficult to obtain. The reason behind this is the oath secrecy to hold on to information that is deemed ethically private in the medical sector. If real data is observed the findings may be varied due to the different claim experiences. Real data claim amounts is inclusive of administration and contingency costs that may have been incurred. Furthermore, the amount of claim may be erroneous since some claims that may be made in one year may not be paid till the next calendar year. The claim amount is recorded in the exiting year, yet the claim number is in the correct year of entry of claim.

Age is an important factor in determining the cost of medical insurance. It is important that age be considered in the premium computation. This helps to assist in obtaining accurate credibility premiums. This is because aged individuals incur lots of expenditure in medical expenditure. The extra cost due to age factor may be imputed in to the model as an extra percentage of the credibility premium.

An individual's health is a variable of time and may occur seasonally probably due to outbreaks or environmental changes. At one season, claim experiences may be higher than another. This fluctuations may lead to inflated premium amounts.

The Buhlmann and Buhlmann straub procedures is faced with the problem of outliers that distort the mean and variance functions. This in turn affects the accuracy of the credibility premiums.

5.2 Recommendation

This study recommends that observed data should be smoothened of any outliers in order to increase accuracy of the credibility premiums. This is by reducing the effect of outliers.

5.3 Further Research

In the models under study above, the assumption of homogeneity within the cohorts has been made. In most cases, if we assume that the years' claim total is heterogeneous in nature, we need to account for heterogeneity in the model in calculation of premiums. This is because the claim experience for individuals is not the same in the different cohorts. This is normally referred to as over-dispersion. This is a more interesting field that much research can be done. This involves the estimation of the over-dispersion parameters and factoring this into the required credibility mode.

5.4 Application

Credibility theory has had lots of application in mortality studies, auto insurance, group insurance and on the study of effects of chance of variation in surplus of insurance companies. This theory has a wide range of application in fields mainly dealing with heterogeneous populations or where the population may be clustered and claim experience for a period of time is recorded.

REFERENCES

Bailey, A.L. (1945): A generalized theory of credibility. Proceedings of the Casualty Actuarial Society 32, 13-20.

Bowers, etal, (1986): Risk theory, *society of actuaries*.

Bowers, etal (1997): actuarial mathematics, society of actuaries.

Buhlmann H, (1967), experience rating and credibility, Astin bulletin, 4: 199-207

Buhlmann H, (1969), experience rating and credibility, Astin bulletin, 5: 157-165

Buhlmann, H. and Gisler, A. (2005). A Course in Credibility Theory and its Application. Springer-Verlag, Berlin.

Buhlmann, H. E and straub E, (1970), credibility for loss ratios, schweizerischer mathematica, 70: 111-133.

Gooverts and Hoogstad (1987). *Surveys of actuarial studies*, No. 4, Rotterdam, Holland.

Hachemeister, C. (1975): Credibility for regression models with application to trend. In Credibility: Theory and applications (ed. P.M. Kahn),129-163. Academic Press.

Hogg, R.V and Klugmann, S .A (1984), Loss distributions, John Willey and sons, New York.

Jewell, W. (1974). Credible means are exact Bayesian for exponential families. Astin Bulletin, 8, 1, 77–90**.**

Jewell, W.S. (1974c): Regularity conditions for exact credibility. ASTIN Bulletin 8, 336-341.

Longley-Cook, L.H. (1962): An introduction to credibility theory. Proceedings of the Casualty Actuarial Society 49 , 194-221.

Mowbray, A., (1914). How extensive a payroll is necessary to give dependable pure premium?

,proceedings of the Casualty Actuarial Society, 1, 24–30.

Norberg, R. (1979). The credibility approach to experience rating, Scandinavian Actuarial journal, pp 181-221.

Sundt, B. (1984), an introduction to non-life insurance mathematics, verlag versicherunswirtschaft.

Vincent Goulet (1998), Principles and Application of credibility theory, journal of actuarial practice, vol .6

Waters, H.R (1987), An introduction to credibility theory, special note: vol .2

Whitney, A. (1918). The theory of experience rating. Proceedings of the Casualty Actuarial Society, 4, 274–292.

APPENDIX

SIMULATION OF PORTFOLIO FOR CLAIM DATA EXPERIENCE

```
##we need to load package “actuar”

#wit <- runif(10, 500, 1000)

#(wit <- runif(25, rep(0.5 * wit, each = 5), rep(1.5 * wit, each =
5)))

#(ratios <- simul(list(contracts = 5, ratio = 5), model.freq =
expression(contracts = rgamma(3, 2),ratio = rpois(weights *
contracts)),weights = wit))

weights<-matrix(c(277.0956 , 414.6941 , 684.7297,  742.1999 ,
448.3508 ,1316.4721 ,1126.2869, 866.2118 , 864.6405 , 781.1477 ,
925.6066,  614.6476,  638.2569,  555.3141,1169.2085 , 528.7861 ,
697.9496,  645.8991, 1136.7010,  672.5972,  642.1303,921.7317,
924.7789,  639.7468 , 721.8057),5,5,byrow=TRUE)

ratios<-matrix(c(541,874 ,1427, 1434,931,1093 ,912,732 ,721 ,680,
1304, 856, 918 , 749 , 1622, 983  ,1336, 1176 , 2100 , 1217, 1502
, 2129 , 2087 , 1562, 1592), 5, 5, byrow=TRUE)

contract<-c(1,2,3,4,5)

weights<-matrix(c(277.0956 , 414.6941 , 684.7297,  742.1999 ,
448.3508 ,1316.4721 ,1126.2869, 866.2118, 864.6405, 781.1477,
925.6066, 614.6476,  638.2569,  555.3141,1169.2085, 528.7861,
697.9496, 645.8991, 1136.7010, 672.5972, 642.1303,921.7317,
924.7789, 639.7468, 721.8057), 5, 5, byrow=TRUE)
```

```
ratios<-matrix(c(541,874 ,1427, 1434,931,1093 ,912,732 ,721
,680,1304,856,918 ,749 ,1622, 983  ,1336,1176 , 2100 , 1217, 1502
, 2129 , 2087 , 1562, 1592),5,5,byrow=TRUE)

ratio1<-structure(ratios[1,],.Names = c("ratio.1","ratio.2",
"ratio.3", "ratio.4", "ratio.5"))

ratio2<-structure(ratios[2,],.Names = c("ratio.1","ratio.2",
"ratio.3", "ratio.4", "ratio.5"))

ratio3<-structure(ratios[3,],.Names = c("ratio.1","ratio.2",
"ratio.3", "ratio.4", "ratio.5"))

ratio4<-structure(ratios[4,],.Names = c("ratio.1","ratio.2",
"ratio.3", "ratio.4", "ratio.5"))

ratio5<-structure(ratios[5,],.Names = c("ratio.1","ratio.2",
"ratio.3", "ratio.4", "ratio.5"))

ratio<-
matrix(c(ratio1,ratio2,ratio3,ratio4,ratio5),5,5,byrow=TRUE)

weights1 = structure(weights[1,], .Names = c("weight.1",
"weight.2", "weight.3", "weight.4", "weight.5"))

weights2 = structure(weights[2,], .Names = c("weight.1",
"weight.2", "weight.3", "weight.4", "weight.5"))

weights3 = structure(weights[3,], .Names = c("weight.1",
"weight.2", "weight.3", "weight.4", "weight.5"))

weights4 = structure(weights[4,], .Names = c("weight.1",
```

```
"weight.2", "weight.3", "weight.4", "weight.5"))

weights5 = structure(weights[5,], .Names = c("weight.1",
"weight.2", "weight.3", "weight.4", "weight.5"))

weight<-
matrix(c(weights1,weights2,weights3,weights4,weights5),5,5,byrow=T
RUE)

creddata<-cbind(contract,ratio,weight)

claims<-weight*ratio          #aggregate claim amounts

colnames(creddata)=c("contract","ratio.1","ratio.2", "ratio.3",
"ratio.4", "ratio.5","weight.1", "weight.2", "weight.3",
"weight.4", "weight.5")

## Fitting of a Buhlmann model to the creddata set

fit <- cm(~contract, creddata, ratios = ratio.1:ratio.5)

summary(fit)                  # more information

fit$means                     # (weighted) averages

fit$weights                   # total weights

fit$unbiased                  # unbiased variance estimators

predict(fit)                  # credibility premiums

## Fitting of a Buhlmann-Straub model with weights.

fit <- cm(~state, creddata, ratios = ratio.1:ratio.5,
```

```
    weights = weight.1:weight.5, method = "iterative")

summary(fit)                  # more information

fit$means                     # (weighted) averages

fit$weights                   # total weights

fit$unbiased                  # unbiased variance estimators

predict(fit)                   # credibility premiums

##totalclaim

totalcl<-weight*ratio

totalcl

##individual sum totals for all claims per contract

contractsum<-
c(sum(totalcl[1,]),sum(totalcl[2,]),sum(totalcl[3,]),sum(totalcl[4
,]),sum(totalcl[5,]))

contractsum

##total weights

wgts<-
c(sum(weights[1,]),sum(weights[2,]),sum(weights[3,]),sum(weights[4
,]),sum(weights[5,]))

wgts

##individual mean entities have been obtained by:
```

```
ind.means<-contractsum/wgts

ind.means
```