DEPARTMENT OF ACTUARIAL STUDIES RESEARCH PAPER SERIES

Assessing the Costs of Adverse Selection

by

Piet de Jong and Shauna Ferris

piet.dejong@mq.edu.au

Research Paper No. 2004/02
February 2004

Division of Economic and Financial Studies
Macquarie University
Sydney NSW 2109 Australia

The Macquarie University *Actuarial Studies Research Papers* are written by members or affiliates of the Department of Actuarial Studies, Macquarie University. Although unrefereed, the papers are under the review and supervision of an editorial board.

Copies of the research papers are available from the World Wide Web at:

http://www.actuary.mq.edu.au/research_papers/index.html

Abstract This article develops and discusses a framework for estimating losses due to adverse selection in insurance. It is shown that expected losses depend only on means, variances and covariances of insurance factors and rates of uptake of insurance. Percentage loadings required to avoid losses are displayed. Application is made to Huntingdon disease where individuals either have the genetic marker or otherwise. Correlated information, such as family history, is also incorporated and it is seen how such information limits losses and decreases required loadings.

Keywords: Adverse selection, genetic testing, premium loading.

1 Introduction

Macdonald (1999) makes a persuasive case for actuaries to quantify the consequences of adverse selection using data and models. Adverse selection occurs when the rate of uptake of insurance is higher for higher risk groups but it is not possible to charge differential premiums based on risk group. This may may occur because the insurance company is ignorant of the relevant characteristics or if there is a prohibition on using such characteristics. This paper develops and discusses a framework for assessing the costs of adverse selection drawing on standard statistical arguments.

The further sections of this paper are as follows. The next section gives the substantive background to the problem. Section 3 outlines the general features of our proposed assessment procedure and contrasts the procedure with existing approaches. Section 4 displays percentage loadings that must be applied to avoid underwriting losses. Section 5 displays the connection of the loadings to the "variance loading" principle. Application of the formulas to life insurance contracts is made in section 6. Situations where the basic underwriting characteristic is binary are discussed in section 7 and section 8. Section 9 considers the use of ancillary information to limit underwriting losses or modify loadings. Finally in section 10 the use of binary ancillary information in the context of binary risk classification is dealt with. Proofs are contained in an appendix.

2 Historical background

In 1706, in London, Charles Povey set up the Trader's Exchange. He invited subscribers to pay a few shillings into a life insurance fund. Each subscriber could nominate a life to be insured. In the event of a death, the subscriber could make a claim. At the end of each quarter, a specified sum would be divided equally among the claimants.

> "Since policy–holders were not required to have an insurable interest in their nominees and could therefore choose whom they pleased, it can have surprised no one except Povey himself that, as he complained in July 1707, 'many Impositions were put upon the Office in its Infancy, by Peoples Subscribing upon the Lives of unhealthy Persons, and upon such too that [it] was morally impossible that they should live to a twelve Months end." (Dickson 1960, p. 24)

Naturally, the large number of deaths reduced the benefit payment per claimant – making the scheme correspondingly less attractive to new subscribers. Within a few

years, the Trader's Exchange mutual life insurance society had been dissolved or abandoned.

This may be regarded as an archetype of an "adverse selection spiral". When all people are charged the same premium – regardless of their risk – then the insurance pool creates a transfer of funds from the group of low–risk participants to the group of high–risk participants. If the low risk participants find this unacceptable, the system may not be viable.

Mr Povey realised this. In the future, he decided, it would be desirable to require each subscriber to "give a reasonable Account of the Health of the Person whose Life they intend to Subscribe upon" (Dickson 1960, p. 24). In other words, underwriting appears necessary for the development of a healthy life insurance business.

Other life offices soon attempted to identify high–risk applicants and charge differential premium rates. By 1725, they were charging a bit extra for publicans, a bit less for people who had already survived smallpox, a bit more for married women of child–bearing age, and so on.

Insurers (and others such as public health officials) were soon collecting and pooling data in order to identify and assess mortality risks. Over the next 300 years, insurers developed underwriting rules based on rating factors such as age, sex, occupation, family history, medical history, lifestyle (smoking, alcohol), and so on.

However, risk classification has often been surrounded by controversy. In some cases, there is strong opposition to the use of certain rating factors to classify risks. The main issues are examined in the next few subsections.

2.1 Is risk classification fair?

Actuaries often suggest that risk classification systems should be actuarially fair. The American Academy of Actuaries suggests that a system will be fair if:

> "Differences in prices among classes reflect differences in expected costs with no intended redistribution of subsidy among classes."

However, the general public might not agree that "actuarial fairness" is really fair in any ethical sense. This is particularly likely to be an issue when risk classification is based on factors which are beyond the control of the individual (i.e. "not his fault") (Chuffart 1995).

Most people would agree that it is fair to charge extra premiums for smokers or racing car drivers. After all, they choose a lifestyle which increases their mortality risk. However, many people do not think that it is fair to charge extra premiums to people who are blind, or people who carry a genetic mutation which increases risk. There is a natural tendency to feel that "There, but for the grace of God, go I." Figure 1 shows the result of a 1990 survey in the USA (Skipper and Black 2000, p. 644).

In recent years there have been a number of fierce controversies about risk classification systems which discriminate against people who are already disadvantaged, often through no fault of their own. For example:

- In the United States, some health insurers discriminate against women who have been victims of domestic violence. They argue that these women have a higher-than-average probability of making claims in the future, so it is "actuarially fair". (see, for example, Hellman (1997)). But several state legislatures have passed laws to prevent such discrimination.

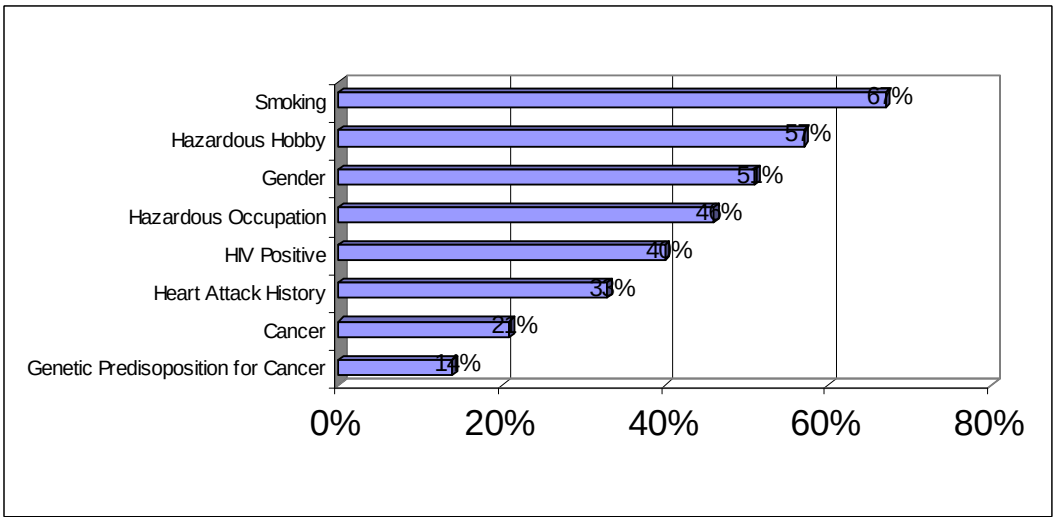

Figure 1: Percentage support for the use of different underwriting factors

- In the United States, many insurers have begun to use credit history in underwriting for motor vehicle insurance and homeowners insurance. There is statistical evidence to show that people with a poor credit rating have higher loss ratios, so it is actuarially fair. But many are concerned that this will create an additional financial burden (in the form of higher premiums) for low-income people who are already struggling. The American Academy of Actuaries noted that "using credit history is often perceived to be in conflict with what society considers as fair, particularly if the individual's score is affected by catastrophic events such as divorce, medical problems, or loss of a job."

- In the Netherlands, the Health Council has opposed the use of genetic testing for life insurance underwriting, stating that "We find it unacceptable that people affected from birth with a genetic predisposition should be faced with additional social obstacles, and that their relatives should also be at a disadvantage in this way."

If insurers are prevented from using certain rating factors, then this may create cross-subsidies between groups - the low risk members of society will subsidize the high-risk members of society. From a public policy perspective, this may be regarded as a desirable outcome.

However, there may be limits to the generosity of the low-risk members of society. Although surveys show that some people would be willing to pay higher premiums to subsidize the more unfortunate members of the community, they don't want to pay more than 5% or 10% extra (Chuffart 1995).

2.2 Is risk classification an unacceptable breach of privacy?

In some cases, life insurers may ask insurance applicants very personal questions – and people resent this intrusion into their privacy. This is particularly likely to be a problem when people may suffer from discrimination if certain personal information is revealed.

As an example, consider the impact of AIDS on insurance. In some countries, insurers began to ask people about their sexual preferences and practices – information

which many people might be reluctant to reveal, especially when there is a social stigma associated with certain sexual preferences.

The use of genetic test data in insurance is also raising concerns, particularly in countries where people with genetic defects were the victims of eugenicists in the 1930s. (For example, in Germany, people with Huntington's were killed in accordance with the government's health strategy (Harper 1992)).

2.3 Does risk classification have undesirable consequences?

If insurers use genetic test data in underwriting, then some people may decide it would be better to avoid or delay testing. This may mean that they do not obtain an accurate diagnosis of their health risks, and hence they might not obtain the best possible treatment.

This became an issue in the UK in the 1980s. At least one insurer refused to provide insurance to people who had taken an HIV test – even if the test result was negative. The insurers argued that people probably would not bother to take a test unless they believed themselves to be at risk – possibly because of high–risk practices such as unprotected sex with strangers, or needle–sharing. A government survey found that some people (a very small minority) would avoid taking an HIV test, because it might affect their ability to obtain insurance. The insurance industry was subject to a considerable amount of adverse comment in the press (Leigh 1996). The issue has also been raised in relation to genetic testing, in several countries.

2.4 Is it equitable?

Ideally, risk classification should sort people into homogeneous groups – so that people in the same group have the same risk, and can be charged the same premium.

However, risk classification is not an exact science and in some cases can lead to the formation of heterogeneous groups, leading to some inequity. Naturally, the people who are disadvantaged are likely to object to this unfair treatment.

As an example, consider the controversy about unisex rating in pension schemes. On average, women live longer than men. So pension funds would usually provide a lower annual payment to women (in return for the same lump sum). This may provide equity between groups (men and women) – but it does not necessarily provide equity for individuals within each group. This may make the risk classification process difficult to explain, especially when people know that they are not "typical" of other members of the group to which they have been assigned.

Woman: Why are you paying me such a low annual payment compared to Joe?

Insurer: Because you're a woman and women have a higher life expectancies than men.

Woman: I have had two heart attacks already! How can you say that I have a higher life expectancy than Joe?

Heterogeneity is particularly likely to arise when there is no causal connection between the rating factor and the risk – often when a rating factor is used as a proxy for some other causal risk factor which is difficult to measure.

In the past, life offices often used race as a rating factor. Mortality rates for blacks were higher than mortality rates for whites. In this case, race was being used as a proxy

for social disadvantage – black people were likely to be poorer and hence suffered all the extra mortality risks associated with poverty.

As an example, consider the use of sex as a rating factor in motor vehicle insurance. It is true that on average, women drivers have lower claims costs than men. However, this difference does not derive from an innate superiority of driving skills arising from possession of an X chromosome. Instead, it reflects the fact that women drivers tend to have a lower exposure to risk – in general, they don't drive as far as men, and they don't drive on the same type of journeys as men. That is, sex is being used as a proxy for exposure to risk. Clearly, rating by sex improves equity for some people (those who fit the stereotype) but it imposes an unfair financial penalty on others (e.g. men who have a low exposure to risk).

2.5 Is risk classification really necessary?

Insurers usually argue that risk classification is necessary – if they are not allowed to classify risks, then an adverse selection spiral might develop, pushing up premium rates and making insurance unaffordable for most people. Insurers can provide examples such as the Traders' Exchange (mentioned above), the MIRAS problems in the UK, the introduction of non–smoker discounts in the USA, and the difficulties arising from community rated health insurance in Australia.

However, others argue that the risk of an adverse selection spiral has been overstated. This spiral is only likely to occur under specific circumstances, e.g. when the elasticity of demand for insurance products is high.

In practice, despite dire predictions of disaster, the elimination of rating factors does not always lead to disruption of the insurance market. In some cases, the elimination of a rating factor might only lead to a relatively small increase in premium rates, which can be absorbed by the market.

Some have suggested that insurers might be "crying wolf" about the risks of adverse selection, seeking to preserve their "Right to Underwrite". For example Hall (1999) writes:

> "The problem with this argument [i.e. re adverse selection] is that it can easily be overblown. In my experience, actuaries are sometimes like the boy who cried wolf when it comes to adverse selection. Adverse selection has failed to materialize in several areas of health insurance, despite doomsday predictions. A case in point is the new federal law, HIPAA, which requires small-group insurers to guarantee-issue all their policies and to shorten pre-existing condition exclusions somewhat. In an extensive series of interviews I have conducted over the past year, most pricing actuaries expected at most about a 1-2% increase in costs as a result of the small-group part of HIPAA, and they viewed this as too little and too uncertain an effect to bother adding it to their rates. They were content instead to determine whether measurable adverse selection occurred at all and then pick up the costs in subsequent renewal rate increases if needed. These actuaries had learned from experience with the earlier state law versions of small-group laws that they had greatly overestimated the extent of adverse selection from guaranteed issue and rating restrictions. In fact these laws caused no detectible drop in enrollment of increase in premiums."

In the UK, insurers tried to persuade the government to allow the use of genetic test data in underwriting. The insurance industry argued that if they were not allowed to use genetic test data, then they might suffer from adverse selection, increasing premiums for all and possibly reducing the availability of some types of insurance. However, when pressed, they found it difficult to estimate the probable cost. As Macdonald (2000, p. 4) put it:

> "The insurers could not say whether lack of access to genetic test results might cost the industry £1 or £1,000,000,000."

In the same vein, Thomas (2001) writes:

> "...there is potential actuarial justification for permitting insurers' access to genetic test results. The potential justification is that such access is essential to the viability of private insurance markets. But there is currently no evidence, in any class of insurance, that this either is or will become the case. The most recent actuarial studies suggest that even under the most unfavourable assumptions, ignoring genetic tests and family history would have little effect on most insurance markets."

3 Our approach

So here is our problem: How can we measure the monetary benefit of using risk classification? Accurate risk classification avoids the problem of adverse selection since each risk is charged the actuarially fair value. Inaccurate risk classification implies subsidies from low risk to high risk groups. Cross subsidies create incentives for adverse selection and excessive contamination of the risk pool with riskier policies. The cost of adverse selection is the extra premium income required to cover adverse selection. Estimating the costs requires an assessment of the extra premium. The cost may increase over time as more and more heterogeneity in risk is introduced into the insurance pool.

To quantify the adverse selection losses and required increases in premiums to prevent them, we use the precise connection between adverse selection costs and covariances or correlations between the rate of uptake in insurance and the costs of providing the insurance. The rate of uptake of insurance measures the propensity of individuals with different characteristics to take up insurance or take up larger amounts of insurance. The cost of providing the insurance company may vary with these same characteristics.

Correlation is measured across contracts in the common insurance pool. The fact that the costs are related to this correlation is perhaps not surprising. However, this paper displays the precise connection of how the connection varies with different types of contracts. The equations can be used estimate costs of adverse selection since costs are quantified in terms of readily interpretable constructs.

Macdonald (1999) develops and discusses two approaches for quantifying adverse selection losses: a Markov approach and the random future lifetime approach. The Markov approach (Hoem 1969; Hoem 1988; Norberg 1995; Macdonald 1997; Subramanian and Lemaire 1999) follows risks through time as they transit between different states corresponding to different characteristics. States are classified according to genetic composition, whether genetic testing has been undertaken, the results of such genetic tests, insurance status, and so on. Examples are given in Macdonald (1997) which leads to the

suggestion that 10% is a reasonable order of magnitude for the additional cost of adverse selection. Random lifetime models (Bowers, Gerber, Hickman, Jones, and Nesbitt 1986; Gerber 1990) parametrize the survival function in terms of frailty value associated with each individual. For example the force of mortality (or hazard) may be assumed to be a standard force multiplied by frailty value which varies with individuals. For both the Markov and random lifetime models, a detailed picture is built up of mortality and insurance rates of different individuals. The detailed model is simulated to arrive at likely costs of adverse selection. The approach is "bottom up" in that adverse selection costs are quantified using detailed assumptions about risks and individual behavior.

The approach of this paper departs from the existing approaches. Our approach is a "top down" and shows that means, variances and correlations of insurance values to rates of insurance are, to borrow statistical terminology, "sufficient" for estimating adverse selection losses. Other quantitative aspects are irrelevant. Of course detailed modelling will give practical insight into the assignment of numerical values to these "sufficient" statistics. The approach of this paper smooths over the details of why differences exists and instead focusses on the measuring of those differences via simple statistics and translating these summary statistics into adverse selection losses. Expressions for adverse selection losses under different scenarios are examined. Loss mitigation strategies using ancillary ancillary information are examined.

4 Adverse selection losses

Throughout the rest of this paper g denotes one or more parameters which serve to classify risks. For example g may include the sex of a person, their age, or gene characteristics. All risks for a given value of g are assumed to be statistically homogenous. In given situations g may be unknown to either the individual seeking insurance, insurance company, or both. Practical interest centers on situations where, for each insurance contract, g is unknown (or unavailable) to the insurance company but known to insurance applicant.

Insurance applicants act on their knowledge of g by buying varying amounts of insurance depending on g and captured through a rate of insurance r_g. Variation in r_g can be due to variation in the likelihood of individuals to buy insurance, or variation in the amount of insurance.

Put μ_g as the expected payout for a risk of type g. In the absence of knowledge of g, the insurance company may be inclined to charge $\bar{\mu} = \mathrm{E}(\mu_g)$, the expected value of μ_g, averaging over all g in the population. This will lead to losses if r_g, the rate insurance, varies with g. To avoid losses, the insurance company is assumed to add a proportionate loading to the premium, independent of g, reflecting ignorance of it.

Theorem 4.1 *Suppose μ_g is the expected value of a benefit for a risk with characteristics g. Further suppose r_g is the rate of insurance given g and the insurance company charges $(1+\lambda)\mathrm{E}(\mu_g)$. Then the expected loss is*

$$\mathrm{cov}(r_g, \mu_g) - \lambda\{\mathrm{E}(r_g)\mathrm{E}(\mu_g)\} \ . \tag{1}$$

If $\lambda = 0$ then the loss (1) is $\mathrm{cov}(r_g, \mu_g)$. The expected loss is zero if $\lambda = \lambda^*$ with

$$\lambda^* \equiv \frac{\mathrm{cov}(r_g, \mu_g)}{\mathrm{E}(r_g)\mathrm{E}(\mu_g)} = \upsilon_{r_g}\upsilon_{\mu_g}\mathrm{cor}(r_g, \mu_g) \ , \tag{2}$$

$$(1+\lambda^*)\mathrm{E}(\mu_g) = \mathrm{E}(\mu_g) + \frac{\mathrm{cov}(r_g, \mu_g)}{\mathrm{E}(r_g)} \ , \tag{3}$$

and where υ_{r_g} *and* υ_{μ_g} *are the coefficients of variation of* r_g *and* μ_g*, respectively, and* cor *denotes correlation.*

The loading λ^* is called "adverse selection" loading and is the proportional increase in premium required to ensure, on average, zero underwriting losses. The premium $(1+\lambda^*)\mathrm{E}(\mu_g)$ is called the "adverse selection loaded" premium. Thus the proportionate increase in premium is large if the correlation and coefficients of variation are large. The increase in premium is paid for by all insuring individuals paying more than their "fair" share.

Expression (2) for λ^* indicates that sufficient statistics for avoiding losses are the means, variances and correlation of μ_g and r_g. While explanations of the size of these sufficient statistics will be useful for assigning actual values to these quantities it must be stressed that adverse selection losses do not depend on such explanations.

The expression for λ^* in (2) is relative to a population of risks, for example males aged 45. Deeper drilling down to for example 45 year old male, non smokers, will lead to a different value of λ^*, which may be both higher or lower depending on the the size of the coefficients of variation and the correlation.

Further noteworthy features of the result are as follows. First, no distributional assumptions are made and the results apply equally to discrete or continuous distributions. Second, the loading λ^* is large if the coefficients of variation of r_g or μ_g are large or if their correlation is large. This seems to accord with intuition. For example, on average r_g may be very small but the variation moderate in which case υ_{r_g} is large leading to large proportionate increases in the premium. Next, the loss is positive if and only if the uptake r_g is positively related, in a linear sense, to μ_g since correlation is a measure of linear association. Hence underwriting losses depend only on the degree of linear association between r_g and μ_g even though they may, for example, be perfectly non–linearly associated.

Insight into the likely size of λ^* in (2) is gained by making reference to the background random variables r and g such that $\mathrm{E}(r|g) = r_g$, $\mathrm{E}(\mu|g) = \mu_g$ and

$$\mathrm{cov}(r, \mu) = \mathrm{cov}(r_g, \mu_g) + \mathrm{E}\{\mathrm{cov}(r, \mu|g)\} \ ,$$

where $\mathrm{cov}(r, \mu|g)$ is the conditional covariance between r and μ given g. Thus

$$\lambda^* = \frac{\mathrm{cov}(r, \mu) - \mathrm{E}\{\mathrm{cov}(r, \mu|g)\}}{\mathrm{E}(r)\mathrm{E}(\mu)} \approx \frac{\mathrm{cov}(r, g)\{\mathrm{cov}(g)\}^{-1}\mathrm{cov}(g, \mu)}{\mathrm{E}(r)\mathrm{E}(\mu)} = \mathrm{cor}(r, g)\mathrm{cor}(\mu, g)\upsilon_r\upsilon_\mu \ , \tag{4}$$

where $\mathrm{cov}(g) \equiv \mathrm{cov}(g, g)$ and the approximation is exact under appropriate regularity such as normality or g binary. Hence λ^* is effectively $\mathrm{cov}(r, \mu)/\{\mathrm{E}(r)\mathrm{E}(\mu)\}$ if within each group g the covariation between r and μ is small. Similarly, adverse selection costs, λ^* are effectively zero if the majority of the covariation between r and g arises from within groups. Note from (4) that λ^* is less than $\mathrm{cov}(r, \mu)/\{\mathrm{E}(r)\mathrm{E}(\mu)\}$ provided within group correlation is on average positive.

5 Covariance loading

Since adverse selection losses are related to the covariance between rates of insurance and actuarial values it seems reasonable to load premiums by a factor proportional to the covariance. This is reminiscent of the variance premium principle introduced by Buhlmann (1970).

Theorem 5.1 *Suppose the notation and conditions of Theorem 4.1 and suppose premium*

$$\mathrm{E}(\mu_g) + \beta \mathrm{cov}(r_g, \mu_g)$$

is charged for each unit of insurance. Then the expected loss is

$$\{1 - \beta \mathrm{E}(r_g)\} \, \mathrm{cov}(r_g, \mu_g) \ .$$

The loss is zero if $\beta = \{\mathrm{E}(r_g)\}^{-1}$ *in which case the premium is the adverse selection loaded premium* $(1 + \lambda^*)\mathrm{E}(\mu_g)$.

Not surprisingly, fair covariance loading, that is loading to yield zero loss, thus produces the same premium as the zero loss premium given in Theorem 4.1. The result has the reasonable implication that under covariance loading, all round higher rates of insurance serve to reduce the excess on each individual contract.

6 Life insurance and lifetime annuities

The results of Theorem 4.1 apply to life insurance contracts. For example μ_g may be the benefit of 1 payable on the death of a person with characteristics g, often denoted A_g. Alternatively μ_g can denote the value of an immediate annuity of 1 payable until the death of a person with characteristics g, often denoted $\ddot{a}_g$ with $A_g = 1 - d\ddot{a}_g$ where d is the rate of discount.

6.1 Whole life insurance with annual premiums

Whole life insurance benefit A_g may be paid for with annual premiums. Early death thus results not only in early payment of the benefit but also early termination of the annuity payments. The cost associated with adverse selection is thus expected to be higher. The next result substantiates this intuitive feel and provides practical expressions for the evaluation of exact losses.

Theorem 6.1 *Suppose whole life insurance with expected value* A_g *is paid for with lifetime annual premiums*

$$(1 + \lambda)\frac{\mathrm{E}(A_g)}{\mathrm{E}(\ddot{a}_g)} \ , \tag{5}$$

where $\ddot{a}_g$ *is the expected value of an immediate annuity on a life* g *and* λ *is a loading. If* r_g *is the rate of insurance given* g*, then the expected loss to the insurance company is*

$$\left\{ \frac{1 + \lambda}{1 - \mathrm{E}(A_g)} - \lambda \left(1 + \frac{1}{\lambda^*} \right) \right\} \mathrm{cov}(r_g, A_g) \ , \tag{6}$$

where λ^* is defined in (2) with $\mu_g = A_g$. If $\lambda = 0$ the expected loss is

$$\frac{\mathrm{cov}(r_g, A_g)}{1 - \mathrm{E}(A_g)} \ . \tag{7}$$

The loading to yield zero loss is $\lambda = \lambda^\dagger$ where

$$1 + \frac{1}{\lambda^\dagger} = \{1 - \mathrm{E}(A_g)\} \left(1 + \frac{1}{\lambda^*}\right) \ . \tag{8}$$

Thus it is clear that adverse selection loadings are greater if the benefit $\mathrm{E}(A_g)$ is paid for with annual payments rather than a single payment. Note that if loadings are zero, the ratio of losses under a single up front premium to annual premiums $1 - \mathrm{E}(A_g)$, unrelated to r_g.

6.2 Separate loading of benefit and premiums

An alternative approach to deriving a zero loss annual premium for whole life insurance and under adverse selection is to separately load the benefit $\mathrm{E}(A_g)$ and the premium stream. Thus the annual premium becomes

$$\frac{(1 + \lambda_A)\mathrm{E}(A_g)}{(1 + \lambda_{\ddot{a}})\mathrm{E}(\ddot{a}_g)} \ .$$

If λ_A and $\lambda_{\ddot{a}}$ are determined so as to yield zero loss on A_g and $\ddot{a}_g$, respectively, then from Theorem 5.1 the annual premium is

$$\frac{\mathrm{E}(A_g) + \frac{\mathrm{cov}(r_g, A_g)}{\mathrm{E}(r_g)}}{\mathrm{E}(\ddot{a}_g) + \frac{\mathrm{cov}(r_g, \ddot{a}_g)}{\mathrm{E}(r_g)}} \ . \tag{9}$$

It must be stressed that the rate r_g in the above setup is the same in both the numerator and denominator and reflects the rate of whole life insurance financed by annual premiums. Thus $\mathrm{cov}(r_g, A_g) = -d\mathrm{cov}(r_g, \ddot{a}_g)$ indicating that the loadings in the numerator and denominator are of opposite sign.

Theorem 6.2 *Suppose benefit A_g is paid for with annual premiums (9). Then the expected loss to the insurance company is zero.*

A noteworthy feature of the zero loss annual premium (9) is that, following the proof, can be written as

$$\frac{1}{\mathrm{E}\left\{\frac{r_g}{\mathrm{E}(r_g)}\ddot{a}_g\right\}} - d \ ,$$

which is the usual annual premium except that the average $\mathrm{E}(\ddot{a}_g)$ has been replaced by a weighted average with weights r_g.

7 Adverse selection losses for binary situations

This section specializes the results of the previous section when g is binary: $g = 1$ or 0. These specialized results lead to additional insights into sources and magnitudes of adverse selection losses and the effect of loss mitigation strategies.

An example binary situation is Huntingdon disease: a person either has the genetic marker for the disease $g = 1$, or otherwise, $g = 0$. For a particular insurance applicant, the value of g may be known or unknown to either the individual, the insurance company, or both. Even if it is known, the insurance may be prohibited from discriminating on the basis of g.

As in previous sections μ_g denotes the expected cost given a risk of type g. If the insurance company knows g and can charge on the basis of g then μ_1 or μ_0 will be charged depending on the value of g.

Theorem 7.1 *Suppose the notation and conditions of Theorem 4.1 and that g is binary, $g = 0$ or 1 with $p = \mathrm{E}(g)$. Then*

$$\lambda^* = \frac{p(1-p)}{(p+\frac{1}{i_r})(p+\frac{1}{i_\mu})} = \frac{\phi(p,i_r)-p}{p+\frac{1}{i_\mu}} \ , \qquad \phi(p,i_r) \equiv \frac{(1+i_r)p}{1+pi_r} \ , \tag{10}$$

$$(1+\lambda^*)\mathrm{E}(\mu_g) = \mu_0 + (\mu_1-\mu_0)\phi(p,i_r) \ , \tag{11}$$

where $i_r \equiv (r_1/r_0) - 1$ and $i_\mu \equiv (\mu_1/\mu_0) - 1$.

The result demonstrates that the presence of risks $g = 1$ in the insurance pool changes premiums in three ways:

- The difference in cost $\mu_1 - \mu_0$ of insuring risks of the type $g = 1$ compared to "normal" risks $g = 0$.
- The proportion p of risks of type $g = 1$.
- An "adverse selection" effect, $\phi(p, i_r) - p$, due to differential rates of insurance for $g = 0$ and $g = 1$ risks.

Expression (10) indicates how the three sources combine to arrive at the total adjustment. The loading $p(\mu_1-\mu_0)$ over and above the "normal" premium μ_0 in (10) is required given the presence of proportion p of individuals type $g = 1$ in the insurance pool. It is the "direct" increase in premiums associated with having the higher risk individuals $g = 1$ in the insurance pool. The final term in square brackets in (10) is the "adverse selection" loading necessitated by the fact that $g = 1$ individuals insure at a differential rate to the $g = 0$ individuals. The adverse selection cost approaches zero as $r_1 \to r_0$ or $p \to 0$ or $p \to 1$.

Figure 2 plots $\phi(p, i_r)$ as a function of $0 < p < 1$ and for different $0 \le i_r \le 100$. The 45 degree line plot corresponds to $r_1/r_0 = 1$ or equivalently $i_r = 0$ and hence there being no difference in rates of insurance between the two groups. In this case, for given difference $\mu_1-\mu_0$, the entire increase in premium varies directly with p, the proportion of $g = 1$. The first line above the 45 degree line corresponds a rate r_1 double that of r_0, the next line above that is a rate r_1 triple that of r_0, and so on up to the highest line which represents a rate $r_1 = 100r_0$. The difference in a particular curve from the 45 degree line represents the adverse selection cost. Figure 2 illustrates that adverse selection costs can be a major component when p is low and i_r is large.

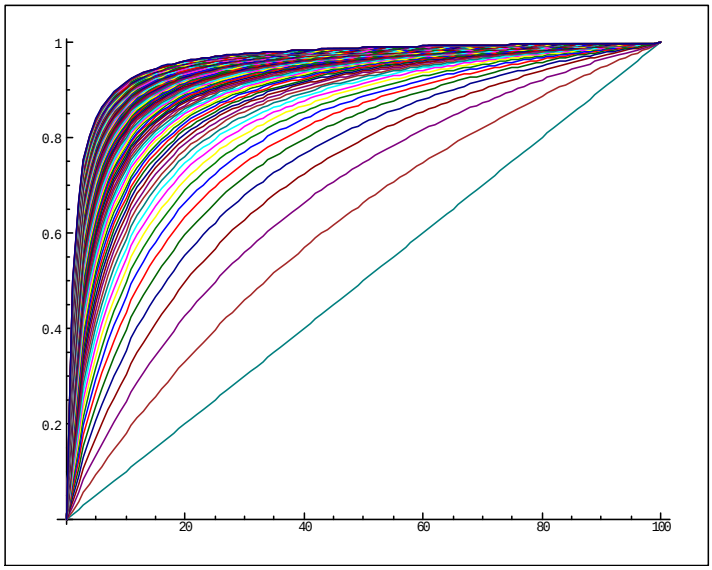

Figure 2: Direct plus adverse selection costs $\phi(p, i_r)$ plotted against $0 \leq 100p \leq 100$ for $i_r = 0, 1, \ldots, 100$.

8 Whole life insurance with annual premiums for binary situations

Theorem 8.1 *Suppose g is binary, $g = 0$ or $g = 1$ with $p = P(g = 1)$. Then under adverse selection with insurance rates r_0 and r_1, the per unit sum assured annual premium for whole life insurance with zero loss to the insurance company is*

$$\frac{\pi_0 - di_{\ddot{a}}\phi(p, i_r)}{1 + i_{\ddot{a}}\phi(p, i_r)}$$

where π_0 is the normal annual premium $A_0/\ddot{a}_0$, $i_{\ddot{a}} = \ddot{a}_1/\ddot{a}_0 - 1$ and $\phi(p, i_r)$ is the function plotted in Figure 2.

Note that if the $g = 1$ group is subject to higher mortality then $i_{\ddot{a}}$ is negative. Thus the numerator is inflated whereas the denominator is deflated leading to higher premiums, as expected.

9 Using ancillary information

Let h be available information on each risk and that h can be used by the insurance company to set rates. In practice h will be correlated with g, but it is assumed that given g, μ_g does not depend on h. In other words knowledge of g renders h obsolete. It is assumed all information in h is included in g and that some attributes in g are not in h. Thus the potential insured are assumed to know more about the risk than the insurance company. This may be due to legal restrictions on insurance companies from using certain information. The value of h is used to determine an extra to the normal premium. Instead of charging $\mathrm{E}(\mu_g)$ the company charges $\mathrm{E}(\mu_g) + x_h$ where x_h is an extra premium depending on h and it is implicitly assumed that h can be used for ratemaking. Such extras will limit losses to a degree that h is predictive for g or μ_g.

Given available information h there are two ways in which h may be used to set rates

- The population may be divided according to h with an analysis as in §4 except that all expressions are conditional h. Thus adverse selection costs are shared within each h group.

- The population can be treated as a whole except that premiums are calculated as $\mathrm{E}(\mu_g|h)$. A common adverse selection percentage loading λ_h^* is then imposed all premiums to yield $(1+\lambda_h^*)\mathrm{E}(\mu_g|h)$. The premium for a specific h group may go up or down relative to $(1+\lambda^*)\mathrm{E}(\mu_g)$ depending on the net effect of the loading λ_h^* and adjustment moving from $\mathrm{E}(\mu_g)$ to $\mathrm{E}(\mu_g|h)$.

In this section we consider the second alternative.

Theorem 9.1 *Suppose the above notations and conditions but the insurance company charges* $(1+\lambda)\mathrm{E}(\mu_g|h)$. *Then the expected loss is*

$$\left\{1-\rho_h-\lambda\left(\rho_h+\frac{1}{\lambda^*}\right)\right\}\mathrm{cov}(r_g,\mu_g)\ , \qquad \rho_h \equiv \frac{\mathrm{cov}\{r_g,\mathrm{E}(\mu_g|h)\}}{\mathrm{cov}(r_g,\mu_g)}\ . \tag{12}$$

If $\lambda=0$, *the expected loss is* $(1-\rho_h)\mathrm{cov}(r_g,\mu_g)$. *The expected loss is zero if* $\lambda=\lambda_h^*$ *where*

$$\lambda_h^* = \frac{(1-\rho_h)\lambda^*}{1+\rho_h\lambda^*} \qquad \Rightarrow \qquad 1+\lambda_h^* = \frac{1+\lambda^*}{1+\rho_h\lambda^*}\ . \tag{13}$$

If $\lambda^*\neq 0$ *then* $\lambda_h^*=0$ *if and only if* $\rho_h=1$.

Note that in the above setup the loading λ_h^* works to increase all premiums by the same percentage amount after adjusting for h. This contrasts with the procedure where each group defined by h is treated separately and adverse selection costs are measured and spread across each separate group.

The loading λ_h^* is the adverse selection loading based on h. The premium $(1+\lambda_h^*)\mathrm{E}(\mu_g)$ is the adverse selection loaded premium given h and, if $\rho_h>0$, is less than the "raw" adverse selection loaded premium $(1+\lambda^*)\mathrm{E}(\mu_g)$. Obviously if $h_2\supseteq h_1$ then $\rho_{h_2}\geq\rho_{h_1}$ implying $\lambda_{h_2}^*\leq\lambda_{h_1}^*$. If μ_g is perfectly predictable from h in the sense $\mu_g=\mathrm{E}(\mu_g|h)$ then $\lambda_h^*=0$. Thus the role of h is to decrease the loading to an extent depending on the proportion of $\mathrm{cov}(r_g,\mu_g)$ explained by h.

The quantity ρ_h can be negative which occurs if

$$\mathrm{cov}\{r_g,\mu_g-\mathrm{E}(\mu_g|h)\}>\mathrm{cov}(r_g,\mu_g)\ .$$

Hence removing the effect of h serves to increase the covariation between r_g and μ_g. A negative ρ_h implies a higher adverse selection loading. Of course particular groups may still have lower overall premium if $\mathrm{E}(\mu_g)-\mathrm{E}(\mu_g|h)$ is very negative.

If h is not used and there is no adverse selection loading the loss is $\mathrm{cov}(r_g,\mu_g)$. If h is used without adverse selection loading, the loss is reduced to $(1-\rho_h)\mathrm{cov}(r_g,\mu_g)$. If $\rho_h=1$ no loss remains and hence no loading is required.

Figure 3 displays the effect of $\sqrt{\rho_h}$ on λ_h^* for different values of λ^*. Each curve corresponds to a given value of λ^* with $0\leq\lambda^*\leq 2$. When $\rho_h=0$, $\lambda_h^*=\lambda^*$ indicated by the value on the y-axis. As $\sqrt{\rho_h}$ increases to 1, λ_h^* decreases with $\lambda_h=0$ when $\rho_h=1$ for all λ^*, as indicated on the extreme right of the graph.

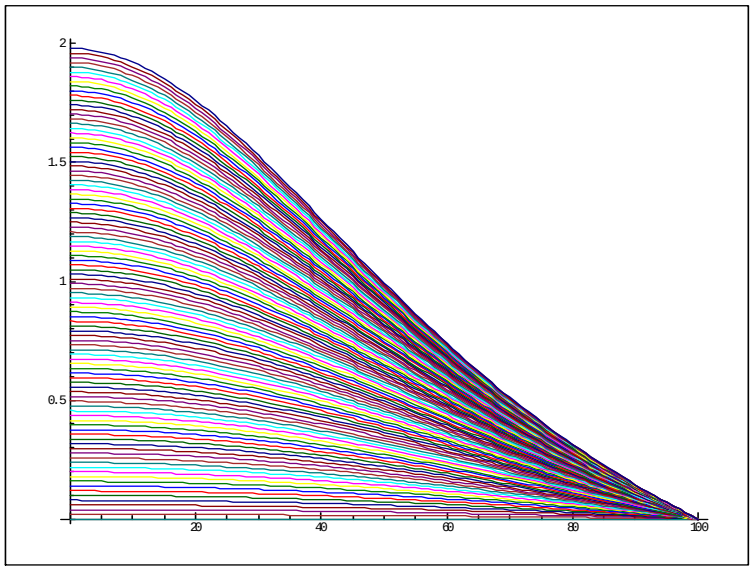

Figure 3: Graph of λ_h^* as a function of $\sqrt{\rho_h}$
where $0 \leq 100\sqrt{\rho_h} \leq 100$ and for $0 \leq \lambda^* \leq 2$.

The graph indicates the values of $\sqrt{\rho_h}$ required to for example halve the raw adverse selection loading. For example if $\lambda^* = 1$ indicating a 100% loading, in order to halve this to 50% requires $\sqrt{\rho_h} \approx 0.6$. Denoting $\alpha = \lambda_h^*/\lambda^*$ as the required reduction then $\rho_h = (1-\alpha)/(1+\alpha\lambda^*)$ is required.

In certain cases the premium $\mathrm{E}(\mu_g|h)$ is linear in h. This occurs for example if h is binary or if (μ_g, h) is normally distributed.

Theorem 9.2 *Suppose* $\mathrm{E}(\mu_g|h)$ *is linear in* h. *Then*

$$\rho_h = \frac{\mathrm{cov}(r_g, h)\{\mathrm{cov}(h)\}^{-1}\mathrm{cov}(h, \mu_g)}{\mathrm{cov}(r_g, \mu_g)} \ , \tag{14}$$

where $\mathrm{cov}(h) \equiv \mathrm{cov}(h, h)$. *Hence* ρ_h *is the proportion of* $\mathrm{cov}(r_g, \mu_g)$ *explained by* h.

If $\mathrm{E}(\mu_g|h)$ is not linear in h then the right hand side of (14) can serve as the definition of ρ_h. With this definition (13) holds provided loss is minimized in the mean square error sense.

10 Setting premiums using dichotomous information

This section considers adverse selection loading when correlated binary information is used to set premiums. For example in the case of Huntingdon disease, family history may be used set rates. Family history is well known to partially predict the presence of the Huntingdon disease gene. A family history of the disease predisposes a person towards the presence of the gene, $g = 1$, even though a $g = 0$ has about a 50% chance of occurring when $h = 1$, and $g = 1$ is possible if $h = 0$, that is if there is no family history of the disease.

Theorem 10.1 *Suppose the conditions and the notation of the previous sections and that g and h are binary with $\mathrm{E}(g) = p$, $\mathrm{E}(h) = q$ and $\mathrm{E}(gh) = s$. Then the adverse selection loaded premium is*

$$(1 + \lambda_h^*)\{\mu_0 + (\mu_1 - \mu_0)\mathrm{E}(g|h)\} \tag{15}$$

$$= \frac{\{\mu_0 + (\mu_1 - \mu_0)\left[\phi(p, i_r) + (1 + \lambda^*)\{\mathrm{E}(g|h) - p\}\right]}{1 + \rho_h \lambda^*} \ , \tag{16}$$

where λ^* and λ_h^* are given in (10) and (12), respectively, and

$$\rho_h = \frac{(s - pq)^2}{p(1 - p)q(1 - q)} \ . \tag{17}$$

and hence ρ_h is the squared correlation between g and h.

Note that $\mathrm{E}(g|h)$ is the conditional probability of $g = 1$ given h. Thus the adverse selection loaded premium for the case $h = 1$ is the base premium μ_0 plus the fraction s/q of $\mu_1 - \mu_0$, with the sum "loaded up" by the factor $1 + \lambda_h^*$. The fraction s/q is the conditional probability of $g = 1$ given $h = 1$. The same computation applies for $h = 0$ except that the fraction is the conditional probability $(p - s)/(1 - q)$ of $g = 1$ given $h = 0$. Alternatively, from the right hand side of (15), the premiums can be viewed as the raw loaded premium plus or minus $(1 + \lambda^*)(\mu_1 - \mu_0)$ times the difference between the conditional and unconditional probability of $g = 1$, and the sum "discounted back" by the factor $1 + \rho_h \lambda^*$. Note that if $\rho_h \to 0$, the premiums reduce to (10). If $\rho_h \to 1$, then the two premiums approach μ_0 and μ_1 for $h = 0$ and $h = 1$, respectively.

APPENDIX: PROOFS

Item numbers below refer to the relevant Theorem.

4.1 and 9.1 The proof is for the general situation in Theorem 9.1 from which the results in Theorem 4.1 follow as a special case. If g and h are continuous with joint probability density $p(g, h) = p(h|g)p(g)$ then the probability density of a randomly selected unit of insurance is $r_g p(g)/\mathrm{E}(r_g)$ and the expected loss associated with a randomly drawn unit of insurance is

$$\int_{g,h} (\mu_g - \pi_h) \frac{r_g p(h|g) p(g)}{\mathrm{E}(r_g)} \, dhdg = \frac{\mathrm{E}\{r_g(\mu_g - \pi_h)\}}{\mathrm{E}(r_g)}$$

$$= \frac{\mathrm{cov}(r_g, \mu_g - \pi_h)}{\mathrm{E}(r_g)} + \mathrm{E}(\mu_g - \pi_h) \ ,$$

where $\pi_h = (1 + \lambda)\mathrm{E}(\mu_g|h)$ is the premium and all expectations and covariances are with respect to the joint distribution of g and h. Multiplying by the total units of insurance $\mathrm{E}(r_g)$ and simplifying yields the loss

$$\mathrm{cov}(r_g, \mu_g) - (1 + \lambda)\mathrm{cov}\{r_g, \mathrm{E}(\mu_g|h)\} - \lambda \mathrm{E}(r_g)\mathrm{E}(\mu_g) \ .$$

Further simplification yields the first expression in (12). Equating the term in curly brackets of (12) to zero and solving for λ leads to the stated expression for λ_h^* in (13).

5.1 The result follows immediately upon identifying $\beta \mathrm{cov}(r_g, \mu_g)\{\mathrm{E}(\mu_g)\}^{-1}$ with λ^* of Theorem 4.1.

6.2 The annual premium (9) can be written as

$$\frac{\mathrm{E}(r_g A_g)}{\mathrm{E}(r_g \ddot{a}_g)} = \frac{\mathrm{E}(r_g)}{\mathrm{E}(r_g \ddot{a}_g)} - d \ ,$$

where d is the rate of discount and we have used $A_g = 1 - d\ddot{a}_g$. Then the expected loss is

$$\int_g \left[A_g - \left\{ \frac{\mathrm{E}(r_g)}{\mathrm{E}(r_g \ddot{a}_g)} - d \right\} \ddot{a}_g \right] r_g p(g) \ dg = \mathrm{E}(r_g A_g) - \mathrm{E}(r_g) + d\mathrm{E}(r_g \ddot{a}_g) = 0 \ ,$$

which completes the proof.

7.1 A direct proof is to note that $r_1 p / \{r_1 p + (1-p) r_0\} = \phi(p, i_r)$ is the relative frequency of $g = 1$ policies in the pool of all policies, implying the expected cost of insurance is the stated expression for $(1 + \lambda^*)\mathrm{E}(\mu_g)$ which can be used to solve for λ^*. However it also useful to deduce the result from the more general expressions given in Theorem 4.1. Since g is binary, $\mathrm{cor}(r_g, \mu_g) = 1$. Hence from Theorem 4.1 the loading λ in $(1+\lambda)\mathrm{E}(\mu_g)$ required for zero loss is $\lambda = \lambda^*$ where $\lambda^* = \sigma_{r_g} \sigma_{\mu_g} / \{\mathrm{E}(r_g)\mathrm{E}(\mu_g)\}$ where both r_g and μ_g are binary. Elementary calculations show

$$\mathrm{E}(r_g) = r_0 + (r_1 - r_0)p \ , \qquad \sigma_{r_g} = |r_1 - r_0| \sqrt{p(1-p)} \ ,$$

with similar results for μ_g. Since $r_1 - r_0$ and $\mu_1 - \mu_2$ have the same sign,

$$(1 + \lambda^*)\mathrm{E}(\mu_g) = \{\mu_0 + (\mu_1 - \mu_0)p\} \left[1 + \frac{p(1-p)(r_1 - r_0)(\mu_1 - \mu_0)}{\{r_0 + p(r_1 - r_0)\} \{\mu_0 + p(\mu_1 - \mu_0)\}} \right] \ ,$$

which simplifies to the expression given in (10). Further

$$\lambda^* = \frac{\mathrm{cov}(r_g, \mu_g)}{\mathrm{E}(r_g)\mathrm{E}(\mu_g)} = \frac{(r_1 - r_0)(\mu_1 - \mu_0)p(1-p)}{\{r_0 + p(r_1 - r_0)\}\{\mu_0 + p(\mu_1 - \mu_0)\}} \ ,$$

which simplifies to stated expression for λ^*.

8.1 Using the covariance loading principle the zero loss annual premium is

$$\frac{A_0 + (A_1 - A_0)\phi(p, i_r)}{\ddot{a}_0 + (\ddot{a}_1 - \ddot{a}_0)\phi(p, i_r)} \ .$$

Dividing both numerator and denominator by $\ddot{a}_0$ and simplifying using $A_1 - A_0 = -d(\ddot{a}_1 - \ddot{a}_0)$ yields the required result.

9.1 Put π as the premium (5). Since $\mathrm{E}(A_g) = 1 - d\mathrm{E}(\ddot{a}_g)$ where d is the rate of discount, it follows,

$$\pi = (1 + \lambda) \left\{ \frac{1}{\mathrm{E}(\ddot{a}_g)} - d \right\} \ .$$

Hence the expected loss associated with a contract on the life of person of type g is

$$A_g - \pi \ddot{a}_g = 1 - d\ddot{a}_g - (1 + \lambda) \left\{ \frac{1}{\mathrm{E}(\ddot{a}_g)} - d \right\} \ddot{a}_g$$

$$= 1 + d\lambda \ddot{a}_g - (1 + \lambda) \left\{ \frac{1}{\mathrm{E}(\ddot{a}_g)} \right\} \ddot{a}_g \ .$$

The total loss, adding over all contracts is

$$\int_g \left[1 + \lambda d \ddot{a}_g - \left\{ \frac{1 + \lambda}{\mathrm{E}(\ddot{a}_g)} \right\} \ddot{a}_g \right] r_g p(g) \ dg = \mathrm{E}(r_g) + \lambda d \mathrm{E}(r_g \ddot{a}_g) - (1 + \lambda) \frac{\mathrm{E}(r_g \ddot{a}_g)}{\mathrm{E}(\ddot{a}_g)}$$

$$= \lambda d \left\{ \mathrm{cov}(r_g, \ddot{a}_g) + \mathrm{E}(r_g)\mathrm{E}(\ddot{a}_g) \right\} - \lambda \mathrm{E}(r_g) - (1+\lambda)\frac{\mathrm{cov}(r_g, \ddot{a}_g)}{\mathrm{E}(\ddot{a}_g)} \ .$$

The first two terms in the last expression simplify to

$$-\lambda \left\{ \mathrm{cov}(r_g, A_g) + \mathrm{E}(r_g)\mathrm{E}(A_g) \right\} \ .$$

Substituting into the previous equation, rearranging and simplifying yields (6).

9.2 If $\mathrm{E}(\mu_g|h)$ is linear in h then

$$\mathrm{E}(\mu_g|h) = \mathrm{E}(\mu_g) + \mathrm{cov}(\mu_g, h)\{\mathrm{cov}(h)\}^{-1}\{h - \mathrm{E}(h)\} \ ,$$

and hence

$$\mathrm{cov}\{r_g, \mathrm{E}(\mu_g|h)\} = \mathrm{cov}(r_g, h)\{\mathrm{cov}(h)\}^{-1}\mathrm{cov}(h, \mu_g) \ .$$

Substituting into the definition of ρ_h given in (12) yields (14). Further $1 - \rho_h$ is the proportion of $\mathrm{cov}(r_g, \mu_g)$ unaccounted for by h and hence ρ_h is the proportion of $\mathrm{cov}(r_g, \mu_h)$ explained by h.

10.1 Since g is binary, $\mu_g = \mu_0 + (\mu_1 - \mu_0)g$ implying $\mathrm{E}(\mu_g|h) = \mu_0 + (\mu_1 - \mu_0)\mathrm{E}(g|h)$ and hence the left side of (15). Writing $(1 + \lambda_h^*) = (1 + \rho\lambda^*)^{-1}(1 + \lambda^*)$ in this expression yields the right hand side from the left side in (15). The expression (17) follows from

$$\rho_h = \frac{\mathrm{cov}(r_g, h)\{\mathrm{cov}(h)\}^{-1}\mathrm{cov}(h, \mu_g)}{\mathrm{cov}(r_g, \mu_g)} = \frac{(r_1 - r_0)(\mu_1 - \mu_0)\{\mathrm{cov}(g, h)\}^2}{(r_1 - r_0)(\mu_1 - \mu_0)\mathrm{cov}(g)\mathrm{cov}(h)} \ ,$$

where $\mathrm{cov}(g, h) = \mathrm{E}(gh) - \mathrm{E}(g)\mathrm{E}(h) = s - pq$, $\mathrm{cov}(g) = p(1-p)$ and $\mathrm{cov}(h) = q(1-q)$. Simplifying the above expression yields (17).

References

Bowers, N., H. Gerber, J. Hickman, D. Jones, and C. Nesbitt (1986). *Actuarial Mathematics*. Itasca, Illinois: Society of Actuaries.

Buhlmann, H. (1970). *Mathematical methods in risk theory*. Berlin: Springer–Verlag.

Chuffart, A. (1995). Genetic underwriting. Paper presented to the Actuarial Life Convention, Glasgow, December 1995.

Dickson, P. (1960). *Sun Insurance Office Ltd, 1710–1960: the history of two and a half centuries of British insurance*. Oxford: Oxford University Press.

Gerber, H. (1990). *Life Insurance Mathematics*. Berlin: Springer-Verlag.

Hall, M. (1999). Restricting insurers use of genetic information: A guide to public policy. *North American Actuarial Journal 3*(1).

Harper, P. (1992). Huntington disease and the abuse of genetics. *American Journal of Human Genetics 50*, 460–464.

Hellman, D. (1997). Is actuarially fair insurance pricing actually fair?: A case study in insuring battered women. *Harvard Civil Rights - Civil Liberties Law Review 32*, 355.

Hoem, J. (1969). Markov chain models in life insurance. *Blatter der Deutschen Gesellschaft fur Versicherungsmathematik 9*(2), 91–107.

Hoem, J. (1988). The versatility of the Markov chain as a tool in the mathematics of life insurance. *Transactions of the 23rd International Congress of Actuaries R*, 171–202.

Leigh, S. (1996). The freedom to underwrite. Paper presented to the Staple Inn Actuarial Society at the Institute of Actuaries, London.

Macdonald, A. (2000). Human genetics and insurance issues. In I. Torrance (Ed.), *Bio-ethics for the new millenium*. St Andrew Press.

Macdonald, A. S. (1997). How will improved forecasts of individual lifetimes affect underwriting? *British Actuarial Journal 3*(15), 1009–1025.

Macdonald, A. S. (1999). Modeling the impact of genetics on insurance. *North American Actuarial Journal 3*(1), 83–101.

Norberg, R. (1995). Differential equations for moments of present values in life insurance. insurance. *Insurance: Mathematics & Economics 17*, 171–180.

Skipper, H. and K. Black (2000). *Life and Health Insurance* (13 ed.). Prentice Hall.

Subramanian, K., J. and A. S. Lemaire (1999). Estimating adverse selection costs from genetic testing for breast and ovarian cancer: The case of life insurance. *Journal of Risk and Insurance 66*(4), 531–50.

Thomas, R. (2001). Genetics and insurance. Response to the Human Genetics Commission Public Consultation, available on website http://www.guythomas.org.uk/pdf/HGC27Feb.pdf.