

australian
actuarial journal



2012
VOLUME 18
issue 2



**Actuaries
Institute**

Acting Editor

David Pitt BEc, BSc, ATCL, PhD, FIAA

Associate Editors

Anthony Asher BBusSc, PhD, FIA, FIAA

Tim Kyng BSc, M Stats, MEc, FIAA, F Fin

Sachi Purcal BEc, MA, PhD, ASA

Bridget Browne BEc, FIAA, FIA, Actuaire (IA)

Rade Musulin BA, ACAS, MAAA

Rob Thomson PhD, FASSA

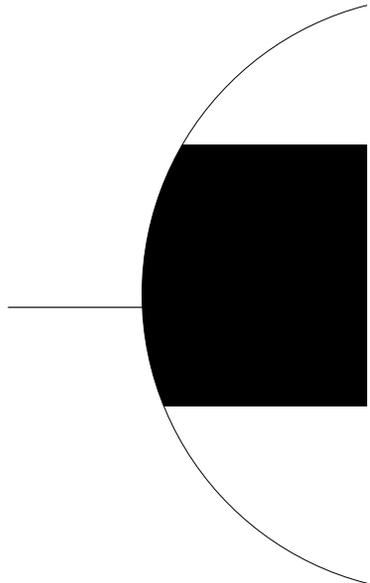
ISSN 1442-3065

© Actuaries Institute of Australia 2012

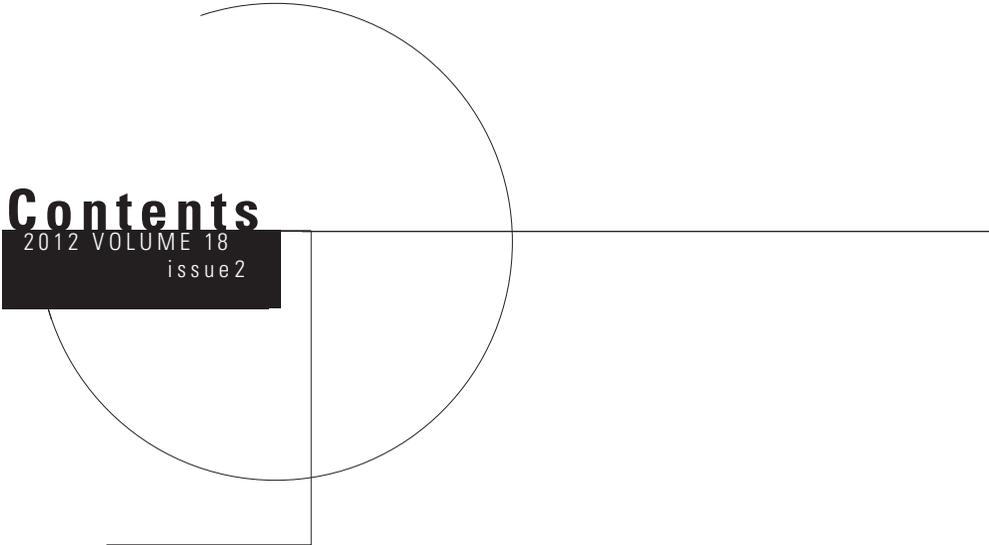
The Australian Actuarial Journal is published by
Actuaries Institute of Australia
Level 7 Challis House
4 Martin Place Sydney NSW 2000 Australia
Telephone: +61 (0) 2 9233 3466

For subscription information and guidelines for authors please see inside of back cover.

The opinions expressed in this journal are not necessarily those of the Editor, the Associate Editors, the Advisory Panel, or the Actuaries Institute of Australia. The Editor, the Associate Editors, the Advisory Panel and the Actuaries Institute of Australia do not accept any responsibility for those opinions or for the accuracy of any statements appearing in the journal.



australian actuarial journal



Contents

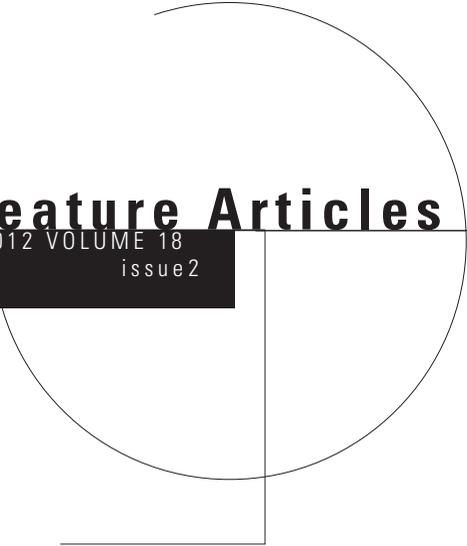
2012 VOLUME 18
issue 2

Feature Articles

- Mispriced Risk in Insurance and Financial Markets: Causes and Consequences
– *S Ferris* 125
- On the Use of Limited-Value Average in Actuarial Modelling –
J Li, B Uditha and Y Sin Lee 191

Thank you to Reviewers

The Australian Actuarial Journal relies upon the goodwill and effort of many people to review papers prior to publication. The Editor is grateful for the help of the reviewers of papers throughout 2012.



Feature Articles

2012 VOLUME 18
issue 2

Mispriced Risk in Insurance and Financial Markets: Causes and Consequences

*S Ferris**

Abstract

Theoretically, insurance should be beneficial to society. But in practice, insurance systems may become dysfunctional, creating an excessive increase in systemic risk which is detrimental to society. In this paper we use a historical example to illustrate the typical features of a dysfunctional insurance market and construct a model of risk-taking behaviour. We draw parallels with the more recent problems in the credit risk insurance market, which contributed to a sharp increase in systemic risk prior to the Global Financial Crisis in 2008/2009. Construction of a model facilitates assessment of some proposals for reforms.

Keywords: moral hazard, asymmetric information, adverse selection, systemic risk, underwriting cycle, utmost good faith, information costs

* Contact: Shauna Ferris, FIAA. Email: shauna.ferris@mq.edu.au.
Thank you to Rade Musulin and other anonymous reviewers.

1. Introduction

Ideally, insurance should be socially beneficial. When insurance markets work well, no individual needs to suffer a catastrophic loss; instead, all members of the insurance pool bear a small share of the any losses. Risks can be traded, which allows markets to set a price for risk. Given this price, each individual and each corporation will be able to determine its own optimal risk management strategy, ie. decide whether to take a risk, ameliorate a risk, avoid a risk, or transfer a risk.

However, under certain circumstances, insurance markets can become dysfunctional. When risks are underpriced, insurance markets may create an undesirable increase in risk-taking, which may be detrimental to the entire community. In response, governments often tighten regulation in order to ameliorate the ill-effects of these imperfect markets.

In this paper we will

- describe the typical characteristics of a dysfunctional insurance market;
- develop a model for the behaviour of market participants; and
- assess some of the solutions which have been adopted in insurance markets in the past.

Our objective is to build a model which will allow us to develop a greater insight into the structure of a dysfunctional market. Building this model is, we hope, the first step towards developing solutions.

For the purposes of illustration we will look at the nineteenth-century marine insurance market in England. The key features of this market were:

- Moral hazard - Ship owners had a great deal of control over the level of risk on each voyage; and there were financial incentives to take more risk.

- Underpricing of risk - For a number of reasons (described in more detail in Section 3), marine insurance was often readily available at relatively cheap premium rates.
- Adverse Selection - The availability of cheap insurance encouraged ship owners to take on more and more risk, by overloading and under-manning their ships.
- Underpricing and the Underwriting Cycle– There were a number of obstacles which prevented the underwriters from making an accurate assessment of risk. Furthermore, competition in the insurance market forced premiums down to unprofitable levels.
- Negative Externalities - Although ship owners often prospered by overloading their ships, this led to a sharp increase in the number of shipwrecks. Thousands of sailors died. Millions of pounds of cargo were destroyed. Several insurance companies failed.
- Regulation - There was a strong public outcry, and demands for more regulation of the industry. There was considerable controversy about the various proposals for reform.

Similar problems may arise in financial markets. Over the last twenty years we have seen an explosion in financial innovation. A plethora of new products has facilitated the transfer of all sorts of financial risks. In some cases, these risk transfer markets did not work efficiently. In particular, credit risks were often seriously underpriced. This underpricing led to an increase in risk-taking. The rampant growth in subprime lending ultimately destabilised the global economy.

We will draw some analogies between the British marine insurance market in the nineteenth century and the American credit risk markets in the twenty-first century. At first glance, these two markets seem quite different. But on closer inspection, it is clear that there are some striking similarities. We will argue that both cases can be modelled using the same approach.

Insurers have had about 200 years to devise solutions to problems in the marine insurance industry. Over the years, systems have evolved to create insurance markets which provides the economic advantages of spreading risk, without producing such severe adverse consequences for society. Is it possible that some of the reforms which were devised in the nineteenth century, to deal with dysfunctional insurance markets, might be useful in dealing with contemporary problems in financial markets?

2 Outline

Section 3 gives a brief description of the marine insurance market circa 1860.

Section 4 describes the characteristics of a dysfunctional insurance market, from the perspective of the customer (the insured). We assume that the customer has at least some control over the level of risk. We model the customer's risk-taking decisions in the absence of insurance.

Section 5 describes the dysfunctional insurance system from the perspective of the insurer. The insurer must determine the rating structure and premium rates. The system will become dysfunctional if there are obstacles which prevent accurate assessment and pricing of risk. What are these obstacles?

Section 6 discusses some reforms which were designed to make the marine insurance system work in a manner which would be more beneficial to society. Many of these reforms were quite controversial and strongly opposed by influential ship owners. Some proposed that ship owners should be required to retain some risk, ie. they should have some "skin-in-the-game".

Section 7 examines the "skin-in-the-game" rules, using our model to estimate the impact of such rules under equilibrium conditions.

3 Background: A Dysfunctional Marine Insurance Market

During the 1860s and 1870s, hundreds of British ships were lost at sea, and thousands of British sailors lost their lives. A British Member of Parliament, Samuel Plimsoll, argued that most of these deaths were preventable. The ships were sinking because the ship owners were neglecting the most basic safety precautions.¹

Plimsoll pointed out that the problem was at least partly caused by the marine insurance system. Ship owners could buy insurance at Lloyd's, which would protect them from the financial effects of shipwrecks. Hence they had little incentive to ensure the safety of their vessels – and quite a large financial incentive to take on more risk, in order to increase profits. The availability of cheap insurance increased risk-taking throughout the industry – ultimately at great cost to the public, in loss of life and property.

Inefficiencies in the insurance market meant that high-risk ventures could obtain insurance at low premium rates. This created perverse incentives. Some ship owners found it profitable to deliberately overload their ships, while relying on insurance to cover the consequential losses. This led to an adverse selection spiral. Prudent ship owners found it difficult to compete with irresponsible ship owners: so prudent ship owners had an incentive to lower their own standards. The overall level of risk in the market spiralled inexorably upwards.

The level of shipwrecks increased. Insurers suffered losses; several insurers failed, resulting in losses for their policy holders. Over time, premium rates increased sharply to cover the higher level of risk; this pushed up the costs of transporting goods, which had an impact on the wider economy.

And of course there was considerable “collateral damage”. Many sailors lost their lives in the shipwrecks, leaving behind destitute

¹ Plimsoll, *Our Seamen*, 1873. Henceforth we will refer to this source as **Plimsoll**.

widows and orphans. More than 8000 lives were lost at sea off the English coast during the period from 1861 to 1870.²

In his polemic *Our Seamen*, published in 1873, Plimsoll provided a thoughtful analysis of the ill-effects of a dysfunctional insurance system. He realised that insurance was in the public interest – but only if properly regulated to eliminate perverse incentives.³

4 The Ship owners' Perspective

In order to understand the mechanics of this dysfunctional market, we need to model the ship owners' decision-making processes.

Each businessman must manage risk. Risk managers must address the following issues:

- **Control:** What risk factors are under my control?
- **Profitability:** How will different risk management strategies affect the profitability of my business?
- **Risk-Return Tradeoff:** Would it be better to forego some profits in order to reduce risk?
- **Diversification:** Can these risks be managed by diversification?
- **Risk Transfer:** Would it be better to transfer some risk to others, eg. by buying insurance or using derivative contracts?
- **Competition:** How will profitability be affected by our competitor's management of risk?

After considering these issues, the businessman will choose the optimal level of risk for his business.

² Plimsoll, page 56, citing Board of Trade statistics.

³ Plimsoll was certainly not the first person to recognise the deficiencies of the marine insurance system. Parliamentary enquiries on wrecks in the timber trade to examine the causes of shipwrecks were set up in 1839 and 1843. "*The investigation drew attention to ten determining factors, including defective construction, inadequate equipment, imperfect state of repair, improper and excessive loading, incompetence of masters, drunkenness among officers and crew, and marine insurance which inclined shipowners to disregard safety*" (emphasis added) cited in Boisson (1999).

Unfortunately, the optimal level of risk for his business may have detrimental consequences for other members of the community.

4.1 Control

First year actuarial students learn that risks are only insurable if they are *fortuitous* – ie. if the occurrence of the event is largely a matter of chance. If the event is not fortuitous – if the insured has a great deal of control over the risk – this may create a moral hazard.

In the 1860s, the owner of a ship had a great deal of control over the risk of shipwreck.

Going to sea was a dangerous business. A certain level of risk was inevitable. Given the state of marine engineering at that time, ships might be expected to founder in extreme weather – such losses were unavoidable. But ship owners could minimise the risk by taking sensible precautions.

Plimsoll analysed the data on shipwrecks from the 1860s. He found that many of the shipwrecks should never have happened, because they were entirely preventable. Plimsoll found that many ships were sinking in relatively moderate weather ie. when the wind was force six or under (no more than a strong breeze).⁴

Why were they sinking?

Plimsoll believed that the ship owners were deliberately taking on more and more risk. The ship owners had control over many of the factors which affected risk: the level of loading; the construction and maintenance of ships; and the number of crew on board.

Overloading: Experienced sailors had rules for safe load limits, eg. by measuring the ratio of freeboard to the depth of the hold. But Plimsoll observed that more and more ships were travelling very low in the water, due to excessive cargos. In stormy weather, the seas

⁴ Plimsoll page 1-12

might easily wash over the deck, and pour into the hold of the ship. The ship would then founder.

Maintenance: A prudent ship owner would carefully overhaul his ship after each voyage. But a government investigation into shipwrecks found that hundreds of deaths were caused by defects in the ships or equipment, including poor maintenance and repairs.

Under-manning: Plimsoll also noted that ship owners were cutting costs by reducing the size of the crew. A small crew could handle the ship safely during fair weather, but would not be able to cope during a storm.⁵

Construction: Ship owners were deliberately constructing vessels which were very likely to be unstable. Ship owners had realised that longer ships could carry more cargo with little extra expense – and make higher profits. As a result, well-built ships were being deliberately modified to be longer than carry more cargo – even when it was clear that this would make them dangerously unstable.⁶

Moral hazard exists whenever the insured has a great deal of control over the level of risk. Of course, moral hazard is present to some degree in many types of insurance contract: eg. a householder who has insurance might be rather careless about locking his back door. But in such cases the insured's behaviour has a relatively small impact on the level of risk. Moral hazard is a much more serious problem when the insured's actions can double or triple or quadruple the risk.

⁵ Plimsoll page 57

⁶ Plimsoll, page 67

Analogy to credit markets

In the US home loan market, the lender has a great deal of control over the level of risk for any loan. The lender controls:

- The level of “overloading”, ie. the risk of default is increased by lending more than the borrower can afford to repay (measured by the size of the loan relative to the borrower’s income); and the risk of loss-given-default is increased by allowing high loan-to-value ratios.
- The level of “manning” ie. the lenders could cut costs by skipping checks on income levels and collateral values, eg. by offering low-doc loans and omitting independent property valuations.
- The “construction of the vessel”, ie. risk can be increased by changing the design of the loan product. Products which included features such as adjustable interest rates, low initial repayments, and negative amortisation are inherently unstable.

Factors which are under the lender’s control have an enormous impact on the probability of default. Over the last few years, we have seen the wide gap between the default rates on traditional loans, and the default rates on various types of high risk loans. For example, in August 2009 the Mortgage Bankers Association National Delinquency Survey reported that the delinquency rate on subprime loans was 25.35%, compared to 6.41% for prime loans; the foreclosure rate for subprime loans was 4.13%, compared to a rate of 1.01% for prime loans.⁷

In order to model the behaviour of our insurance system, we will need to estimate a *Risk Function*. We will assume that the probability of a shipwreck (q) depends on the Load (L). The Load variable represents all the risk factors which are under the control of the ship owner.

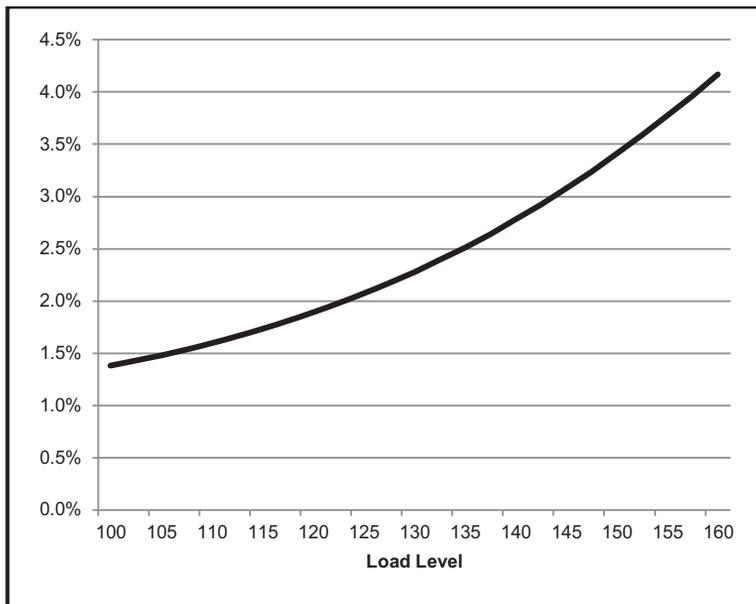
$$\text{Probability of Shipwreck} = q(L)$$

In our model, we will assume that a higher Load increases the risk of shipwreck (ie. q is an increasing function of L). The graph below

⁷ Mortgage Brokers Association (2010)

shows a hypothetical Risk Function. The derivation of this Risk Function will be explained in more detail below.

Figure 1 Probability of shipwreck as a function of load



4.2 Profitability: Financial incentives to take more risk

Moral hazard exists whenever the insured has a great deal of control over the level of risk. When these people also have a strong financial incentive to take on more risk, then the moral hazard increases. Therefore our model of a dysfunctional system will also need to include a Wealth Function, which measures the financial incentives to take risk.

As Plimsoll pointed out, in the nineteenth century, the ship owners could make enormous profits by overloading their ships. This situation was exacerbated by leverage, because each voyage had a high level of fixed costs.

“When you consider how small an addition to the fair load of a ship will augment the profits of a trip 25, and even 50 per cent, you will easily see how great was the temptation, especially in settled weather, to add the extra weight.

When freights run low, the margin for profit over expenses is small; it may take nine-tenths of the cargo to pay the costs; an addition, then of only 10 per cent to the weight of the cargo will double the profit, and 20 per cent, which will still leave the ship in trim difficult to find fault with, will treble the earnings; and when we consider the enormous advantage this gave to the reckless, and the temptation to even those who disapproved of the practice to follow it in self-defence, it is really wonderful to me that the practice should now be, as it undoubtedly is, confined to only a section of the trade.”⁸

Analogy to credit markets

Bankers also have to make a decision about how much Load (risk) they are willing to take, and of course bank profits are highly leveraged. A bank may be able to increase its return on equity by taking on more risk – ie. by making riskier loans at higher interest rates. Obviously some banks will be tempted by this opportunity – especially if they can easily insure themselves against the default risk.

In our model, the *Wealth Function* will show the relationship between Load (risk-taking) and Wealth.

In our model, Expected Wealth is calculated as follows:

We have a ship worth S .

We can borrow L , which will be used to buy goods.

If the voyage is successful, we can sell the goods for a profit of m per unit load and repay the loan L . Our net wealth will be $S + mL$.

⁸ Plimsoll page 41; see also page 28 for similar comments

On the other hand, if the ship is wrecked then we lose the ship, and we must still repay the loan L .⁹

Mathematically:

Wealth if successful = $S + mL$ with probability $1-q(L)$

Wealth if shipwrecked = $-L$ with probability $q(L)$

Expected wealth = $[1-q(L)](S+mL) + q(L)(-L)$

We will assume that each shipwreck leads to 100% loss: nothing is salvaged. We could make a more realistic model by building in a parameter for loss-given-default, but at this stage we would prefer to keep the model as simple as possible.

Wealth Maximisation

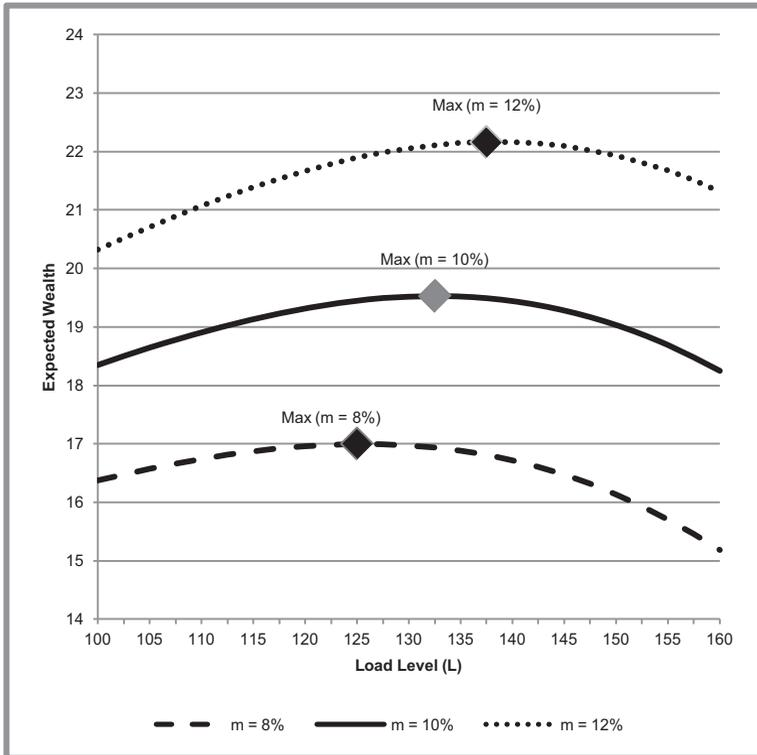
Our model suggests that the expected wealth W will vary depending on the values of the Load L , the value of the ship S , and the profit margin m .

A ship owner who wishes to maximise his expected wealth can solve the equation to find the value of L which maximises the expected wealth for given values of the profit margin m and the value of the ship S .

The solid line in the following diagram shows the expected value of wealth for $m = 10\%$ and $S = 10$. In this case the ship owner would maximise his expected wealth by choosing a Load of about 133.

⁹ We could allow for interest on the loan but this is an unnecessary complication – we could just adjust m and L to allow for interest.

Figure 2: Expected Wealth as a Function of Load, Varying Profit Margin m



The optimum Load Level depends on both the profit margin m and the value of the ship S .

As you would expect, the higher the profit margin m , the greater the potential rewards for overloading (for a given value of S). If the profit margin was 8%, then the optimum Load would be about 126 units. If the profit margin was 12%, then the optimum Load would be about 138 units.

Similarly, the value of the ship S affects the loading policy. If the ship has a low value, the ship owner has less to lose if the ship is lost at sea, and hence the ship owner might be more willing to take risks by overloading. For example, in the mid-1800s, the shipping industry

was disrupted by the development of new technology: iron ships replaced wood. This meant that the old, less efficient wooden ships were no longer very valuable. Since this reduced the value of S , it seems likely that the owners of the older wooden ships were more likely to take risks by overloading (other things being equal).

Analogy to credit markets

If we apply this model to credit risk, then S may be regarded as roughly analogous to the bank's capital, and m might be regarded as roughly analogous to the interest margin on a loan.

The model suggests that (i) lenders with low levels of capital would be more likely to take more risk and (ii) lenders will take more risk if there is a larger spread between the interest rates charged on prime loans and subprime loans.

This fits in with some observed behaviour in financial markets. For example, during the Savings and Loans crisis of the 1980s, many S&Ls were allowed to continue to operate even when their (realistic) capital levels had been severely eroded. With nothing to lose, many of the S&Ls adopted high-risk lending strategies, moving out of residential lending and into commercial property development and junk bonds.

The above model assumes that the ship owner would be required to repay the loan L even if the ship foundered. Suppose that the ship owner declares bankruptcy when his ship is wrecked; suppose that bankrupts are not required to repay any debts. We can modify the model so that the wealth is 0 after a shipwreck (instead of $-L$). If we allow for such an "insolvency put option", then the optimum Load Levels would naturally be much higher (other things being equal); and the optimum Load Level becomes very sensitive to both the profit margin m and the value of the ship S .

4.3 Risk-Return Trade-offs: Allowing for Risk Aversion

Of course, most ship owners would be risk-averse. Therefore they would not simply aim to maximise expected profits – instead, they would choose a load level which would maximise their expected

utility. Each ship owner will have his own risk-return preference which will determine the choice of Load. Hence, in order to model risk-taking behaviour, we need a *Utility Function* for the ship owner.

We will assume that our ship owners have an exponential utility function, where

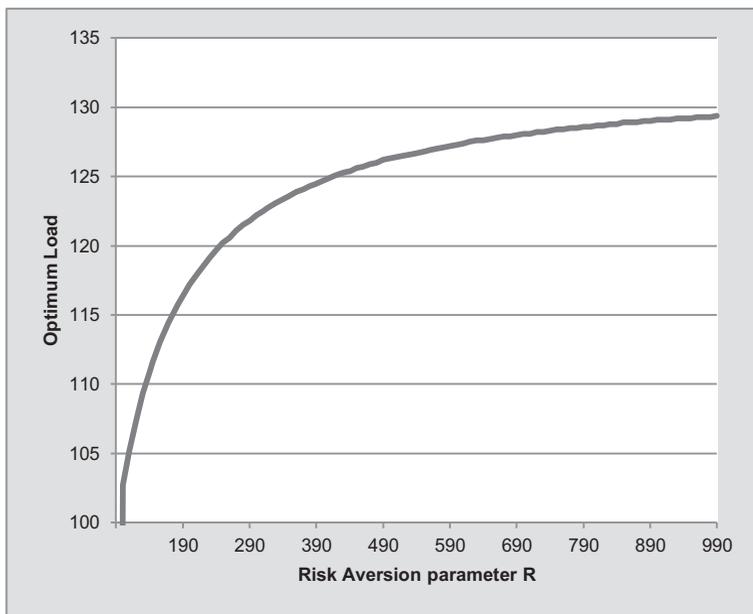
$$U(w) = 1 - \exp(-w / R) ,$$

where the parameter R reflects the ship owner's propensity to take risk.

In the following diagram we show the expected utility for ship owners who have exponential utility functions for different values of R.

When the ship owner is risk-neutral (R approaching infinity), he will choose a Load of about 133 (ie. he will simply maximise profit regardless of risk) and the shipwreck probability will be about 2.4%.

Those with higher levels of risk aversion will choose lower Loads, ie. preferring lower risk and lower profits. For example, a ship owner who is highly risk averse (R=100) will choose a load of just 103, with shipwreck probability 1.4%.

Figure 3: Optimum Load as a function of Risk Aversion R

Based on Plimsoll's observations, most *uninsured* ship owners were quite risk averse: they diligently repaired their ships after every voyage and avoided overloading.

Analogy to credit markets

Prior to the development of the securitisation market, most banks were reasonably prudent. Most banks would apply quite strict underwriting criteria for home loans, and the overall level of default was correspondingly low – even during recessions. Although the banks could have earned higher rates of interest by making sub-prime loans, they normally refrained from doing so (with some exceptions - there are always a few banks which have a stronger risk appetite).

As explained below, the development of credit risk transfer markets allowed diversification of risk, which shifted the risk-return trade-off.

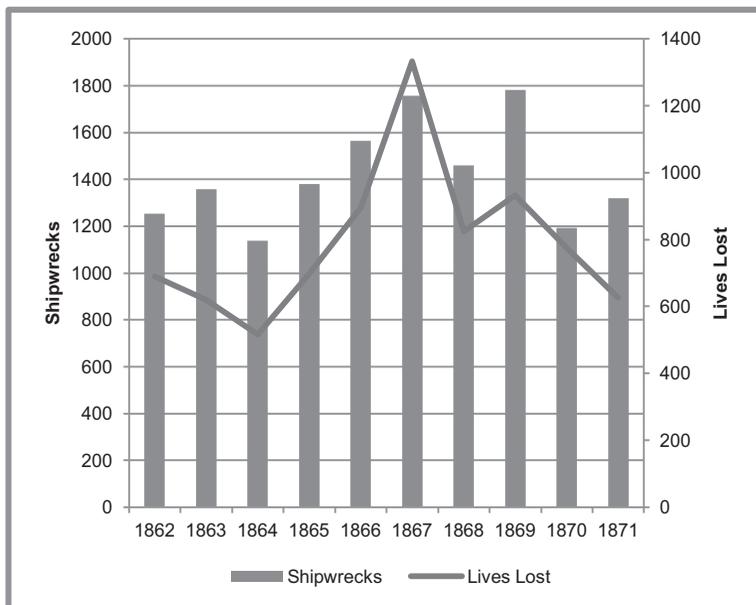
4.4 Diversification Benefits: Correlated Risks (The Weather)

Could a ship owner use diversification to manage his risk? Diversification would be most beneficial if the risks were independent. But was this a realistic assumption?

In shipping, the risk of shipwreck is not constant – it varies from one year to the next, depending on external factors such as the weather. These environmental factors affect all the ships which are at sea in a given geographical area.

The shipwreck statistics indicate that there was considerable variation in the number of wrecks from year to year. The following graph shows the number of wrecks and the number of lives lost at sea between 1862 and 1871 (Note: Plimsoll's *Our Seamen* was published in 1873).¹⁰

¹⁰ Plimsoll page 56

Figure 4: Number of Shipwrecks and Lives Lost 1862-1871

Plimsoll noted that there was an interaction between the load level and the weather. An overloaded, undermanned, and poorly maintained ship would have a good chance of survival, **as long as the weather was favourable** – in Plimsoll’s words, “*as long as the sea was as smooth as a mill-pond the whole way*”.

In good weather, safely-loaded and overloaded ships would *both* have a very good probability of survival; in poor weather, the overloaded ships would have a much higher probability of wreck.

This interaction creates a correlation between risks. Allowing for the interaction between Load and Weather, this correlation is much stronger for overloaded ships. Adopting the terminology of investment theory, since all risks are exposed to the same weather-related risks, there is a certain level of “non-diversifiable” or systematic risk. The overloaded ships have a higher exposure to these systematic risks – analogous to a high-beta share which has a stronger correlation to market forces than a low-beta share.

Analogy to credit markets

In the credit risk markets, economic conditions are analogous to “the weather”. Interest rates, home prices, and unemployment rates are factors which will affect the probability of default for *all* outstanding loans; and they are beyond the control of any individual lender. These economic conditions vary over time and affect the risk of default on all loans.

However, there is an interaction between loan type and economic conditions. When the “economic weather” is favourable, all types of loans (both prime and subprime) will have low default rates. But default rates will diverge when the economy falls into recession. Prime loans will continue to have quite low default rates. But default rates on sub-prime debt will spike sharply upwards whenever interest rates rise, home prices fall, and unemployment increases.

In order to model diversification benefits, we will need a better *Risk Function*, one which allows for both environmental factors (“the Weather”), and individual risk factors (“the Load”). Let $f(L,W)$ be the probability of shipwreck for Load L and Weather W .

For our model, we will assume that there is an interaction between Load and Weather so that:

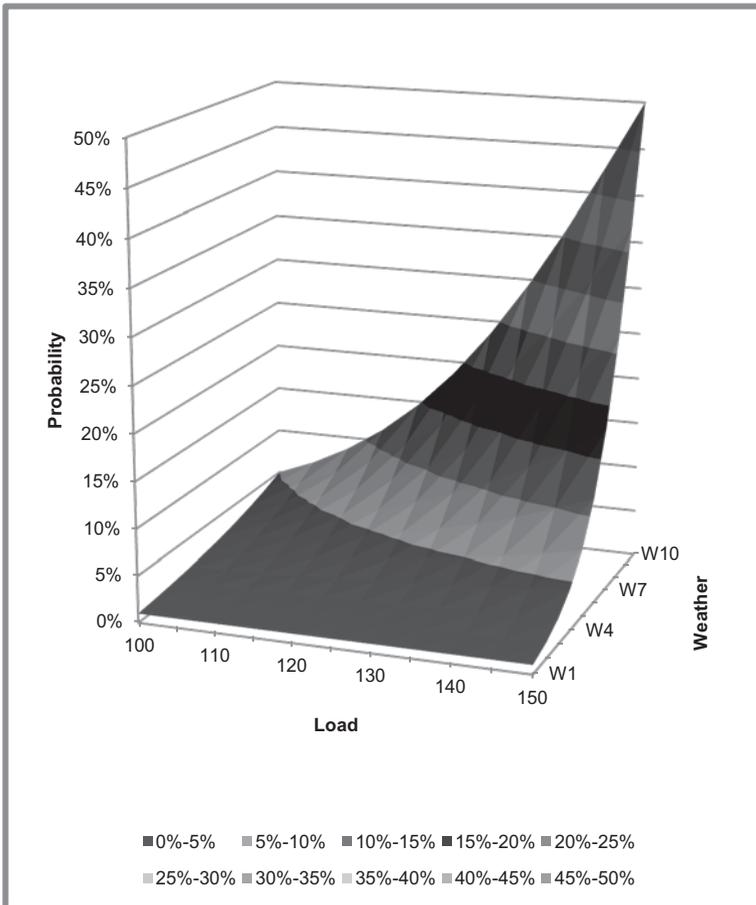
- When the Weather is favourable, the risk of loss is low for all ships, whatever Load they are carrying. In our model, we assume a 1% chance of shipwreck for all ships during very good weather.
- When the Weather is poor, then overloaded ships are especially vulnerable to shipwreck; but ships which are in a good state of repair and carrying a safe Load are much more likely to survive despite such bad weather. For the purposes of illustration, our model assumes that during a very severe storm the safest ships have a 5% chance of shipwreck, but ships which are severely overloaded have a 50% chance of shipwreck [$q(100,10) = 5\%$ and $q(150,10) = 50\%$].

- We could model the probability of shipwreck using a logistic model which allows for such interactions, ie.

$$f(L,W) = \frac{\exp(a + bL + cW + dLW)}{1 + \exp(a + bL + cW + dLW)}$$

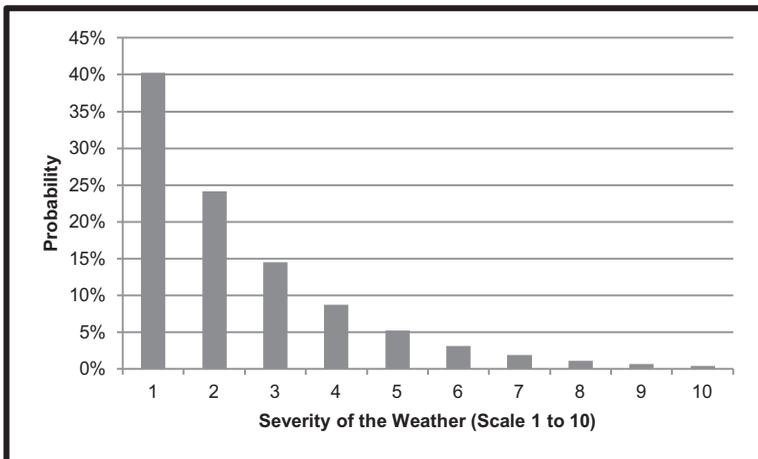
where a, b, c, and d are fitted parameters of the model.

Figure 5: Probability of Shipwreck as a function of Load & Weather



To make the model complete, we would also need to know the probability distribution for the weather. How often do severe storms occur? Naturally this would vary depending upon the route of the ship and the season. For the purposes of illustration, we have adopted the following hypothetical distribution for the weather on a particular trade route. Weather is treated as a discrete random variable which takes values from one to 10 (in Plimsoll's time, a scale of one to 10 was used to measure wind strength).

Figure 6: Distribution of the Weather



If we know the impact of Weather-Load interactions (given by $f(W,L)$) and probability of bad weather (given by $Pr(W=w)$), then we can derive the hypothetical Risk Function $q(L)$. Figure 1, given in section 3.1 above, shows this function.

We can also model the impact of diversification.

If a ship owner was able to diversify his risks, this would increase his “efficient frontier”, ie. his set of possible risk-return alternatives. This might alter his choice of Load levels – he might be willing to take more risk on each individual voyage, as long as the variation in total profits (for his whole fleet) could be reduced by diversification.

Based on our model, a risk-averse non-diversified ship owner would choose a Load Level of 103. But a ship owner who diversified his risk among 20 equally loaded ships would maximise his utility by choosing a Load Level of 125 (using the same base-case assumptions).

So the effect of diversification would be an increase in risk-taking. This matches Plimsoll's observations – he identified certain wealthy fleet owners as being particularly prone to overload all of their ships. Ship owners who owned just one ship would be reluctant to overload.

Of course, the choice of optimum load level is strongly affected by the probability distribution of the weather, which may be very difficult to assess. If ship owners assumed that the probability of bad weather was somewhat lower than it really is, (and hence overestimated the benefits of diversification), then they might easily take on too high a level of risk.

Similarly, the choice of optimum load level is strongly affected by the assumptions about the risk of shipwreck in severe weather. If ship owners underestimate the vulnerability of overloaded ships to bad weather, they would be more likely to overload. Again, this probability might be difficult to assess accurately. Suppose that in the past, very few ships were severely overloaded. Suppose that severe weather is relatively uncommon. There may be very little data on the past experience of overloaded ships in poor weather, which could lead to either under-pricing or over-pricing.

Analogy to credit markets

It is clear that many financial institutions which invested in structured products overestimated the benefits of diversification, and hence had a poor understanding of the tail risks of their portfolios. They misjudged the correlations between the different products they owned – and this led them to take on too much risk. The difficulties were exacerbated by a dearth of historical data on the behaviour of structured products during the downside of the economic cycle.¹¹

¹¹ For a critique of the models and data used by the Credit Rating Agencies, see Financial Stability Forum(2008) ; the SEC (2008) ; and Baily, Elfendorf, and Litan (2008)

This is not the first time the financial markets have overestimated the benefits of diversification. During the 1980s, Drexel Burnham encourage investors to buy diversified portfolios of “high yield” bonds (ie. junk bonds). Researchers claimed that diversification could be used to reduce portfolio. However, when the junk bond market collapsed, it became clear that the correlations between these junk bonds had been underestimated. Diversification did not protect the investors from large losses.

4.5 The Impact of Risk Transfer Arrangements

The establishment of an insurance market changes the risk-return trade-offs. Hence, the availability of insurance may well lead to an increase in risk-taking.

This effect was explicitly recognised in the earliest English legislation governing marine insurance, ie. in Elizabethan times (1601).

*“...by means of which policies of assurance it comethe to passé, upon the loss or perishing of any shippe there followeth not the undoing of any man, but the losse lighteth rather easily upon many than heavily upon fewe, and rather upon them that adventure not than upon them that doe adventures, **whereby all merchants, speciallie the younger sort, are allured to venture more willingly and more freele**”.*¹² (emphasis added)

This increase in risk-taking might be beneficial to society. Plimsoll noted that marine insurance had positive effects for society: “By giving greater security to our trade, has tended to increase our mercantile marine, to the great advantage of the nation.”¹³

4.5.1 Bottomry – The Earliest Asset-Backed Securities?

In the earliest days of marine insurance, “bottomry” was commonly used to spread risks. The ship owner or merchant would

¹² Lowndes page 33 citing the Elizabethan Statute of 1601

¹³ Plimsoll page 42

borrow money to buy goods, promising to repay the loan with interest at a specified rate if the voyage was successful. However, if the ship was wrecked, then the merchant did not have to repay the loan or the interest. That is, it was a non-recourse loan. The ship owner would still benefit from any profits made on the sale of goods, over and above the interest rate paid on the bottomry loan, but his losses would be limited.

The interest rate charged on the loan would of course reflect the dangers inherent in the voyage (the route taken, the seasonal weather effects, etc). A higher rate would be charged when there was a higher risk of shipwreck.

Analogy to credit markets

We can see the similarity to more modern methods of finance. These days, the home-lender will create a Special Purpose Vehicle which provides a source of funds for lending. The investors in the SPV receive bonds which pay a fixed rate of interest – secured by the home loan repayments. If the loans default, then the SPV investors suffer a loss. They do not have recourse to any payment from the home lender. The rate of interest earned by the SPV bond-holders should theoretically reflect the risks inherent in the portfolio of loans.

4.5.2 Marine Insurance at Lloyd's

Over time a more efficient marine insurance market developed, largely centred on Lloyd's in London.

As an example, consider the risk-averse ship owner described previously. In the absence of any insurance, he will choose a Load of 103. How will his decision be affected by the existence of insurance?

This depends on the structure of premium rates.

Case 1: Risk Premium

If the insurance premium is equal to the risk premium (ie. the expected value of losses), and the ship owner can insure the full value of the ship and the load and the expected profit on sale of goods, then the premium will be $P(L) = q(L) * (S + L + mL)$.

The ship owner's wealth is:

Wealth = $S + mL - P(L)$ if the voyage is successful

Wealth = $S + mL - P(L)$ if the voyage ends in shipwreck

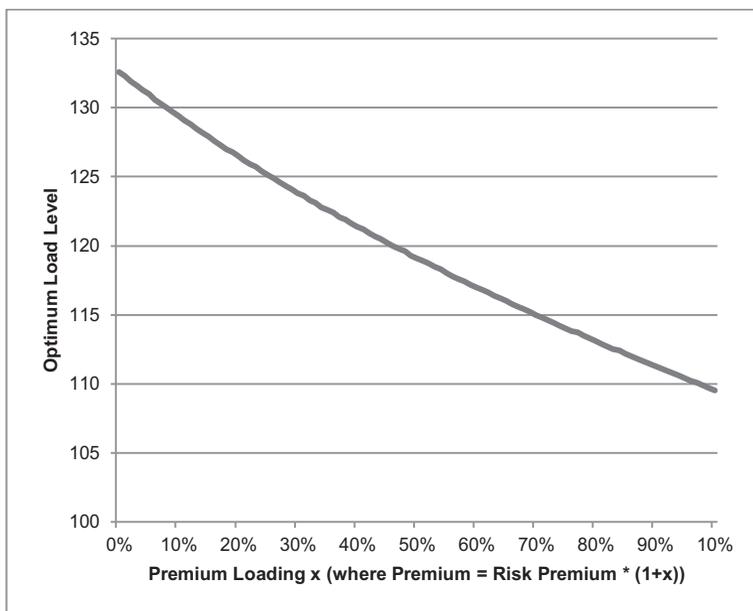
Expected Wealth = $S + mL - P(L)$ with certainty

In this case, the insurer has eliminated the risk to the ship owner, without making any charge for the transfer of risk. That is, the market price of risk is 0. Based on our model, in these circumstances every ship owner, even those who are most risk-averse, should choose the Load Level which maximises profit. The optimum Load Level will jump to 133 for every ship owner.

Case 2: Expected Value Pricing plus a Risk Margin

Of course, insurers would normally include a risk loading in their premium rates. The simplest type of risk loading is a constant percentage loading (ie. the premium charged is $(1+x)$ * the risk premium). How would this percentage loading affect the ship owners decision about risk-taking? The following diagram shows the optimal load levels (ie. optimal in terms of maximising the utility function) for different percentage loadings on the risk premiums. For example, if the risk premium is loaded by 30%, the optimum Load falls to 124; a premium loading of 60% would reduce the optimum Load to 117.

Figure 7: Optimum Load Levels for an Insured Ship as a function of the Premium Loading x



The insurer's risk aversion would affect the size of the premium loading – which would in turn affect the level of risk-taking in the shipping industry.

Of course, these days insurers would normally use more sophisticated methods to determine the appropriate risk margin. These methods usually require a reasonable estimate of the distribution of outcomes. It would be an interesting exercise to assess the impact of different theoretical pricing systems on the optimum load levels. However, in the marine insurance markets of the 1860s, the pricing of insurance was much less sophisticated. For reasons outlined in Section 5 below, marine insurers had great difficulty in estimating accurate risk premiums for different risks.

Case 3: Pricing based on Average Industry Loss Ratios

In many cases, the Lloyd's underwriters tended to charge the same premium rate for all ships, as long as they were registered in the Lloyd's Book.¹⁴ Since the same premium was charged regardless of the quality of construction, then was no incentive for the ship owners to manage their risks. This created a "race to the bottom" in the constructing and maintenance of ships.

"Underwriters, it is notorious, as a rule, make no distinction in their premiums; all who come up to the standard in Lloyd's Book are rated alike; that is, perhaps, a necessary result of the wholesale system of dealing to which they are driven by competition. ...A new ship, then, seldom or never falls short of this minimum; but equally seldom, not to say never, exceeds it...That is to say, the present system of insurance causes that the minimum of efficiency becomes, in the vast majority of cases, likewise the maximum. The ship owner who insures his ship up to her value has no pecuniary interest in doing more to her, in any direction, beyond coming up to a sort of minimum standard."¹⁵ (emphasis added)

We can incorporate this approach into our model by setting a premium rate which is unrelated to the shipwreck risk for any individual ship. Suppose that premium rate is based on the average shipwreck rate for all ships, which maybe denoted as q . The premium charged is simply q multiplied by the sum insured. Assuming that the sum insured covers the ship plus the value of all the cargo (including the expected profit margin on the cargo).

$$\text{Premium} = q * (S + L + mL)$$

Then the insured ship owner will have

$$\text{Expected Wealth} = S + mL - \text{Premium with certainty}$$

¹⁴ The premiums would of course reflect the length of the trip, the route, seasonal weather factors, the nature of the cargo, etc.

¹⁵ Lowndes (1884) page 108

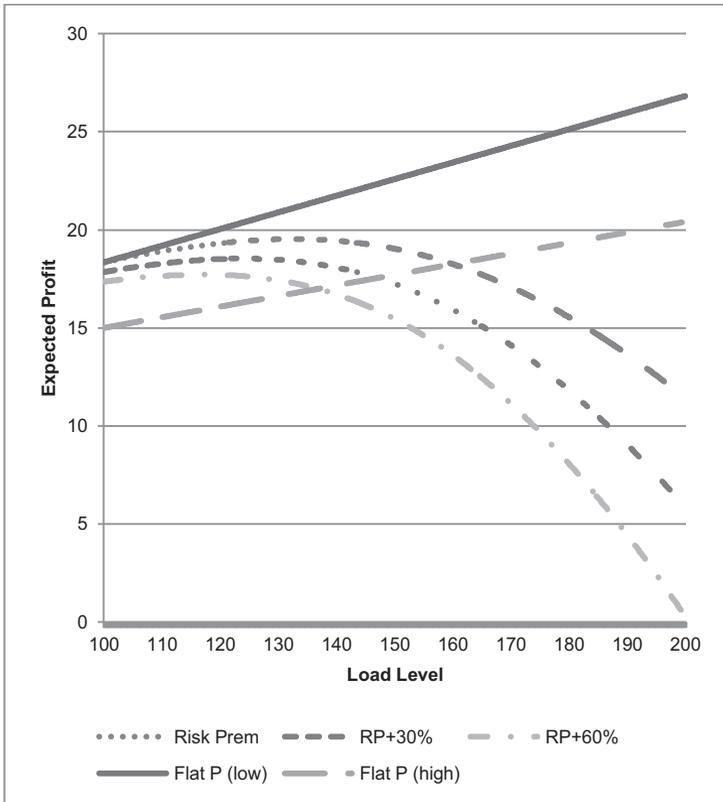
This means that the ship owner has every incentive to load up as much cargo as the ship can physically bear.

The graph shows the effect. The two straight lines represent the expected profit when a flat premium rate is applied. One line represents a low value of q (industry average risk equal to Load of 100), the other represents a high value of q (industry average risk equal to Load 160). In both cases the profit-maximiser will choose the highest possible Load, ie. as much as he can physically fit onto the ship.

This fits with the behaviour which Plimsoll observed in the 1860s. Some ships were indeed laden to the absolute physical limit – ie. so heavily laden that the deck was barely above the waterline. When seas were rough, waves could easily wash over the decks and pour into the hold, with disastrous results.

For comparison, the graph also shows the Expected Profits when the premium rate is the risk premium appropriate for the ship's own load, plus a profit loading of $x\%$. Three curved lines represent this premium structure, with loadings of 0%, 30%, and 60%. In these cases, the profit-maximiser will choose Load levels of 133, 124, and 117 respectively.

Figure 8: Impact of Different Premium Rates on Expected Profit & Optimum Load



4.6 Competition and the Adverse Selection Cycle

Whenever insurers charge the same premium rates for different risks, there is the risk of creating an adverse selection cycle.

Despite the financial incentives to overload, perhaps some ship owners would initially resist this temptation. After all, each shipwreck was likely to lead to the deaths of all of the crew members. But it would be difficult for these prudent ship owners to compete with the more irresponsible ship owners. If other ship owners were overloading, they would be making higher profits and/or would be

able to charge lower prices. This would force the more prudent ship owners out of the market – or else force them to change their strategy to match their competitors: they would be forced to increase loads, and/or cut costs on construction and maintenance.

As a result, the average level of risk in the market would gradually increase.

An increase in risk-taking would eventually lead to an increase in claims costs. But it might take some time for the insurers to react to this change in the systemic risk. There might be a few years of good weather, so that claim rates might not increase immediately. Even when bad weather led to an increase in claims, it might be difficult to distinguish between the “weather effect” and the “load effect” – insurers might assume that the increase in claims was simply a random fluctuation, rather than a result of increased risk-taking. Hence, they might delay premium rate increases. Those insurers who were slow to adapt to the new situation would keep premium rates low, hence winning more new business. These insurers would ultimately suffer larger losses, possibly leading to insolvency (ie. a phenomenon known as “the winner’s curse”).

Eventually, of course, insurers would be forced to increase premium rates simply to remain in business. These rate rises would affect all ship owners, both those who overloaded and those who did not. The increase in premium rates would probably push any prudent ship owners out of business.

In other words, an adverse selection cycle would develop, ultimately leading to increasing risk, increasing claims costs, and increasing premiums.

This pattern of escalating premium rates was indeed observed in the marine insurance market in the mid 1800s. Over a period of about twenty years, there was an enormous increase in the number of claims; and premium rates doubled.¹⁶ This created problems for the

¹⁶ Masters, page 87

ship owners. Indeed, some of the more responsible merchants and ship owners petitioned the government to impose more regulation (including restrictions on overloading), because they could no longer afford the ever-increasing premium rates.¹⁷

Furthermore, as claim costs escalated, insurers faced declining profitability. During the 1860s, several marine insurers became insolvent – especially new insurers which were not as competent in assessing risks, and also keen to cut premiums to gain market share (explained in more detail below).

Analogy to credit markets

The adverse selection cycle is also apparent in the credit risk markets over the last decade. Initially, there were only a few lenders who specialised in subprime lending and low doc loans. But the subprime lenders reported handsome profits (apparently with low risk); and they gradually increased their share of the market.

Eventually, in order to defend their market share, the more established lenders also relaxed their underwriting standards and moved into the subprime market, making low-doc loans with high loan to value ratios. The subprime loan market grew, increasing the overall credit risk in the economy.

Gerardi, Lehnert, Sherlund and Willen (2008) describe the rapid deterioration in underwriting standards from 1999 to 2007.¹⁸ The adverse selection cycle was driving a decline in underwriting standards.

4.7 The Growth of Speculation

Normally, insurers will only sell insurance to those who have an insurable interest. This reduces moral hazard.

However, when marine insurance was mispriced, speculators became keen to invest in insurance policies. During the 1860s and

¹⁷ Plimsoll page 97

¹⁸ Gerardi, Lehnert, Sherlund and Willen,(2008)

1870s, there was an upsurge in “wager policies”, ie. purchase of insurance by those who had no insurable interest.

If the speculators had no special knowledge - no inside information - then this was merely gambling.

However, if the speculators did have inside information about the seaworthiness of any vessel, then wager policies could be a highly profitable form of investment for the speculators – but of course ruinous for the insurers who sold such policies.

If the speculators were able to influence the likelihood of a shipwreck, then over-insurance was a much more serious risk. A ship owner could load up a ship with worthless cargo and insure it for far more than it was worth. The ship owner and his friends might also buy additional “wager policies”. Then it would become extremely profitable to scuttle the ship (many poorly-paid sailors were susceptible to bribery; they could scuttle the ship while it was conveniently close to shore, ensuring no loss of life).¹⁹

Analogy to credit markets

From 2000 to 2008, there was an explosion in speculative trading of credit risks. If an investment banker had inside information about the probability of default – information which was not available to the rest of the market – then this could be very profitable indeed.

As an example, consider Goldman Sachs’ ABACUS deal. Goldman created synthetic CDOs where the return was linked to the performance of a particular set of mortgage backed securities. Goldman had good reason to believe that these particular mortgages had a high probability of default (because these mortgages were selected by a hedge fund which was shorting the securities). However Goldman neglected to share this information with the investors who were buying the ABACUS securities. The investors suffered losses of more than \$1 billion. Goldman was later fined \$550 million for providing incomplete information.²⁰

¹⁹ See Jones (2006), page 314 for Plimsoll’s allegations and public opinion on this issue

²⁰ *Goldman Sachs to pay Record \$550 million to settle SEC Charges related to subprime Mortgage CDO*; US Securities and Exchange Commission Litigation Release No 21592/ July 15,2010

At the Financial Crisis Inquiry, Lloyd Blankfein admitted that Goldman Sachs had “*sold bundles of troubled mortgages at the same time it placed bets with Goldman’s own money that their value would fall.*” The Chairman of the Commission suggested that this was “*a little bit like selling a car with faulty brakes and then buying an insurance policy on the buyer of those cars.*”²¹

Speculation can be highly profitable when markets are inefficient.

In the next section, we examine market inefficiencies in pricing risk.

5 Inefficiencies in Pricing Risk

Under-pricing of insurance will naturally tend to cause an increase in risk-taking. So why was underpricing so common in the marine insurance markets in the 1860s?

Based on Plimsoll’s observations of the marine insurance market, the obstacles to accurate underwriting were:

- Asymmetric information;
- Cost-benefit considerations in underwriting (the impact of diversification);
- Unreliable ratings from collective risk assessment agencies;
- An influx of naive investors into the insurance market; and
- Difficulties in enforcing utmost good faith requirements.

5.1 Asymmetric Information

Asymmetric information is always an issue for insurers – an applicant for insurance is very likely to know more about the risk than the insurer knows. However, this problem is naturally exacerbated

²¹ Sorkin (2010)

when the applicant for insurance has a great deal of control over the risk factors.

For example, in the 1860s, the wealthy ship owners generally had control over the construction of their ships. At the time, there were no regulations to impose any minimum standards on the ship-builder, so the quality of construction was quite erratic.²²

The underwriter would naturally wish to be sure that the ship's construction was sound. But this was difficult, unless the underwriter actually inspected the ship regularly while it was under construction. Once the construction was complete, it would be difficult for any external inspection to find any hidden flaws.

For example, Plimsoll reported that some ship-builders used "devils". These were bolts made of cheap iron instead of expensive copper. The visible part of the bolt (the head and an inch of the shaft) was made of copper, but the rest was iron. These bolts slowly corroded, with disastrous results. As soon as the ship ran into heavy weather, it would come apart at the seams.²³

If the underwriter sent a surveyor to inspect the ship after construction was complete, these faulty bolts would be hidden from view – making it impossible to assess the risk accurately.

Of course the ship owner would be well aware of the heightened risk arising from the use of such poor quality materials; and he would be sure to fully insure the risk (and perhaps even over-insure).

The greater the ship owner's control over the risk factors, the greater the potential for information asymmetry.

²² Plimsoll page 51

²³ Plimsoll page 64

Analogy to credit markets

In credit markets, the quality of loans is controlled by those who originate the loan. Each mortgage broker would, of course, have “official” underwriting standards – but would the broker’s employees actually adhere to these standards, when they were under considerable pressure to make more and more loans? Evidence from a number of lawsuits suggests that many of the largest mortgage originators gradually relaxed their underwriting standards – but failed to reveal this to those who were buying the mortgages. (As an example, see the SEC’s complaint against Countrywide’s senior executives).²⁴

Theoretically, independent agencies were appointed to check a sample of the loans included in a CDO package, to ensure that they met the stated underwriting standards. But would these inspectors be willing and able to detect poor underwriting? Muolo and Padilla (2008) have described a number of shortcomings in the review process – including pressure to work quickly, and to approve a high proportion of loans.²⁵

5.2 The Impact of Diversification on Underwriting Standards

As Plimsoll noted, in the Lloyd’s market, each insurer took only a small share of any risk. Typically, each individual insurer would take no more than 5% of the risk on any ship, so that 20 different insurers shared the risk. An individual insurer would only suffer a maximum loss of say £100 or £150 on any risk.

Since the amount at risk was so low, it was not worthwhile for any individual insurer to assess any individual risk. The underwriter could check Lloyd’s Register for basic information about the construction and maintenance of the ship. But any more detailed assessment would have caused inordinate delays in underwriting. Huebner (1920) described the situation:

“... the underwriter at Lloyd’s has comparatively little opportunity to examine the risk as he would do in most other

²⁴ SEC (2009)

²⁵ Muolo and Padilla (2008)

branches of insurance... Concerning many factors relating to stowage, the amount of load, the size and efficiency of the crew, and numerous other facts equally vital to the safety of the vessel and cargo at sea, these publications [Lloyd's Register] can offer little assistance.

Nor would it be to the interest of the insurer at Lloyd's to make such an examination, assuming that he could do so. Not only would his limited time and the large number of proposals made to him daily render this impossible, but the mere fact that probably half a hundred other persons have underwritten the same policy will make it seem foolish that he alone should undertake the examination. To retain his business he must be quick in accepting or rejecting proposals on the spot, and cannot afford to tarry, since it is the broker's business to secure insurance for his patrons as quickly as possible.

Moreover, the amount of the total risk to which he subscribes is comparatively small and usually limited to an amount which will not make it worth his while to pursue a detailed examination.²⁶ *(emphasis added)*

Each insurer would underwrite hundreds of risks each year, but would be responsible for only a small amount of each risk. The insurers realised that some of these risks would be poor risks – but overall, the insurer assumed that the total claims cost would be acceptable, on average. This works well as long as the insurer's “mix” of risks is reasonably stable – but the insurers would suffer a loss if the mix of risks in the market was shifting over time, eg. if ship owners were gradually taking on more and more risk.

Analogy to credit markets

John Talbot (2009) has argued that diversification has been a major factor in causing the global financial crisis.

“I think that this is the mistake that these very large institutional investors made with regard to mortgages and other

²⁶ Huebner (1920), page 25

*assets and with regard to the pricing of risk. They assumed that by being properly diversified they would minimise their risk, but **their diversification strategy itself required that they hold so many assets that they did not have time to evaluate risk and return for each.** Rather, [assuming that the market was efficient] they allowed the market to properly price the assets for them. In such a passive, highly diversified world, in which few are doing fundamental analysis, it almost ensures that the market itself will become corrupted.”²⁷ (emphasis added)*

In an analysis of the causes of the GFC, the Financial Stability Forum found out that many investors simply did not take the trouble to assess risk adequately – especially when the risks were complex and it would have been rather time consuming and expensive to conduct the analysis. When the issuers of securities became aware that the investors were not making detailed risk assessments, they took advantage of the situation to sell riskier securities.²⁸

5.3 Collective Risk Assessment (Lloyd’s Register)

Clearly, it was impractical for each underwriter to inspect each ship. Lloyd’s underwriters found a solution: collective risk assessment. Effectively, this was a “rating agency” for ships.

Some time prior to 1760, the leading Lloyd’s underwriters formed the *Society for the Register of Shipping*. The Society appointed inspectors (“surveyors”) to evaluate the seaworthiness of ships. The surveyors reported their result to the Society, which then published this information in Lloyd’s Register.²⁹

However, it required constant vigilance to protect the integrity of the Register.

The Lloyd’s underwriters realized that it was important to ensure that the risk classification of any ship should be determined

²⁷ Talbot (2009) Page 41

²⁸ Financial Stability Forum 2008, Page 8

²⁹ See Martin (1876) Chapter XVIII for a history of Lloyd’s Registry of Shipping.

objectively and impartially. But this turned out to be surprisingly difficult to achieve in practice.

It appears that initially, the Society relied on subscriptions from underwriters as the sole source of funding for the inspections.³⁰ This did not work very well. Underwriters were reluctant to pay subscriptions if they could obtain the information for free, eg. by borrowing a subscriber's book. The publishers attempted to protect their revenue stream by imposing punitive fines on any subscribers who shared information with non-subscribers. However this was difficult to enforce.³¹ (Obviously, copyright problems have been around for quite some time, long before the Internet was invented.)

The free exchange of information undermined the Society's business model, so that revenues fell. The low level of revenues made it difficult for the Society to fund the very thorough inspections which were desirable to ensure accurate underwriting.³²

Analogy to credit markets

Credit rating agencies provide a cost-effective collective method for assessing credit risks. But the issuer-pays business model introduces some conflicts of interest which have undermined the reliability of ratings.

There have been some proposals for a change to a subscriber-based system. This approach would probably face the same problems as the Underwriters Register, ie. the problem of "free riders". Given the free flow of information via the Internet, it would be very difficult to maintain the confidentiality of any ratings – thereby undermining the subscription-based business model.

Since the Society was underfunded, the Register's surveyors were underpaid – and this created problems, as the surveyors sometimes supplemented their income with additional payments from the owners of the ships they surveyed (creating a conflict of interest). As Martin (1876) commented:

³⁰ Golding and King-Page (1952) page 179

³¹ Lay page 163

³² Martin (1976) page 343

*“The system, it is needless to say, was bad in principle, and its defects were fully acknowledged by the leading members of Lloyd’s.”*³³

The ship owners also undermined the rating process. The ship owners felt that the Lloyd’s surveyors were much too strict. In 1798 they developed their own register, called “The New Register Book of Shipping” (commonly known as the Ship owners Book). This publication was controlled by the ship owners and merchants; no underwriters were on the committee.³⁴ The results are not at all surprising.

*“To curry favour with the great body of ship owners, the conductors of the New Register book during the first few years of its publication, placed nearly the whole of the shipping in the lists under the category A1....Such classification was an absurdity, and of no value whatever.”*³⁵

So from 1798 to 1833 there were two competing Registers. Neither organisation was financially viable, and under the circumstances it is not surprising that the quality of surveys deteriorated.³⁶

Clearly, this situation was unsatisfactory. In 1824 a joint committee of underwriters, merchants and ship owners was formed, with the aim of devising a more objective and reliable method of classifying ships. It took ten years to reach agreement, but this eventually resulted in the Lloyd’s Register of British and Foreign Shipping (established in 1834).

The business model of the new Register was different: the ship owners would be charged a standard fee for each inspection. These fees accounted for about two-thirds of the Register’s income. Subscriptions covered the remainder of the costs. Standards of risk

³³ Martin page 357

³⁴ Martin page 333

³⁵ Martin page 335

³⁶ Lay page 181. In 1836 a Select Committee of the House of Commons reported that “the systems adopted by the old registers were not sound and did not promote a high standard in ship construction”. Golding et al page 182

classification improved substantially under the new regime – with flow-on effects to ship-building standards.

However, since the Society was now financially dependent on the fees paid by the ship owners, it became vulnerable to competition from other risk classification agencies. If competitors charged lower fees and/or provided more lenient classifications, the Society might lose revenue. And this is exactly what happened.

For example, in 1862, Liverpool ship owners established their own Underwriters Register for Iron Vessels (Liverpool was the centre of steam shipping). According to one American observer, the competition for business led to a sharp decline in classification standards at both Registers, inevitably followed by a decline in ship-building standards.

“... the stronger the competition to get business from Lloyd’s, the farther into the ground the trade of iron shipbuilding was run. By the time this book [ie. the Liverpool Register] had absorbed the greater part of the iron tonnage of the kingdom, its character had fallen to a low state. Twenty-year ships in numbers were sent to sea, and never heard from afterward.

*Lloyd’s attenuated to compete with the Red Book, reduced their requirements for strength, and the opposition followed suit, until the consequences of this deteriorating rivalry attracted the attention of the world...”*³⁷

Lloyd’s eventually managed to eliminate the local competition: in 1885 the Liverpool Register was amalgamated with Lloyd’s Register. But the Society was still vulnerable to competition from ratings agencies based overseas. At the 1873 Royal Commission into Unseaworthy Ships, the secretary of the Lloyd’s Registry complained about the French-based Bureau Veritas.

“...the Bureau Veritas stepped in, and when they found that we made concessions, they gave further concessions; for instance if we gave a vessel an eight years class, they would give nine; and

if we gave it ten, they would give eleven. That goes on until it brings into existence an inferior class of ships to what would otherwise be produced.”³⁸

In summary: when the rating agencies relied on fees from the ship owners, then they were vulnerable to competitive forces. Competition between the risk classification agencies led to a deterioration in the classification standards.

Analogy to credit markets

Credit Rating Agencies (CRAs) have been widely criticised for their role in causing the subprime debt crisis. They did a poor job in evaluating the risks underlying Asset Backed Securities and Collateralised Debt Obligations.

The CRAs have a issuer-pays business model. The CRA's compete with each other to attract clients. Since the CRAs depends on revenue from the issuer, there is an inherent conflict of interest which can (and did) undermine the integrity of risk classification.

Theoretically, US legislation requires each agency to have policies and procedures which will “*prevent the credit rating agency from being influenced to issue or maintain a more favorable credit rating in order to obtain or retain business of the issuer or underwriter.*” However, when the Securities and Exchange Commission investigated the performance of the CRAs, they found evidence that these policies and procedures were not effective. Competitive factors apparently did indeed influence rating decisions.³⁹

5.4 An Influx of Inexperienced Risk Takers

The Lloyd's market was quite competitive. Insurance brokers would enter the Underwriting Room, determined to place a risk. If one underwriter demurred, then the broker would simply move on to

³⁸ Martin (1876) page 354

³⁹ SEC (2008)

the next underwriter.⁴⁰ The forces of competition led to some underpricing of risks.⁴¹

Throughout the first part of the nineteenth century, Lloyd's had a virtual monopoly in the British marine insurance market. There were legal obstacles which made it difficult for companies to provide marine insurance. As a result, the Lloyd's brokers and underwriters had developed a considerable amount of expertise in underwriting marine risks – and the expertise was all concentrated in one Underwriting Room.⁴²

However, the situation changed after the Joint Stock Companies Registration Act was passed in 1844. The immediate effect result was “*the springing up of a whole host of marine insurance companies*”. Dozens of new companies were formed in the 1840s – and most of them failed within a very short time. Many were blatant swindles: the directors simply collected premiums for a short while and then quietly disappeared (insurance company prudential legislation was not particularly stringent at that time). For a while, it seems, this dampened the ship owners enthusiasm for non-Lloyd's insurance policies.

However, during the 1860s there was a resurgence of interest in marine underwriting. At the time, the British insurance market was growing rapidly. This growth was a result of the American Civil War (1860-1864). Ship owners switched their insurance from American to British insurers. For a while, insurance was a booming market; and higher profitability attracted more insurers into the market.

There was an influx of inexperienced underwriters, who drove down premiums in an attempt to buy market share. Since they lacked the skills necessary to make an accurate assessment of risks, they often offered cheap insurance to poor risks.

⁴⁰ Plimsoll page 17

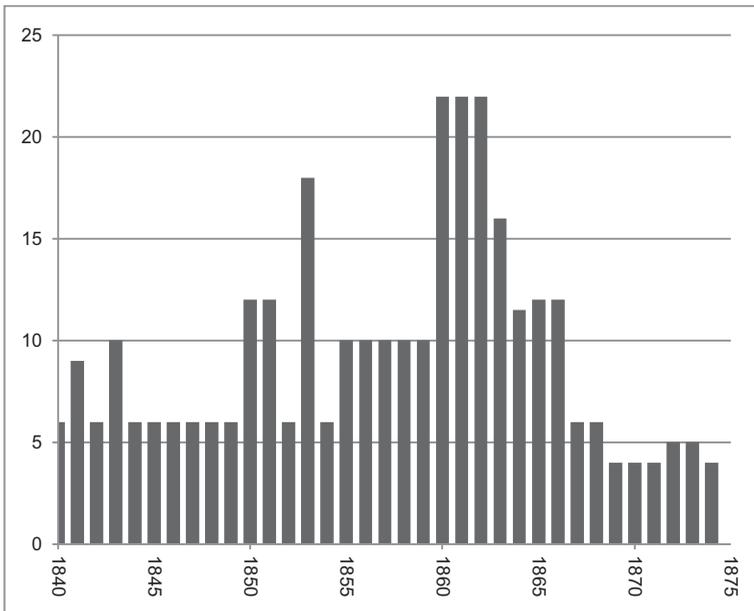
⁴¹ Martin (1876) page 235

⁴² Martin (1876) page 240

After the end of the Civil War, there was excess capacity, which led to cut-throat competition in pricing. The new companies undercut the established Lloyd's underwriters, who were forced to cut their own rates in order to survive.⁴³ As you would expect, the price cutting war eventually led to disaster. Martin (1876) noted that all the new companies which were established in the five years from 1865 to 1869 subsequently went broke.⁴⁴

The following graph shows the dividend and bonus rates for one insurer during this period. This illustrates the boom-bust cycle: very high profitability in the early 1860s, followed by a sharp decline in profits after several new insurance companies entered the market.⁴⁵

Figure 9: Total Dividends + Bonus Rate 1840-1875 of the Alliance Marine Insurance Company



⁴³ Martin (1876) page 322

⁴⁴ Martin (1876) page 320

⁴⁵ Martin (1876) page 305-306

Ship owners were able to buy marine insurance at very cheap rates during the late 1860s and early 1870s.

This is an early example of an underwriting cycle. A typical underwriting cycle starts in times of high profitability. High profits attract new entrants into the market, ie. an increase in capacity. Competition pushes premiums down; new entrants to the market lack expertise in underwriting, and/or may be willing to slash premiums in order to grab market share. Our model suggests that this underpricing of risk will provide an incentives for more risk-taking.

Analogy to credit markets

Economists have attempted to analyse the causes of the GFC. Goodhart (2008) points out that there was a long period of relative economic stability prior to 2007. Perhaps this created some complacency among investors. Central bankers became concerned about under-pricing of risk during this period. Risk spreads – the differential between yields on risky and safe securities – declined to historically low levels.⁴⁶

There was strong demand for mortgage backed securities, which seemed to offer high returns (relative to government bonds), with little risk. This demand fuelled the expansion in subprime lending. Much of the demand for risky structured finance products came from rather unsophisticated investors, who had little understanding of the risks they were taking.

5.5 Unenforceable Disclosure Requirements

Asymmetric information is a problem which has plagued many insurance markets. Historically, this problem can be addressed by “utmost good faith” requirements – ie. anyone purchasing insurance must disclose all relevant information about the risk to the insurer. Theoretically, if the insured fails to reveal the any pertinent information, then the insurer can take legal action to avoid the claim.

⁴⁶ Goodhart (2008)

Theoretically, the threat of such legal action should be enough to prevent the policyholders from misrepresenting the facts about load levels and crew numbers.

Plimsoll pointed out the enormous chasm between theory and practice. He claimed that in practice, the “utmost good faith” rules did little to prevent overloading. Underwriters seldom took legal action against their policyholders, and seldom won their cases in court.⁴⁷

Firstly, the costs of taking legal action often exceeded the benefits to the insurers. At Lloyd’s, the risks were diversified – each ship owner only took a small proportion of each risk. As Plimsoll said, “*The loss to each individual underwriter is too small to make it worth his time and trouble[to take legal action]*”.⁴⁸

Secondly, there was an imbalance of power in the legal system. The most over-loaded ships often belonged to wealthy individuals who owned fleets of ships. It was difficult for an underwriter to succeed against wealthy, influential, and very determined adversaries.⁴⁹

Thirdly, even if the underwriters decided to fight the claim, their chances of success were small. In most cases, all evidence of wrongdoing was at the bottom of the sea.⁵⁰ Sometimes the crew members would survive, and theoretically they could be called upon to give evidence. But in practice, they might be very reluctant to co-operate.

*“They may be unwilling witnesses: the captains, the officers, the men all know that if they give evidence they destroy their chances of future employment even by owners who disapprove the practices of the owners of the lost ship.”*⁵¹

⁴⁷ Plimsoll Page 12

⁴⁸ Plimsoll page 12-23

⁴⁹ Plimsoll page 25

⁵⁰ Plimsoll page 27

⁵¹ Plimsoll page 31

Even if the underwriters could find witnesses who were willing to state that (in their opinion) the ship seemed to be overloaded, they might not win the case. The ship owners might offer their own (well-paid) expert witnesses to refute any such accusations. At that time there was no objective standard for safe load levels – in 1853 the Board of Trade had attempted to develop such standards, but without success. “*They could not get any two persons to agree as to the method of developing the load-line.*” In the absence of any regulatory minimum safety standards, it was difficult to prove that the ship was unsafe.⁵²

Martin (1876) agrees that legal remedies were quite ineffective in dealing with insurance fraud and non-disclosure.⁵³

Analogy to credit markets

After the subprime debt market collapsed, it became clear that many investors were given a great deal of mis-information about the riskiness of the products they invested in. At present there are dozens of cases proceeding through the courts, alleging misrepresentation and/or lack of disclosure; and of course the Securities and Exchange Commission is also pursuing investigations.

However, it may be difficult for the investors to prove wrong-doing to the standards required in a court of law; and even if they do so, it may not be possible for the investors to obtain compensation (eg. if the assets have been moved to another company or another country).

It remains to be seen whether the results of these lawsuits will be sufficient to deter future mis-representation, deception, non-disclosure and fraud in securities markets.

Efficient insurance markets rely on full disclosure. Efficient credit risk markets also rely on full disclosure. If these disclosure standards are unenforceable in practice, then obviously these markets will not function efficiently.

⁵² Plimsoll page 35

⁵³ Martin p 270

6 Proposals for Reform

6.1 The Need for Regulation

The marine insurance market was dysfunctional, because it had ill-effects for many members of the community: for the sailors, the insurers, the ship owners, and the general public.

The Sailors

Plimsoll pointed out that the sailors and their families were the main victims of the dysfunctional marine insurance system. Over the decade from 1862 to 1871, about 8000 lives were lost at sea within ten miles of the English coast. The ship owners did not suffer any financial penalties as a result of the deaths of the sailors they employed. These days, an economist might describe these as “negative externalities”.

The Insurers

As noted above, many insurers (especially those who were less experienced in underwriting) became insolvent during the 1860s and 1870s. One could argue that the insurers deserved little sympathy – they were in the business of taking risks, and if they misjudged those risks they should bear the consequences. However, the insurers were often the victims of fraud, misrepresentation, and non-disclosure.

The ship owners

Some wealthy ship owners profited from the availability of cheap insurance – especially if they were guilty of overloading. But the more responsible ship owners were suffered. As noted previously, in the 1860s, some of the merchants and ship owners petitioned the government to impose more regulation (including restrictions on overloading), because they could no longer afford the ever-increasing premium rates which arose from excessive risk-taking.⁵⁴

⁵⁴ Plimsoll page 97

Ultimately, the ship owners could only stay in business by passing on these insurance costs to the general public, ie. by charging higher prices on any goods transported by sea.

Analogy to credit markets

As a result of the sub-prime debt crisis, three groups have suffered financially:

- Subprime borrowers lost their homes when they could not afford to make their mortgage repayments. The number of foreclosures increased astronomically.
- Investors (including many banks and pension funds) suffered large losses on their mortgage backed securities.
- Risk margins increased sharply, so that many businesses were compelled to pay higher interest rates when borrowing money (and in some cases, finance was simply no longer available).

Restrictions on lending led to a slowdown in economic activity, with devastating effects on the whole community.

What sort of legislation would be effective in preventing these problems? In the 1860s, proposals included:

1. Restrictions on excessive risk-taking.
2. Transparent and reliable standards for the measurement of risk.
3. Changes to the business model for the ratings agencies.
4. Compulsory risk retention.

6.2 Restrictions on excessive risk-taking

Plimsoll's primary concern was the safety of sailors. Hundreds of lives were lost at sea every year. *Our Seamen* is full of heart-rending accounts of grieving mothers, heart-broken widows, and destitute orphans. Public outrage was stirred by his distressing accounts of

human suffering. The regulation of shipping soon became a political issue.

Plimsoll, who was a Member of Parliament, suggested that the simplest way to protect the sailors and passengers was the imposition of compulsory safety standards for all ships. He proposed:

- Loadlines (commonly known as Plimsoll lines). He suggested that a loadline should be painted on the side of each ship. This loadline would indicate the safe level of freeboard. No ship would be permitted to leave port unless the loadline was clearly visible.⁵⁵
- Surveys. Each ship should be surveyed by the Board of Trade, to ensure that it was seaworthy and well-maintained.

However, Plimsoll found it very difficult to persuade Parliament to pass the proposed legislation. Although many people suffered from the defects of the marine insurance system, a minority (ie. some wealthy and influential ship owners) were making quite a lot of money. Plimsoll suggested that this small minority had a disproportionate influence on the legislation.

“Although they (those who recognise the necessity of the proposed legislation) are many, and those who profit by these practices are few, there is this difference on the part of the latter – it is their business to resist change; they profit by things as they are; they are determined, energetic, and sleeplessly vigilant.

*You must remember that large fortunes are being made by them; they are the most energetic and pushing men in the trade; and it would not be a matter of surprise if three of them had even got into Parliament.”*⁵⁶

There were heated arguments in Parliament.

⁵⁵ Plimsoll was by no means the first one to think of this solution. Loadlines were in common use in other countries, and indeed had been used for centuries. Plimsoll simply adopted proposals which had been suggested to him by some of the more responsible ship owners.

⁵⁶ Plimsoll page 105

The ship owners complained about excessively burdensome legislation.⁵⁷

The government and the Board of Trade stated that it was inappropriate for the government to interfere with free enterprise. The government had no role in assessing the seaworthiness of ships: that should be left to the judgment of the merchants and ship owners.⁵⁸

The government also argued that any regulation to introduce compulsory safety standards for ships would be counterproductive. It would be expensive and hence would be detrimental to the shipping industry. British ships would not be able to compete with foreign ships, which were unfettered by any such regulations.⁵⁹

The government set up a *Royal Commission in Unseaworthy Ships* in 1873. After considering submissions from all concerned, the Commission's report firmly rejected most of Plimsoll's proposals.

However, Plimsoll was a very determined man with a flair for rhetoric and a gift for politics. In 1873, he published *Our Seamen*, which won a great deal of public support for his cause. He campaigned in marginal seats, so that Members of Parliament who opposed him became worried about losing their seats. Eventually Plimsoll and his allies pushed through some reforms – but it was a very long slow process.

In 1873 the government reluctantly agreed to pass the Merchant Shipping Act, which gave the Board of Trade the power to inspect ships (but only in response to a complaint): if the ship was unseaworthy, then it would not be allowed to leave port. Within the next two years, the Board of Trade stopped 558 ships from going to sea.⁶⁰

⁵⁷ Jones (1922) page 92

⁵⁸ Peters (1975) page 97

⁵⁹ Peters (1973) page 123

⁶⁰ Peters page 80

Legislation imposing compulsory loadlines (set by government authority) was not passed until 1890, after almost 20 years of agitation by Plimsoll and his allies.

Analogy to credit markets

In recent years, there have been calls for stricter regulation of “safety standards” in home lending. For example, John C Dugan, the Comptroller of the Currency has suggested that certain high-risk lending practices should simply not be allowed. He suggests the imposition of compulsory minimum underwriting standards on all mortgage loans: no more low-doc loans; no more negative amortisations loans; no more inflated property values; no more excessive LTVs. All mortgage originators would be required to meet underwriting standards set by the federal regulator. And no one would be allowed to sell or transfer a mortgage without providing a warranty that such standards had been satisfied.⁶¹

The UK’s Financial Services Authority has also set out new standards for responsible lending – including compulsory affordability tests and the elimination of stated-income loans.⁶²

Of course, if the rules are too strict, this might limit the ability of low-income people to buy a home.

Negative externalities can also be remedied (to some extent) by changing the cost-benefit equation, so that the decision-makers incur higher costs when they make decisions which are detrimental to society. For example, financial penalties may be imposed. In the marine insurance industry, this could be achieved by requiring the ship owners to pay a specified amount in compensation for the death of any crew member or passenger.⁶³ In nineteenth century England, this was not politically feasible.

⁶¹ Dugan (2010) and Dugan (2009)

⁶² FSA (2010a)

⁶³ It would be interesting to consider the effect on aviation safety standards, if airlines were required to pay much higher amounts in compensation for any deaths.

6.3 Transparent and Reliable Standards for the Measurement of Risk

When insurers provided cheap insurance, then ship owners had an incentive to take on too much risk. So the problem would be alleviated if insurers could assess risk more accurately. The development of clearly defined, objective safety standards was a key part of this process.

As noted previously, during the 1860s there were few standards for construction, maintenance, or loading of ships. Those who opposed legislative reform argued that it was simply too difficult to develop uniform standards for load levels and construction.

*“There were so many types of ships built in so many different ways that it was impossible to frame a law covering them all without doing an injustice to some and probably ruining a number of owners”.*⁶⁴

In this respect, Lloyd’s Register was an important ally in the battle for reform.

While the politicians were debating these issues in Parliament, the Lloyd’s Register was quietly and efficiently developing a comprehensive set of safety standards, in everything from the construction of ships, to the strength of cables, to the frequency of inspection of ships, and so on – including loadline standards. The Register also set up a Technical Committee which revised the rules periodically and kept up to date with new developments in marine engineering.

The standards were transparent – handbooks were published which set out, in great detail, the requirements for an A1 classification. Often, the standards developed by Lloyd’s were subsequently adopted by the regulatory authorities.

⁶⁴ Masters (1955) page 201

These days, there is an International Association of Classification Societies (which include Lloyd's Register). The IACS promotes the development of consistent technical standards for the classification of ships.

Analogy to credit markets

In 2008 the Securities and Exchange Commission examined the processes used by the credit rating agencies to determine ratings for structured finance products. Many of the products were quite complex, so that it was difficult to develop standard models for risk assessment. The CRAs did attempt to develop some expertise in modelling these risks, but the SEC later identified numerous deficiencies in the models used, eg. underestimation of correlations, infrequent updating of model parameters, etc. The CRAs simply did not have enough historical data, and enough experienced staff, to deal with the new types of complex products.

The SEC also found that these rating process was not transparent – “significant aspects of the rating process were not always disclosed”.⁶⁵

Better models, with better transparency, would improve the accuracy of the credit ratings process.

6.4 Changes to the business model for the ratings agencies

When risks are highly diversified and information costs are high, it makes sense to use a collective system to assess risks, ie. a ratings agency. But this system will not work unless the standard-setting organisations can be trusted to perform their duties with integrity. If an organisation is financially dependent on the industry it monitors, and there is competition between organisations to attract business, then there is obviously a temptation to compromise integrity – ie. to have a “race to the bottom” in safety standards.

In 1824 a Committee had considered the problem of funding surveys. They realised that a subscription-based system could never

⁶⁵ SEC (2008)

cover the cost of a really excellent risk classification system; and a fee-based system would compromise independence. Hence, the Committee decided that government support was a necessity (perhaps funded by a small increase in the duty on marine insurance).⁶⁶ However, this solution was not feasible - the government politely declined to offer any financial assistance to the Register.

If competition between agencies was the problem, then perhaps the creation of a monopoly might be a solution? During the nineteenth century, other British-based Registers were set up from time to time. Lloyd's responded by taking over or merging with competitors. During the nineteenth century, Lloyd's Register gradually consolidated its position as the leading classification agency in the UK, dominating all its competitors. By the 1880s, about 90% of merchant ships built in the UK were surveyed and classified by Lloyd's.⁶⁷ According to one historian, as Lloyd's competitive position grew stronger, it was better able to enforce its own standards – often over the vehement protests of some ship owners.⁶⁸

However, over the next century, Lloyd's Register continued to face competition from international classification agencies; and competition led to a deterioration in underwriting standards. Unfortunately, this was a problem which was not easy to solve. A historical study of classification societies reveals that they have always struggled to maintain the integrity of their ratings.⁶⁹

The problems which beset the industry in 1860 were still causing trouble in 1980: ship owners were still shopping around for the best rating; classification societies were still relaxing their standards in order to keep customers; insurers were still complaining angrily about the poor reliability of the information provided.

⁶⁶ Martin (1876) page 343

⁶⁷ *Annals of Lloyd's Register 1834-1884* (author unknown), 1884

⁶⁸ Lay (1925) page 180

⁶⁹ Boisson (1994)

Theoretically, the classification agencies should avoid “racing to the bottom”. The International Association of Classification Societies rules state that :

“Competition between Societies must be on the basis of services (technical and field) rendered to the marine industry - it must not lead to compromises on safety of life and property at sea or to the lowering of technical standards.”⁷⁰

However, experience suggests that this sort of rule is rather difficult to enforce.

Ultimately, the European Union decided that government intervention was the only solution. In 2009, the European Union passed new regulations for the supervision of classification agencies. The European Maritime Safety Agency is now responsible for inspecting each recognised classification agency (including Lloyd’s). Each agency must comply with a set of standards, designed to maintain the integrity of the classification systems.⁷¹

Analogy to credit markets

In 2004, the International Organisation of Securities Commissions issued the Code of Conduct Fundamentals for Credit Ratings Agencies.⁷² The Code of Conduct suggests that the CRAs should avoid conflicts of interest, maintain independence, and ensure the integrity of their ratings.

This raises some questions:

- Is a Code of Conduct likely to be successful in maintaining standards of integrity, with a issuer-pays business model, in a competitive environment?
- Have professional standards been successful in maintaining standards in other areas, such as auditing and actuarial valuations?
- Should regulators supervise ratings agencies more closely? For example, banking and insurance regulators will not allow the use of

⁷⁰ IACS(2006)

⁷¹ EMSA (2012)

⁷² IOSC (2004)

internal risk models for capital requirements, unless the financial institution can produce convincing evidence that the model is likely to be reliable. This would normally include checks on corporate governance; an established model based on best-practice; inclusion of all relevant risk factors; a well-defined process for choosing parameters for the model; back-testing and validation; adequate quality and quantity of data; periodic reviews and updating; thorough documentation; adequate resourcing, eg. availability of experienced and qualified staff to run the model; proscription of any remuneration systems that might create a conflict of interest; and limitations on subjective revision of model results (eg. reasons for any such revisions must be documented). Should similar standards be imposed on credit rating agencies?

6.5 Compulsory Risk Retention (“Skin-in-the-Game” Rules)

Moral hazard can be reduced by compulsory risk retention. The insured is required to keep some “skin-in-the-game”.

A study of the history of marine insurance indicates that risk retention rules were common. Many countries passed laws which limited the amount of insurance that ship owners could buy. For example, in French maritime law, dating back to the 16th century, ship owners were required to retain 10% of the risk on any cargo. Spanish law forbade insurance of more than seven-eighths of the value. In Holland, ship owners were not allowed to insure more than two-thirds of the value of the ship (as determined by a public official).

Plimsoll proposed that similar rules should be applied to the British marine insurance industry. Any insurance contracts which breached these limits would be legally void and unenforceable.

However, Plimsoll was never successful in persuading the Parliament to pass such legislation.

Analogy to credit markets

The US House of Representatives has passed the Wall Street Reform and Consumer Protection Act. This Act requires companies that sell mortgage-backed securities to retain at least 5% of the risk, i.e. to “keep some skin-in-the-game”.

Other countries are considering similar requirements. For example the UK’s FSA has a proposal to change capital standards for banks, to prevent a bank from investing in any securitisation unless the originator retains an economic interest.⁷³

There are some doubts as to whether this is feasible. Will issuers find a way around the rules?⁷⁴

And what level of risk retention is desirable? Is 5% enough? George Soros has suggested 10%⁷⁵, and others believe the rate should be even higher for riskier securities, maybe even 20%.⁷⁶

7 “Skin-in-the-Game” Rules

Moral hazard will be reduced when the policyholder has some “skin-in-the-game”. But how much “skin” is enough?

The model we have developed in this paper could be used to assess the impact of compulsory risk retention rules. However, the model relies on accurate information about all of the following relationships:

- The *Risk function*: risk must be expressed as a function of (i) factors which are under the control of the risk originator (“the Load”); and (ii) systemic factors which vary over time and affect all risks (“the Weather”);
- The *Weather distribution*: the probability distribution of the systematic factors which affect risk;

⁷³ FSA 2010b

⁷⁴ Dugan 2010

⁷⁵ George Soros 2009

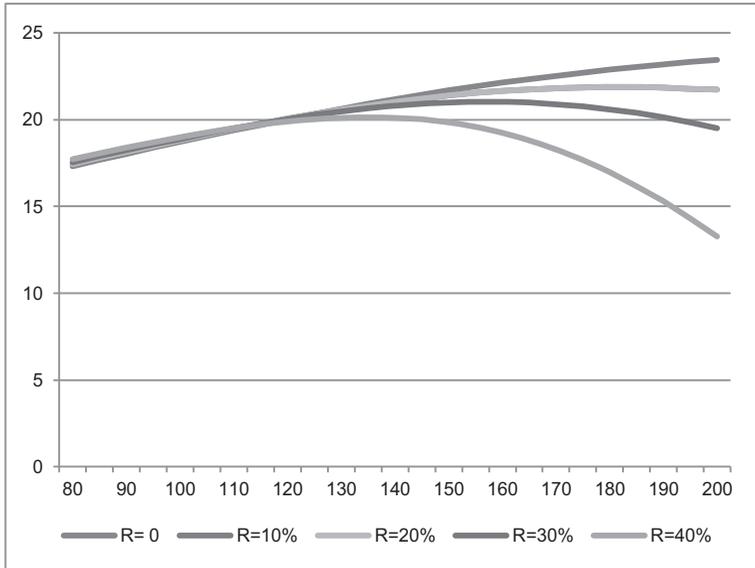
⁷⁶ Saunders 2010

- The *Profit function*: the payoffs for the risk originator, which will vary depending on the profit margins available on risky assets (m) and capital requirements (S);
- The *Utility function*: which measures the level of risk aversion for the risk originators; and
- The *Premium Rate Structure*: the cost of buying insurance for different levels of risk.

Given this set of assumptions, and a specified level of compulsory risk retention, we can use the model to determine the theoretically optimum risk-taking strategy for the person who originates the risk. The regulators would then need to make a policy decision about the level of risk-taking which is desirable (ie. in the public interest); and then set the regulations to achieve the desired effect.

For example, let's apply the model to marine insurance. Suppose that the regulations require the ship owner to retain $R = 10\%$ of the risk (ie. he can't insure more than 90% of the combined value of the ship and cargo). Given all of the above data, we can solve to determine the Load Level L which will maximise the ship owner's expected utility.

The following diagram shows the equilibrium Load levels for different risk retention requirements R , in a market with very poor underwriting (so that the same premium rate is charged regardless of risk). When there is no "skin-in-the-game" ($R = 0$), the optimum Load is "off the scale". If the ship owner is required to retain 10% of the risk ($R = 10\%$), the optimum Load reduces to 180. If $R = 20\%$, the optimum Load reduces to about 155.

Figure 10: Expected Utility as a function of Load and Retention

[In this example we used $S = 10$, $m = 10\%$, Premium Rate = $q(130) \cdot 1.20$ which is the risk premium for a load of 130 with a 20% expense/profit loading for the insurer.]

If the ship owner's optimum Load Level is too high – ie. the government considers that it would involve a socially undesirable level of risk-taking – then the compulsory minimum retention level R could be increased (or vice versa if the proposed retention level is quashing economic development).

The model could also be used dynamically, to solve for an equilibrium state. In practice, the insurer's premium rates will change in response to experience. If the insurer's expected profits are unduly low, then the insurers will raise their premiums: this will flow through to cause a reduction in the ship owner's risk-taking preferences. If the insurer's expected profits are unduly high, then competition will force a reduction in premiums: this will flow through to cause an increase in the ship owner's risk-taking preferences. The dynamic version of the model allows for this feedback effect. Equilibrium will be achieved when the insurance industry is making enough profits to reward the shareholders for the risks which they are taking.

The actuarial profession could make a useful contribution to public policy by developing better models to predict the impact of any proposed “skin-in-the-game” regulations.

Conclusion

A risk transfer market becomes dysfunctional when:

- the insured have control over the level of risk (creating a moral hazard);
- the insured have a strong financial incentive to take risks;
- risk factors interact and risk factors fluctuate over time (which makes it difficult to assess tail risks and correlation);
- there is information asymmetry;
- the legal system fails to enforce “utmost good faith” principles;
- information costs exceed the benefits of accurate underwriting (because risks are thinly spread among many insurers);
- collective risk assessment agencies have conflicts of interest built into their business models;
- the underwriting cycle creates downward pressure on premium rates; or
- (all for the reasons given above) underwriters are unable to classify risks accurately.

These factors combine to create mispriced risks, which ultimately leads to increasing systemic risk.

If market forces creates an incentive for too much risk-taking, then the traditional solutions are (a) impose direct restrictions on risk taking; (b) improve underwriting standards by developing better risk assessment models; (c) development of more reliable collective risk assessment agencies; and (d) minimising moral hazard by imposing “skin-in-the-game” requirements.

If regulators do choose to impose compulsory risk retention rules, then it would be desirable to build some models which will assist in the development of an effective policy.

References

- Baily, Elfendorf, and Litan (2008). *The Great Credit Squeeze: How it Happened, How to Prevent Another*, Economic Studies at Brookings.
- Bates, William W. (1893). *American Marine: The Shipping Question in History and Politics*, Houghton Mifflin and Company, Boston
- Boisson, Phillippe (1994). Classification societies and Safety at Sea, Marine policy, Volume 18, Issue 5, September 1994, pages 363-377
- Boisson, Phillippe (1999). *Safety at Sea : Policies, Regulations and International Law*. Preface by William A. O'Neil. Paris, Edition Bureau Veritas, 1999, cited at the website of the International Marine Organisation,
- <http://www.imo.org/KnowledgeCentre/ReferencesAndArchives/Pages/TheOriginsOfIMO.aspx>
- Dugan, John C (2009). *The Need for Minimum Mortgage Underwriting Standards*, speech to the Special Seminar on International Banking and Finance, November 18, 2009
<http://www.occ.gov/ftp/release/2009-143a.pdf>
- Dugan, John C (2010). *Securitisation, Skin-in-the-game Proposals, and Minimum Mortgage Underwriting Standards*, Speech before the American Securitisation Forum February 2, 2010
- European Maritime Safety Agency (EMSA) (2102). Assessment of Classification Societies, at
<http://www.emsa.europa.eu/implementation-tasks/visits-and-inspections/assessment-of-classification-societies.html>, downloaded February 2012.

Financial Service Authority (2010a). New Standards for Responsible Lending to Protect Mortgage Customers, http://www.fsa.gov.uk/pages/consumerinformation/product_news/mortgages/responsible_lending/index.shtml, downloaded 16/7/2010

Financial Service Authority (2010b). Capital Requirements Directive / Basel 2, <http://www.fsa.gov.uk/pages/About/What/International/Basel/index.shtml>, downloaded 16/7/2010

Financial Stability Forum (2008). Report of the Financial Stability Forum on Enhancing Market and Institutional Resilience

Gerardi, Kristopher S; Lehnert, Andreas; Sherlund, Shane M.; and Willen, Paul S. (2008). *Making Sense of the Subprime Crisis*, Federal Reserve Bank of Boston Public Policy Discussion Paper 09-01. <http://www.bos.frb.org/economic/ppdp/2009/ppdp0901.pdf>, downloaded April 2010

Golding CE and King-Page D (1952). *Lloyd's*, McGraw-Hill Publishing Company, London

C.A.E. Goodhart (2008). *The Background to the 2007 Financial Crisis*, Journal of International Economics and Economic Policy, Volume 4, Number 4, p331- 340

Huebner, Solomon Stephen (1920). *Marine Insurance*, D. Appleton-Century, New York

International Association of Classification Societies (IACS) (2006). *Classification Societies, What, Why and How?*, www.iacs.org.uk

International Organisation of Securities Commissions (2004). *Code of Conduct Fundamentals for Credit Rating Agencies*,

IOSCO, available at

<http://www.iosco.org/library/pubdocs/pdf/IOSCOPD180.pdf>

Jonsson, Asgeir (2009). *Why Iceland? How One of the World's Smallest Countries became the Meltdown's Biggest Casualties*, McGraw Hill

Jones, Clement (1922). *British Merchant Shipping*, Edward Arnold and Company, London.

Jones, Nicolette (2006). *The Plimsoll Sensation: The Great Campaign to Save Lives at Sea*, Little Brown, London.

Lay, H.G. (1925). *A Text Book of the History of Marine Insurance Including the Functions of Lloyd's Register of Shipping*, Post Magazine, London

Lindsay, W.S. (1876). *History of Merchant Shipping and Ancient Commerce*, Sampson Low Marston Low and Searle, London.

Lloyd's Register of Shipping, (1884). *The Annals of Lloyd's Register 1834-1884*, Lloyd's Register, London

Lowndes, Richard (1884). *Insurable Interest and Valuations*, Stevens and Son, 119 Chancery Lane, London.

Martin, Fredrick (1876). *The History of Lloyd's and of Marine Insurance in Great Britain*, Macmillan and Company, London.

Masters, David (1955). *The Plimsoll Mark*, Cassell.

Mortgage Brokers Association(MBA), *Delinquencies Continue to Climb, Foreclosures Flat in Latest MBA National Delinquency Survey*, Press Release 20 August 2009, available at <http://www.mbaa.org/NewsandMedia/PressCenter/70050.htm> downloaded April 2010

- Muolo, Paul and Padilla, Matthew (2008) .Chain of Blame: How Wall Street Caused the Mortgage and Credit Crisis, John Wiley and Sons, New Jersey.
- Peters, George H (1975). *The Plimsoll Line*, Barry Rose (Publishers), London 1975
- Plimsoll, Samuel (1873). *Our Seamen: An Appeal*, Virtue and Co, London. This book is out of copyright and is available via the Internet Archive.
- Sanders, Anthony B, Oral Testimony to the House Financial Services Committee, March 23, 2010
- Sorkin, Andrew Ross (2010). “*Wall St. Ethos under Scrutiny at a Hearing*”, New York Times, 14 January 2010
- Soros, George (2009). *My three steps to financial reform*, Financial Times, 17 June 2009
- Talbot, John R (2009). *The 86 Biggest Lies they Tell on Wall Street*, Scribe, Melbourne 2009
- United States Securities and Exchange Commission (SEC) (2008). *Summary Report of Issues identified in the Commission Staff's Examinations of Select Credit Rating Agencies*, by the Staff of the Office of Compliance, Inspections and Examination and Office of Economic Analysis, July 2008, available at <http://www.sec.gov/news/studies/2008/craexamination070808.pdf>
- United States Securities and Exchange Commission (SEC) (2009). *Complaint for Violations of the Federal Securities Law: Demand for Jury Trial*, United State District Court, Central Court of California, filed June 4 2009. Securities and Exchange Commission, Plaintiffs vs Angelo Mozilo, David Sambol, and Eric Sieracki, Defendants.

United States Securities and Exchange (2010a). *Commission SEC Charges Goldman Sachs With Fraud in Structuring and Marketing of CDO Tied to Subprime Mortgages*, Media Release 16 April 2010.

United States Securities and Exchange Commission (SEC) (2010b). *Goldman Sachs to pay Record \$550 million to settle SEC Charges related to subprime Mortgage CDO*; US Securities and Exchange Commission Litigation Release No 21592/ July 15 2010

On the Use of Limited-Value Averages in Actuarial Modelling

*Uditha Balasooriya, Jackie Li, and Yoong Sin Lee**

Abstract

In actuarial modelling, certain statistical tests, such as Pearson's chi-square test, are commonly used for evaluating the goodness of fit. Besides these standard tests, occasionally the limited-value averages (LVAs) are also examined for such purposes and are compared, in a rather casual manner, with the corresponding limited-value expectations under the specified distribution law. In fact, as there are often coverage limits in practice, the LVAs are probably more relevant to the dollar values involved in the losses. In this article, we explore the application of a formal statistical setting of the LVAs test for goodness-of-fit testing. We apply it to the well-known hurricane loss data and also another set of individual claims data. Two simulation experiments are carried out to study the limiting chi-square property and the power of the test. A formal proof of the limiting property is provided in the appendix. Our results suggest that this LVAs test is potentially suitable for wider use in actuarial practice.

Keywords: modelling loss data, goodness-of-fit tests, limited-value averages, limiting chi-square property

* Contact: Jackie Li, PhD, FIAA. Nanyang Business School (Actuarial Science), Nanyang Technological University, Singapore. Email: jackieli@ntu.edu.sg.

1 Introduction

An important step in general insurance loss modelling is to find appropriate distribution laws to represent the individual losses. To evaluate the goodness of fit of a particular distribution law, it is desirable to perform one or more standard statistical tests. Very often, more than one distribution exhibits good fit to the data. Then further analysis is needed to select the most suitable option. Recently, criteria based on predictive errors such as the global distance measure and L^1 , L^2 norms have also been employed in the literature, eg. Balasooriya and Low (2008). In actuarial practice, however, Pearson's chi-square test is still commonly used in fitting loss distributions, due to its prolonged history and straightforward computation.

In addition to usual goodness-of-fit tests, sometimes the limited-value averages (LVAs) are also used in actuarial studies (eg. Hogg and Klugman, 1984), in which the LVA is the simple average of the observations limited at a given upper value. Perhaps, the LVAs are more relevant to the dollar values involved in the losses, as there are often policy coverage limits and reinsurance arrangements in place. Under this method, the observed LVAs are compared with the corresponding limited-value expectations under the specified distribution law. But in many applications, it only involves an inspection of the two sets of values to evaluate the fit intuitively, without having regard to a formal test of significance. Accordingly, in this article, we discuss a formal statistical setting of the LVAs test and explore its application for goodness-of-fit testing in loss modelling.

Our simulation results show that under usual actuarial settings, the LVAs test generally performs better than Pearson's chi-square test in terms of the power of the test. The LVAs test contains more information over each interval subdivided, which appears to explain its superior performance. Not only is the LVAs test intuitively favored as being more relevant, it is also shown to possess more desirable features, eg. the limiting property. All these findings suggest that this test is potentially a useful addition to the actuarial toolkit.

Section 2 provides the derivation of the LVAs test under a typical actuarial setting. Section 3 sets forth the testing procedure and applies it to two data sets, hurricane losses (Hogg and Klugman, 1984) and individual claims (Currie, 1992). In Section 4, we carry out simulation studies on the limiting chi-square property of the LVAs test. Section 5 presents some simulation results with regard to the power of the test. In Appendix A, a formal proof of the limiting property is given. Appendix B provides the limited-value moments for selected distributions.

2 A Test of Fit Using Limited-Value Averages

Let $X (> 0)$ be a continuous random variable with density function $f(x)$ and distribution function $F(x)$. Consider the random variable $W_a(X)$ with its value limited above by a :

$$W_a(X) = \begin{cases} X & \text{if } X \leq a; \\ a & \text{if } X > a. \end{cases}$$

It can also be represented as

$$W_a(X) = \min(X, a) \text{ or } W_a(X) = X I_{(0,a]}(X) + a I_{(a,\infty)}(X),$$

where $I_S(y)$ is an indicator function equal to one when y is in the set S and zero otherwise. The latter form of representation is particularly convenient and is utilised extensively in what follows. Lee (1990) describes some other applications of this representation. The limited-value expectation of X is then just the expectation of $W_a(X)$:

$$E(X; a) = E(W_a(X)) = \int_0^a x dF(x) + a(1 - F(a)) = \int_0^a x dF(x) + aG(a),$$

where $G(x) = 1 - F(x)$.

We also need to consider the limited-value second moments. More generally, the limited-value k^{th} moment of X is

$$E(X^k; a) = E(W_a(X)^k) = \int_0^a x^k dF(x) + a^k G(a).$$

If $F(x)$ can be expressed in a simple functional form, it is often more convenient to use another integration formula for this quantity:

$$E(X^k; a) = k \int_0^a x^{k-1} G(x) dx .$$

This formula can be derived using integration by parts.

Consider a random sample X_1, X_2, \dots, X_n of size n from $F(x)$. Then the LVA of the sample, all limited above by a , is given by

$$Y(a) = \frac{1}{n} \sum_{i=1}^n W_a(X_i) .$$

Note that $\{ W_a(X_i), i = 1, 2, \dots, n \}$ is a set of independent and identically distributed (iid) random variables, each distributed like $W_a(X) = X I_{(0,a]}(X) + a I_{(a,\infty)}(X)$. The LVA $Y(a)$ is thus a sum of iid random variables, and so its mean and variance are

$$E(Y(a)) = E(X I_{(0,a]}(X)) + a G(a) = \int_0^a G(x) dx ;$$

$$\text{Var}(Y(a)) = \frac{1}{n} \left[E(X^2 I_{(0,a]}(X)) + a^2 G(a) - E(Y(a))^2 \right] = \frac{1}{n} \left[\int_0^a 2xG(x) dx - E(Y(a))^2 \right] .$$

Now let $a < b$ and consider the two random variables $Y(a)$ and $Y(b)$, in which

$$\text{Cov}(Y(a) , Y(b)) = \text{Cov} \left(\frac{1}{n} \sum_{i=1}^n W_a(X_i) , \frac{1}{n} \sum_{i=1}^n W_b(X_i) \right) = \frac{1}{n} \text{Cov}(W_a(X) , W_b(X))$$

To find the covariance on the right, we need to derive the product moment. First, we have

$$\begin{aligned} W_a(X) W_b(X) &= \left(X I_{(0,a]}(X) + a I_{(a,\infty)}(X) \right) \left(X I_{(0,b]}(X) + b I_{(b,\infty)}(X) \right) \\ &= X^2 I_{(0,a]}(X) + aX I_{(a,b]}(X) + ab I_{(b,\infty)}(X) . \end{aligned}$$

By adding and subtracting $a^2 I_{(a,\infty)}(X)$ to above, we have, after some algebra,

$$W_a(X) W_b(X) = W_a(X)^2 + a (W_b(X) - W_a(X)),$$

which gives

$$E(W_a(X) W_b(X)) = E(W_a(X)^2) + a [E(W_b(X)) - E(W_a(X))].$$

By writing

$$A_1(a) = E(W_a(X)) = \int_0^a x dF(x) + a G(a) = \int_0^a G(x) dx; \quad (1)$$

$$A_2(a) = E(W_a(X)^2) = \int_0^a x^2 dF(x) + a^2 G(a) = 2 \int_0^a xG(x) dx, \quad (2)$$

it can be seen that

$$E(Y(a)) = A_1(a);$$

$$\text{Var}(Y(a)) = \frac{1}{n} (A_2(a) - A_1(a)^2);$$

$$\text{Cov}(Y(a), Y(b)) = \frac{1}{n} [A_2(a) + a (A_1(b) - A_1(a)) - A_1(a) A_1(b)]$$

Therefore, a set of LVAs $[Y(d_1), Y(d_2), \dots, Y(d_p)]$, $d_1 < d_2 < \dots < d_p$, (in theory, it is possible to have $p = \infty$, but in practice, a finite p is sufficient) has a mean vector

$$[A_1(d_1), A_1(d_2), \dots, A_1(d_p)]$$

and a covariance matrix \mathbf{B} of order $p \times p$ with elements ($d_i \leq d_j$)

$$b_{ij} = \frac{1}{n} [A_2(d_i) + d_i (A_1(d_j) - A_1(d_i)) - A_1(d_i) A_1(d_j)]. \quad (3)$$

Moreover, this set of LVAs can be represented as an average of random vectors:

$$[Y(d_1), Y(d_2), \dots, Y(d_p)] = \frac{1}{n} \sum_{i=1}^n [W_{d_1}(X_i), W_{d_2}(X_i), \dots, W_{d_p}(X_i)].$$

By the multivariate central limit theorem, this random vector tends to follow a multivariate normal distribution with the same mean vector

and covariance matrix as above, when the number of observations n increases to infinity. The random vector

$$\mathbf{Z} = \left[Y(d_1) - A_1(d_1), Y(d_2) - A_1(d_2), \dots, Y(d_p) - A_1(d_p) \right]'$$

may then be treated as approximately multivariate normal with covariance matrix

$$\mathbf{B} = [b_{ij}]$$

as defined in (3), and so the quadratic form

$$T_0 = \mathbf{Z}'\mathbf{B}^{-1}\mathbf{Z}$$

follows approximately the χ^2 distribution on p degrees of freedom. Let \mathbf{C} be the inverse matrix of \mathbf{B} :

$$\mathbf{C} = [c_{ij}] = \mathbf{B}^{-1}. \quad (4)$$

We may also write T_0 as

$$T_0 = \sum_{i,j} c_{i,j} \left[Y(d_i) - A_1(d_i) \right] \left[Y(d_j) - A_1(d_j) \right] \quad (5)$$

For testing the null hypothesis that the random sample X_1, X_2, \dots, X_n comes from a completely specified distribution with distribution function $F(x)$, we compute the test statistic T_0 and compare its value against the upper 100α percentile of the χ^2 distribution on p degrees of freedom. A high value of T_0 reveals significant departure from the null hypothesis or a poor fit.

In practice, the observed LVAs are not compared against a completely specified distribution. More often, the hypothesis is that the observations come from a distribution with its parameters estimated from the same set of data. In this case, one could calculate the minimum quadratic form based on the LVAs, i.e. the minimum value attained by T_0 in (5) with an appropriate choice of the parameter values. In Appendix A, we show that the test statistic so constructed follows, for very large samples, the χ^2 distribution on $p - r$ degrees of freedom, where r is the number of parameters estimated.

The minimization of the quadratic form T_0 is relatively complicated. This is even so for the analogous, simpler case of Pearson's chi-square test based on frequency counts:

$$X^2 = \sum_{i=1}^p \frac{(O_i - E_i)^2}{E_i},$$

where O_i is the observed, and E_i the expected, frequency count in interval i for $i = 1, 2, \dots, p$. In conducting this chi-square test, the calculation of the minimum X^2 is usually not carried out in practice. Instead, some efficient estimates such as the maximum likelihood estimates (MLEs) of the parameters are obtained. The E_i 's in the X^2 formula are then computed based on these estimates.

This procedure can also be followed for the calculation of the quadratic form statistic. The MLEs can first be found for the parameters and then the quadratic form statistic

$$T = \sum_{i,j} \hat{c}_{i,j} \left[Y(d_i) - \hat{A}_1(d_i) \right] \left[Y(d_j) - \hat{A}_1(d_j) \right] \quad (6)$$

can be obtained where $\hat{c}_{i,j}$ and $\hat{A}_1(d_i)$ are computed according to the fitted distribution with its parameters estimated.

The value of T based on the MLEs is to be compared to the χ^2 distribution on $p - r$ degrees of freedom, although in theory it would not have exactly a limiting χ^2 distribution. We observe later in Section 4 that this procedure makes it easier to reject the hypothesis than when the minimization of the quadratic form is used, and the rejection rate is actually higher than the nominal rate. When the null hypothesis is not rejected, the fit is, on average, better than indicated by the nominal significance level. Moore (1984, 1986) provide a discussion on the distribution of the chi-square test statistic under this condition.

The computation of T requires the knowledge of the first two limited-value moments of the distribution under test. For some common distributions used in modelling insurance losses, the limited-value expectations are given in Hogg and Klugman (1984). Appendix B presents the formulae we derive for the limited-value h^{th} moments $E(X^h; x)$ for those distributions. Moreover, Luong and Thompson

(1987) discuss a more general framework of minimum-distance methods for goodness-of-fit testing.

Truncated Observations

Often the smaller losses are omitted from the data (eg. policy excess), and the quadratic form test suggested above needs to be modified. Suppose the point of truncation is w so that values less than w are not observed. Let the truncated random variable be denoted by X^* and its density function by f^* where $*$ refers to the truncation. Then we have

$$f^*(x) = \begin{cases} f(x)/G(w) & \text{if } x > w ; \\ 0 & \text{otherwise .} \end{cases}$$

The limited-value k^{th} moment of this distribution with the limit at a ($> w$) is

$$\begin{aligned} E(X^{*k}; a) &= \int_w^a x^k f^*(x) dx + a^k G^*(a) = \frac{\int_w^a x^k f(x) dx + a^k G(a)}{G(w)} \\ &= \frac{\int_0^a x^k f(x) dx - \int_0^w x^k f(x) dx + a^k G(a)}{G(w)} . \end{aligned}$$

By subtracting and adding $w^k G(w)$ in the numerator we have

$$A_1^*(a) = w + \frac{A_1(a) - A_1(w)}{G(w)} ; \tag{1}$$

$$A_2^*(a) = w^2 + \frac{A_2(a) - A_2(w)}{G(w)} , \tag{2}$$

which correspond to $A_1(a)$ and $A_2(a)$ and should be used in place of these two functions for the calculation of T when the observations are truncated from below. More specifically, given a set of LVAs $Y^*(d_1)$, $Y^*(d_2)$, ..., $Y^*(d_p)$ at the limits $d_1 < d_2 < \dots < d_p$ (all greater than the point of truncation w), we have ($d_i \leq d_j$)

$$b_{ij}^* = \frac{1}{n} \left[A_2^*(d_i) + d_i \left(A_1^*(d_j) - A_1^*(d_i) \right) - A_1^*(d_i) A_1^*(d_j) \right]; \quad (3^*)$$

$$\left[c_{ij}^* \right] = \left[b_{ij}^* \right]^{-1}; \quad (4^*)$$

$$T_0^* = \sum_{i,j} c_{ij}^* \left[Y^*(d_i) - A_1^*(d_i) \right] \left[Y^*(d_j) - A_1^*(d_j) \right]; \quad (5^*)$$

$$T^* = \sum_{i,j} \hat{c}_{ij}^* \left[Y^*(d_i) - \hat{A}_1^*(d_i) \right] \left[Y^*(d_j) - \hat{A}_1^*(d_j) \right]. \quad (6^*)$$

The test can then be carried out as in the case without truncation.

A Test Based on Layers

Let L_j be the average value within the interval (d_{j-1}, d_j) $d_1 < d_2 < \dots < d_p$, with p possibly equal to infinity, defined as

$$L_j = Y(d_j) - Y(d_{j-1}).$$

For a collection of losses, L_j would represent the average layer of loss within the limits (d_{j-1}, d_j) . Consider the random vector (L_1, L_2, \dots, L_p) . By argument similar to that applied to the $Y(d_j)$'s, this vector also has a limiting multivariate normal distribution with its mean vector and covariance matrix that can readily be derived. A quadratic form similar to T_0 and T can be constructed with a limiting χ^2 distribution. A moment of reflection shows that the the set of L_j 's has the advantage of having correlations lower than those between the $Y(d_j)$'s. We will further explore this layer approach in the future research.

3 The Testing Procedure

The following is a summary of the testing procedure:

- a) Decide on the theoretical distribution to fit the data. This may be completely specified with parameter values known or more often with parameters estimated from the same set of observations.

- b) Decide on the truncation point w , ie. the value below which no observations are made. This value may be zero.
- c) Decide on the set of points $w < d_1 < d_2 < \dots < d_p$ at which the LVAs are to be calculated. To obtain good results, these points should be reasonably well-spaced. They need not be fixed, and can depend on the observations. Further discussion on the spacing of intervals is set forth below.
- d) Compute the sample LVAs $Y(d_1), Y(d_2), \dots, Y(d_p)$ at the limits d_1, d_2, \dots, d_p , where $Y(a) = \frac{1}{n} \sum_{i=1}^n \min(X_i, a)$.
- e) Estimate the parameter values if they are not known, as is usually the case. The simpler maximum likelihood method or the minimum quadratic form method may be used.
- f) Compute the limited-value moments, the covariance matrix, and its inverse from (1), (2), (3), and (4) for the case without truncation (ie. $w = 0$) or from (1'), (2'), (3'), and (4') for the case with truncation (ie. $w > 0$). Formulae for computation of $A_1(a)$ and $A_2(a)$ for some commonly used distributions are given in Appendix B.
- g) For fitting a completely specified distribution, compute the T_0 statistic from (5) (without truncation) or T_0^* from (5') (with truncation). For fitting a distribution with its parameters estimated, compute the T statistic from (6) (without truncation) or T^* from (6') (with truncation).
- h) Compare T_0 (or T_0^*) to the χ^2 distribution on p degrees of freedom where the parameter values are fully specified, or compare T (or T^*) to the χ^2 distribution on $p - r$ degrees of freedom if r parameters are estimated.

When the minimum quadratic form is to be computed, steps (e) to (g) will have to be performed repeatedly in succession using some iterative minimising algorithm.

The choice of the intervals or cells has received widespread attention in the literature on the classical chi-square goodness-of-fit test. Given the same hypothesis and the same set of data, different systems of intervals could lead to different conclusions, introducing an element of arbitrariness in the test. The general recommendation is to use an equal-probability system of intervals to achieve not only objectivity but also closer approximation to the χ^2 distribution. There are also suggestions for the appropriate number of cells, a specific one being $2n^{2/5}$ cells with significance level at 5%. More important is the rule of avoiding cells having very small expected frequencies. One explicit suggestion is that the expected frequency of any cell should not be less than 5. Moore (1986) provides a detailed discussion on this topic.

It is not known how these suggestions are exactly related to the quadratic form test. In practice, the chi-square test works satisfactorily with widely varying configurations of intervals, when none of the cells have very small expected frequencies. In actuarial work, where good judgement is involved in so many other steps in an investigation, the pursuit of objectivity is not a major concern. For example, one may like to put more emphasis on the upper tail of the distribution because this region contributes more dollar value. In the quadratic form test, the layer averages, ie. the successive differences of LVAs (end of Section 2) correspond to the cell probabilities in the chi-square test. While the question of spacing is not thoroughly investigated in this exploratory study, it appears that a reasonable system of intervals would have layer averages not too widely different, with none of them being too small.

Numerical Examples

We first illustrate the testing procedure by applying it to a data set of hurricane losses as contained in Chapter 4 of Hogg and Klugman (1984). There are $n = 35$ observations, truncated below at $d = 5,000,000$, as shown in Table 1.

Table 2 shows the results of fitting the Weibull distribution (refer to Appendix B for the form of the density function). The second column presents the end points up to which the LVAs are calculated. The remaining three columns show the expected LVAs, $E(Y(d_j))$, when the parameters (c, τ) are estimated by three different methods, namely the maximum likelihood (ML), the minimum chi-square (MC), and the minimum quadratic form (MQ) methods. The MC parameter estimates are found by minimising $X^2 = \sum_{i=1}^p (O_i - E_i)^2 / E_i$, where O_i and E_i are the observed and expected frequency counts in the i^{th} interval. The MQ parameter estimates are the values that minimise the quadratic form T^* in (6').

Note that both the MC and the MQ methods require an iterative process. For the computation of the MQ parameter estimates, for example, initial values of (c, τ) can be found by some convenient methods such as the method of moments and ML. Given the (c, τ) values, the values of $A_1(d_j)$ and $A_2(d_j)$ in (1) and (2) can be obtained from the formulae in Appendix B as $E(X; d_j)$ and $E(X^2; d_j)$ for the Weibull distribution. The covariance matrix $[b_{ij}^*]$ and its inverse $[c_{ij}^*]$, and hence the quadratic form T^* , can be computed according to the formulae in Section 2. (Note that the observations are truncated from below at 5,000,000.) One of the available computer algorithms can be used to determine a new, improved set of parameter values. The iterative process continues until the minimum value of the quadratic form is reached.

The fitted parameter values of the Weibull distribution and the resulting values of the test statistics are summarised in Table 3. The corresponding results from fitting the lognormal and the Pareto distributions are also presented for comparison.

The Weibull, lognormal, and Pareto distributions are fitted to the data (using 8 limits with 2 parameters estimated) resulting in minimum quadratic form values of 2.1802, 2.2613, and 3.1286 respectively. Thus

all these distributions can be fitted without leading to a rejection of the null hypothesis at any reasonable significance level when the observed values of the test statistics are compared to the limiting χ^2 distributions. The Weibull distribution leads to the smallest quadratic form value and so can be judged as providing the closest fit, as also noted in Hogg and Klugman (1984) in a rather informal way.

Similarly, the three distributions are fitted to a set of individual claims data (using 12 limits in this case) considered in Currie (1992). There are $n = 96$ observations, as shown in Table 1. Table 3 shows that the lognormal distribution leads to the smallest quadratic form value and provides the closest fit. The null hypothesis is not rejected for both the lognormal and Pareto distributions.

Table 1 Hurricane Loss Data (in thousands) in 1981 Dollars, Truncated below at $w = 5,000,000$, and Individual Claims Data

Hurricane Loss Data								
6,766	7,123	10,562	14,474	15,351	16,983	18,383	19,030	25,304
29,112	30,146	33,727	40,596	41,409	47,905	49,397	52,600	59,917
63,123	77,809	102,942	103,217	123,680	140,136	192,013	198,446	227,338
329,511	361,200	421,680	513,586	545,778	750,389	863,881	1,638,000	
Individual Claims Data								
24	26	73	84	102	115	132	159	207
240	241	254	268	272	282	300	302	329
346	359	367	375	378	384	452	475	495
503	531	543	563	594	609	671	687	691
716	757	821	829	885	893	968	1,053	1,081
1,083	1,150	1,205	1,262	1,270	1,351	1,385	1,498	1,546
1,565	1,635	1,671	1,706	1,820	1,829	1,855	1,873	1,914
2,030	2,066	2,240	2,413	2,421	2,521	2,586	2,727	2,797
2,850	2,989	3,110	3,166	3,383	3,443	3,512	3,515	3,531
4,068	4,527	5,006	5,065	5,481	6,046	7,003	7,245	7,477
8,738	9,197	16,370	17,605	25,318	58,524			

Table 2 Expected LVAs (in millions) for Fitting Weibull Distribution to Hurricane Loss Data using ML, MC, and MQ Methods

Interval	End Point	ML	MC	MQ
1	5	5.00	5.00	5.00
2	10	9.75	9.76	9.75
3	25	22.03	22.11	22.03
4	50	38.64	38.89	38.66
5	100	63.90	64.58	63.98
6	250	110.47	112.32	110.80
7	500	149.45	152.68	150.22
8	1,000	181.45	186.10	182.84
9	2,500	202.48	208.29	204.62
10	5,000	206.19	212.24	208.57
11	∞	206.66	212.75	209.09

Table 3 Parameter Estimates of Selected Distributions and Computed Values of T Statistics

Hurricane Loss Data									
Method	Weibull			Lognormal			Pareto		
	c	τ	T^*	μ	σ	T^*	λ	α	T^*
ML	0.000072264	0.52090	2.1817	17.9526	1.6027	2.3203	73709152	1.1571	3.2740
MC	0.000064491	0.52559	2.1908	18.0451	1.6208	2.2877	70734053	1.0445	3.1424
MQ	0.000074999	0.51877	2.1802	17.9913	1.6424	2.2613	68486509	1.0457	3.1286

Individual Claims Data									
Method	Weibull			Lognormal			Pareto		
	c	τ	T	μ	σ	T	λ	α	T
ML	0.004073	0.713241	21.9138	7.021540	1.398584	4.1096	2702.47	1.908407	5.6400
MC	0.002008	0.816562	71.8597	7.041927	1.293009	4.5061	3392.86	2.313451	6.4799
MQ	0.005633	0.668106	19.5294	7.051828	1.359028	3.9063	2643.86	1.881632	5.6364

4 Tendency to χ^2 Distribution – Simulation Studies

Sampling experiments are now carried out to check how well the quadratic form test statistic T in (6) follows the limiting χ^2 distribution. Tables 4A to 4C show the results of these experiments. Both Pearson's chi-square test statistic χ^2 and the quadratic form test statistic T are calculated. In each case, two scenarios are considered: (a) the parameters are estimated from the observations by the ML method; (b) the parameters are estimated by the MC and the MQ methods correspondingly.

For each sample, the range is divided into 10 intervals based on the following quantiles: (0.20, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 0.95, 0.975). In the case of known parameter values, the intervals are fixed according to this set of quantiles. Following the usual practice in conducting the chi-square test, when the parameter values are not known, the intervals are not fixed; rather, they vary with the observations with divisions roughly corresponding to the quantiles above. More specifically, parameter estimates are first obtained under scenarios (a) and (b). The intervals are then determined by the quantiles above, in accordance with the estimated distribution. Moore (1986) provides a detailed discussion of using such random intervals for the chi-square test. We expect the validity of the chi-square procedure to be carried over to the quadratic form test, as revealed by our simulation results.

With this set of intervals/limits, the χ^2 statistic and the T statistic are computed for each sample. Ten thousand repetitions are made. In every repetition, each of the two statistics is obtained under two different scenarios (a) and (b), yielding a total of four quantities, the distributions of which are shown across the sub-tables in Tables 4A to 4C.

From the statistical theory in Appendix A, the T statistic under scenario (b) would have a limiting χ^2 distribution on $10 - r$ degrees of freedom, where 10 is the number of limits according to the quantiles

above and r is the number of parameters of the distribution law under test ($r = 2$ for all the three distributions considered here). Note that, unlike the chi-square test, the number of limits is not diminished by one in calculating the degrees of freedom in the quadratic form test. On the other hand, the χ^2 statistic under scenario (b) would have, according to the classical theory, a limiting χ^2 distribution on $9 - r$ degrees of freedom. The reason for this difference is that while the total frequency count is fixed and therefore forms a constraint in the chi-square test, there is no such constraint involved in the quadratic form test.

The purpose of the simulation exercise is to check whether, or how well, the four statistics follow the limiting χ^2 distributions. For each case, the following 5 quantiles of the χ^2 distribution with an appropriate degree of freedom are used: (0.25, 0.50, 0.75, 0.90, 0.95). Each sub-table in Tables 4A to 4C shows the expected frequencies for a statistic, out of 10,000, falling into 6 intervals formed from the 5 quantiles, assuming this statistic follows the appropriate χ^2 distribution, and the corresponding observed frequencies. For each statistic, the degree of agreement between the observed and expected frequencies is an indication of how close the distribution of this statistic is to the limiting χ^2 distribution.

The overall closeness between the frequency distribution of each statistic and the limiting χ^2 distribution can be measured by another χ^2 value, and such values are shown in the bottom line of each sub-table. (These values can then be compared against the χ^2 distribution on 5 degrees of freedom, corresponding to the 6 intervals. The critical value at 5% significance level is 11.07.) Firstly, there is evidence that the T statistic from the MQ method follows the χ^2 distribution more closely, ie. it converges faster to the limiting χ^2 distribution as the sample size increases, when compared to the χ^2 statistic from the MC method. Secondly, while our figures support the theory that the X^2 statistic calculated from the ML parameter estimates do not follow the χ^2 distribution, the distribution of the T statistic from the ML method appears to be relatively closer to the χ^2 distribution. This result

suggests that the χ^2 distribution may serve as a practical approximation for the T statistic based on the ML parameter estimates.

A number of our figures show significant overall departure of the test statistics from the limiting χ^2 distributions, especially when the sample size is small. For practical applications, however, it is more useful to focus on the upper range. Generally, the T statistic based on the MQ method tends to have a smaller tail than the limiting χ^2 distribution. If the procedure is to reject the null hypothesis when the observed minimum quadratic form exceeds the the upper 5th percentile of the χ^2 distribution, the probability α of type I error would actually be less than 5%. The hypothesis would thus be rejected less often than suggested by the nominal rate. This problem can be greatly alleviated when the sample size is increased to, say, a few hundreds. Note also that for the T statistic based on the ML parameter estimates, the χ^2 distribution provides a reasonable approximation at the upper end, when the samples are simulated from the Weibull and Pareto distributions. Referring to the remark in Section 2, however, the hypothesis tends to be rejected more frequently than suggested by the nominal level in this case.

Table 4A Simulation Results on χ^2 and T Statistics for Weibull Samples ($c = 0.0000725$, $\tau = 0.521$)

E_i	Q_i	$(Q_i - E_i)^2 / E_i$	Q_i	$(Q_i - E_i)^2 / E_i$	Q_i	$(Q_i - E_i)^2 / E_i$	Q_i	$(Q_i - E_i)^2 / E_i$
n = 50	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	2,042	83.91	2,644	8.29	1,906	141.13	2,889	60.53
2,500	2,511	0.05	2,590	3.24	2,575	2.25	2,696	15.37
2,500	2,723	19.89	2,497	0.00	2,810	38.44	2,446	1.17
1,500	1,698	26.14	1,346	15.81	1,656	16.22	1,259	38.72
500	505	0.05	450	5.00	567	8.98	380	28.80
500	521	0.88	473	1.46	486	0.39	330	57.80
Total		130.91		33.81		207.42		202.38
n = 100	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	2,168	44.09	2,550	1.00	2,084	69.22	2,802	36.48
2,500	2,373	6.45	2,512	0.06	2,409	3.31	2,583	2.76
2,500	2,709	17.47	2,532	0.41	2,690	14.44	2,464	0.52

E_i	O_i	$(O_i - E_i)^2/E_i$	O_i	$(O_i - E_i)^2/E_i$	O_i	$(O_i - E_i)^2/E_i$	O_i	$(O_i - E_i)^2/E_i$
1,500	1,663	17.71	1,418	4.48	1,707	28.57	1,309	24.32
500	525	1.25	482	0.65	551	5.20	456	3.87
500	562	7.69	506	0.07	559	6.96	386	25.99
Total		94.66		6.67		127.70		93.94
n = 250	χ^2 from ML		τ from ML		χ^2 from MC		τ from MQ	
2,500	2,118	58.37	2,470	0.36	2,163	45.43	2,609	4.75
2,500	2,445	1.21	2,534	0.46	2,388	5.02	2,566	1.74
2,500	2,658	9.99	2,560	1.44	2,600	4.00	2,575	2.25
1,500	1,628	10.92	1,444	2.09	1,663	17.71	1,363	12.51
500	589	15.84	468	2.05	613	25.54	449	5.20
500	562	7.69	524	1.15	573	10.66	438	7.69
Total		104.02		7.55		108.35		34.15
n = 500	χ^2 from ML		τ from ML		χ^2 from MC		τ from MQ	
2,500	2,172	43.03	2,429	2.02	2,282	19.01	2,528	0.31
2,500	2,538	0.58	2,602	4.16	2,496	0.01	2,627	6.45
2,500	2,658	9.99	2,466	0.46	2,571	2.02	2,476	0.23
1,500	1,512	0.10	1,465	0.82	1,548	1.54	1,422	4.06
500	560	7.20	500	0.00	549	4.80	466	2.31
500	560	7.20	538	2.89	554	5.83	481	0.72
Total		68.09		10.35		33.20		14.09

Table 4B Simulation Results on χ^2 and T Statistics for Lognormal Samples ($\mu = 17.2$, $\sigma = 1.45$)

E_i	Q_i	$(Q_i - E_i)^2$	Q_i	$(Q_i - E_i)^2$	Q_i	$(Q_i - E_i)^2$	Q_i	$(Q_i - E_i)^2$
$n = 50$	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	1,971	111.94	2,660	10.24	1,999	100.40	3,033	113.64
2,500	2,511	0.05	2,508	0.03	2,519	0.14	2,667	11.16
2,500	2,785	32.49	2,493	0.02	2,791	33.87	2,438	1.54
1,500	1,657	16.43	1,407	5.77	1,707	28.57	1,219	52.64
500	533	2.18	419	13.12	495	0.05	348	46.21
500	543	3.70	513	0.34	489	0.24	295	84.05
Total		166.78		29.51		163.28		309.23
$n = 100$	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	2,121	57.46	2,505	0.01	2,083	69.56	2,806	37.45
2,500	2,399	4.08	2,548	0.92	2,481	0.14	2,640	7.84
2,500	2,641	7.95	2,411	3.17	2,673	11.97	2,484	0.10
1,500	1,677	20.89	1,438	2.56	1,693	24.83	1,317	22.33
500	582	13.45	514	0.39	541	3.36	403	18.82
500	580	12.80	584	14.11	529	1.68	350	45.00
Total		116.62		21.17		111.55		131.54
$n = 250$	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	2,140	51.84	2,474	0.27	2,260	23.04	2,698	15.68
2,500	2,447	1.12	2,426	2.19	2,481	0.14	2,529	0.34
2,500	2,583	2.76	2,476	0.23	2,502	0.00	2,442	1.35
1,500	1,677	20.89	1,431	3.17	1,680	21.60	1,389	8.21
500	581	13.12	540	3.20	534	2.31	448	5.41
500	572	10.37	653	46.82	543	3.70	494	0.07
Total		100.10		55.88		50.80		31.06
$n = 500$	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	2,089	67.57	2,433	1.80	2,268	21.53	2,605	4.41
2,500	2,490	0.04	2,459	0.67	2,444	1.25	2,531	0.38
2,500	2,658	9.99	2,479	0.18	2,609	4.75	2,492	0.03
1,500	1,616	8.97	1,457	1.23	1,603	7.07	1,355	14.02
500	573	10.66	516	0.51	543	3.70	487	0.34

E_i	O_i	$(O - E)^2$						
500	574	10.95	656	48.67	533	2.18	530	1.80
Total		108.17		53.06		40.49		20.97

Table 4C Simulation Results on χ^2 and T Statistics for Pareto Samples $\lambda = 70 \times 10^6, \alpha = 0.8$

E_i	O_i	$(O - E)^2$	O_i	$(O - E)^2$	O_i	$(O - E)^2$	O_i	$(O - E)^2$
$n = 50$	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	2,116	58.98	2,198	36.48	1,924	132.71	2,332	11.29
2,500	2,519	0.14	2,563	1.59	2,527	0.29	2,678	12.67
2,500	2,672	11.83	2,568	1.85	2,751	25.20	2,684	13.54
1,500	1,619	9.44	1,626	10.58	1,734	36.50	1,466	0.77
500	510	0.20	582	13.45	535	2.45	464	2.59
500	564	8.19	463	2.74	529	1.68	376	30.75
Total		88.79		66.69		198.84		71.62
$n = 100$	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	2,335	10.89	2,296	16.65	2,027	89.49	2,375	6.25
2,500	2,426	2.19	2,564	1.64	2,521	0.18	2,606	4.49
2,500	2,626	6.35	2,610	4.84	2,637	7.51	2,633	7.08
1,500	1,624	10.25	1,512	0.10	1,711	29.68	1,489	0.08
500	492	0.13	496	0.03	560	7.20	465	2.45
500	497	0.02	522	0.97	544	3.87	432	9.25
Total		29.83		24.22		137.93		29.60
$n = 250$	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	2,295	16.81	2,422	2.43	2,179	41.22	2,497	0.00
2,500	2,472	0.31	2,451	0.96	2,411	3.17	2,422	2.43
2,500	2,586	2.96	2,537	0.55	2,617	5.48	2,576	2.31
1,500	1,598	6.40	1,586	4.93	1,660	17.07	1,538	0.96
500	516	0.51	509	0.16	568	9.25	504	0.03
500	533	2.18	495	0.05	565	8.45	463	2.74
Total		29.17		9.08		84.63		8.48
$n = 500$	χ^2 from ML		T from ML		χ^2 from MC		T from MQ	
2,500	2,403	3.76	2,463	0.55	2,240	27.04	2,510	0.04
2,500	2,474	0.27	2,470	0.36	2,452	0.92	2,467	0.44

E_i	O_i	$(O_i - E_i)^2$						
2,500	2,557	1.30	2,515	0.09	2,663	10.63	2,540	0.64
1,500	1,518	0.22	1,538	0.96	1,562	2.56	1,504	0.01
500	522	0.97	484	0.51	546	4.23	483	0.58
500	526	1.35	530	1.80	537	2.74	496	0.03
Total		7.87		4.27		48.12		1.74

5 The Power of the Test

In this section, we present some simulation results on the power of Pearson's chi-square test and the suggested quadratic form test. In Table 5, the fitting distribution shown in the second column is the distribution under the null hypothesis, and this distribution is fitted to the observations generated from another distribution stated in the first column. For example, in the first row, the null hypothesis is that the observations follow a lognormal distribution, but they have in fact come from a Weibull distribution. The remaining items are the empirical rejection rates at a significance level of 5%.

In our simulation studies, using a sample size of 100, ten thousand repetitions are made for each case. The range is divided into 10 intervals as described in Section 4. For each sample, the X^2 statistic and the T statistic under scenario (b) in Section 4 are calculated and the empirical rejection rates at the 5% significance level are obtained.

Because the distributions of the X^2 and T statistics are not exactly that of χ^2 , it may not be accurate enough to take the χ^2 percentile as the cut-off value for the tests. As such, the cut-off point for each fitting distribution is determined empirically as follows. Samples are generated from the fitting distribution and, for each sample, the two test statistics under the null hypothesis that the generating distribution and fitting distribution belong to the same family are computed. The resulting values could be regarded as observations from the null distributions of the X^2 and T statistics. For each of these two statistics, a rescaled gamma distribution with adjustable parameters is fitted to the simulated values of the statistic, and the null hypothesis cut-off point at 5% is thus determined. This procedure would ensure

that the rejection rates are close to the nominal rates under the null hypothesis. We also use the sample quantiles of the two statistics as a comparison. The generating distributions have the following parameter values:

Weibull: $c = 0.0000725, \tau = 0.521$;

Lognormal: $\mu = 17.2, \sigma = 1.78$;

Pareto: $\lambda = 40 \times 10^6, \alpha = 0.8$.

These generating distributions resemble the distribution of the hurricane losses discussed in Section 3. Other parameters have also been tested and the results are broadly similar.

Table 5 shows that in general, the quadratic form test has appreciably better power than the chi-square test. The power varies considerably over different fitting-generating pairs, indicating perhaps the degree of similarity between the distributions in a pair. In comparing testing methods, it is the relative power, rather than the absolute power, being the major concern.

It should be noted that both the chi-square test and the quadratic form test are general significance tests without specifying any alternatives. Cox and Hinkley (1974) (Chapter 3) call these pure significance tests. In the statistical literature (eg. Moore 1986), there are other tests that assess the goodness of fit of a given distribution against a specified alternative distribution. Such tests usually make use of the characteristics of the given distributions and tend to have greater power.

In theory, it can be argued that the LVAs generally contain more information than frequencies. Consider the empirical distribution function (edf) over the interval formed by two consecutive points in the tests. The frequency within the interval gives the jump of the edf over the interval, but the LVAs would also give some indication about the shape of the edf. This appears to explain the superior performance of the quadratic form test over the chi-square test in our simulation studies.

Table 5 Power of X^2 and T Statistics

Rescaled Gamma				Sample Quantile			
Population	Fit	T	X^2	Population	Fit	T	X^2
W	LN	0.3669	0.2079	W	LN	0.3834	0.2219
W	Pa	0.8096	0.5481	W	Pa	0.8176	0.5643
LN	W	0.7145	0.3795	LN	W	0.7286	0.4008
LN	Pa	0.0749	0.0639	LN	Pa	0.0846	0.0755
Pa	W	0.9359	0.7451	Pa	W	0.9375	0.7619
Pa	LN	0.2028	0.1182	Pa	LN	0.2049	0.1327

(W: Weibull ; LN: Lognormal ; Pa: Pareto)

References

- Balasoorya U. and Low C. K. (2008). *Modeling insurance claims with extreme observations: transformed kernel density and generalized lambda distribution*. North American Actuarial Journal 12(2): 129–142.
- Cox D. R. and Hinkley D. V. (1974). *Theoretical Statistics*. Chapman and Hall, London.
- Currie I. D. (1992). *Loss Distributions*. Institute of Actuaries, London, and Faculty of Actuaries, Edinburgh.
- Hogg R. V. and Klugman S. A. (1984). *Loss Distributions*. John Wiley & Sons Inc., New York.
- Lee Y. S. (1990). *On the representation of loss distribution*. Proceedings of the Casualty Actuarial Society Vol. 77: 204–224.
- Luong A. and Thompson M. E. (1987). *Minimum distance methods based on quadratic distances for transforms*. Canadian Journal of Statistics 15(3): 239–251.
- Moore D. S. (1984). *Measures of lack of fit from tests of chi-squared type*. Journal of Statistical Planning and Inference 10: 151–166.
- Moore D. S. (1986). *Tests of chi-squared type*. D'Agostino R. B. and Stephens M. A. (Eds.), Goodness-of-Fit Techniques, pp. 63–96. Marcel Dekker, New York.
- Rao C. R. (1973). *Linear Statistical Inference and Its Applications*. John Wiley & Sons Inc., New York.

Appendix A

Consider a random sample X_1, X_2, \dots, X_n from the density function $f(x; \theta)$, where the parameter θ may be vector-valued with dimensionality r . Let $\mathbf{Y} = [Y_1, Y_2, \dots, Y_k]'$, $k > r$, be a vector representing the LVAs at the limits $d_1 < d_2 < \dots < d_k$ for this random sample, and $\boldsymbol{\eta}(\theta) = \boldsymbol{\eta} = [\eta_1, \eta_2, \dots, \eta_k]'$ be a vector representing the corresponding population limited-value means.

By the central limit theorem, the random variable \mathbf{Y} has, for large samples, a distribution equivalent to a multivariate normal distribution with mean $\boldsymbol{\eta}$ and known covariance matrix $\boldsymbol{\Sigma}/n$ ($k \times k$). Asymptotically, the density function of \mathbf{Y} is

$$f_{\mathbf{Y}}(\mathbf{y}) \approx (2\pi)^{-k/2} (|\boldsymbol{\Sigma}|/n)^{-1/2} \exp\left\{-\frac{n}{2}(\mathbf{y} - \boldsymbol{\eta})' \boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\eta})\right\}. \quad (7)$$

The log-likelihood function is therefore

$$l = K - \frac{1}{2} \ln(|\boldsymbol{\Sigma}|) - \frac{n}{2} (\mathbf{y} - \boldsymbol{\eta})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\eta}),$$

where K is a constant, ie. not dependent on $\boldsymbol{\eta}$ or $\boldsymbol{\Sigma}$. Now consider the estimation of $\boldsymbol{\eta}$. The efficient score \mathbf{S} ($r \times 1$) is given by

$$\mathbf{S} = \frac{\partial l}{\partial \boldsymbol{\theta}} = n \frac{\partial \boldsymbol{\eta}'}{\partial \boldsymbol{\theta}}^{-1} (\mathbf{y} - \boldsymbol{\eta}),$$

and the information matrix \mathbf{I}_θ ($r \times r$) is given by

$$\mathbf{I}_\theta = \boldsymbol{\Sigma} \text{Var}(\mathbf{S}) = n \frac{\partial \boldsymbol{\eta}'}{\partial \boldsymbol{\theta}}^{-1} \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}'}$$

Let $\hat{\boldsymbol{\theta}}$ be an efficient estimator of $\boldsymbol{\theta}$ with the form

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\alpha} + \boldsymbol{\beta}' \mathbf{S} + \mathbf{R},$$

where $\boldsymbol{\alpha}$ ($r \times 1$) and $\boldsymbol{\beta}$ ($r \times r$) consist of real constants, possibly functions of $\boldsymbol{\theta}$, \mathbf{R} ($r \times 1$) is a remainder of order $O(n^{-1/2})$ or lower, and $\hat{\boldsymbol{\theta}}$ is asymptotically unbiased for $\boldsymbol{\theta}$ with asymptotic covariance matrix \mathbf{I}_θ^{-1} . (Refer to Rao 1973, Chapter 5 for more details.) These requirements lead to

$$\hat{\theta} - \theta \sim n \mathbf{I}_{\theta}^{-1} \frac{\partial \boldsymbol{\eta}'}{\partial \boldsymbol{\theta}} \mathbf{Y}^{-1} (- \boldsymbol{\eta}), \tag{8}$$

where the symbol \sim means the expressions on the two sides are asymptotically equivalent.

By standard estimation theory, the maximum likelihood estimator of θ (with $\boldsymbol{\eta} = \boldsymbol{\eta}(\theta)$) for the density function in (7) will satisfy these requirements. Note that the ML estimates in this case are precisely the parameter values that minimise the quadratic form in (6) and would lead to the minimum quadratic form test statistic T . Let also $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}(\hat{\theta})$ be the estimator of the vector of the limited-value means $\boldsymbol{\eta}$, based on the estimator $\hat{\theta}$. In the notation of this section, $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}(\hat{\theta})$ would be the vector of the expected LVAs at the limits $d_1 < d_2 < \dots < d_k$, evaluated with respect to the distribution having the parameter value $\hat{\theta}$.

We now show that the quadratic form

$$T = n (\mathbf{Y} \boldsymbol{\Sigma} \hat{\boldsymbol{\eta}}) \mathbf{Y}^{-1} (- \hat{\boldsymbol{\eta}}) \tag{9}$$

has a limiting χ^2 distribution with $k - r$ degrees of freedom. First we write, with the help of (8), $\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}$ asymptotically as

$$\hat{\boldsymbol{\eta}} - \boldsymbol{\eta} = \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}'} (\hat{\theta} - \theta) = n \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}'} \mathbf{I}_{\theta}^{-1} \frac{\partial \boldsymbol{\eta}'}{\partial \boldsymbol{\theta}} \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \boldsymbol{\eta}) = \mathbf{B} \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \boldsymbol{\eta}),$$

where

$$\mathbf{B} (k \times k) = n \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}'} \mathbf{I}_{\theta}^{-1} \frac{\partial \boldsymbol{\eta}'}{\partial \boldsymbol{\theta}}.$$

Note that this is essentially a Taylor expansion retaining only the first-order term. Substituting this in (9) yields

$$\begin{aligned} T &= n (\mathbf{Y} - \boldsymbol{\Sigma} \boldsymbol{\eta} - \mathbf{B}^{-1} (- \boldsymbol{\Sigma} \boldsymbol{\eta})) \mathbf{Y}^{-1} (\mathbf{B} - \boldsymbol{\Sigma} \boldsymbol{\eta} - \mathbf{Y}^{-1} (- \boldsymbol{\eta})) \\ &= n (\mathbf{Y} - \boldsymbol{\Sigma} \boldsymbol{\eta}) \mathbf{B}' (\mathbf{E} - \mathbf{I}^{-1} \mathbf{B}) \boldsymbol{\Sigma}^{-1} (- \mathbf{Y}^{-1}) (- \boldsymbol{\eta}) \\ &= n (\mathbf{Y} \boldsymbol{\Sigma} \boldsymbol{\eta}) \mathbf{I}^{-1/2} (\mathbf{B} \boldsymbol{\Sigma}^{-1/2} \boldsymbol{\Sigma}^{-1/2}) \mathbf{Y}^{-1/2} (- \boldsymbol{\eta}), \end{aligned}$$

in which it can be shown, from the expressions for \mathbf{I}_{θ} and for \mathbf{B} , that the matrix $\mathbf{I} \boldsymbol{\Sigma}^{-1/2} \mathbf{B}' \boldsymbol{\Sigma}^{-1/2}$ is idempotent with rank $k - r$. From the

theory of the distribution of quadratic forms in normal variables, we conclude that T asymptotically follows the χ^2 distribution with $k - r$ degrees of freedom.

If Σ is estimated from the data, then T becomes

$$T = n (\mathbf{Y}\hat{\Sigma}\hat{\boldsymbol{\eta}})\mathbf{Y}^{-1}(-\hat{\boldsymbol{\eta}}),$$

where $\hat{\Sigma}$ is an efficient estimator of Σ based on the observations, and it is natural to take $\hat{\Sigma} = n \begin{bmatrix} \hat{b}_{i,j} \end{bmatrix}$ as given in Section 2. In this case, by the well-known Wald's theorem, T again asymptotically follows the χ^2 distribution with $k - r$ degrees of freedom.

It is important to note that $\hat{\boldsymbol{\theta}}$ should be obtained by maximising l , ie. minimising the quadratic form in (6); then it is equivalent to the maximum likelihood estimator based on the density in (7) and is therefore efficient, and T will follow the χ^2 distribution asymptotically. However, if $\hat{\boldsymbol{\theta}}$ is obtained by other methods, eg. maximising the original likelihood function based on $f(x;\boldsymbol{\theta})$, then the T statistic so constructed will not asymptotically follow the χ^2 distribution. As stated in Section 4, it is generally easier to reject the hypothesis of good fit under the latter case. If the sample size n is large and the number k of the LVAs is not too small, the χ^2 distribution would nevertheless serve as a reasonable approximation. Moore (1986) provides a similar discussion on the X^2 statistic.

Appendix B

This Appendix gives the limited-value moments we derive for the eight distributions listed in the Appendix of Hogg and Klugman (1984). If a random variable Y has distribution function $F(y)$, then the limited-value k^{th} moment limited to x is

$$E(Y^k; x) = \int_0^x y^k dF(y) + x^k (1 - F(x)) = k \int_0^x y^{k-1} (1 - F(y)) dy.$$

Although for the purpose of this paper only the first and second moments are needed, general expressions for the moments are presented. Note that the first moment is also provided in Hogg and Klugman (1984).

Burr distribution

$$f(x) = \alpha \tau \lambda^\alpha x^{\tau-1} (\lambda + x^\tau)^{-\alpha-1}, \quad x > 0, \alpha > 0, \lambda > 0, \tau > 0.$$

$$E(X^h; x) = \begin{cases} \frac{h \lambda^{\frac{h}{\tau}} \Gamma(\frac{h}{\tau}) \Gamma(\alpha - \frac{h}{\tau})}{\tau \Gamma(\alpha)} B\left(\frac{h}{\tau}, \alpha - \frac{h}{\tau}; \frac{x^\tau}{\lambda + x^\tau}\right) & \text{if } \alpha > \frac{h}{\tau}; \\ \frac{h \lambda^{\frac{h}{\tau}}}{\tau} \int_0^{\frac{x^\tau}{\lambda + x^\tau}} \frac{x^\tau}{\lambda + x^\tau} t^{\frac{h}{\tau}-1} (1-t)^{\alpha - \frac{h}{\tau}-1} dt & \text{if } \alpha > 0. \end{cases}$$

Pareto distribution

$$f(x) = \alpha \lambda^\alpha (\lambda + x)^{-\alpha-1}, \quad x > 0, \alpha > 0, \lambda > 0.$$

$$E(X^h; x) = h \lambda^h \sum_{j=0}^{h-1} \binom{h-1}{j} (-1)^{h-1-j} \frac{1}{\alpha-1-j} \left(1 - \left(\frac{\lambda}{\lambda+x} \right)^{\alpha-1-j} \right),$$

where for $j = \alpha - 1$ the term $\frac{1}{\alpha-1-j} \left(1 - \left(\frac{\lambda}{\lambda+x} \right)^{\alpha-1-j} \right)$ is to be replaced by $\ln\left(\frac{\lambda+x}{\lambda}\right)$. For $h = 1, 2$, the limited-value moments are as follows.

$$E(X; x) = \begin{cases} \frac{\lambda}{\alpha-1} \left(1 - \left(\frac{\lambda}{\lambda+x} \right)^{\alpha-1} \right) & \text{if } \alpha \neq 1; \\ \lambda \ln\left(\frac{\lambda+x}{\lambda}\right) & \text{if } \alpha = 1. \end{cases}$$

$$E(X^2; x) = \begin{cases} 2\lambda \left(x - \lambda \ln \left(\frac{\lambda + x}{\lambda} \right) \right) & \text{if } \alpha = 1; \\ 2\lambda^2 \left(\ln \left(\frac{\lambda + x}{\lambda} \right) - \frac{x}{\lambda + x} \right) & \text{if } \alpha = 2; \\ 2\lambda^2 \left[\frac{1}{\alpha - 2} \left(1 - \left(\frac{\lambda}{\lambda + x} \right)^{\alpha - 2} \right) - \frac{1}{\alpha - 1} \left(1 - \left(\frac{\lambda}{\lambda + x} \right)^{\alpha - 1} \right) \right] & \text{if } \alpha \neq 1, \alpha \neq 2. \end{cases}$$

Generalized Pareto distribution

$$f(x) = \frac{\Gamma(\alpha + k) \lambda^\alpha x^{k-1}}{\Gamma(\alpha) \Gamma(k) (\lambda + x)^{k+\alpha}}, \quad x > 0, \alpha > 0, \lambda > 0, k > 0.$$

$$E(X^h; x) = \begin{cases} \frac{\prod_{i=0}^{h-1} (k+i)}{\prod_{i=1}^h (\alpha-i)} \lambda^h B(k+h, \alpha-h; \frac{x}{\lambda+x}) + x^h \left(1 - B(k, \alpha; \frac{x}{\lambda+x}) \right) & \text{if } \alpha > h; \\ \frac{\Gamma(\alpha+k)}{\Gamma(\alpha) \Gamma(k)} \lambda^h \int_0^{\frac{x}{\lambda+x}} t^{k+h-1} (1-t)^{\alpha-h-1} dt + x^h \left(1 - B(k, \alpha; \frac{x}{\lambda+x}) \right) & \text{if } \alpha > 0. \end{cases}$$

Transformed Gamma distribution

$$f(x) = \frac{\lambda^\tau \alpha^\tau x^{\tau-1} e^{-(\lambda x)^\tau}}{\Gamma(\alpha)}, \quad x > 0, \alpha > 0, \lambda > 0, \tau > 0.$$

$$E(X^h; x) = \frac{\Gamma(\alpha + \frac{h}{\tau})}{\lambda^{\frac{h}{\tau}} \Gamma(\alpha)} \Gamma\left(\alpha + \frac{h}{\tau}; (\lambda x)^\tau\right) + x^h \left[1 - \Gamma\left(\alpha; (\lambda x)^\tau\right) \right].$$

Gamma distribution

$$f(x) = \frac{\lambda (\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad x > 0, \alpha > 0, \lambda > 0.$$

$$E(X^h; x) = \frac{\prod_{i=0}^{h-1} (\alpha+i)}{\lambda^h} \Gamma(\alpha+h; \lambda x) + x^h \left[1 - \Gamma(\alpha; \lambda x) \right].$$

Loggamma distribution

$$f(x) = \frac{\lambda^\alpha (\ln x)^{\alpha-1}}{x^{\lambda+1} \Gamma(\alpha)}, \quad x > 1, \alpha > 0, \lambda > 0.$$

$$E(X^h; x) = \begin{cases} \left(\frac{\lambda}{\lambda-h} \right)^\alpha \Gamma(\alpha; (\lambda-h)\ln x) + x^h [1 - \Gamma(\alpha; \lambda \ln x)] & \text{if } \lambda > h; \\ \frac{e^{(h-\lambda)\ln x} (\lambda \ln x)^\alpha}{\Gamma(\alpha)} \sum_{i=0}^{\infty} \frac{((\lambda-h)\ln x)^i}{\alpha(\alpha+1)\dots(\alpha+i)} + x^h [1 - \Gamma(\alpha; \lambda \ln x)] & \text{if } \lambda > 0. \end{cases}$$

Lognormal distribution

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2}, \quad x > 0, \sigma > 0.$$

$$E(X^h; x) = e^{\frac{h\mu + \frac{1}{2}h^2\sigma^2}{\sigma}} \Phi\left(\frac{\ln x - \mu - h\sigma^2}{\sigma}\right) + x^h \left(1 - \Phi\left(\frac{\ln x - \mu}{\sigma}\right)\right).$$

Weibull distribution

$$f(x) = c\tau x^{\tau-1} e^{-cx^\tau}, \quad x > 0, c > 0, \tau > 0.$$

$$E(X^h; x) = \frac{h \Gamma\left(\frac{h}{\tau}\right)}{\tau c^\tau} \Gamma\left(\frac{h}{\tau}; c x^\tau\right).$$