RISK MEASUREMENT AND MANAGEMENT OF INSURANCE COMPANIES

VASILIOS SIOKIS

THESIS SUBMITTED FOR THE QUALIFICATION OF DEGREE OF DOCTOR OF PHILOSOPHY

CITY UNIVERSITY BUSINESS SCHOOL
DEPARTMENT OF INVESTMENT,
RISK MANAGEMENT & INSURANCE

AUGUST 2001

CONTENTS

List of tables

List of figures

Acknowledgments

Declaration

Abstract

Introduction

Chapter 1: Risk and capital management and insurance

Chapter 2: Rethinking the solvency margin system in Europe

2.1 Introduction

2.2 Reasons behind the two-stage approach

2.3 The current European solvency margin

2.4 Weaknesses of the solvency margin system

2.5 Risk-based capital systems

2.6 Critique of US-Risk based System

2.7 Empirical tests on the efficacy of the US-Risk based system

2.8 The capital base in selected markets

2.9 Stress tests on the capital base

2.10 The trade-off between security and capital costs

2.11 Capital requirements of the rating agencies

2.12 Hybrid capital and the cost of capital

2.13 Supervision of insurance groups and financial conglomerates

2.14 Conclusions

Chapter 3: Underwriting cycles in the UK

3.1 Introduction

3.2 Market imperfections theory

3.3 Rational expectations theory

3.4 Capacity constrained theory

3.5 Methodology and econometric considerations

3.6 Empirical results

3.7 Conclusions

Chapter 4: Cost of capital and credit risk in the insurance industry

4.1 Introduction

4.2 Early credit risk research

4.3. Credit risk research in the insurance area

4.4 Recent developments in credit risk research

4.5 Insurance ratings and the cost of capital

4.6 Methodology and econometric considerations

4.7 Empirical results

Chapter 5: Securitisation and capital in the insurance industry

5.1 Introduction

5.2 The securitisation process

5.3 Literature review

5.4 Securitisation in the insurance industry

5.5 Origins of insurance securitisation and the problem of risk transfer

5.6 Catastrophe risk management: Capital market trends

5.7 Analysis of catastrophe bond structure and performance

5.8 Conclusions

Conclusions and future research

References

LIST OF TABLES

Table 2.1: Main causes of insolvency in the US from 1969 to 1989

Table 2.2: Solvency margins of European insurance companies

Table 2.3: Comparison of average current and target solvency ratios

Table 2.4: Rise in the capital requirements

Table 2.5: Consequences for the solvency ratio of a 35% fall in stock market prices

Table 2.6: Solvency and ROE by legal form

Table 2.7: Amount of hybrid capital in the EU calculation

Table 2.8: Amount of hybrid capital in the US calculation

Table 3.1: Augmented Dickey-Fuller Unit Root Tests for 8 lines of General Insurance

Table 3.2: Cycle periods by line of insurance, 1980 through 1998

Table 3.3: Results of the Generalized Least Squares Regression Analysis

Table 4.1: ZETA ratio statistics

Table 4.2: Summary statistics of yield spreads

Table 4.3: Summary statistics of yield spread changes

Table 4.4: Relation between credit spreads and credit ratings

Table 4.5: Transformation of S&P's codes

Table 4.6: Fixed effects versus random effects model

Table 5.1: PCS Catastrophe indices

Table 5.2: Excess returns for one-year cat notes

LIST OF FIGURES

Figure 2.1: Frequency of insolvency amongst non-life insurers

Figure 2.2: Structure of the solvency margin system

Figure 2.3: Frequency distributions of insolvency ratios in 1997

Figure 2.4: Rise in the average solvency ratio

Figure 2.5: Growing volume of equity investments relative to insurance volume

Figure 2.6: Frequency distribution of the solvency ratio for UK insurers

Figure 2.7: Frequency distribution of ratings in 1998 (Global)

Figure 2.8: Frequency distribution of ratings in 1998 (Europe)

Figure 3.1: Underwriting profits of General Insurance Industry, 1980 through 1988

Figure 4.1: Credit versus market returns

Figure 4.2: Average credit spreads by rating class and maturity bucket

Figure 5.1: Illustrative CLS structure

Figure 5.2: Structure of the Namazu Re transaction

Figure 5.3: General scheme of the coverage layer

DECLARATION

This declaration grants powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

ABSTRACT

This thesis reviews some fundamental risk measurement and management concepts that insurance companies will face in the following years. The first chapter evaluates the theoretical and practical framework of the different approaches with respect to the determination of regulatory capital held by insurance companies. A critical assessment and substantial interpretation of these approaches is performed. Moreover, a number of new approaches is brought forward in order to add a more thorough and clear way of evaluating the level of the regulatory capital.

Then, we provide evidence of the presence of the underwriting cycle in the UK. The underwriting cycle has been identified in a number of OECD and non-OECD countries and highlights the different stages and maturity of the insurance market. A number of reasons for the presence of this cycle is presented and evaluated in contrast with the reasons behind the underwriting cycle in other countries. The level of profitability of the insurance companies is used to determine the presence of the cycle.

In the third chapter, profitability and cost of capital are connected with the credit rating assigned by credit agencies to insurance companies. The credit risk that insurance companies face is explained by the use of financial ratios that explicitly explain the particular credit rating. The credit rating is implicitly connected with the cost of capital, which in turn is explained by the level of the credit spread between the Treasury yield and European bonds.

Finally, securitisation as an alternative method of minimizing credit and market risk is analyzed. Different structures of securitised deals are presented and evaluated. The benefits of securitisation are presented in a systematic way.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Professor Elias Dinenis, for his invaluable help throughout the period of my studies. He is an exceptional teacher and his guidance has been priceless in both my academic and professional days.

I would also like to thank Eleni for being there in both good and bad times. She has been an inspiration and provided continuous encouragement all the way.

I would like to thank all my friends that supplied not only support but valuable academic assistance as well.

Last but not least, I would like to thank my brother and my parents. I cannot put into words the amount of patience and moral and financial aid that have offered and still offering. A big 'thank you' is the only consolation that I can offer to them.

INTRODUCTION

In the face of new century, firms in the global financial marketplace have been scrambling once again, searching for appropriate tools and managerial approaches to guide their organizations. Adequate risk management systems and expertise require substantial firm-level commitment. Risk exposures must be identified, measured and managed. To do so, financial institutions must have the ability to understand global positions and the exposure inherent in them. Moreover, financial institutions need to exploit new ways of gathering and redistributing capital and managing the cost that is associated with it.

The research presented here provides a structure for better understanding the new dimensions of risk measurement and management performed by insurance companies. The ultimate strategic issue in insurance today is how to invest in unique talents and capabilities to generate sustainable attractive returns in a rapidly evolving and therefore challenging world.

The issues that we utilize here cover a wide range of the insurance companies' activities: regulation, risk measurement, and risk management. The regulatory framework that European insurance companies currently operate is being evaluated and compared with the approaches in different parts of the world (US, Canada, and Japan). This framework has been criticized in the past years for a number of reasons. We identify these reasons and provide a new framework that can be employed in order insurance companies to operate in a more efficient way.

Furthermore, the underwriting cycle that pertains insurance companies' activities is identified for the UK market and proof for the existence of this cycle is provided. An extensive number of reasons with respect to the nature of the underwriting cycle is validated and a by-line analysis of the cycle is performed. The macroeconomic activity of underwriting cycles can provide a specific linkage between underlying macroeconomic conditions and the credit risk assumed in different asset portfolios.

These credit risk and credit rating issues are raised in chapter three. The state of the market is one of the major drivers of systematic credit risk, especially as lower credit classes are much more sensitive to macro-economic factors. The cost of capital for insurance companies is explicitly associated with the credit rating that credit agencies assign to insurance companies by employing an analysis of the close relationship between corporate bond credit spreads and credit rating classes. Since most business decisions important enough to occupy the attention of senior decision-makers will involve sunk costs and the danger of loss, credit risk represents one of the major threats for an insurance company's viability.

Finally, the securitisation process employed by insurance companies is investigated as a different approach in order to access the capital markets. Securitisation can provide extra capital for the insurance companies allowing them to assume higher market and credit risk levels. The whole process is analyzed and proof of its efficiency is provided. Capital markets have shown themselves willing to provide capital when the price is right. The breaking down of the barriers between banking and insurance will enable securitisation to flourish. On the other hand, government solutions to the problems of the private market do not really come to grips with the problem of capital adequacy. By and large, these plans are not adequately capitalized to bear the losses that would be created by major catastrophes. With the reforms suggested in our analysis, however, and a consequent increased level of support from capital markets, there seems to be no reason why a purely private insurance/capital industry cannot provide catastrophe insurance products which provide full insurance to all policyholders and a reasonable level of profit for all concerned.

CHAPTER 1

RISK AND CAPITAL MANAGAMENT IN INSURANCE

Over the past several years, there have been rapid advances in the risk measurement and management capabilities of financial institutions. Increasingly sophisticated tools have evolved to measure market risk (value at risk measurement tools), credit risk (expected and unexpected loss measurement tools) and insurance risk (dynamic financial analysis tools). There have been also advances in using these evolving risk metrics to help guide executive management in their strategic decision-making. The framework through which this is accomplished typically has two parts.

Firstly, risk is related to the amount of capital the firm requires to achieve a sufficient level of protection against adverse circumstances. Secondly, risk is used to adjust the returns from business activities to determine whether activities are value-adding or value-destroying. The first part should reflect a debtholder's perspective on risk (i.e., is there sufficient capital to cover "worst case" risks?). The second part should reflect a shareholder's perspective on risk (i.e., are we getting a sufficient return for the systematic risk being taken?).

The purpose of this thesis is to analyze the role of capital in the modern insurance markets. It is crystal-clear that the optimum level of capital held by a financial institution is subjective, hence difficult to replicate and validate. Moreover, the fiduciary nature of insurance companies imposes an extra burden in the calculation of the appropriate level that has to be held. As a result, there is a number of different approaches in the calculation of regulatory capital that essentially try to measure the riskiness of the insurance company. Furthermore, the level of regulatory capital can be employed as an early insolvency flag for the public. The public may be aware of the estimated levels of regulatory and economic capital only to the extent that the firm is able to attain these levels on an ongoing basis, in which case it would be reflected in the publicly reported actual capital level. A close examination of the current approaches is, therefore, of interest in order to establish a clear picture of the ways that insurance companies calculate their regulatory capital and possible implications that might be caused.

The level of regulatory capital is also closely related to the stage of the insurance business and underwriting cycles. Numerous studies have identified the cause and nature of the underwriting cycle across different markets. It is self-evident that regulation will be tightening up in a falling market in order to protect policyholders and shareholders from possible insolvencies. Thus, it is of interest to identify and measure the current state of the insurance underwriting cycle since it is closely related to the level of regulatory capital held by insurance companies. By determining the nature and the causes of the underwriting cycle in the UK, we can provide a solid framework of analysis with respect to the appropriate level of capital that must be held in order for the insurance companies to operate in an efficient way.

Moreover, the first step in attributing capital is developing a theoretical framework for relating risk to the amount of capital the insurance company needs to hold. Many insurance companies have developed a framework based on their financial statements. In our analysis, we employ financial ratios that are observed from the financial statements of insurance companies with the cost of capital related to the credit rating of them. Since the credit rating of an insurance company implicitly represents the cost associated with the capital of this company and the credit risk inherited in its financial statements, we can identify a number of key financial ratios that provide further insight to the cost of capital and credit risk of insurance companies.

Financial regulation has made increasing use of external credit ratings in recent years. One of the key examples of such applications is the package of rules for determining the required capital with respect to market risk issued by the BIS in 1996. Moreover, the new guidelines with respect to the calculation of regulatory capital will force financial institutions to form their own modeling approaches with respect to credit risk and internal credit ratings. A byproduct of this approach will be the increased need for a more thorough and in-depth checking of financial ratios and their relation to the credit rating and the level of capital held by financial institutions. Thus, the insurance regulatory approach will need to determine a new framework for the calculation of regulatory capital by taking into account how these ratios determine the financial health of the insurance company. Statistical scoring credit scoring methods have been shown to perform quite well. In particular, linear discriminant analysis seems robust even when the underlying statistical hypotheses do not exactly hold, especially when used with large samples. Logit analysis has produced similar results. Moreover, measures of credit quality based on equity price data have been used in analyzing credit risk.

If regulators apply a process for certifying rating agencies' ratings for regulatory use that employs similar criteria to what investors use when determining which rating agencies provide them with credible signals about credit quality, there will be an explicit recognition of the credit rating as a measure of the financial health of a company and its relation to the level of capital held. In an international context, ratings-based capital requirements that rely on this sort of synthetic market discipline would be most effective if regulators cooperate closely, so that the certification criteria are harmonized across borders. Absent such coordination, international financial institutions would have an incentive to book a rated asset in the county that certifies the rating agency with the most benign view of the underlying credit risk.

An important issue in analyzing ratings of insurance companies and their impact on the level of capital is how to best incorporate underwriting cycles. Cycles are forces exogenous to the insurance companies and companies cannot control them. The timing, length and severity of underwriting cycles are at best difficult to predict, which complicates further their incorporation into credit ratings. However, given the current state of the market, rating agencies are supposed to separate out equilibrium and cyclical components of companies when assigning a credit rating. Thus, it is self-evident that an analysis of the underwriting cycle will be of help when assigning credit rating to insurance companies.

Finally, innovative insurance companies will engage themselves to a number of financial engineering activities in order to mitigate credit risk and employ their capital in a more efficient way. By doing this, insurance companies can lower the levels of capital that they need to hold as a buffer against "worst-case" scenarios and utilize it in order to get excess returns to their investments. Today, many large insurance companies use modern risk financing methods to reflect a total financial risk management approach whereas traditional insurance/reinsurance places emphasis on risk transfer. However, lack of insurance capacity and price volatility frustrate insurance companies seeking stability in risk financing. Scarcity of capacity may also generate concern about the availability of reinsurance security. The result is a notable increase in self-insurance. Within an alternative risk transfer segment, this capacity can be accessed via securitisation. It makes sense to unite reinsurance and finance skills in order to address these concerns with a number of initiatives that can be used in combination. Traditionally, the need for liquidity in the event of a major claims experience and the need for financing to smooth the impact of reserve strengthening or claims payout over time to mitigate volatility in reported earnings, have been met through reinsurance or finite risk arrangements. It is easy to imagine insurance companies underwriting insurance policies, effectively acting as originators, sourcing and selling those risks to the capital markets much as interest rate and credit risk are traded today. Insurance risk has always been traded, but the trading activity has been confined to the insurance capital market. The use of financial products can expand trading to broader capital markets. These would include structured financings and insurance derivatives to transfer insurance risks to the capital markets.

CHAPTER 2

RETHINKING THE SOLVENCY MARGIN SYSTEM

IN EUROPE

2.1 Introduction

The regulation of insurance companies has evolved over the years both in the United States, the European Union, Canada, and Japan. The current regulatory framework is wide, covering both prudential and market regulation. Prudential regulation seeks to ensure the financial soundness of insurance companies and to protect consumers from insurance company failures.

As insolvencies are a normal side effect of competitive markets, the justification for prudential regulation and regulatory control rests on consumer protection issues (e.g. Berger et al., 1995):

1. To limit the real costs associated with insurance companies' failures, while at the same time permitting insurance companies to carry out their socially critical functions, including the insuring of risky counter-parties;
2. To limit the losses to the government associated with providing a safety net to regulated entities;
3. To prevent resource misallocation that might result from moral hazard (i.e., to eliminate the chance that insurance companies might engage in riskier investments than those chosen in an unregulated system with no safety net); and
4. To promote macro-economic stability, in particular by limiting the chance of a "systemic" event that, for example, might cause a general collapse of confidence in the financial system (i.e., a severe weather event on the part of a number of general insurers, who might withdraw coverage from a particular geographic area or type of risk).

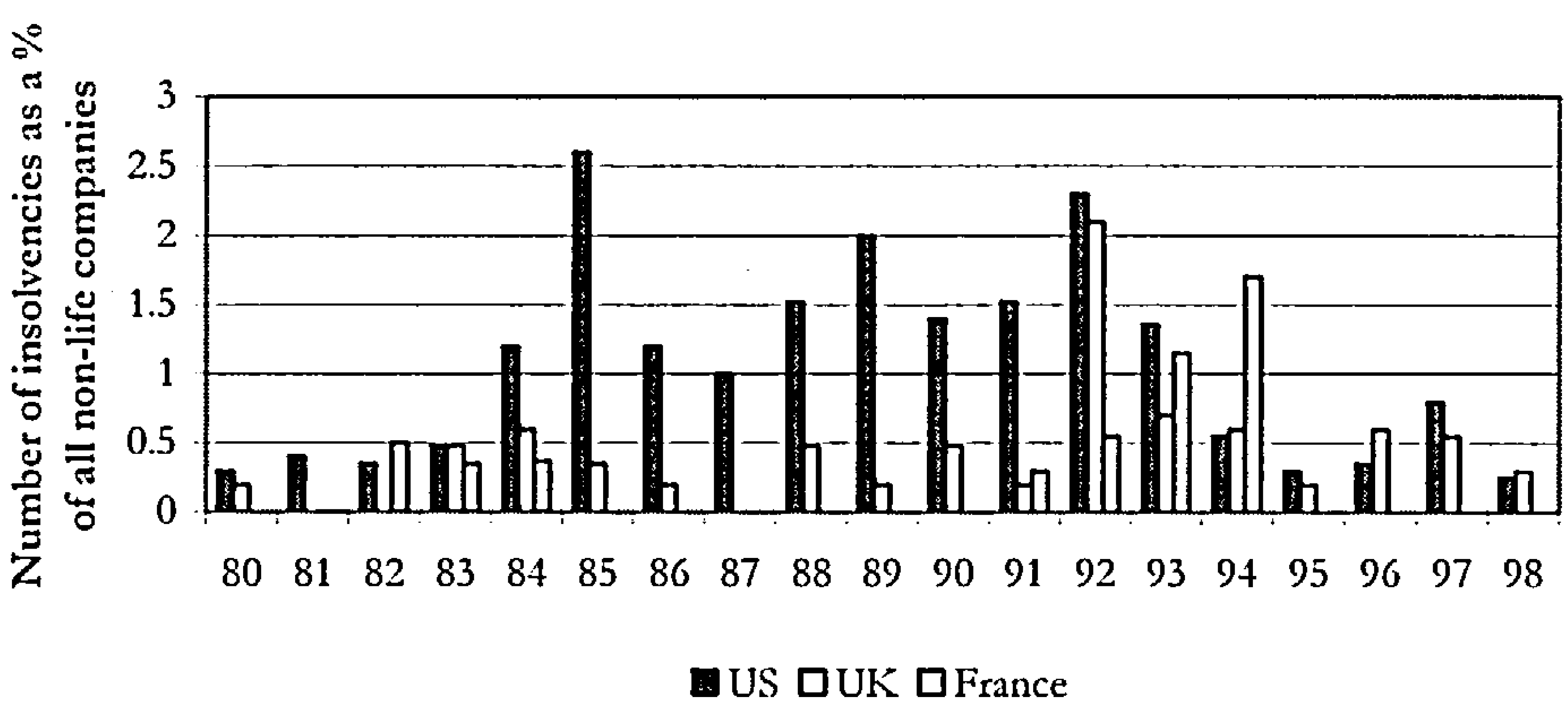
It is for these reasons that prudential regulation is implemented through the requirement that insurance companies maintain a minimum level of capital. The appropriate minimum level of regulatory capital for any financial institution should be consistent with the objectives of regulation. In an unregulated market there is no requirement for a minimum level of capital. The other extreme case is when prudential regulation requires that the probability of insolvency is zero. However, if the maximum insolvency probability is set to zero it is very doubtful that insurance companies could continue to perform their social function of insuring risky counter-parties (Klein, 1995).

If the maximum insolvency probability cannot be set to zero, what should then be the level regulatory capital? Arguably, it should satisfy the following criteria:

1. It should be sufficiently high to provide effective protection and ensure that insurance companies operate at an "investment grade" level of financial soundness;
2. It should not be higher than the "economic" level of capital. Requiring firms to increase their regulatory capital is not costless where it leads to an increase in actual capital. In economic terms, regulatory capital should not be increased beyond the point where the marginal cost of further increases outweighs the marginal benefit from holding capital. The costs of excessive capital requirements may include a perverse incentive effect, since requiring more capital might induce firms to seek higher returns in areas that are high risk or outside their core business.
3. It should be at a level that will generate the appropriate intervention by the authorities.

The US liability crisis of the eighties and Hurricane Andrew in 1992 caused the number of insolvencies in the US to spiral upwards over a ten-year period. In the UK, a string of natural catastrophes at the end of the eighties and a contracting market at the beginning of the nineties caused major losses, resulting in a wave of insolvencies. However, even in years in which market results are satisfactory, the wide spread of results and capital bases can also lead to insolvencies. In contrast to the US and the UK, the German market – which has been heavily regulated since the 1930s – has not suffered a single insolvency. France has experienced insolvencies, particularly at the beginning of the nineties (Figure 2.1).

FIGURE 2.1 Frequency of insolvency amongst non-life insurers



Source: A.M. Best (1991)

Either side of the balance sheet can trigger insolvency:

1. Loss of value of the assets/investments, e.g. due to a stock market crash, interest rate changes or defaults by the issuers of bonds.
2. Underwriting risk (increase in liabilities)
   - Risk of random fluctuation (random increase in claims, although the loss distribution was estimated correctly)
   - Risk of error (calculation of the premium on the basis of an incorrect estimate of loss distribution)
   - Risk of change (the loss distribution changes during the treaty term or run-off period).

Table 2.1 provides data from A.M. Best on the causes of 683 insolvencies in the US between 1969 and 1998. The most common triggers were insufficient premiums or reserves, rapid growth, and catastrophic events. At least 41% of all insolvencies are attributable to underwriting risks.

Regulators have adopted different ways of determining the amount of regulatory capital. The US and Japanese approaches are relatively new measures and they apply various weights to the different aspects of the risks associated with the insurance business. Essentially, they comprise asset risk, credit risk, underwriting loss and loss adjustment expense reserve risk, underwriting premium risk and off balance sheet risk. On the other hand, the European approach is employed within the context of the premium income and the realised losses by the insurance companies.

Table 2.1 Main causes of insolvency in the US from 1969 to 1989

Main causes of insolvency	Insolvencies	In %	
Insufficient reserves/premiums	143	22%	
Too rapid growth	86	13%	Underwriting risks 41%
Catastrophe losses	36	6%	
Overvalued assets	40	6%	
Failure of ceded reinsurance	22	3%	Asset risks 13%
Subsidiaries	26	4%	
Significant change of core business	28	4%	
Fraud	44	7%	
Miscellaneous	44	7%	
Non-identifiable	169	26%	
Total	638	100%	

Source: A.M. Best (1991)

Browne and Hoyt (1995) have also confirmed that underwriting risk is the major factor influencing the frequency of insolvencies. They also find that the number of property-liability insurers is highly and positively correlated with the insolvency rate. Moreover, they identify that the failure rate is higher in the first quarter of the year, suggesting that regulatory oversight with respect to insurer insolvency is not being applied evenly throughout the year.

In a subsequent study on the factors that cause the insolvency of life-health insurance companies, Browne and Hoyt (1999) report that life-health insurer insolvencies are positively related to increases in the average yield on long-term interest rates, personal income per capita, unemployment, the stock market, and to the number of insurers, and negatively related to real estate returns. Whereas in their 1995 study they find evidence of increased failures in the first quarter of the year for property-liability insurers, in the 1999 study they identify a seasonal pattern of insolvencies among life-health insurers, which is highest in the second quarter of the year.

In September 2000, the European Commission put forward a proposal for a directive on changes to the solvency margin system for insurance companies. These changes, which are discussed below, are expected to be approved as a Directive by the European Parliament later this year or early 2001. It will then be incorporated into national insurance legislation within the countries across the EU, and the wider European Economic Area (EEA), in 2001. The need for change was clear when redrafting the Third Insurance Directives in 1992 for life and for non-life insurance (92/49/EEC and 92/96/EEC): the solvency margin requirements for life and non-life insurance companies had not been changed since the 1st Insurance Directives (the Establishment Directives): 1973 for non-life insurance and 1979 for life insurance. Hence, a reappraisal after a 20-year interval was long overdue, in part because of the cumulative effect of inflation, and in part because of new approaches for setting capital adequacy which more accurately reflects the risks facing an insurance enterprise. Some changes were made to solvency margins for credit insurances in the Credit Insurance Directive of 1988, and there were also a few minor indirect changes in the Third Insurance Directives, such as the use of subordinated debt as admissible capital. But there were no fundamental changes. The Third Directives, however, obligated the European Commission under Article 25 of the Third Non-Life Insurance Directive and under Article 26 of the Third

Life Insurance Directive to review and report on possible changes to the solvency margin system. In April 1994, the EU Insurance Committee requested the Conference of the Insurance Supervisory Authorities of the European Union countries to set up a working party to look into how the solvency margin system might be updated. This working party, consisting of representatives of insurance supervisory authorities from across the EEA countries, was formally set up in May 1994, under the chairmanship of Dr. Helmut Muller.

The conference reported back to the Insurance Committee of the European Commission in April 1997, putting forward an analysis of the current system, detailing its strengths and weaknesses, and suggested areas for reform (Conference of the Insurance Supervisory Services (1997)). Reflecting a wider concern with the need to look at capital adequacy across the financial services sector as a whole, the issue of capital standards for insurance companies was endorsed in the Financial Services Action Plan at the Cologne Summit of heads of government in 1996, which committed the European Commission to introduce a directive by mid-2000. Subsequently, the EU Insurance Committee decided to pursue a two-stage approach to the problem. The first stage was to restrict itself to introducing some improvements to the existing solvency margin system; this is the proposal for a Directive announced in September 2000. A second stage of the review was to explore alternatives to the current solvency margin system in the context of the wider solvency assessment of insurance companies. This second stage report is expected to take three to four years to complete.

2.2 Reasons behind the two-stage approach

There were several reasons why a two-stage approach was adopted, apart from the time exigencies to mid-2000 deadline under the Financial Services Action plan. First, capital adequacy standards for insurance companies have been undergoing major change outside of Europe. In the United States, risk-based capital systems had been introduced in 1992 for life insurance and in 1994 for non-life insurance, and similar risk-based capital systems have subsequently been introduced in Canada, Australia and Japan. These risk-based capital systems for insurance have modeled themselves on existing banking standards which were risk-capital based. Hence, one future option for Europe is to consider adopting some form of risk-based capital system. Second, Basel Committee on Banking Supervision at the Bank of International Settlement which had established the widespread use of risk-based capital standards for the banking sector has begun to review its own capital standards, which had been in place since 1988. A new capital adequacy framework was introduced by the Basel Committee in June 1999 (Bank of International Settlements, 1999). It has initiated a discussion process between regulators, banks and other specialists in order to come up with a more robust system which can more accurately reflect the underlying risks faced by banks and to take into account the improvements of new risk management techniques and computer-based modelling approaches. In its terms of reference, it has explicitly recognised the limitations of an over-reliance on the present risk-based capital system.

Third, under the umbrella of the Financial Stability Forum, set up by the G7 Finance Ministers in February 1999, there is renewed pressure to ensure that any capital standards for insurance, banking and securities sectors should have a common underlying structure, even though financial services enterprises face different types of risk and have different demands for capital. While the main reason for the Forum is to underpin the stability of the global financial system by having sound and transparent regulation, there is a subsidiary concern with ensuring level competitive playing fields between financial institutions internationally, with the growing overlapping of financial services. Examples are: the Alternative Risk Transfer (ART) products, which combine investment banking and insurance products; and bank guarantees, credit insurance and credit derivatives have common features. There is also a wish to minimise potential regulatory arbitrage, whereby corporate customers can exploit anomalies in capital requirements between financial services providers. In addition, with the emergence of financial conglomerates, double gearing concerns have been brought more into focus. Double gearing can occur in a financial group, whereby it can count capital twice, both in a holding company and in one of its subsidiaries, thus reducing its overall capital requirements compared to a single financial services provider.

Finally, a capital adequacy system has to be underpinned by consistent accounting and actuarial valuation procedures, since the measurement of capital depends critically on how one measures assets and liabilities, since capital is, by definition, the difference between assets and liabilities (to policyholders and other non-capital providers). The International Accounting Standards Committee (IASC) is currently engaged in a process of trying to establish greater harmonisation in the published accounts of listed financial services firms. There are specific proposals for insurance companies, as there are for banks and securities firms. So far the proposals by the IASC for insurance have not found a general consensus, in particular there are differences on the degree to which fair value accounting should be used, in contrast to the more traditional deferral and matching approach (EU Insurance Committee, 2000). Moreover, there are also differences between the IASC and the FASB, the US accounting standards body, on what an international insurance accounting standard should be.

It was because of the uncertainty created by these various background developments that the decision to move cautiously on a new capital adequacy standards for EEA insurers was seen to be, and indeed is, a sensible way forward.

2.3 The current European solvency margin

The European solvency margin system was set up under the 1st Insurance Directives in the 1970's as a first step in seeking to create a common market in insurance. The move towards a solvency margin system was seen in general as an improvement over deposit-based systems, which still existed in a number of European countries, whereby insurance companies set aside funds in trust and were only recoverable in the event of financial distress. Solvency margins were more flexible and allowed insurance companies to invest their funds more efficiently, as well as not having these funds tied-up. It was clear at the outset that the solvency margin system did not capture all the risks that an insurance company would face. The overriding concern was to have a system, which could be easily implemented and monitored by the regulatory authorities. The fact that there were significant accounting and actuarial differences between countries in how assets, particularly investments, were valued and how liabilities, particularly technical provisions, were valued encouraged the adoption of a simple standard.

The overall structure of solvency margin systems for life and non-life insurance in the EEA are broadly similar. There is an absolute minimum level of capital, the minimum guarantee fund, which is the legal minimum which an insurance company must hold at all times, including at its inception. This minimum guarantee fund for life insurance was set at 0.8 million ecu (now the euro), with some possible reductions for mutuals. For non-life insurance companies there are four categories of minimum guarantee fund, reflecting different types of business and their inherent riskiness. The minimum guarantee funds range from 0.2 million euro to up to 1.4 million euro. Although these are minimum requirements, in practice national supervisors have normally required much higher levels when new companies have been established. When a new insurance company is set up, there needs to be sufficient funds not just to absorb the risks of the business but to also fund anticipated business growth.

The required solvency margin is the higher threshold and hence is more binding. The required solvency margin reflects the scale of business: larger insurance companies requiring more capital than smaller ones. For non-life insurance, the required margin of solvency is expressed as the <u>higher</u> of two results: a) 16% of the gross written premiums up to a level of 10 million ecu (euros) plus 16% of the gross premiums in excess of 10 million ecus, during the last year; b) 26% of the gross average of claims up to 7 million ecu plus 23% of the gross claims in excess of 7 million ecu. The inclusion of the claims basis was in part to loosely capture the riskiness of claims over time and in part to penalise companies with consistently poor underwriting performance. The averaging of claims for most classes of business for the last three years, while for natural hazards insurances, there is an averaging over seven years.

Although the premium and claims systems were both on a gross of reinsurance basis, some reduction for reinsurance was allowed. Non-life insurance companies are able to reduce the required margin of solvency by the ratio of net claims incurred to gross claims incurred, but with a lower limit of 50%. The 50% limit was set to capture some of the credit risks which insurance companies face when buying reinsurance, but also it was to discourage insurance companies from fronting activities.

The solvency margins for life insurance are lower than for non-life insurance, reflecting the fact that life insurance is, in general, a less risky business. For life insurance business which contains investment risk, the minimum required margin of solvency is 4% of the gross technical provisions (i.e. the actuarial or mathematical reserves) plus 0.3% of the capital at risk, where capital at risk is an amount equal to the difference between the maximum payments under the policies underwritten and the actuarial (mathematical) provisions. The 4% figure sought to capture the investment risks and the expense risks faced by a life insurance company, while the capital risk percentage captures mortality risks. For term insurances under five years, there were some reductions to the 0.3%. For life insurance contracts where there is no investment risk carried by the insurance company, which related mainly to equity-linked life contracts, the solvency margin requirement is 1% of the gross technical provisions and the 0.3% capital at risk, again some reduction to 0.3% for term insurances under 5 years.

Hence, there is an explicit recognition within the life insurance solvency margin standards for investment risks and these were captured, if in an approximate way, by the 3% of the gross technical provisions requirement. On the other hand, there is no explicit recognition at all of investment risk in the solvency margins for non-life insurers.

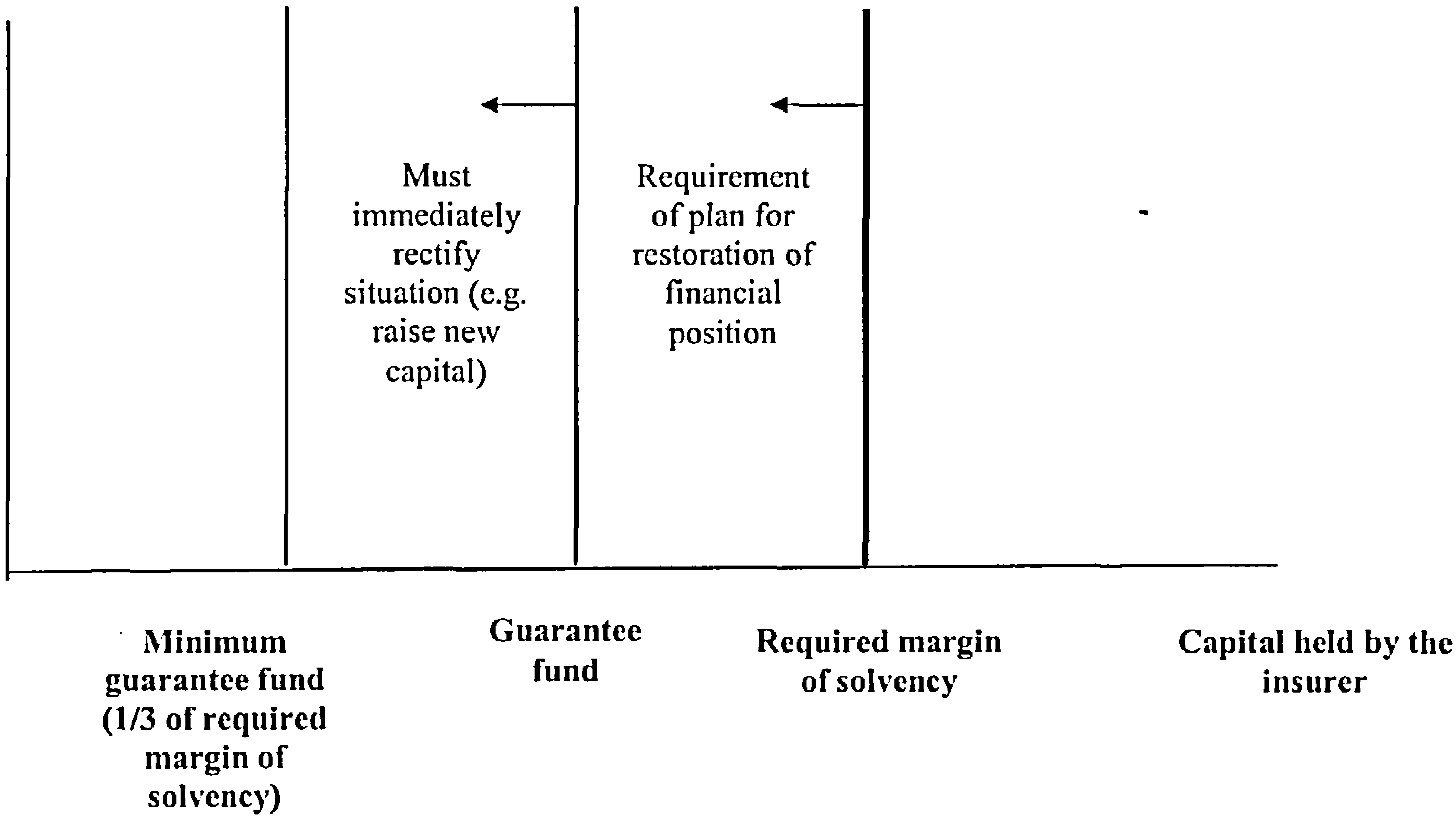
As with non-life insurance, there is scaling down adjustment for life insurance, based on the ratio of net technical provision to gross technical provisions, with a maximum reduction of 85%. The higher value of 85% for life insurance companies reflects their lower need for reinsurance compared to non-life insurance companies, and because the consequences of the fronting of business are much less.

In addition to the required solvency margin, there is a guarantee fund, which is one third of the required solvency margin. The guarantee fund was set up as a specific threshold that would automatically trigger regulators to intervene, if the capital of an insurer fell below this level. Hence, the required solvency margin, the guarantee fund and the minimum guarantee fund are a set of technical thresholds. Regulators within Europe have always had the powers to intervene well before the capital level fell below the required solvency margin, and this has occurred on a number of occasions. There has always been flexibility in European system, since it is well recognised that the solvency of an insurance company depends on a wider set of factors, other than capital adequacy alone.

The second important issue in the European system was that the insurance company must at all times hold capital in excess of the required margins of solvency. Only certain types of capital are admissible in determining the level of capital that is deemed to be held for this purpose. In addition, the valuation of assets and liabilities have to be carried out on a conservative basis. Hence, the 'true and fair' valuation of capital held by insurance companies is well above the admissible level that regulators recognise as capital to meet these regulatory threshold levels. Insurance supervisors take a worst scenario, a liquidation scenario, when measuring the capital available to meet policyholder demands above the levels set aside in the technical provisions, and equalisation provisions. Indeed, the solvency margin is the word used to describe the level of capital held by an insurer when valued under the conservative criteria required by regulation.

Figure 2.2 illustrates how the structure of the European solvency margins operates. The solvency margin, which is the qualifying capital of the insurance company, must be in excess of the required solvency margin, the highest of the three thresholds, at all times. The value of the solvency margin will fluctuate from time to time, for example, as assets values change, due to movements in the stock market prices and interest rates, and as liability values change. In general, the required margin of solvency was set at a relatively low level to provide for flexibility in the light of to changing economic and market forces. But the quid pro quo has always been that supervisors have the discretion to demand that an insurance company rectifies its capital positions before the required solvency margin threshold is reached.

Figure 2.2 Structure of the solvency margin system



Source: Dickinson (1997)

Table 2.2 shows information collected by the Conference of Insurance Supervisory Authorities in their report to the EC Insurance Committee. It details the solvency margins of a large sample of European insurance companies in 1995, and shows the various types of admissible capital that was held.

Table 2.2 Solvency margins of European insurance companies

	Life insurance		**Non-life insurance**	
Solvency margin (breakdown)	Limited companies	Other (mutuals)	Limited companies	Other (mutuals)
Paid-up capital + reserves + profits brought forward	63,543	16,094	50,342	21,267
Profit reserves (life insurance only)	18,563	5,908	None	None

Hidden reserves	19,534	9,049	7,478	4,939
Unpaid capital	347	59	243	15
Future profits (life insurance only)	4,791	1,015	None	None
Supplementary contributions (mutual undertakings)	None	None	None	3,674
Subordinated loans and cumulative preferential shares	801	150	640	2,956
Other	568	20	1,231	193
Total solvency margins	**149,198**	**67,665**	**70,297**	**34,669**
Total required solvency margins	46,246	14,150	22,361	5,737
Number of enterprises	952	198	1,569	630

Source: EU Insurance Committee (2000)

Apart from paid-up capital and reserves and profits brought forward, some of the hidden reserves in the assets and liabilities have been allowed by supervisors. These hidden reserves in assets relate mainly to insurers in a number of continental European countries were current or market values of their investment holdings are not used. Similarly, for life insurance, it can be seen that some of the future profits have also been allowed as part of the solvency margin, reflecting the slow emergence of profits when conservative actuarial valuation bases are required to be used.

2.4 Weaknesses of the solvency margin system

The solvency margin system has a number of weaknesses. The first of these is that there is inadequate recognition of investment risks. As noted above, the required solvency margin for non-life insurance companies demands no extra capital to be held to cover these risks. This assumption has never been accurate, and this weakness has increased over time as non-life insurance companies have held riskier investment, viz. longer dated bonds, loans and equities, in order to increase their overall profitability, in the face of poorer underwriting results. Similarly, for life insurers, the recognition of an investment risk is only an approximate one, with a value of 3% of technical provisions required as capital to reflect these risks, irrespective of the characteristics of the investment portfolio or any significant maturity or interest rate mismatching of assets and liabilities. To some extent this weakness in the solvency margin was addressed by a reconsideration of the investment regulations in the 3rd Directives, but this is only a crude recognition. Similarly, the currency matching requirements of assets and liabilities also capture some aspects of the investment risks. However, these are clear weaknesses in the current system. Under risk-based capital systems in the United States and elsewhere, and indeed in the banking sector, asset risk is a key component of capital adequacy requirements.

Another weakness of the solvency margin system is its inadequate treatment of reinsurance. It is clear that ceding insurers carry credit risk and, if this credit risk is significant, there should be extra capital to absorb this. The 50% reinsurance constraint for non-life insurers and the 85% constraint for life insurers are inefficient risk controls. The current system tends to penalise smaller insurance companies rather than capture the credit risks attendant on buying reinsurance. In addition, the quality of the reinsurance supplier is not taken into account in the capital requirements, although it must be said that the adequacy of reinsurance suppliers is separately checked by the supervisory authorities, as part of the wider solvency assessment process. Similarly, non-proportional reinsurances are penalised compared to proportional reinsurances, since there is a greater risk transfer per unit of premium paid under non-proportional reinsurance contracts.

In addition, the risks associated with underwriting in non-life insurance are only approximately recognised by the claims criteria. Riskier lines of insurance, including long term liability insurances, are inadequately captured in the capital requirements. Paradoxically, insurers that set aside stronger technical provisions are penalised, since this increases the values for claims incurred more than they would be if reserving had been less conservative, thus penalising the more prudent insurers. Even so, regulatory authorities separately look at the determination of technical provisions to ensure their adequacy. For life insurance, another weakness in the solvency margin system is a failure to capture any mismatch in the duration and interest rate risks in the assets and liabilities. Similarly, a rapid growth of business has no direct impact in the solvency margin calculation, since the solvency margin requirement is retrospective in nature; evidence shows that rapid growth can overstretch resources of an insurer and a poorer quality of business is likely to be underwritten.

The above outlines the main weaknesses of the current solvency margin system. There is much less fine tuning than under a risk-based capital system. However, the European solvency margin system is simple and flexible, and is not costly to administer. This is its main strength. It also presupposes that other aspects of solvency assessment are important other than capital.

2.5 Risk-Based Capital Systems

Risk-based capital approach will be one alternative to the solvency margin system, in the deliberations of the second stage of the EU solvency review. Hence it would seem appropriate to evaluate the US experience in using a risk-based capital system, since it has been now been in operation for several years. It is worth observing that the motivation for the introduction of a risk-based capital system in the United States arose from pressure in the federal government to have a more formalised system of capital adequacy, following the failures of one or two major US insurance companies.

Previously, state insurance supervisors (commissioners), and the National Association of Insurance Commissioners (NAIC), already had for many years informal guidelines relating to capital adequacy. They clearly realised the importance of capital but were aware that there are a variety of other factors that can contribute to the insolvency of an insurance company, apart from capital.

The NAIC had in place an early warning system (IRIS) based on a multi-factor model to determine potential insurance insolvency, with capital being one of the factors. But it was political pressure from the federal government that led to working parties being set up (under the auspices of the NAIC) to propose with a more formal system, even though regulation of insurance was the responsibility of state governments.

The key feature of a risk-based capital system is that the main types of risk that a financial services company faces are separately identified and measured and a specified minimum amount of capital is required to cover these risks. Providing the individual risk exposures are measured accurately, and the capital requirements for each risk class are allocated appropriately, a risk-based system tailors the capital requirement of a company to the risks that it takes. These minimum capital levels are then aggregated across the company as a whole to obtain the overall minimum level of capital to be held. However, because all risks will not occur at the same time (i.e. the principle of the law of large numbers or portfolio diversification), some method of scaling down of the summation of these individual capital requirements is necessary. This scaling-down process is critical and depends on the knowledge of how all the individual risks are correlated between themselves.

Risk-based capital is used to determine the acceptable minimum level of capital that an insurance company must hold as part of its solvency assessment. This minimum standard should safeguard insurers from financial crises, allow them to fulfil their obligations at all times and in the long term, and above all prevent them from ever going into liquidation (Cummins et al., 1993). Risk capital charges (i.e. the required margins for adverse deviation) are calculated for the different aspects of risks faced by the insurance company. It is not the purpose here to go into detail on the various risk factors which make up the risk-based capital system in the United States. The main risk factors for non-life insurance are: R_0, asset risk (guarantees and contingent liabilities from affiliates); R_1, asset risk (fixed income bonds and short-term investments); R_2, asset risk (stocks, real estate and participations); R_3, credit risk (50% of RBC of ceded reinsurance and other receivables); R_4, loss reserves risk; and R_5, written premium risk (NAIC, 1993). Off-balance sheet risks are split into various groups. Non-controlled assets, guarantees for affiliates and contingent liabilities are included in R_0. Risks from strong company growth are divided into requirements relating to loss reserve growth in R_4 and premium growth in R_5.

These charges are then summed to give the total RBC (before covariance adjustment). The total RBC charge after the covariance adjustment is equal to the result of the following formula:

$$R_0 + \sqrt{R_1^2 + R_2^2 + R_3^2 + R_4^2 + R_5^2}$$

where R_0 to R_5 are RBC charges. The covariance adjustment is then simply the difference between the straight sum of R_0 to R_5 and the result of the formula above.

The reason for incorporating a covariance adjustment in the RBC formula is that the total RBC for an insurer should generally be less than the simple sum of the RBC amounts for each risk element, because of the benefits of diversification and because the separate risk elements are not perfectly correlated (Hooker et al, 1995).

Having a formula in place to compute uniform minimum capital standards for insurers is only part of the risk-based capital system. Regulators also need some legal basis to act. In addition to defining a minimum threshold for each insurance company, the US system has introduced a detailed set of regulatory responses that are triggered if the capital base of an insurance company (measured by conservative valuation criteria under state regulation) falls below these thresholds.

US regulatory action levels

(percentage of the total capital / RBC ratio)

Below 70%	Regulators *must* seize the company. This is called the *Mandatory Control Level.*
70-100%	Regulators *may* seize the company. This is called the *Authorised Control Level.*
100-200%	The company must submit a plan of action to regulators. The *Regulatory Action Level* is set at 150%, below which the regulators will perform an examination of the company and issue a corrective order. Above 150%, but below 200%, is known as the *Company Action Level,* where the company's actions alone are deemed to be sufficient, without the need for a regulator's corrective order.
250%+	Typical value for the vast majority of companies.

Source: Hooker et al. (1995)

In the United States, a risk-based capital system for life insurance and health insurance was introduced relatively quickly, since there were no major disagreements within the insurance industry about the determination of the risk groupings and the details of implementation. The fact that sufficient flexibility existed in the actuarial valuation basis for solvency assessment was a key factor in achieving this consensus between the regulatory authorities and the insurance industry. However, problems arose in respect of non-life business, not least in determining the risk factors for loss reserves (outstanding claims provisions) and the underwriting risks attached to new business: the liabilities side of the balance sheet. In addition, the correlations between assets and liabilities were also more complex. For example, the system would have to cope with determining the correlation between capital market movements and the underwriting cycle if it were to provide a credible measure of capital adequacy.

For most US non-life insurers almost two-thirds of their aggregate risk, and hence their minimum risk-based capital requirements, are considered to arise from their current and past underwriting operations (underwriting risks on new business and risks associated with claims provisions and their associated settlement expenses); investment risks are significant but less important. In contrast, most of the risks for a typical life insurer, and hence its capital requirements, are related to risks from its investments and to interest rate risks from the maturity mismatching of assets and liabilities.

Japan introduced a risk-based capital system for non-life insurers in 1997. This system is similar to the US system and is designed to act as an early warning system for the regulatory authorities. The solvency margin is calculated as a ratio of the admitted capital base and the total risk. The adjusted capital is compared to the insurer's total risk. The total risk is calculated using the following formula:

$$\text{Total risk} = \frac{1}{2}\left(\sqrt{R_a^2 + (R_b + R_c)^2} + R_d\right) + R_e$$

where R_a is the underwriting risk, R_b is the interest rate risk, R_c is the asset-management risk (price/downside risk, credit risk, derivatives), R_d is the management risk and R_e is the natural catastrophe risk.

Similar to EU Directive, underwriting risk is determined using the premium or loss index, whichever is higher. The risk factors vary for six different lines of business. Natural catastrophe risk is defined using the actual catastrophe exposure. As in the Us system, there are different threshold levels giving the regulatory body the power to intervene.

2.6 Critique of US Risk-Based System

There are a number of criticisms that can be levelled at the US risk-based capital system. Several types of risk are not incorporated into the US system, for example:

- there could be a substantial adverse movement in the market value of the company's investments, as opposed to normal investment volatility;
- the nature, currency and term of the company's assets could be inappropriate, given the nature, currency and term of the liabilities;
- the company could be exposed to an accumulation of risk, either from a natural disaster or from a combination of economic conditions; and
- there could be a significant loss due to exposure to investment trading, e.g. in derivative instruments, for speculation rather than hedging.

In addition, the cross balance sheet correlations between assets and liabilities are not captured in non-life insurance. One clear omission is the failure to explicitly recognize any potential for currency mismatching. Moreover, the calculation of a number of the risk factors appears to be arbitrary (Dickinson, 1997). For example:

- the charge against reinsurance ceded in the credit risk factor; and
- the allowance for excessive growth in the off balance sheet risk factor.

The risk framework is also essentially retrospective in nature, being based on statistical data at both the industry and company level over a preceding time period. For example, the factors for the 1994 year-end reflect the historical experience of the U.S. insurance industry in the period 1984-93, including the severe adverse development that occurred in general liability and the very severe loss ratios in medical malpractice and reinsurance. Another limitation is that the risk-based capital system is based on a "snap-shot" of a company at a given point in time. It is a static rather than dynamic approach to solvency testing (Cummins et al., 1993). A modern approach to solvency testing is dynamic financial analysis, usually implemented by cash flow simulation (Cummins et al., 1999). A cash flow model can take into account patterns of loss reserve runoffs and asset cash flows and can incorporate external economic information such as yield curves and inflation rates. Thus, it can provide information on a company's ability to withstand potentially adverse economic developments that cannot be captured by a static system.

Criticism of the static nature of the current solvency model and its focus on the past as evidenced by its use of annual report data has led to a discussion of future oriented models. The cash flow models that are often used in this context are based on the principle that the economic value of a company is determined by the discounted value of all future cash flows. Cash flows are forecast over a certain planning horizon and the cash values added. The cash flow forecast is heavily dependent on the assumptions made regarding the future development of the business.

Given the interaction of risks on the asset and liability sides, this model is also a regular topic of discussion in the context of asset-liability management. For regulatory purposes the focus is on honouring existing obligations, however. Cash flow models that are drawn up at the request of regulators tend to simulate the liquidation of the insurer (run-off) and not a continuation in the underwriting of new risks, as in the case of asset-liability management.

Cash flow or dynamic financial analysis (DFA) models are currently being used by life insurers in some of the states of the US and, since 1999, by Canadian non-life insurers as well. The Canadian DCAT model (dynamic capital fund adequacy testing), which is based on the "going-concern" principle, carries out stress tests on non-life insurers regarding their capital base.

These tests apply to:

- loss frequency and severity
- pricing
- under-reserving
- inflation
- change in interest rates
- premium volume
- expense increase
- reinsurance failure
- depreciation of investments
- government and political action
- off-balance sheet risks

In addition, to plausible base scenarios, the above-mentioned risks are to be modeled and the three scenarios involving the greatest risk published. The additional use of cash flow models for non-life insurers is currently being discussed at an abstract level in the US, although there are currently no concrete plans for their introduction.

The experience during the development of the risk-based capital formula was also dominated by several factors:

- The 1984-86 US liability 'crisis', particularly in relation to medical malpractice;
- The emergence of long-tail asbestos and environmental pollution claims;

- An apparent increase in the frequency of natural catastrophes;
- High interest rates, creating pressures to engage in cash flow underwriting; and
- High inflation rates.

Management competence is also an important concept that is missing altogether as a risk factor, although it would clearly have been difficult to measure it. The history of individual managers may be relevant, but it will be almost impossible to score this on a numerical scale.

An inaccurate risk-based capital formula will have quite obvious results both in terms of the behavior of insurance companies and the costs that are associated with it. There will be distortion of investment, underwriting, and reinsurance decisions of well-managed insurers, leading to less effective diversification, reduction of safety levels for financially sound insurers and higher premium rates for any given level of safety. A poorly designed system could also lead to unjustified damage to the reputations of well-managed insurers, raising the costs of capital for these firms and impending their ability to raise new equity capital. The result could be a reduction in the efficiency of insurance markets and an inefficient shift by buyers towards self-insurance and other risk management alternatives (Cummins et al., 1995). Furthermore, there is an impact that regulatory capital has on both insurance companies and policyholders, and it is associated with the way economic resources will be absorbed in order to restore a certain level of capital. The value of the extra resources that would be absorbed by the regulatory regime in this respect reflects a direct cost imposed on both the regulator and the policyholders. Also, there are costs that would not have been incurred in the absence of risk-based capital, for example the costs of any additional systems, training, and management time that is required by the regulator. Finally, there are indirect costs that are least obvious from a cash perspective. These include the costs of reduced competition, the costs of imposed uniformity and the costs of moral hazard (Richardson and Stephenson, 2000). Cost-effective capital standards will be met only when their economic benefits exceed its economic cost

The trigger points for regulatory intervention in the US system are also quite rigid, in that they give a relatively small amount of discretion to the regulator. It can be argued that policyholders are likely to be disadvantaged if insurance companies are forced to hold excessive levels of capital, since there is a cost in holding capital for an insurer, and in competitive markets costs tend to get passed onto policyholders over time.

2.7 Empirical Tests on the Efficacy of the US Risk-based system

Since its introduction, there have been a number of empirical studies in the United States which have sought to test the efficacy of the risk-based system in identifying and predicting the failure of insurers. Grace, Harrington, and Klein (1993) calculated the ratio of capital held to the risk-based minimum capital in 1990 and 1991 for those US non-life insurers that failed between 1991 and 1993. They found that, while there was a significant inverse relationship between the ratio of actual capital to their risk-based capital minimum (the RBC ratio) and their insolvency risk, in both univariate tests and multiple logistic regressions, relatively few failed companies had RBC ratios that would have triggered regulatory action prior to their failure.

A subsequent study by Cummins, Harrington, and Klein (1995), employing a similar data set and a multiple logistic regression model of insolvency risk, determined that the accuracy of the RBC ratios in classifying failed and surviving insurers could be materially improved by adjusting the weights of the basic RBC components, and by including firm size and organization form as additional variables in the risk-based system. They also found that the NAIC risk-based capital formula classifies small firms more accurately than large firms.

Finally, Cummins, Grace, and Phillips (1999) compared the risk-based capital system with other models for predicting insurer failure. These models were the Financial Analysis and Surveillance Tracking (FAST) audit ratio system and a cash flow simulation model. The FAST system and the older Insurance Regulatory Information System (IRIS) were designed to prioritise insurers for further regulatory action. The IRIS system consists of twelve audit ratios with published ranges that are deemed acceptable by the regulators. The FAST system consists of twenty-nine ratios and corresponding scores for each ratio (NAIC, 1995). The ultimate output from the FAST system was that the overall FAST score equal to the sum of the individual insurer's audit ratios multiplied by the corresponding scores.

Logistic regression was used to test the models for a large sample of solvent and insolvent property-liability insurers, using data from the years 1990-1992 to predict insolvencies over three-year prediction horizons. They found that the FAST system was superior to RBC as a static method for predicting insurer insolvencies. Furthermore, they found that the cash flow simulation variables add significant explanatory power to the regressions and lead to more accurate solvency prediction than the ratio-based models taken alone.

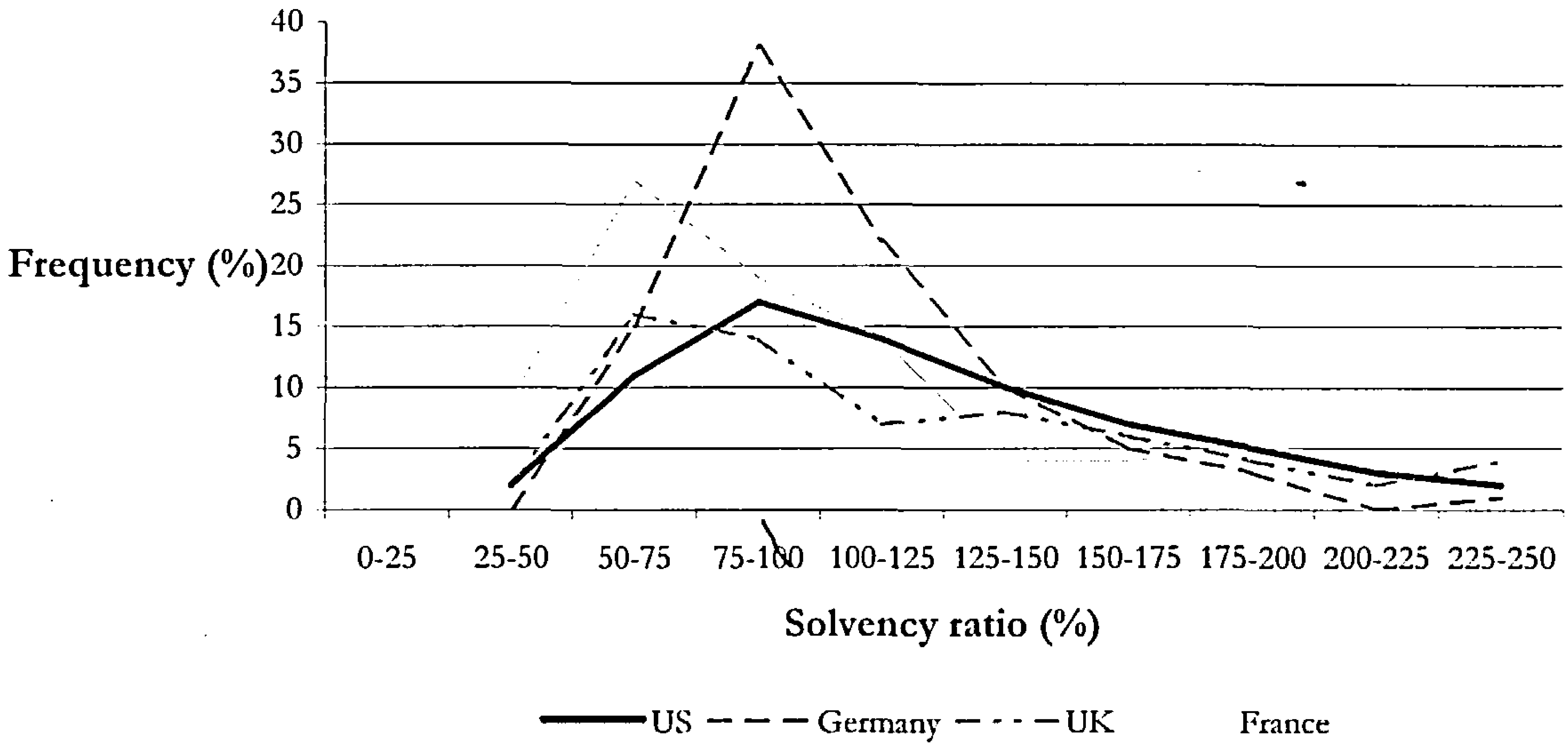
The general conclusion of these studies is that capital levels of insurers, even when measured under a complex risk-based system, are not sufficient alone to predict or prevent the failure of insurance companies.

A.M. Best, the rating agency, undertook a study of the causes of all non-life insurance company failures over the period 1969 to 1998 (A.M. Best, 1991). The main causes are listed in Table 1.1. As can be seen there were a wide variety of underlying causes. These findings do not weaken the case for a reliance on capital standards. Capital still acts as a line of last defense to absorb all risks, and even if an adequate level of capital does not prevent failure of an insurer, it can reduce its likelihood, or delay its occurrence so that supervisors have time to find an orderly wind-up.

2.8 The capital base in selected markets

The average values and distribution curves of the solvency ratio (capital funds expressed as a percentage of net premiums) show that insurers in the US, UK, Germany and France currently hold far more capital funds than required by the regulators (see Figure 2.3). A comparison can be drawn between the solvency ratios for the US and the UK as both countries use similar accounting principles when calculating capital funds and consequently have similar average solvency ratios. A comparison of the frequency distribution shows, however, that US insurers are more heavily capitalised than their European counterparts. This can be traced back to the large number of small but well capitalised US insurers which even together represent only a small market share.

FIGURE 2.3 Frequency distribution of solvency ratios in 1997



Source: A.M. Best (1999)

Table 2.3 shows the median ratio of the effective solvency margin to the target for the US and UK. It reveals that the average insurer holds around four times more capital funds than the amount required by the regulators. This shows clearly that the minimum statutory standards in force constitute no effective restriction to the activities of the average insurer.

Table 2.3 Comparison of average current and target solvency ratios

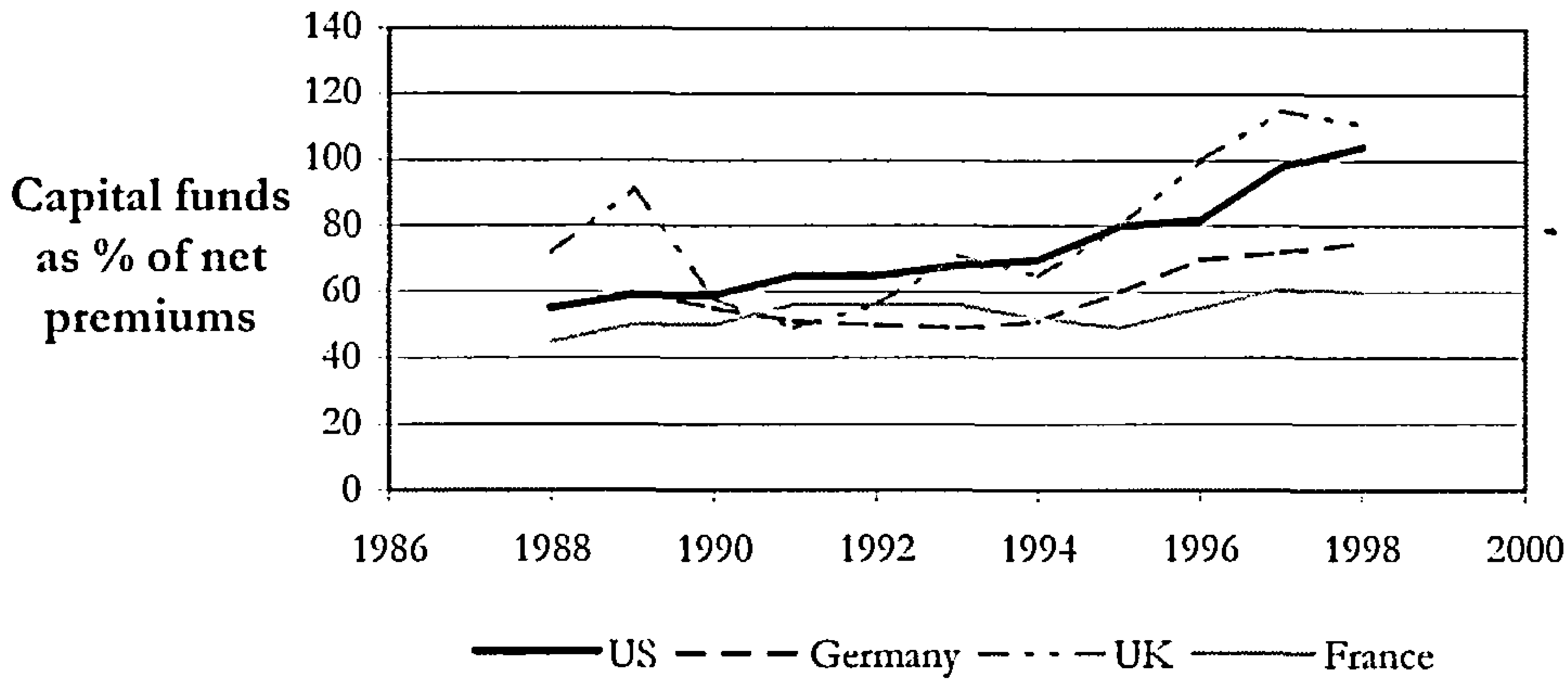
	1990	1994	1997
US	N.A.	345%	389%
UK	400%	278%	469%

Source: A.M. Best (1999)

From the beginning of the nineties onwards the solvency ratio increased dramatically in all the markets. In the UK and US, where capital funds are calculated largely at market value, the increase was the most striking. In the 1990-1997 period it soared by 168% and 199% respectively (see Figure 2.4).

The rallying capital markets in combination with relatively good insurance results boosted market capitalisation in relation to premium volume. However, only part of the rise in the solvency ratio translates into improved risk security, since premium volume alone does not give a sufficiently accurate picture of an insurer's overall risk. Falling premium rates reduce premium income in relation to risk exposure. Furthermore, the trend towards higher cover limits and rising retentions by commercial policyholders increases the volatility of the portfolio. A further rise in risk exposure is the outcome of the growing significance of asset risk. In summary, all these effects serve to increase the capital funds requirement in relation to premiums. In the US this can be seen clearly in the rise in the amount of the risk-based capital required per dollar of premiums earned (Table 2.4).

FIGURE 2.4 Rise in the average solvency ratio



Source: A.M. Best (1999)

Table 2.4 Rise in the capital requirements

	1994	1995	1996	1997	1998
Solvency ratio	72%	76%	82%	92%	103%
RBC/net premiums (average)	21%	26%	26%	28%	29%
Proportion of shares in total assets invested	17%	18%	19%	21%	23%
Shares as a percentage of net premiums	40%	46%	46%	57%	62%

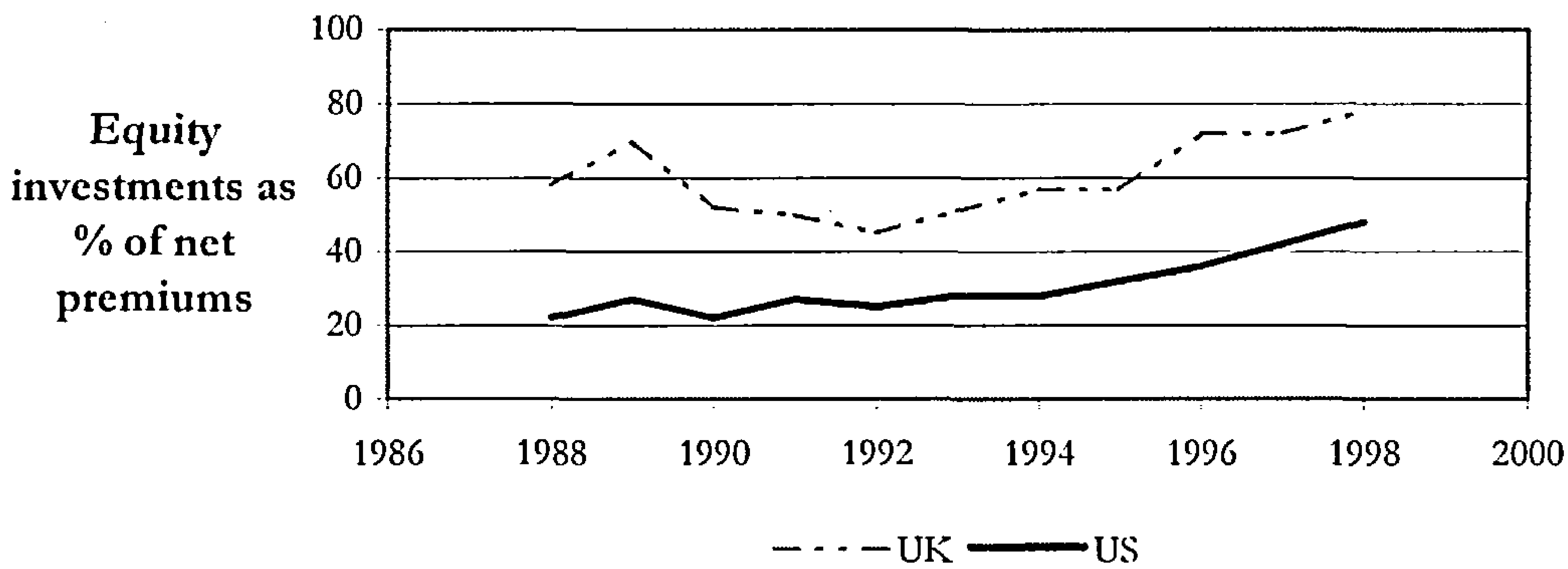
Source: A.M. Best (1999)

Insurers' capital bases have increased at a faster pace than risk exposure, however, with the result that substantially more capital funds are available to cover technical and investment risks. Consequently, the European and North American insurance markets find themselves with excess capacity today.

2.9 Stress tests on the capital base

The main reasons behind the accumulation of capital funds were the rallies on the US and European equity markets. On the other hand, this has also significantly increased investment risk over the past few years. Figure 2.5 shows the rise in the value of equity investments expressed as a percentage of net premiums in the UK and US. Today insurers are exposed to investment risk on equity investments amounting to more than 55% of their net premium volume.

FIGURE 2.5 Growing volume of equity investments relative to insurance volume



Sources: A.M. Best (1999)

An interesting question arises here on how would the capitalisation of the various markets will react to a stock market crash. Table 2.5 shows the simulated effects on the solvency ratio of a 35% downward correction of equity markets.

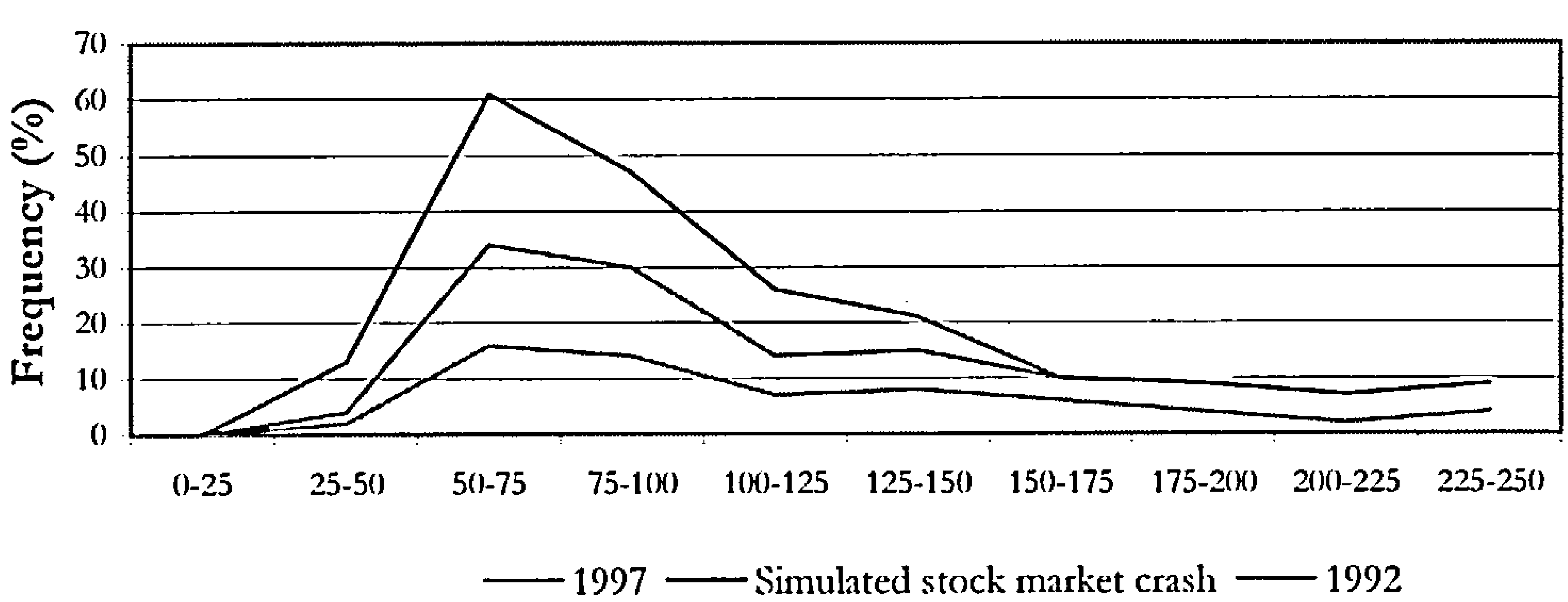
Table 2.5 Consequences for the solvency ratio of a 35% fall in stock market prices

Average market data	Germany	UK	US
Equities/Capital funds 1998	59.2%	68.3%	46.2%
Solvency ratio 1998	159.8%	114.7%	103.3%
Change in solvency ratio (% points)	-31.2%	-25.9%	-15.7%

Source: A.M. Best (1999)

Capital funds react immediately to a change in the market value of assets without losses having to be realised in the profit and loss account. A 35% market correction would substantially reduce capital funds (at market value) and in turn the solvency ratio. Average solvency, however, would not fall to a critical level in any of the markets examined – although this would not necessarily be the case for every individual company. To take UK as an example, the share of insurers failing to meet solvency

FIGURE 1.6 Frequency distribution of the solvency ratio for UK insurers



requirements would increase from 3.6% to 4.7%. Figure 2.6 illustrates the impact of market crash simulation on the frequency distribution of the solvency ratio for the UK. The curve shifts to the left whenever capital funds are reduced, i.e. the solvency ratio of most companies and therefore of the market average is reduced; however, it would still be higher than five years ago.

The consequences of a large loss event are more difficult to simulate. The influence loss events can have on the solvency margin is in the form of a negative annual result, although the company has a range of financial instruments at its disposal that can be used to try to bring losses under control. While reinsurance covers some of the gross losses, a company can avoid a balance sheet loss and thus a reduction in capital funds by releasing reserves and selling some investments. This was the reason why, despite Hurricane Andrew in 1992 ($17 billion in gross losses), capital funds in the US market grew by 2.6%. Extreme loss scenarios involving hurricanes in Florida or earthquakes in California of a magnitude that takes place only once every hundred years are based on insured gross losses of $60-$70 billions. After reinsurance net losses are likely to still be between $40-$45 billions for direct insurers. However, with the introduction of new financing techniques, like catastrophe bonds and catastrophe options, the insurers are able to mitigate their exposures by reinforcing their capital base with these new tools.

The high average capitalisation of the markets examined does not mean that catastrophes cannot result in individual insolvencies: in 1992 Hurricane Andrew triggered the greatest number of insolvencies ever in the US (63). In weakly capitalised markets natural catastrophes can have far-reaching consequences for the capital base.

2.10 The trade-off between security and capital costs

Generally insurers hold substantially more capital than the amount required by regulators. The main advantage of this buffer is that policyholders can feel secure in the knowledge that their claims will be paid and shareholders can be comfortable that the ability of the company to continue making profits is protected. However, holding capital funds is costly. Incurring unnecessary capital costs lowers shareholder investment returns and raises policyholder premium rate.

A number of stakeholders have perhaps diverging interests regarding the equity base required of an insurer. Policyholders benefit from the knowledge that the insurer can meet claims-paying commitments – yet policyholders do not want capital requirements to become so burdensome that premium rates become excessive. Regulators, on the other hand, aim at protecting the consumer while maintaining the long-run viability of insurance markets. A company's owners, in contrast, are interested in generating a high risk-adjusted return on their investments and so must make a trade-off between protecting the franchise value of their company on the one hand and incurring capital costs on the other. Both staff and management have a vested interest in keeping their company in business and in having leeway for action, while also keeping shareholders happy. Rating agencies, too, are interested in the fulfilment of all obligations, which includes all the claims of investors. Each of these stakeholders has a different view regarding the trade-offs involved in holding capital, and conflicting views about the optimal amount of capital funds can result.

Only a relatively small number of physical assets, such as an office building and computer hardware, are needed for a company to offer insurance protection. Risk capital is not tied to normal business activities and can thus be invested profitably. The net costs of reserves are thus the costs of capital funds minus the investment returns. From the investors' standpoint, the fact that an insurer has capital funds available which it can reinvest in the capital market gives it the traits of an investment fund. The insurer's indirect investment risk in the capital market is leveraged by the underwriting risk. However, an insurance company's investment of capital involves substantial tax disadvantages and agency costs when compared to a direct investment by an investor.

An insurer's cost of capital depends to a certain extent on the company's legal form. Mutual insurance companies generally hold more capital in relation to business volume than stock companies (Table 2.6). In the case of mutual insurance companies the owners and policyholders are one and the same. They are therefore under less capital cost pressure than stock companies, which constantly have to balance the interests of the policyholders and shareholders.

Table 2.6 Solvency and ROE by legal form

	Mutual companies		**Stock companies**	
	ROE	Solvency ratio	ROE	Solvency ratio
France	5.9%	123%	9.1%	72%
Germany	8.6%	64%	5.4%	59%
US	8.7%	94%	14.2%	82%

Source: A.M. Best (1999)

Table 2.6 shows the average solvency ratio of the markets for the 1995-1997 period in which mutual insurance companies play the most important role. A wider survey of 2424 European non-life insurers confirm these results and show that the average current/target solvency of mutual insurers is 5.4:1, in contrast to 3.7:1 for public limited companies (European Commission, 1999).

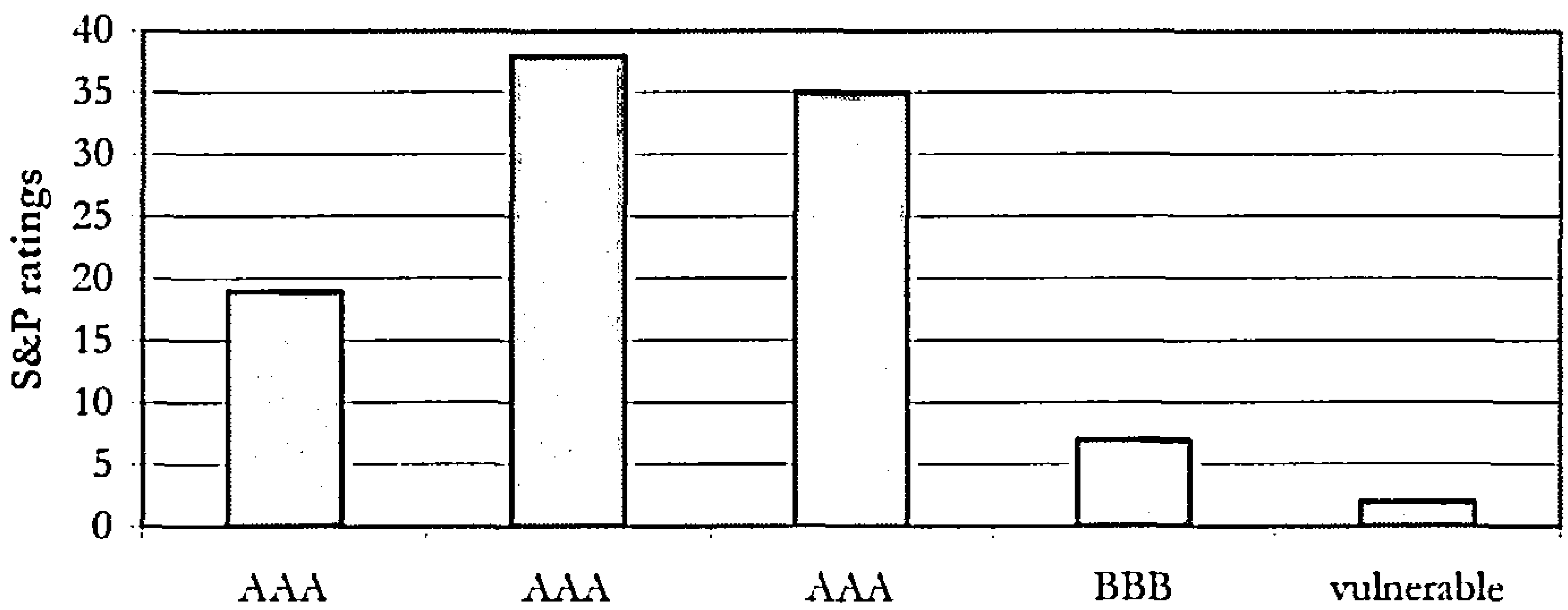
On the basis of the above-mentioned effect, the direct consequence of a wider capital base is a lower return on equity. This is confirmed for the US and France in a comparison of the legal forms of companies shown in Table 2.6. The conflicting aims of solvency and return on equity have not been so obvious over the past few years as insurers have been generating above-average returns on the back of the rallying stock markets.

2.11 Capital requirements of the rating agencies

The capital requirements imposed by the rating agencies are becoming increasingly important to the success of insurance companies. In a market characterised by growing international competition, a top rating is a must. The capital requirements needed in order to be assigned a top rating are normally higher than those laid down by state regulators. Figures 2.7 and 2.8 show the distribution of ratings assigned by Standard & Poor's (global) and A.M. Best (US) for 1998. Despite their basically generous capital base, most companies are assigned ratings in the middle categories. In this way the rating agencies make an important contribution to the establishment of a high standard by which to measure a company's capital base.

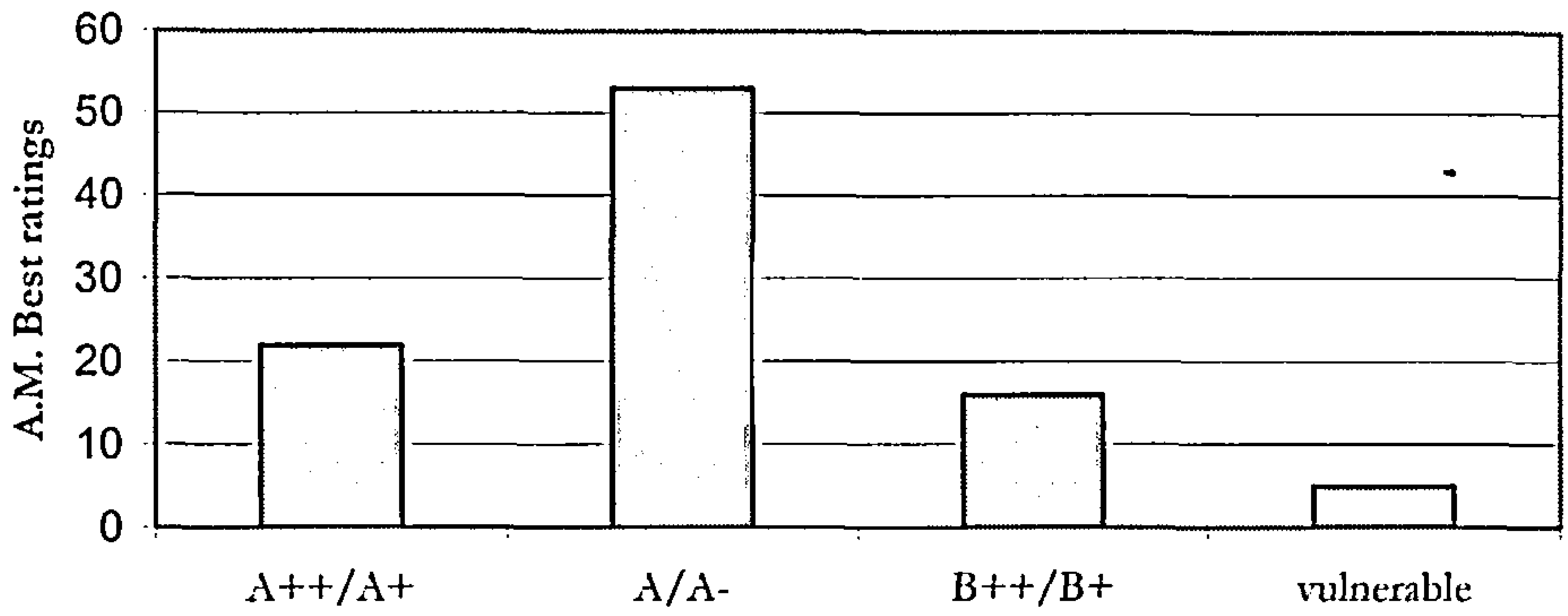
Rating agencies aim at encouraging security in the fulfilment of payment obligations to policyholders and investors. In addition to solvency, criteria such as liquidity, financial

FIGURE 2.7 Frequency distribution of ratings in 1998 (Global)



strength and management quality are factored into complex models. Given the different models, it is difficult to find an empirical agreement between the solvency ratio or risk-based capital ratio and the ratings. An analysis of 878 US companies revealed only a very minor correlation between the risk-based capital ratio and the ratings assigned by A.M. Best, which are based on publicly accessible information.

FIGURE 2.8 Frequency distribution of ratings in 1998 (Europe)



Over the past few years the rating agencies have contributed to the increase in the solvency standard. In their models the rating agencies use among other factors the relative assessments to the market average. Since the market average has risen, so too have the demands for each rating category. However, critics of the ratings base their arguments on the inherent conflicts of interests between the rating agencies and the insurers who sometimes request the rating to be drawn up. On the other hand, various surveys have shown that even private ratings do not sufficiently forecast insolvencies (Cummins et al., 1999).

2.12 Hybrid capital and the cost of capital

Hybrid capital ranks somewhere between equity and debt. This form of capital usually guarantees a constant return if the insurer generates a positive annual surplus, and serves as risk capital in the event of an insolvency. By issuing hybrid capital the insurer can target a reduction in its insolvency risk without having to expand its equity base by issuing shares with voting rights. In addition, the interest payments are tax deductible.

The cost of raising this capital thus lies between that for equity and debt. Hybrid capital represents an interesting investment instrument for institutional investors as it enables them to generate higher returns on a form of fixed-income security. However, non-life insurers have as yet made little use of this financing possibility.

In the event of bankruptcy, subordinated debt ranks for repayment behind all the other borrowings. The issue of subordinated debt does not reduce the right for "normal" creditors. Subordinated debt carries a higher default risk than a "normal" loan or bond and thus increases the security of higher ranked liabilities. The advantage this form of funding has over the issuance of shares is that it avoids diluting the voting rights and dividends paid out. The ratios used in financial analysis, such as earnings per share or ROE, are affected by changing the numerator, as additional interest payments reduce the amount of profits earned. There is no dilutive effect on the denominator.

Participation rights bear more similarity to capital funds, although have limited terms to maturity and no voting rights. Participation rights come in various forms but the interest they generate is usually dependent on corporate earnings. They also have no right to any residual claims in the event of an insolvency. Interest income is normally tax deductible. This form of security does not dilute capital funds, for example, in the ROE calculation.

In the case of surplus notes (US), regulators have to give step-by-step approval, not only of their issue but also of their interest payments and redemption. In this way a reduction in the assets can be prevented if the risk-based capital requirements have not been fulfilled. As surplus notes are classified as capital funds, in contrast to subordinated loans or participation rights, there is a dilution of capital funds, e.g. in the calculation of the ROE.

Provided certain regulations are complied with, hybrid capital can be combined with the available capital funds in the calculation of a company's solvency both in the EU and in the US. The rating agencies also include a certain share of hybrid capital in their calculation of the capital funds. The term to maturity plays an important role in this respect. The general rule is that the more of an equity character the hybrid capital takes, the more of it can be classified as capital funds. In Tables 2.7 and 2.8, there is a breakdown of the amount of the hybrid capital that can be included in the calculation of the solvency margin in the EU and in the calculation of the risk-based capital in the United States.

Table 2.7 Amount of hybrid capital in the EU calculation

Fixed term (up to 25%)	Undefined term (up to 50%)
• Subordinate to other liabilities	• Subordinate to other liabilities
• Fully paid in	• Fully paid in
• Original term at least five years; amortised over the last five years or, with the approval of the regulators, one year before maturity	• May not be redeemed without the approval of the regulators • The issuer must have the option of postponing the payment of interest
• Premature redemption only in the event of liquidation or with the approval of the regulators	• Both the nominal value and unpaid interest must be able to absorb losses

Source: European Commission (1999)

Table 2.8 Amount of hybrid capital in the US calculation

Surplus notes	Capital notes
• Capital funds in accordance with SAP* • Capital funds in accordance with GAAP** • Full inclusion in the RBC ratio • Subordinate to other liabilities • Interest rate and redemption only with the approval of the regulators	• Debt capital in accordance with SAP • Capital funds in accordance with GAAP • Inclusion of up to 25% of capital in the RBC ratio • Subordinate to other liabilities • The regulators can cancel interest rate payments if the RBC ratio falls under 100%

* Statutory Accounting Principles ** Generally Accepted Accounting Principles

Source: NAIC (1995)

Hybrid capital is still more the exception than the rule. Participation rights equalled 0.6% of equity at market value in Germany in 1997. Subordinated debt amounted to even less. In the US, surplus notes have become slightly more important in the financing of mutual insurers (around 3% of capital funds), as this represents their only external access to risk capital. The attraction of hybrid capital can increase very rapidly if, for example, following a stock market crash, an insurer's capital base is reduced and the cost of capital rises.

2.13 Supervision of insurance groups and financial conglomerates

Currently only individual companies are subject to solvency regulation in the EU. This enables insurance groups to use the same capital funds several times at different levels; this is termed double gearing. If a parent company increases the capital funds of a subsidiary, for example, the underwriting capacity of the subsidiary increases, although the consolidated funds of the group have not changed. The directive on the supervision of insurance groups that came into effect at the end of 1998 is meant to address this problem (Directive 97/78/EC). Supervision will remain at the individual company level, but regulators will require intra-group relationships to be monitored as well. The EU Directive has to be transformed into the national law of the EU member states by June 2000 and to come into effect by the 2001 business year at the latest.

The additional supervision mainly involves the insurer having to provide its adjusted solvency ratio. In addition, regulators should check whether intra-group business (e.g. loans, guarantees) poses a risk to the insurer's solvency. If the adjusted solvency margin is negative, the regulators must intervene. Since a number of European states had already introduced provisions relating to the supervision of insurance groups before this directive was brought into effect, three different options of calculating adjusted solvency have been specified in the EU directive. Each individual member state has to decide which of the three methods it wants to employ. In the case of the requirement deduction method, the sum of the elements eligible for the solvency margin are subtracted from the parent company's capital funds. The disadvantage of this method is that any surplus funds of a subsidiary or participation are not taken into account. The deduction and aggregation method is aimed at correcting this disadvantage. A third method is the calculation of adjusted solvency using the consolidation-based accounting method. If the company in question is a wholly owned subsidiary, this third variation gives the same result as the deduction and aggregation method.

Moreover, the advancing convergence of banking and insurance through mergers and acquisitions represents a great challenge to regulators in the US and Europe regarding the capital requirements of conglomerates. The European Commission plans of drafting a directive relating to the supervision of financial conglomerates by the end of 2000. One core theme is likely to be the problem of double gearing. However, it is difficult to abolish double gearing since different regulators are responsible for insurers, banks and securities firms, all of whom, given the different risks involved in the individual businesses, have different solvency requirements.

In drawing up a draft directive, the European Commission will take the recommendations of the Joint Forum into account. The Joint Forum was established in early 1996 under the aegis of the Basle Committee on Banking Supervision (Basle Committee), IOSCO (International Organisation of Securities Commission) and IAIS (International Association of Insurance Supervisors). During the course of 1999 the Joint Forum published a number of papers on the supervision of financial conglomerates. The recommendations put forward by the Joint Forum include various measurement techniques for determining the amount of capital funds financial conglomerates should hold and suggestions on how to approach the problem of double gearing. The paper does not recommend replacing the system for the supervision of individual companies for the supervision of individual companies, but rather the introduction of an additional supervisory element.

In the US, discussions revolving around the supervision of financial conglomerates apply more to the jurisdiction of the authorities that they do to the substance of the recommendations. These discussions are closely linked to the revision of the Glass-Steagall Act according to which the strict division between banking and insurance services are being relaxed. As insurance is supervised by the individual states and the banks by a federal authority, there are basic conflicts regarding jurisdiction that will have to be overcome before joint supervision can be introduced. A model is currently being discussed which leaves the supervision of insurance companies to the individual states, while making conglomerates subject to additional supervision by the Federal Reserve system. The development of integrated models to measure all the risks to which a conglomerate is exposed is also at its early stages. As it will most likely be difficult to develop a universally valid model, one possible solution would be for companies to take on the responsibility for their own risk management and seek regulatory approval for their risk management models.

Closely connected with the introduction of an international supervisory standard is the discussion of a universal accounting standard. In addition to the proposals made by the European Commission, the IASC (International Accounting Standards Committee) is endeavouring to introduce one universal accounting standard. However, the IAS would be in competition with US GAAP. The introduction of an international accounting standard would influence the amount of capital reported by a financial conglomerate.

2.14 Conclusions

In comparing the current European solvency margin system with the US risk-based system, one must conclude that the European model is inadequate. How should one proceed to draw the best from the risk-based models which are conceptually superior, without having the complexities, rigidities and costs of these systems? There is no simple answer to this question. One solution that suggests itself is a more collaborative approach between regulators and company management. Under such a system the regulatory authority would define risk models that would be acceptable, but would leave the detailed modelling to the companies themselves, with the oversight of an approved internal audit function within the company. Hence it would be possible to combine sufficient rigour without undue cost and inflexibility. An internal audit team combining actuarial, accounting, underwriting and economic expertise would seem appropriate for non-life insurance; perhaps an actuary alone would be sufficient for life insurance, since the task is less complex. The role of the regulatory authority would be to approve the models and to monitor the output of the models against benchmark standards.

Although the current European solvency margin system may appear crude compared to the US risk-based system, with weak theoretical foundations, its application is very straightforward. The dilemma between objective correctness and functionality is resolved in favor of the latter (Farny, 1997). The US risk-based capital model is significantly more tailor-made to an individual insurer risk profile, although its theoretical foundations are far from strong. It is also harder to apply in practice and imposes high internal compliance costs on insurers. There is a general skepticism as to whether the US-style risk-based model could be transferred to Europe. And this skepticism can be found among European supervisors in the Muller Report. Moreover, the US system is not only complex but is also very prescriptive even for insurance supervisors, since it has a detailed set of regulatory responses that are triggered if the capital base of an insurance company falls below the minimum risk-based capital level. One can argue that this degree of prescription builds undue rigidities into the regulatory process.

Regulatory capital is of course just one of a number of regulatory tools, alongside other supervision techniques and regulations covering conduct of business rules and client money. Ideally, the marginal costs and marginal benefits of all these tools applied in combination need to be judged, to try to decide on the most efficient mix.

Additionally, the imposition of capital rules in isolation is of little value if the regulator cannot be sure that the insurer has adequate systems to monitor and measure the risks that the capital standards are intended to limit, and that the firm's management are honest and competent. As well as choosing the right balance of regulatory tools, regulators need to take account of other mechanisms which are not part of the regulatory armoury, but which can also help them achieve their objectives. Market discipline is one important example: if the market is in a position to judge that a firm is weakly capitalised or poorly run, it may penalise the firm in various ways, such as discounting the firm's share price. And less tangibly, evidence of weak management can damage a company's reputation, which may make it harder for it to write new business. All these factors impose incentives, in different degrees on different companies, to operate in a sound and efficient manner, and to hold capital as a cushion against future losses.

These are some of the lessons that the Stage 2 of the EU solvency exercise. But it is clear that insurance supervisors and the insurance industry will also need to work together to construct a workable and cost-effective system for solvency of assessment, including capital adequacy, if the new computer-based risk modelling techniques are to be used. The fruitful interaction of regulators and the insurance industry has a good precedent in Europe: the development of the 3rd Insurance Directives. But what is also clear is that the current rethinking of capital adequacy and solvency within the banking sector by the Basel Committee on Banking Supervision will also have a key influence, as banking regulation itself moves away from its own static risk-based capital rules. A new solvency framework, including capital standards, is likely to emerge which has common features across the financial services sector, not least because of the need for level competitive playing-fields, as product and corporate convergence within the sector continues. However, the framework must be able to take into account the particular characteristics of the risks facing non-life insurers and those facing life insurers, if it is be effective.

CHAPTER 3

UNDERWRITING CYCLES IN THE UK

3.1 Introduction

It is the prevailing wisdom of the insurance industry the world over that the industry is subject to an underwriting cycle. The study of cycles in general insurance is of major importance. In the last ten years, the study of cycles has progressed significantly. The industry results tend to follow a cycle consisting of alternating uniform periods of rising and then falling underwriting profits.

A study of insurance price cycles is of interest for a number of reasons. Firstly, since insurance profit is a component of total earnings of insurers (along with investment income), a study of temporal behavior is of interest to those seeking to value insurance stocks.

Secondly, cyclical behavior of insurance earnings might, under some circumstances, be transmitted into temporal movements in insurance stock prices. If confirmed, such movements would carry challenging implications for the weak form of the efficient markets hypothesis.

Moreover, if temporal price movements represent adjustments to changing market equilibria, they may present valuable insights into the competitive processes at play in the insurance market. Furthermore, it provides a case study of the price linkages between two markets; the insurance product market and the capital market.

The underwriting cycle is characterized by periods of intensely competitive insurance pricing that result in low premiums and sizeable insurer underwriting losses, followed by periods of much higher insurance prices (see Venezian, 1985; Cummins and Outreville, 1987; and Winter, 1994). Although variations in profits suggest that a market mechanism may be operating, industry observers usually interpret the cycle as a supply-side phenomenon.

The typical explanation is that the insurance industry causes the cycle more or less on its own, through periods of destructive competition followed by cutbacks in supply. More sophisticated versions relate the recurring phases of the cycle to key operating ratios such as the premiums-to-surplus ratio, which is said to represent capacity.

An underwriting cycle can be viewed as a repeated sequence of "hard" and "soft" markets in the industry (Harrington and Danzon, 1994; Niehaus and Terry, 1993). The sequence may be observed in the prices, profitability, and supply data. In a "hard" market, the supply of insurance coverage shrinks amid high and rising insurance prices and profitability. In a "soft" market, the availability of insurance coverage expands as prices and profits tumble. The underwriting cycle does not necessarily synchronise with the general business cycle. In fact it is much more regular than the general business cycle (Webb, 1992).

The existence of an underwriting cycle has been recognised by researchers. In addition, they have been numerous studies and debates on the causes of the underwriting cycle. Nevertheless, there is no generally accepted view of what the causes are. One school of thought suggests that the causes are irrational behaviour like competitor-driven pricing, naïve rate-making processes, and capacity constraints. Another school of thought, which is related to the rational expectations/institutional intervention hypothesis, however, does not agree that insurance markets and insurers are irrational. Instead, it suggests that the underwriting cycle is created by external factors and market characteristics that are outside the control of insurers. These factors include data collection, regulatory, policy renewal and accounting lags, interest rates, stock markets, and the general business cycle.

3.2 Market imperfections theory

The first school of thought is based on the premise that insurance markets operate irrationally. Venezian (1985) attributes the cycle to the imperfection of extrapolative methods used in the naïve rate-making process. He finds that the United States underwriting profit data follow a second order autoregressive process with a cycle period of about six years.

The argument of irrational behavior suggests that insurance markets are destabilised by phenomena such as extrapolative forecasting and so-called "cash flow underwriting", which can result in prices considerably higher or lower than competitive levels due to erroneous estimates of losses or investment income.

Venezian (1985) describes a ratemaking model in which past loss levels are used (via time trending) to extrapolate estimates of losses in future periods. These forecasted losses, in turn, are used to set premiums. By incorporating reasonable estimates of experience and policy projection periods used in the United States, Venezian predicts that a cycle with average periods ranging from four to nine years should exist.

Venezian's model is substantiated by empirical tests. Parameters needed to measure the cycle period are obtained by estimating the following autoregressive model with ordinary least squares:

$$\Pi_t = a_0 + a_1\Pi_{t-1} + a_2\Pi_{t-2} + \omega_t \quad (3.1)$$

where Π_t: the underwriting profit in period t, and

ω_t: a random error term

The profit margins for each line of insurance, adjusted for linear trend, were used as dependent variables in regressions that included up to four lagged values as independent variables. In the majority of the cases the first two terms give the highest adjusted r-squared, and generally these two terms do as well as four terms in predicting underwriting profit margins.

A cycle will be present if $a_1 > 0$, $a_2 < 0$, and $a_1^2 + 4a_2 < 0$. In this case, the characteristic of the second order difference equation in underwriting profits will have complex roots, implying that profits follow a cyclical pattern. The period of the cycle is obtained from the following formula:

$$T = \frac{2\pi}{\cos^{-1}\left(a_1/2\sqrt{-a_2}\right)} \quad (3.2)$$

The cycle will be damped (i.e., have a tendency to die down over time) if $\sqrt{-a_2} < 1$. If $\sqrt{-a_2} > 1$, the cycle will be explosive. Even a dumped cycle will be maintained over time if random shocks occur.

Cummins and Outreville (1987) show that simple lags in data collection or price regulation may be sufficient to produce cyclical performance even in rational expectations setting. That is, insurance prices typically are based on annual data which are not available for use until several months after the close of the "experience" period. Certainly, as technological advances in data base management occur, this delay is shortened. Nevertheless, delays are currently experienced in tabulating and analyzing data, and the slow emergence of information on losses in long-tail lines dictate that projections are made based on lagged loss observations.

The authors developed an alternative model that is consistent with observed profit cycles in insurance. Specifically, they showed that cycles in reported underwriting profits are consistent with a simple rational expectations model of insurance price determination, provided that institutional lags and reporting practices are taken into account. The rational expectations hypothesis implies that economic agents forecast economic variables without systematic error, i.e., that their subjective expected values of these variables are the same as the actual of objective expected values, conditional on all information available at the time the forecasts are made.

A second-order process can be created by combining informational and regulatory lags with renewal lags and calendar-year reporting practices. Recall that rates are assumed to change at the beginning of each year and to remain in effect for one year. Also assume that policy terms are one year in length and that policies are renewed evenly throughout the year. The cycle is apparent in the sense that it has nothing to do with the underlying economic and statistical characteristics of insurance profits but rather reflects the institutional factors and accounting practices.

While the empirical findings of Cummins and Outreville are consistent with the rational expectations/institutional intervention hypothesis, they also may be consistent with other hypotheses such as Venezian's extrapolative expectations hypothesis. Information on ratemaking procedures, regulatory constraints, lengths of policy terms, and accounting procedures would be helpful for further development of the model. More precise and detailed information on actual rather than reported profits also would be useful.

Regulatory lags arise in countries in which insurers are required to submit rates for approval prior to use. This requirement further extends the delay between the experience period and the effective use of revised rates. This delay can be shortened by simplifying the regulatory process; nevertheless, regulatory rate approval is required in varying degrees across countries and across lines of business.

Policy renewal lags exist because the insurance price cannot be adjusted simultaneously to reflect information as it becomes available. Most property-liability insurance policies have a set premium for the entire policy period (e.g., for an entire year). Furthermore, when new rates are approved, typically a lag in changing to the new rate level occurs.

Moreover, financial reporting practices may give rise to apparent underwriting cycles in a rational marketplace. Loss estimates for each year would reflect all information available at the end of that year. Nevertheless, calendar-year data are used typically in financial statement reporting of losses, and financial statement data are used in cycle studies. These data are reported on an incurred basis, meaning that losses are matched to the coverage period during the calendar year. Likewise, premiums are based on accrual accounting; earned premiums include premiums attributed to policies issued within the first day of the preceding year to the last day of the reported year. A mismatch exists between the informational content of the reported premiums and reported losses.

Cummins and Outreville (1987) also claim that Venezian's (1985) hypothesis implies a certain degree of irrationality on the part of insurers and is inconsistent wit the rational expectation hypothesis advocated by modern economic theory developed for other types of financial markets. Furthermore, Venezian's (1985) hypothesis may not be able to explain the presence of underwriting cycles in countries where extrapolative trending procedures are not used.

Lamm-Tennant and Weiss (1997) also find results that support the rational expectations/institutional intervention hypothesis. They utilise a generalised least squares regression model to analyse the changes of premiums with respect to changes in lagged losses, interest rates, average stock price, real gross domestic product (GDP), concentration ratio, regulations, policy periods, and catastrophic losses of nine developed countries. Specifically, their results indicate that data collection, regulatory, policy renewal and accounting lags, interest rates, real GDP, and stock markets are closely related to the underwriting cycle. They find that the length of the cycle period is largely determined by interest rates, rate regulations, and catastrophic loss growth.

In determining the presence of the underwriting cycle Lamm-Tennant and Weiss (1997) estimate equation (3.1) individually for twelve countries using the average loss ratio, the overall combined ratio, and by-line loss ratios for six lines of business as the dependent variables. A linear time trend is added to each equation to control for declining expense ratios over time. The equations are estimated using ordinary least squares. The period of the cycle, if a cycle is observed, is estimated from equation (3.2).

After establishing the presence of underwriting cycles in the twelve countries, the rational expectations/institutional intervention hypothesis is tested empirically by utilising a generalised least squares (GLS) model which controls for autocorrelation within countries and heteroscedasticity across countries. The specification of the GLS model is:

$$\Delta P_{it} = \alpha + \sum_{j=1}^{J} \beta_j \Delta x_{jt} + \sum_{i=1}^{n-1} c_i D_i + \varepsilon_{it} \qquad (3.3)$$

ΔP_{it} are the changes in aggregate premiums for country i and time period t,

ε_{it} $= \varrho \varepsilon_{i,t-1} + \mu_{it}$

μ_{it} $\sim N(0, \sigma^2_{iu})$

n The number of countries, and

D_i a dummy variable equal to one for country i and zero otherwise

The independent variables (Δx_{jt}) are lagged losses, interest rates, average stock price, real gross domestic product (GDP), concentration ratio, regulations, policy periods, and catastrophic losses of the sample countries.

Furthermore, they regress the underwriting cycle period on the regulatory and market characteristics of the sample countries to identify the direct impact of the independent variables on the length of the underwriting cycle. The regression model is stated as follows:

$$CycPer_{ij} = \alpha_0 + \beta_1 Dis_{ij} + \beta_2 Per_{ij} + \beta_3 Cat_{ij} + \beta_4 Reg_{ij} + \beta_5 Res_{ij} \qquad (3.4)$$
$$+ \beta_6 CVLoss_{ij} + \sum_{k=1}^{K} \beta_k D_{ijk} + \varepsilon_{ij}$$

where the dependent variable is the cycle period in country i and line j, and the independent variables are averages for the interest rate, policy period, catastrophe loss growth, premium regulation, reserve discount, coefficient of variation of the loss ratio, and dummy variables for line (ocean and inland marine). A dummy variable is included for ocean and inland marine insurance, because these lines are less regulated and more international in scope than lines where the bulk of business is written on domestic risk.

In their final analysis, Lamm-Tennant and Weiss (1997) attempt to predict the presence of cycles using market characteristics and institutional/regulatory features for the sample countries using a logit model.

In general, their results suggest that underwriting cycles are present in all countries and in at least one line. Differences in the presence and length of the underwriting cycle are evident across countries and across lines of business. For example, when comparing a long-tail business such as liability insurance to a short-tail line of business such as fire and allied lines, the cycle is considerable longer for liability insurance. (8 years versus 5 years).

The changes in current aggregate premium levels are significantly related to changes in past loss levels. These results are consistent with Venezian (1985). Also, overall premium changes are related to the concentration measure, change in stock index, premium regulation, and policy as hypothesised, and are statistically significant. Finally, on the predictability of the cycle, the results of the logit regression model conform to the expectations. Liability lines are more likely than property lines to exhibit an underwriting cycle.

Niehaus and Terry (1993) used time series causality tests to examine hypotheses about the determinants of insurance premiums and causes of the underwriting cycle. Their results suggested that market imperfections play an important role in insurance pricing. Consistent with the capital market imperfection hypothesis, the evidence suggested that past values of surplus affect premiums.

Niehaus and Terry consider a time series regression where the dependent variable is premiums written in year t, and the independent variables are current and future loss payments on policies written in year t, past loss payments, past values of surplus, and other control variables:

$$\text{Premiums}_t = f(\text{loss payments on policies written in year t}, \qquad (3.5)$$
$$\text{past losses, past surplus, control variables}) + \varepsilon_t$$

According to the perfect markets hypothesis, premiums on a pool of policies should be explained by the loss payments on those policies, but not by past losses or past surplus. The capital market imperfection hypothesis implies that past values of surplus help explain premiums, and Venezian's (1985) hypothesis suggests that past values of losses may help explain premiums.

However, in order to measure loss payments on policies written in year t, we need a loss variable that associates losses paid to the year in which the policy was written. Such loss data are not publicly available. Therefore, the losses paid by insurers in a calendar year is used as a proxy. The losses paid in a calendar year are not all from policies written in the same year. For lines with a short claim tail, the losses paid in a calendar year are likely to be from policies written in the reporting year or the year before. For lines with a long claim tail, losses paid could reflect losses written many years earlier. Thus, the losses paid variable includes measurement error, especially in lines with a long claim tail.

Linear equations of the following form were estimated:

$$P_t = \alpha + \sum_{s=0}^{M} \lambda_s L_{t+s} + \sum_{s=1}^{M} \pi_s L_{t-s} + \sum_{s=1}^{M} \gamma_s S_{t-s} + \sum_{s=1}^{M} \theta_s P_{t-s} + \beta R_t + \varepsilon_t \qquad (3.6)$$

where P_t: Premiums written in calendar year t,

S_t: Aggregate policyholder surplus at time t,

L_t: Losses paid in year t,

R_t: Annual return on treasury bills in year t,

ε_t: The error term

Although most studies use aggregate data, Venezian (1985) and Fields and Venezian (1989) argue that individual lines should be examined separately. Two problems prevent examination of individual lines. First, losses paid can be calculated only for aggregate series because the change in the loss reserve is not reported by line. Second, surplus, which is central to the capital market imperfection hypothesis, is not allocated by line. Consequently, aggregate data are examined which tend to reflect the experience of the largest lines: auto and workers' compensation.

Consistent with time series evidence, lagged values of premiums are included in the model as control variables. Including lagged values of the dependent variable on the right hand side of the equation is also consistent with the usual implementation of causality tests. To conserve degrees of freedom, a lead and lag structure of two years is assumed for all variables; that is, $M = 2$ in equation (3.6).

Equation (2.6) is estimated using both first differences and percentage changes. The results suggested that the greatest weight should be placed on the percentage changes rather than the first differences of the series. Under the perfect market hypothesis, premiums would be positively related to current and future loss payments. Both approaches (first differences and percentage changes) provide little support for this prediction. The coefficients of future losses are insignificantly different from zero, and the coefficient on contemporaneous losses (L_t) is significant at the ten percent level only when percentage changes are used. In these equations, the coefficients on lagged premiums are highly significant and have opposite signs, suggesting cyclical behavior.

When lagged premiums are not included in the equations, the statistical significance of future losses increases. In addition, the coefficients of the future loss variables alternate in sign, suggesting that the future loss variables in these equations are capturing the cyclical effect that the lagged premium variables capture in the first equations. The negative coefficient on the one year lead loss variable (L_{t+1}), however, is inconsistent with the perfect market hypothesis. The lack of support for the perfect markets hypothesis may be due to the measurement error in losses.

The perfect markets hypothesis also predicts that past information, such as past loss payments and past surplus, would not help predict premiums, but evidence to the contrary was presented by the authors. The coefficient on lagged losses is positive and statistically significant using first differences, suggesting that high losses are followed by higher premiums. This finding is consistent with imperfections in the rate setting process (Venezian, 1985). It is also consistent with the hypothesis that the future loss variables measure expected future losses with error and that this error is correlated with past losses. When percentage changes are used, however, lagged losses are not statistically significant.

3.3 Rational expectations theory

A second explanation of the underwriting cycle builds on the fact that the "underwriting profit" represents a measure of the average price of the contracts traded. Insurance pricing models based on financial theory are unanimous in showing that competitively determined insurance prices are inversely related to interest rates and will therefore change as interest rates change. Using this model, together with the rational expectations and lag features of the Cummins and Outreville model, Doherty and Kang (1988) show that the intertemporal behavior of underwriting returns in insurance markets is quite well explained as a market clearing process in which equilibrium prices change in lagged response to changing interest rates. Their model considers both supply and demand, and the resulting prices and profits arise from the interaction of these two market forces. Relying on capital-asset pricing theory, supply is considered to be a function of interest rates and expected profits. The sign of the interest rates term is expected to be positive, i.e., insurers increase supply when interest rates rise in order to obtain funds to invest (known as cash-flow underwriting). The demand for insurance in the Doherty-Kang model is hypothesised to be a function of price (the inverse of the loss ratio) and aggregate economic activity (income), with the latter representing an index of the amount of insurable goods and services. The equilibrium price is determined in the model by equating the quantity demanded with the quantity supplied.

In estimating the structural system described by the supply and demand functions, Doherty and Kang use a three stage least squares (3LS) model. Due to secular trends in some of the series, the model was run on first differences. The structural equations were estimated for aggregate property liability insurance, automobile physical damage and automobile liability. The results that Doherty and Kang report are supportive. The supply function results indicate a positive partial adjustment from previous output level that is determined by the degree of excess return. It shows the expected significant positive response to expected excess returns in the total industry result and for both auto lines. Evidence of price elasticity of demand was weak or absent.

From their analysis Doherty and Kang showed that the cycle appears to have resulted from the market's continuing attempts to clear. This supports the view that the cycle is evidence of a rational economic response to prevailing economic circumstances. However, this explanation stands in contrast to the traditional professional explanation that the cycle is a disequilibrium phenomenon which reveals the inability of the market to converge on its clearing prices.

3.4 Capacity constrained theory

A third explanation focuses on external shocks to the value of the insurer's equity and therefore to its underwriting capacity. These models sometimes are called "capacity constrained" models.

Berger (1988) develops a model in which insurance capacity depends on the current level of equity that, through retained earnings, is determined largely by pricing decisions made in the previous period. The model developed excludes any consideration of expenses, taxes, investment income, interaction with the capital market, or ratemaking methodology; rather, the hypothesis is that the dynamics of the cycle derive from the fact that profits feed back into surplus with a lag. In order to derive Venezian's second order autoregressive equation in profits, Berger assumed two one-year lags in the structure of the insurance business. First, the firm is assumed to set its underwriting policy for the upcoming year on the basis of end-of-year surplus. The more financially secure the firm is, the more willing it will be to underwrite what would otherwise be considered marginal risks. Secondly, the profit and loss results of the firm's underwriting policy will also be assumed to follow with a one-year lag. The unearned premium reserve was considered to be a component of the firm's initial capitalization.

The market for insurance was modeled by way of standard supply and demand analysis, in which equilibrium price and quantity result from the intersection of market supply and demand schedules. It was assumed that market supply is a function of prior period surplus, since at any given price firms will be more willing to underwrite marginal risks when surplus is high.

Given this assumption, the resultant market price and quantity will also depend on prior period surplus, since the position of the supply function will determine the intersection of the supply and demand functions. If P_t, Q_t, S_t and π_t respectively represent market price, quantity, surplus and economic profits in period t, then:

$$(P_t, Q_t) = f(S_{t-1}) \tag{3.7}$$

It is also assumed that profitability is a function of price and quantity in the prior period:

$$\pi_t = g(P_{t-1}, Q_{t-1}) = g(f(S_{t-2})) = h(S_{t-2}) \tag{3.8}$$

Since $S_t = \pi_t + S_{t-1}$, equation (2.8) implies

$$\pi_t = h\left(h^{-1}(\pi_{t-1}) + \pi_{t-2}\right) \tag{3.9}$$

when h is invertible. This is a second order difference equation in profits.

When h is linear $(h = aS + b)$, equation (3.9) becomes

$$\pi_t = \pi_{t-1} + a\pi_{t-2} \tag{3.10}$$

This equation will generate a cycle when $a < -1/4$, with a period of $2\pi/\cos^{-1}\left(1/2\sqrt{-a}\right)$. The value $a = -1$ yields a period of six years.

The individual firm's problem is to maximize the following objective function:

$$\max_{Q} J(\pi, \varrho) = PQ - C(Q) - \varrho(Q, S) \tag{3.11}$$

where costs $C(Q)$ satisfy $C', C'' > 0$, and $\varrho(Q, S)$ is an increasing function of the probability of ruin. This objective function simply embodies the trade-off which the firm faces between the expected profits which may result from increased volume, versus the increased probability of ruin which may also result. The function $\varrho(Q, S)$ is not the probability of ruin, but is only the "dollar equivalent" of it.

The specification is essentially behavioral in nature; i.e., showing just how much the firm values expected profits relative to the possibility of bankruptcy. It is assumed that $\partial P/\partial Q > 0$, $\partial \varrho/\partial S < 0$, and $\partial^2 \varrho/\partial Q \partial S < 0$. This means that the probability of ruin increases with volume, declines with surplus, and that the marginal increase in the probability of ruin due to an increase in volume declines as surplus increases. Differentiating the objective function with respect to Q, the first order condition is (assuming price-taking firms):

$$P = C'(Q) + \frac{\partial \varrho(Q, S)}{\partial Q} \tag{3.12}$$

The marginal profit is $P - C'(Q)$, which is positive since $\partial P/\partial Q > 0$. The term $\partial P/\partial Q$ is seen to represent the risk premium, and the condition $\partial^2 \varrho/\partial Q \partial S < 0$ means that the risk premium declines as surplus increases. The second order condition

$C''(Q) + \partial^2\varrho/\partial Q^2 > 0$ is always satisfied when $\partial^2 P/\partial Q^2 > 0$, which is also be assumed to be true. This assumption of a convex ruin function is justified by the presence of selective underwriting in these markets, in that an increase in the quantity of insurance sold requires a loosening of underwriting standards, and therefore raises the marginal increase in the probability of ruin due to an additional unit of business. Equation (3.12) is the supply function of the firm, which aggregates to form the market supply function. Equation (3.7) is the intersection of market supply with demand.

Recall again the usual "story" told about the cycle, that when profits are high firms loosen underwriting standards and take on less desirable risks. However, whereas it is often asserted that firms' expectations (apparently incorrect) regarding the profitability of the business leads to the cycle, in the model developed, it is the improved financial position of firms, coinciding with increased profitability, which leads to the cycle. Since profits over the cycle are zero, the market is in long run equilibrium, so that capital movements which normally are of a long-term nature will not serve to mitigate the cycle. Although historically short-term capital movements have not been a factor, future institutional changes could allow this to become prevalent, resulting in the eventual elimination of the cycle.

Winter (1994), Gron (1992, 1994), and Cummins and Danzon (1992) show that, given limited liability and costs to raising external capital, sudden shocks to insurers' liabilities can generate price and quantity effects such as those observed over the insurance cycle. The "capacity constraint" theory posits that cycles are caused by impediments to capital flows that result in alternating periods of excessive and inadequate capacity in the industry.

According to this scenario, the underwriting cycle is most prominent on long-tail lines (usually liability lines) because forecasting horizons are longer and anticipated investment income is more substantial for these lines.

By joining the interest rate and capacity constrained models, Doherty and Garven (1995) provide a different empirical approach. Absent capacity effects, interest rate changes should produce changes in underwriting returns of the opposite sign. However, the capacity effects of the same interest rate changes will affect insurers differently according to the interest sensitivity of their asset and liability portfolios and according to differences in their respective costs of raising new capital.

The first cross-sectional difference can be measured by asset and liability durations and the second difference by organizational factors such as ownership structure, size, whether public or privately traded, and so on.

Since evidence for the insurance cycle is usually presented as a time series for underwriting returns, Doherty and Garven use the expected underwriting return in order to check the sensitivity of underwriting returns to interest rate changes.

They show that the time series of underwriting returns could be explained simply from the spot equilibrium prices required to deliver to the insurer a fair rate of return on equity. By showing this, the notion that the cycle is a purely monetary phenomenon would be evident. In order to test this hypothesis, Doherty and Garven estimate equation (3.13) by using generalized least squares (GLS).

They use as an independent variable the average settlement delay, k, times the weighted average of monthly spot Treasury-bill rates (rather than daily rates that were unavailable over the entire period of their analysis, 1939-1988).

$$r_{ut} = \alpha_0 + \alpha_1(kr_{ft}) + \varepsilon_t \qquad (3.13)$$

The estimated slope coefficient exceeds negative unity and that is consistent with the presence of capacity constraints. However, the interpretation requires some caution since the data available for estimating the average settlement delay, k, were not ideal.

In order to establish further evidence of the capacity constraints, the authors include a squared term, $(kr_f)^2$. The inclusion of the squared term was made in order to predict the different (asymmetric) responses to rising and falling interest rates. Responses to rising and falling interest rates are unlikely to be symmetric. As interest rates rise and bond prices fall, insurers are faced with the question of whether to float new equity. Raising new equity involves explicit transaction costs (e.g., underwriting fees) as well as the adverse selection costs arising from information asymmetry between insurance management and external investors. Similar transaction costs arise if the insurer responds not by raising equity but by increasing its ceded reinsurance.

If the insurer finds itself under-leveraged due to a fall in interest rates, it is less costly to adjust to its desired capital structure. The distribution of equity through dividends is less costly than the raising of new issues. Similarly, the reduction in the net value of reinsurance ceded is likely to be less costly than an increase. The extreme case of asymmetry is the case of mutuals where the cost of raising capital in the short run (i.e.

above immediate earnings) is, theoretically, infinite. However, the mutual is perfectly free to distribute equity to its policyholders whenever it sees fit to do so.

Thus, this asymmetry can be detected by including the aforementioned squared term as follows:

$$r_{ut} = \alpha_0 + \alpha_1(kr_{ft}) + \alpha_2(kr_{ft})^2 + \varepsilon_t \qquad (3.14)$$

where $\alpha_1 < 0$ and $\alpha_2 > 0$. The results show evidence of this predicted asymmetry, with underwriting profit being significantly more responsive to falling rates than to rising rates.

To provide a more rigorous test the authors turn to individual firm data in order to see whether cross-sectional differences in the responses of different firms to changing interest rates correlate with differences in equity duration and access to external capital. A two-pass regression procedure was adopted. In the first pass, the GLS regression described by the following equation was estimated for each of the firms in the sample:

$$r_{ujt} = \alpha_{0j} + \alpha_{1j}(kr_{ft}) + \varepsilon_{jt} \qquad (3.15)$$

where r_{ujt} is the period t underwriting return for firm j, and the term kr_{ft} corresponds to the period t product of the average settlement delay and the annualized weighted average of monthly returns on 1-year Treasury bills in the year surrounding time t.

In the absence of capacity constraints, the α_{1j} coefficients in the first pass should not differ significantly from negative unity.

With differences in duration and differences in access to capital and reinsurance markets, however, there should be cross-sectional variation in the α_{1j} coefficients. Accordingly, the parameter estimates α_{1j} were then used in a second-pass OLS regression in order to test for cross-sectional differences in the responsiveness of α_{1j} to a number of firm-specific variables.

The second-pass equation was specified in the following manner:

$$\alpha_{1j} = \beta_{0j} + \sum_{i=1}^{n} \beta_{ij} X_{ij} + \mu_j \tag{3.16}$$

where

X_{1j} = $SIZE_j$ natural logarithm of firm j's size, measured in terms of admitted assets;

X_{2j} = $QDUR_j$ mean value of equity duration for firm *j* during the period 1980-86;

X_{3j} = $REINS_j$ the slope coefficient determined from the OLS regression of the reinsurance variable against returns on 1-year Treasury bills;

X_{4j} = $PUBLIC_j$ 1 if firm j or its parent is a publicly traded stock corporation, or 0 otherwise;

X_{5j} = $PRIVATE_j$ 1 if firm j or its parent is a privately held stock corporation, 0 otherwise;

X_{6j} = $GROUP_j$ 1 if firm j is an insurance group, 0 otherwise;

X_{7j} = $GROUPRE_j$ the product of $GROUP_j$ and $REINS_j$.

The results of the cross-sectional tests support the model. The model predicts that those firms for which changes in capital structure are most costly will exhibit greatest difficulty in adjusting to equilibrium prices following changes in interest rates. Firms with high equity duration and more costly access to reinsurance and capital markets will show the greatest frictional disturbances in their insurance prices following a change in interest rates.

Overall, the time-series analysis that the authors employ confirms that cycles do seem to be dampened from the equilibrium path, which points to the presence of capacity constraints. Asymmetric responses to rising and falling interest rates provide further support for the presence of capacity constraints. The cross-sectional analysis revealed that price disturbances are more pronounced in those firms for which leverage adjustments are most costly. Specifically, firms with more costly access to new equity, or with less flexible access to reinsurance markets, show more evidence of frictions in responding to changing capital market changes. These results support the general class of capacity constrained models of insurance cycles.

Grace and Hotchkiss (1995) provide evidence of a long-run link between the general economy and the insurance underwriting cycle as measured by the combined ratio. Time series methods are employed to examine the property-liability insurance industry to determine effects on the insurance cycle of external factors such as shocks to real income, inflation, and the short-term interest rate.

By using cointegration techniques, they estimate the long-run relationship between the general economy as measured by real gross domestic product, the short-term interest rate, and inflation on the underwriting cycle. To test the theory that the combined ratio (the sum of the ratio of expenses before taxes to premiums written and the ratio of losses and loss adjustment expenses to premiums earned) is tied to the general business cycle in the long run, they test whether the combined ratio and real gross domestic product are cointegrated. Finding that the real gross domestic period and the combined ratio are cointegrated would suggest that (economic) factors are at work tying the movement of the combined ratio cycle to a more wide-ranging national business cycle. Cointegration as an indicator of long-run relationships was introduced by Granger and Weiss (1983) and has been used extensively to examine a variety of relationships. In order to determine whether economic series are cointegrated, each series must be stationary. A series X_t is said to be integrated of order one, I(1), if it is stationary in its first difference. By using the Augmented Dickey-Fuller (ADF) test, Grace and Hotchkiss determine whether the series are cointegrated. The results from cointegrating regressions between the combined ratio and each of the other three series (real GDP, short-term interest rate, and Consumer Price Index) indicate that the combined ratio is cointegrated with each of the other series as well as all four series being cointegrated. Consequently, in the long run, the real gross domestic product, the combined ratio, the short-term interest rate, and the consumer price index to be tied together as there exist forces that tie the movement of the combined ratio with the movement of the national business cycle, the movement of short-term interest rates, and the movement of prices. This determination of cointegration indicates that an equilibrium relationship exists between the four series and that a more structured model should take the form of an error-correction model to account for this equilibrium relationship.

Grace and Hotchkiss use vector autoregression to allow the data to determine the dynamic structure of the relationship. The resulting vector autoregression specification took the form of an error-correction model:

$$Y_t = FY_{t-1} + \Theta m_{t-1} + G\xi_t \qquad . \qquad (3.17)$$

where m_t is the cointegrating regression residual, which controls for the pertinent information regarding the ability of the series to achieve long-run equilibrium.

The estimated parameters showed a number of important relationships. The majority of the non-lagged regression parameters were not significant since most of the behavior seemed to be explained by past behavior. The coefficient of the error-correction term (m_t) represents the short-run dynamic behavior of the dependant variable. Taken together, the error-correction term's coefficients imply that the combined ratio, the short-term interest rate, and the consumer price index all respond in the short run to changes in the long-run relationship described in the cointegration regression, while real gross domestic product does not. However, due to the fact that the vector autoregression estimates are often difficult to interpret, the authors turned to impulse response functions to describe the behavior of the system.

The impulse response function is used to simulate the impact of a shock to one of the series on the outcome of the other series included in the vector autoregression. The response was measured in terms of combined ratio standard deviations. Grace and Hotchkiss find that a shock to the real gross domestic product of one standard deviation initially causes the combined ratio to decline and then to increase before eventually dying out. Shocks to the short-term interest rate and the consumer price index increase the combined ratio, whose response remained positive before dying out.

These responses of the combined ratio to shocks in external factors have a number of interpretations. First, there is a pure income effect when real gross domestic product experiences a shock. A positive shock to real gross domestic product is interpreted as an increase in total income, leading to increased demand for all normal goods, thus increasing the revenue and profits in the property-liability industry. Second, since in a competitive market insurance premiums will reflect discounted expected losses, there is a direct and positive relationship between the competitively determined combined ratio and the interest rate. This is consistent with theory and empiricism (Cummins, Harrington and Klein, 1992).

Finally, the response of the combined ratio to a shock in the consumer price index illustrates a number of possible effects. First, there is the direct effect of an increase on claim costs once policies are sold. Second, an increase in prices of other goods competes with insurance for expenditures. It is likely that the inflationary impact on claims expenses dominates the effect of increasing prices of other goods as the short-run demand for insurance is relatively price inelastic. The negative impact of a positive shock implies that the income effect dominates the substitution effect when the price of other goods increases.

3.5 Methodology and econometric considerations

The time series of the economic data will receive a great deal of attention, particularly in regard to the stationarity of the data. A variable is said to be stationary if its mean and variance are constant over time and its covariance are functions only of the lag length and not of time alone. More mathematically, the sequence Y_t, $t = 1,2,...$ is stationary if $E(Y_t) = \mu$, $var(Y_t)$, and $cov(Y_t, Y_{t-s}) = \sigma_s$ for $s \geq 1$.

We say that a series is integrated of order k if it needs to be differenced k times to become stationary. Thus, if a variable Y_t requires differencing once in order to achieve stationarity, then Y_t is intergrated of order one, which is denoted $Y_t \sim I(1)$. The first difference of Y_t, denoted $\Delta Y_t = Y_t - Y_{t-1}$, is therefore stationary and we may write $\Delta Y_t \sim I(0)$. The process Y_t is then said to contain a unit root.

The most popular empirical test to detect the presence of non-stationarity in a time series is the Dickey-Fuller (DF) test. This is a test of the hypothesis $\beta = 0$ against the alternative $\beta < 0$ in the equation

$$\Delta Y_t = \beta Y_{t-1} + u_t \qquad (3.18)$$

where u_t is a stationary random disturbance. Clearly, if $\beta = 0$, then $\Delta Y_t = u_t$, which is stationary, and hence $Y_t \sim I(1)$. If $\beta > 0$, the process Y_t is explosive, but $\beta < 0$ ensures that $Y_t \sim I(0)$, since $Y_t = (1 + \beta)Y_{t-1} + u_t$ and $(1 + \beta) < 1$ in this case. The statistic used for testing the null hypothesis $H_0 : \beta = 0$ against the alternative

$H_1 : \beta < 0$ in equation (3.18) is the ratio $t = \beta/\sigma(\beta)$, where $\sigma(\beta)$ denotes the standard error of β. Because Y_t is non-stationary under the null hypothesis, this statistic no longer has the conventional t-distribution, and so the critical values derived by Fuller (1976) must be used. The DF test can also be conducted with a constant term and a time trend in equation (3.18), which in turn constitutes a different set of critical values. The augmented Dickey-Fuller (ADF) test includes lagged dependent variables in equation (3.18) in order to "whiten" the residuals and is also based on the usual t-statistic using the appropriate critical values in Fuller (1976).

An important concept that arises from the modeling of integrated time series is that of co-integration. Consider a vector X_t of n random variables, each of which has been found to be $I(k)$. These variables are said to be cointegrated if a linear combination of these variables is integrated of a lower order than k, that is, if $\alpha' X_t \sim I(k-b)$, where $b \geq 1$. The n-vector α is known as the cointegrating vector and provides information about the stable (long-term) relationship between the non-stationary elements of X_t. In empirical investigations many economic time series have been found to be $I(1)$, and so tests for co-integration focus on whether there exists a linear combination of the variables which is $I(0)$ (stationary).

The results of the unit root testing for each coverage line are presented in Table 3.1. The Bayesian information criteria were used to determine the order of each equation. The results of the unit root tests partition the general insurance industry into two components - stationary and non-stationary.

Table 3.1

Augmented Dickey-Fuller Unit Root Tests for 8 lines of General Insurance

Stationary variables	*t-value*	*N*	Lags
Accident	-3.2235	17	Level
Marine	-6.0415	17	Level
Miscellaneous	-17.9140	17	Level
Third party	-3.9273	17	Level
Transport	-10.2794	17	Level
Sector	-4.1563	17	Level
Non-stationary variables	*t-value*	*N*	Lags
Aviation	-4.3995	17	1
Motor	-3.6298	17	1
Property	-4.6360	17	1

The test and analysis of UK underwriting cycle is performed in two stages. First, tests are performed to determine whether the underwriting cycle exists in UK and in different lines of business. In the second stage, we analyze the relationship between premium changes and market/institutional features of the country and the relationship between cycle period lengths and these same features.

The first stage consists of estimating equation (3.1) using the annual aggregate and by-line underwriting profit margins as the dependent variables. The equations are estimated using ordinary least squares. The fitted line measures the long-term equilibrium path that exists between the variables. The deviations from the line (error terms) represent the short-term movements about equilibrium. Further, a linear trend variable t will be added to equation (3.1) to control for the downward trend in expenses over time. However, as in Cummins and Outreville (1987), the resulting coefficient of this trend variable will not be discussed because it is just a control variable. The time frame of the test will be from 1980 to 1998, as permitted by data availability.

The strength of this model is that it can provide a simple yet formal way to identify the underwriting cycle. A caveat to this model is that it can also be applied under the assumptions of the Venezian-type hypothesis. Therefore, significant results obtained from this model can only prove the existence of a cyclical behaviour. The causes of the behaviour cannot be distinguished. The period of the cycle, if a cycle is observed, is estimated from equation (3.2). Independent regressions are run for each line of business for which data are available.

After establishing the presence of underwriting cycles in the market, we investigate the causes of the underwriting cycle under the rational expectations/institutional intervention hypothesis. We examine the relationship between premium changes and the variables associated with this hypothesis (Lamm-Tennant and Weiss, 1997). Underwriting cycles in the United States have been associated with wide swings in insurance prices or premiums from year to year. If losses really are exogeneous, then the manifestation of the underwriting cycle would be linked directly to premiums such that the variables hypothesized to determine underwriting cycles will act directly through premium changes. In fact, previous research on underwriting cycles attempts to determine whether cost-related factors can explain premium changes (e.g. Cummins, Harrington, and Klein, 1992). For example, premiums are affected by discount rates since discounted expected losses are incorporated in the premium. Premiums will be directly affected also by the cost (and supply) of capital, while data used to determine premiums (specifically expected losses) incorporate directly any lags attributable to regulation, data collection, and accrual accounting (i.e., the smoothing of earned premiums and incurred losses over adjacent years). As such, a pooled cross-section time series will be estimated.

The specification of the generalized least squares (GLS) equation is as follows:

$$\Delta P_t = \alpha + \beta_1 \Delta Loss_{1,t} + \beta_2 \Delta Loss_{2,t} + \beta_3 \Delta Loss_{3,t} + \varepsilon_t \qquad (3.19)$$

where

ΔP_t: $\ln(\text{Premiums written})_t - \ln(\text{Premiums written})_{t-1}$

$\Delta Loss_{1,t}$: $\ln(\text{Claims paid})_{t-1} - \ln(\text{Claims paid})_{t-2}$

$\Delta Loss_{2,t}$: $\ln(\text{Claims paid})_{t-2} - \ln(\text{Claims paid})_{t-3}$

$\Delta Loss_{3,t}$: $\ln(\text{Claims paid})_{t-3} - \ln(\text{Claims paid})_{t-4}$

ε_t: a random error term

All in all, equations (3.17), (3.18), and (3.19) will be estimated for the general insurance industry as a whole. The by-line results of each sector will also be estimated as they are essentially more meaningful and desirable (Lamm-Tennant and Weiss, 1997).

3.6 Empirical results

Table 3.2 summarises the cycle periods for UK general insurance business and five major lines. The cycle periods estimated from the underwriting profits ranges from 4.1082 years for the third party liability lines to 5.1295 years for the accident & health lines. In general, the results suggest that the underwriting cycles are present in all lines.

Table 3.2

Cycle periods by line of insurance, 1980 through 1998

Line	a_1	a_2	$a_1^2 + 4a_2$	Period	R-Bar-squared
Sector	1.1770	-0.7414	-1.5803	**4.8694**	0.69371
	(5.5819)	(-3.3864)			
Accident & Health	0.8741	-0.5051	-1.2563	**5.1295**	0.41718
	(3.3623)	(-1.9664)			
Marine	1.4447	-0.7470	-0.9007	**4.2113**	0.83802
	(7.1634)	(-3.6742)			
Miscellaneous	1.1791	-0.6318	-1.1370	**4.6304**	0.62393
	(4.7256)	(-2.4065)			
Third party	1.1274	-0.4319	-0.4567	**4.1082**	0.61406
	(4.0864)	(-1.5369)			
Transport	1.3247	-0.7093	-1.0822	**4.4359**	0.73540
	(5.6586)	(-2.6935)			

To examine the causes of the underwriting cycle under the rational expectations/institutional intervention hypothesis, we performed the tests for the sector and for all lines separately. Table 3.3 presents the results of the generalized least squares regression analysis of equation (3.19) for years 1980 through 1998. The results are based on overall premium changes for the UK market and for the five lines of business. These lines were included in the analysis so that we can investigate whether the independent variables affect the premium changes of different lines differently.

Table 3.3

Results of the Generalized Least Squares Regression Analysis

Variable	Expected sign	Sector	Accident	Marine	Misc	Third party	Transport
Intercept	N.A.	-0.1447	-0.2102	-0.1653	-0.19613	-0.0306	-0.077972
		(-7.7579)	(-7.0337)	(-4.5461)	(-3.2135)	(-0.44083)	(-3.2782)
$\Delta Loss_{1,t}$	+	0.4048	0.3575	0.1121	-0.02006	0.1588	0.15849
		(2.2904)	(2.2969)	(0.70233)	(-0.1024)	(0.62685)	(1.1120)
$\Delta Loss_{2,t}$	+	-0.2118	0.0529	0.0341	-0.00215	-0.2688	0.086289
		(-0.93153)	(0.29964)	(0.16211)	(-0.0108)	(-0.86276)	(0.42497)
$\Delta Loss_{3,t}$	+	0.2349	-0.0216	0.6959	0.07920	-0.4802	0.27102
		(1.4332)	(-0.11855)	(3.6712)	(0.44616)	(-1.5616)	(1.4137)
R-bar squared		0.3557	0.4669	0.6007	0.28547	0.3063	0.23196

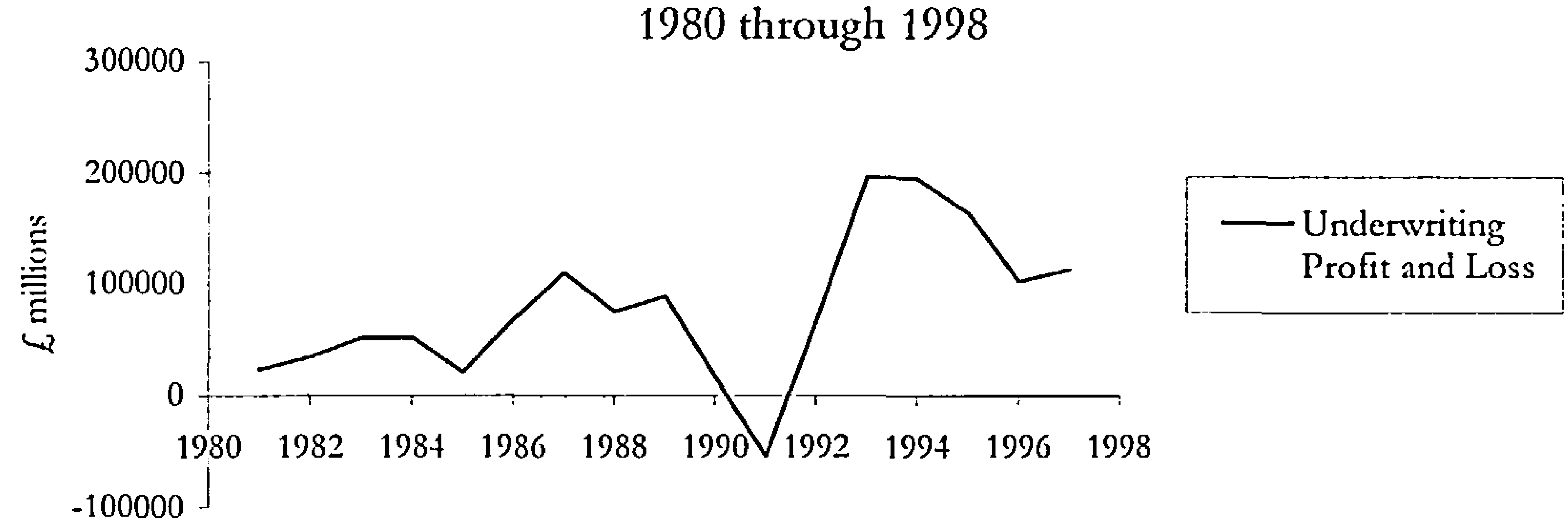
The results reported in Table 3.3 show that changes in the overall premium level are significantly related to the changes in one-year lagged losses. Similarly, the changes in the premiums of the Accident & Health business are significantly related to the changes of one-year lagged losses, while changes in the Marine premiums are only significantly related to the three-year lagged losses. The explanatory power of the model for the aggregate premiums is about thirty five percent and for the by-line premiums ranges from twenty five percent to sixty percent.

Generally, the results of our study are rather different from the results of Lamm-Tennant and Weiss (1997). For example, variables like changes in the one- and two-year lagged losses are significant in their study but are generally less so in our study. We use the underwriting profits from the UK general business to verify the correctness of the length of the cycle period calculated. Figure 3.1 plots the trend of the underwriting profits of the UK general insurance industry from 1980 to 1998. The cycles are 1980 to 1985, 1986 to 1991, and 1992 to 1996 respectively, which gives a cycle length of 5 years. This is very consistent with the cycle period of 4.8694 calculated for the market using equations (3.1) and (3.2). This proves that the methodology proposed by Cummins and Outreville (1987) to determine the existence of the cycle and cycle period is fairly accurate and acceptable.

Figure 2.1

Underwriting Profits of General Insurance Industry

1980 through 1998



Among the five major lines tested in UK, all of them exhibit a cycle (Table 3.3). The cycle length is very similar to that of the overall insurance business and ranges between 4.1082 years (for Third Party Liability line) and 5.1295 years (for Accident & Health line). The explanatory power of the model based on the aggregate sector data and the individual lines is good. The R-squared ranges from 0.417 for Accident & Health to 0.735 for Transport.

3.7 Conclusions

The main objective of this study is to investigate the presence and causes of the underwriting cycle in the United Kingdom and reveals several interesting findings.

First, the results of the second-order autoregressive model largely support the existence of the underwriting cycle in the UK because underwriting cycles are found in the aggregate sector and all the lines that were tested.

Second, the analysis of premium changes provides some support for the rational expectations/institutional intervention hypothesis although it is not able to gather enough evidence for the hypothesis that composite data collection, regulatory, policy renewal and accounting lags (Cummins and Outreville, 1987) have caused the underwriting cycle.

Third, the results generally differ from those found for the developed countries by Lamm-Tennant and Weiss (1997). This could be due partly to the fact that economic developments in these countries are different from the UK. It could also be due partly to the different level of regulatory control prevailing in these countries.

As UK is deemed to continue as the largest insurance and reinsurance financial centre, our findings pertaining to the underwriting cycle in UK would be useful to the existing insurers as well as those seeking to invest in the UK insurance market. One of the important findings from this study for the existing insurers and prospective entrants is that although the underwriting cycle does exist in the UK, the causes of it are different from those found in other countries. Therefore, they should take into account the differences when they enact measures to circumvent the detrimental effects of the underwriting cycle in the UK.

One of the shortcomings of this study is that, due to the lack of data and information, the analysis of premium changes is not able to provide support for the hypothesis that the underwriting cycle is caused by the institutional lags as advocated by Cummins and Outreville (1987). Further, since only five lines of the insurance industry are included in this study, future research can extend the underwriting cycle test and analysis of premium changes to more lines when data are available. The cycle length of each line could be further analysed by using the cycle period analysis model proposed by Lamm-Tennant and Weiss (1997). Future research in this area could be particularly interesting if there are differences in the institutional structure and regulatory oversight of the industry in the UK.

CHAPTER 4

COST OF CAPITAL AND CREDIT RISK IN THE INSURANCE INDUSTRY

4.1 Introduction

Over the past decade, financial institutions have developed and implemented a variety of sophisticated models in order to capture the market risk in their transactions. Much more recently, important steps were taken towards the modelling of credit risk faced by these institutions. Many institutions have applied the tools of financial engineering to the problems of credit risk management. New and powerful techniques have been developed to estimate the credit exposures of individual financial transactions and of entire portfolios, to incorporate credit risk into the pricing of different instruments, and to manage credit risk efficiently by separating it from other risks and selectively transferring it to other institutions.

Historically, credit risk analysis is an expert system that relies, above all, on the subjective judgement of trained professionals. However, the detection of company operating and financial difficulties is a subject which has been particularly susceptible to financial ratio analysis. Traditional credit scoring systems can be found in virtually all types of credit analysis. The idea is essentially the same: pre-identify certain key factors that determine the probability of default and combine or weight them into a quantitative score. In some cases, the score can be used as a classification system: it places a potential borrower into either a good or a bad group, based on a score and a cut-off point.

Tools from statistics and operations research, such as survival analysis, neural networks, mathematical programming, deterministic and probabilistic simulation, and game theory, have all contributed to the progress in credit risk measurement. Financial and option pricing theories have been extensively used in the construction of appropriate models.

The ability to classify and identify financial distress is important to regulators, legislators, shareholders, auditors, and even the general public. Moreover, insolvency prediction and credit risk models can help in identifying whether a company is in danger of failing and can also help auditors and regulators decide whether the company is a "going concern".

Classic credit analysis, a system carefully nurtured by banks over many years and at great expense, provides a model that non-bank financial institutions continue to emulate. To that extent, fund managers and insurance companies that lack core credit skills turn to bankers to supply the missing expertise. Moreover, insurance regulators, following the path of banking regulation, have established certain risk-based capital (RBC) standards in order to provide a cushion against unexpected increases in liabilities and decreases in the value of assets.

Credit risk models are important today because they provide the decision maker with insight or knowledge that would not otherwise be readily apparent or that could be marshaled only at prohibitive cost. In a marketplace where margins are fast disappearing and the pressure to lower costs is unrelenting, models give their users a competitive edge. In any large financial institution that has a wide variety of exposures, operates in many geographic regions, and has a large and varied workforce, quantitative models can inject a useful degree of objectivity. Moreover, credit risk models are also used to assist in releasing the value of financial assets that would otherwise be hidden from equity investors. Since structured finance products reallocate credit risk in such a way that the subordinated pieces offer a combination of equity risk and equity return, credit risk models may be used in the stratification or construction of such portfolios.

4.2 Early credit risk research

In recent decades, a number of objective, quantitative systems for scoring credits have been developed. In univariate accounting-based credit-scoring systems, the credit analyst compares various key accounting ratios. The univariate approach enables an analyst starting an inquiry to determine whether a particular ratio differs markedly from the norm for its industry. In reality, however, the unsatisfactory level of one ratio is frequently mitigated by the strength of some other measure. A firm, for example, may have a poor profitability ratio but an above-average liquidity ratio. Another limitation of the univariate approach is the difficulty of making trade-offs between such weak and strong ratios.

Although univariate models are still in use today in many institutions, most academics and an increasing number of practitioners seem to disapprove of ratio analysis as a means of assessing the performance of a business enterprise. Many respected theorists downgrade the arbitrary rules of thumb (such as company ratio comparisons) that are widely used by practitioners and favour instead the application of more rigorous statistical techniques. In some respects, however, these latter techniques should be viewed as a refinement of traditional ratio analysis rather than as a radical departure from it.

One of the classic studies of ratio analysis and bankruptcy was performed by Beaver (1967). Beaver found that a number of indicators could discriminate between matched samples of failed and non-failed firms for as long as five years prior to failure. In a subsequent study, Deakin (1972) utilised the same 14 variables that Beaver analysed but applied them within a series of multivariate discriminant models. Although Deakin achieved a high classification accuracy in the development sample (more than 95 percent for the first three years prior to failure) there was substantial deterioration in the classification accuracy in the hold-out sample one year prior. The significance of this finding is that it is premature to conclude from test results from a development sample that a valid empirical relationship has been detected.

In general, ratios measuring profitability, liquidity, and solvency appeared to be the most significant indicators in univariate studies. The order of their importance was unclear, however, because almost every study cited a different ratio as the most effective indicator of impending problems. An appropriate extension of the univariate studies, was to build upon the findings by combining several measures into a meaningful predictive model.

Altman's Z-score model was a multivariate approach built on the values of both ratio-level and categorical univariate measures (Altman, 1968). These values are combined and weighted to produce a measure (a credit risk score) that best discriminates between firms that fail and those that do not. Such a measure is possible because failing firms exhibit ratios and financial trends that are very different from those of companies that are financially sound.

The Z-score model was constructed using multiple discriminant analysis (MDA), a multivariate technique that analyses a set of variables to maximise the between-group variance while minimising the within-group variance. This is typically a sequential process in which the analyst includes or excludes variables based on various statistical criteria. It should be noted that if the groups are not very different at the univariate level, a multivariate model will not be able to add much discriminatory power.

Discriminant analysis provides a procedure for assigning sample cases to predetermined populations and then determining the accuracy of the classification procedure. Assuming that the status of a firm is a function of multivariate normal variables allows tests for significance between group mean-profiles. Additionally, if the variances of the financial variables of the distressed firms equal the variances of the financial variables of the solvent firms, and if the covariances between the financial variables of both groups are also equal, a linear classification model is optimal.

When discriminant analysis is used to classify companies into groups, correct classification may be due to three different factors:

1. Real differences between groups;
2. Sampling errors; and
3. Intensive search for the variables that give the best results for the sample used.

The objective of validating the discriminant model is to determine that the results are due to real differences between group means. Initially a discriminant function is constructed which combines a set of variables in such a manner as to maximise the differences between two group means, and that minimises the likelihood of misclassification. The original discriminant function is:

$$Z = V_1X_1 + V_2X_2 + \dots + V_nX_n \tag{4.1}$$

where $V_1, V_2, \dots, V_n$ are the discriminant coefficients and $X_1, X_2, \dots, X_n$ are the independent variables.

To arrive at a final profile of variables, the following procedures are utilised:

(a) Observation of the statistical significance of various alternative functions, including determination of the relative contributions of each independent variable;

(b) Evaluation of inter-correlations among the relevant variables;

(c) Observation of the predictive accuracy of the various profiles; and

(d) Judgement of the analyst.

From the original list of 22 variables, the final Z-score model chosen was the following 5-variable model:

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.999X_5 \quad (4.2)$$

where

$$X_1 = \frac{\text{Working Capital}}{\text{Total Assets}},$$

$$X_2 = \frac{\text{Retained Earnings}}{\text{Total Assets}},$$

$$X_3 = \frac{\text{Earnings before interest and Taxes}}{\text{Total Assets}},$$

$$X_4 = \frac{\text{Market value of equity}}{\text{Book value of total liabilities}}, \text{ and}$$

$$X_5 = \frac{\text{Sales}}{\text{Total Assets}}.$$

The Z-score model's overall classification accuracy was 95 percent one year before bankruptcy on the development sample, and 82 percent two years before. Classification accuracy is one of the outputs examined in ascertaining whether a model will perform well in practice. This accuracy is expressed as Type I error (the accuracy with which the model identified failed firms as weak) and Type II error (the accuracy with which the model identified healthy firms as such). Overall accuracy is a combination of Type I and Type II errors. Generally, Type I error is viewed as more important than Type II error, because the inability to identify a failing company will cost far more than the opportunity cost of rejecting a healthy company as a potential failure.

Because the results based on the development sample suffer from sample bias, secondary sample testing is extremely important. One type of testing is to estimate parameters for the model using only a subset of the original sample and then classify the remainder of the sample based on the parameters established. A simple t-test is then applied to test the significance of the results. Five different replications of the suggested method of choosing subsets (16 firms) of the original sample were tested. The five replications include:

(a) Random sampling;

(b) Choosing every other firm, starting with the first firm;

(c) The same test, but starting with the second firm;

(d) Choosing firms 1 through 16; and

(e) Choosing firms 17 through 32.

All the results showed that the discrimination function was statistically significant. Additional tests using secondary samples (completely independent of the development sample) were performed. Type II errors ranged from 15 to 20 percent in the secondary samples.

In order to score privately held companies, Altman (1993) revised the original Z-score model by substituting book value for market value when calculating the ratio X_4. He arrived at the following Z'-score model:

$$Z' = 0.717X_1 + 0.847X_2 + 3.107X_3 + 0.420X_4 + 0.998X_5 \tag{4.3}$$

The univariate F-test for the book value of X_4 is lower than the level of the market value, but the scaled vector results show that the revised book value measure was still the third most important contributor. Indeed, the order of importance (i.e., X_3, X_2, X_5, X_4, and X_1) was retained in the private firm model.

In 1977, Altman, Haldeman, and Narayanan (1977) presented a second-generation model with several enhancements to the original Z-score approach. Their purpose was to construct a measure that explicitly reflected recent developments involving business failure. Because the average size of bankrupt firms has increased dramatically, the new study focused on larger firms, with an average of $100 million in assets two years prior to failure. In addition, the new study reflected the most recent changes in financial reporting standards and accepted accounting practices. It also incorporated refinements in discriminant statistical techniques.

The new model, which was named ZETA, was effective in classifying bankrupt companies up to five years prior to failure, with over 90 percent accuracy one year prior and over 70 percent accuracy up to five years prior to failure. The inclusion of retailing firms in the same model as manufacturers did not appear to affect the results negatively. Twenty seven variables were selected for inclusion in the analysis, based on their widespread use in credit analysis. The variables were classified as measures of profitability, coverage and other earnings relative to leverage, liquidity, capitalisation ratios, earnings variability, and miscellaneous.

The model not only classified the test sample well but also proved the most reliable in various validation procedures. These are its seven variables:

- Return on assets, measured by earnings before interest and taxes / total assets.
- Stability of earnings, indicated by a normalised measure of the standard error of estimate around a 5- to 10-year trend in X_1.
- Debt service, measured by the familiar interest coverage ratio, that is, earnings before interest and taxes / total interest payments.
- Cumulative profitability, measured by the firm's retained earnings (balance sheet / total assets).
- Liquidity, measured by the familiar current ratio.
- Capitalisation, measured by common equity / total capital.
- Size, measured by the logarithm of the firm's total assets.

When ZETA was developed, the ratios that were included in the model were carefully chosen as ones that would not be expected to change over time. With respect to their ability to identify distressed companies, table 4.1 shows the means and F ratios of the model variables in the development sample, and the same ratios for 480 bankruptcies over the period 1981-1993 compared with randomly paired companies that did not fail.

Table 4.1

ZETA Ratio Statistics

Development Sample (1977) and 1981-1993 Sample

	1977 data			1981-1993 data		
Variable	Bankrupt group mean	Non-bankrupt group mean	F ratio	Bankrupt group mean	Non-bankrupt group mean	F ratio
Return on assets	-0.0055	0.1117	54.3	-0.08223	0.09253	316.36
Stability of earnings	1.6870	5.7840	33.8	0.88471	3.83302	247.28
Debt service	0.9625	1.1620	26.1	0.87261	1.09928	156.72
Cumulative profitability	-0.0006	0.2935	114.6	-0.21484	0.21139	559.55
Liquidity	1.5757	2.6040	38.2	1.13783	2.20532	190.48
Capitalisation	0.4063	0.6210	31.0	0.47803	0.58502	44.69
Size	1.9854	2.2220	5.5	1.63024	2.01598	40.07

Source: ZETA Services, Inc.

It is seen that the ratios continue to demonstrate consistency and contrast between failed firms and non-failed firms. The ratio means are of the same order of magnitude, and the F ratio continues to be statistically significant at the 0.001 level for all of the variables. F ratios in the latter period are much higher, which reflects the much greater sample size (480 vs. 53 in 1977).

4.3 Credit risk research in the insurance area

Early studies on financial distress in the property-liability industry lacked methodological and statistical verification and were mostly descriptive (Denenberg, 1967; Evans, 1968; Nelson, 1971). Also, Kenny's (1967) tests, such as the surplus ratio ("2 for 1" rule) and other measures of performance, were criticised as "rules of thumb".

Following the path of studies in the other types of corporations, a number of empirical studies have compared statistical models that use insurers' financial data to predict insolvencies in the property-liability insurance industry.

Trieschmann and Pinches (1973), using MDA, performed the first study on predicting financially distressed property-liability insurers. In the case of property-liability insurance firms, the objective is to classify them by using a set of reasonably independent financial variables into groups called distressed and solvent. In their study, discriminant analysis was employed to classify property-liability insurance firms into one of two groups – distress or solvent – based on their financial characteristics. In this study, the discriminant model identified that the variables which discriminate and identify the insurance firms with a high potential for financial distress were:

- The Agents Balances / Total Assets ratio
- The Stocks-Cost (preferred and common) / Stocks-Market (preferred and common) ratio
- The Bonds-Cost / Bonds-Market ratio
- The Loss Adjustment Expenses Paid + Underwriting Expenses Paid / Net Premiums Written ratio
- The Combined ratio

➢ The Premiums Written Direct / Surplus ratio

In order to determine the relative contribution of each variable to the final multivariate model, they adjusted the discrminant coefficients for differences in the units of measure of the original variables. This analysis showed that the Agents Balances / Total Assets ratio was the most important variable followed by the Premiums Written Direct / Surplus ratio.

Pinches and Trieschmann (1974) used the same sample to examine the efficiency of univariate versus multivariate financial ratio models for solvency surveillance. The MDA outperformed the univariate models in identifying financially distressed insurers. Cooley (1975), using prior probabilities for populations of solvent and insolvent firms, as well as the relative misclassification costs in prediction, found the impact of both to be substantial.

In a subsequent study, Pinches and Trieschmann (1977) examined the impact of three separate factors influencing classification results obtained from discriminant analysis – multivariate normality, equality of the variance/covariance matrices and misclassification error rates – by using data from their previous study (1973). They also illustrated that many different results are possible from the same MDA model. They examined the six variables that found they discriminate and identify the insurance firms with a high potential for financial distress for both univariate and multivariate normality. They found that the distributions for five of the six variables employed were skewed to the right, and all six of the variable distributions are leptokurtic (more peaked with higher tails than a normal distribution). Since the data were not univariate normal, they obviously could not be considered to be multivariate normal.

Furthermore, in order to test for the equality of the variance/covariance matrices, a χ^2 statistic was employed. The conclusion was that the two dispersion matrices were not equal. The next step was to employ a quadratic classification procedure instead of the linear rule. The result was that the linear rule did slightly better overall. The result was that linear and quadratic rules could produce significantly different classification results that are directly related to the differences in the dispersion matrices, the number of predictor variables, and the separation between the two groups.

Finally, in order to test for any classification error rates, they employed the jack-knife procedure. The essence of this procedure is to omit each observation sequentially, calculate a classification function based on the remaining N – 1 observations, and then classify the omitted observation. They found that the percentages for both the linear and the quadratic rule dropped, but the results from their previous analysis were also upward biased, so the actual probability of misclassification should have been between reclassification results and jack-knife results.

The Best's ratings were also viewed as surrogates for degrees of insolvency. Harmelink (1974) used MDA to predict the degree of insolvency among property-liability firms as measured by a decline in Best's policyholder's ratings. After eliminating the highly correlated variables and classifying the financial ratios into groups such as performance ratios, debt-related ratios, asset-liquidity ratios and turnover ratios, he found that a combination of seven variables was discriminating better than any larger set of variables.

The seven variables were:

1. Net income to total assets
2. Combined loss and expense ratio

3. Cash and investment assets to total liabilities
4. Net worth to net premiums earned
5. Total assets to net premiums earned
6. Investment income to total assets
7. Organization type (stock or mutual)

The reported results indicated that multiple discriminant analysis was an effective technique in predicting the event of interest, at least up to 4 years in advance of the event (by the fifth year prior to the decline in rating, the results were no longer significant).

Ambrose and Seward (1988) incorporated Best's general policyholder rating and financial size rating with variables created from a firm's readily available financial information. The rating variables were then used to alter the prior probabilities of classification under multivariate linear discriminant analysis. They found that the insolvency prediction abilities of Best's ratings and sets of financial ratios were statistically equivalent. Their results validated the practice of evaluating insurer health using Best's ratings. They also found that the predictive ability could be improved by using a two-stage prior probability approach. In the first stage, MDA analysis based on Best's ratings is used in order to determine the levels of Type I and II errors. In the second stage, these levels are used as the prior probabilities of solvency when classifying the same sample on the basis of the set of financial ratios.

However, a number of issues need to be raised here. Most studies in the property-liability industry have used MDA while ignoring its potential problems, which include violation of the normal distribution assumptions on the variables, unequal covariance matrices, and the lack of a screening-out procedure for insignificant variables through significant tests on the single-univariate coefficients (thus, standard t-tests of significance are not applicable).

Moreover, the model is linear whereas the path to bankruptcy may be highly non-linear (the relationship between the X_i's is likely to be non-linear as well). Second, the model is essentially based on accounting ratios. In most countries, accounting data appear only at discrete intervals (e.g., quarterly) and are generally based on historic or book value accounting principles. It is also questionable whether such models can pick up a firm whose condition is rapidly deteriorating (e.g., in the recent Asian crisis). Indeed, as the world becomes more complex and competitive, the predictability of simple Z-score models may worsen. A good example is Brazil. When fitted in the mid-1970's the Z-score model did a quite good job of predicting default even two or three years prior to bankruptcy (Altman, Baidya, and Dias, 1977). However, more recently, even with low inflation and greater economic stability, this type of model has performed less well as the Brazilian economy has become more and more open (Sanvicente and Bader, 1996).

The insurance regulatory information system (IRIS), developed by the National Association of Insurance Commissioners (NAIC) during the 1970s, classifies insurers with four or more of eleven financial ratios outside of specified ranges as priority firms for immediate regulatory scrutiny. Thornton and Meador (1977) concluded that the IRIS tests were not reliable indicators for insolvency prediction. Hershbarger and Miller (1986) used MDA to examine the ability of the IRIS ratios to discriminate between sound, priority, and insolvent insurers. They concluded that the IRIS test includes a number of ratios that have very little ability to distinguish between solvent and insolvent companies.

The NAIC also calculates a broader set of ratios known as the Financial Analysis and Tracking System (FAST) and recently adopted risk-based capital standards for both property-liability and life insurers. This system consists of a series of ratios that are multiplied by various balance sheet and income statement variables to compute risk-based capital "charges" for the principal risks facing insurers. The sum of the charges, reduced by a covariance adjustment, equals the insurer's risk-based capital. Grace, Harrington, and Klein (1993) find that, although the ratio of actual capital to RBC is negatively and significantly related to the probability of subsequent failure, relatively few companies that later failed had ratios of actual capital to RBC within the NAIC's ranges for regulatory action. Cummins, Harrington, and Klein (1995) confirm that the predictive accuracy of the RBC ratio is very low, even when the components of the ratio, rather than the overall ratio, are used as predictors. They have also reported that predictive power can be significantly improved by adding controls for insurer size and organizational form.

Eck (1982) employed a regression model in order to detect the financial distress of insurance firms. The regression results pointed out that the failed firm was characterised by high commissions, salaries, and dividends, low underwriting expenditures, poor receivables management, and underwriting losses. The test for the relative importance of the variables that were used in the analysis, showed that the combined ratio was the most important variable and the written premiums / net worth ratio was the least important. This reinforces the idea that loss reserve should not be disregarded but that other financial factors (such as underwriting expenses and agents' balances) should also be considered.

Harrington and Nelson (1986) employed another regression-based methodology to detect firms in financial distress. The procedure was based on the notion that equilibrium in the property-liability insurance market is likely to involve a tradeoff between the amount of insurance written relative to surplus and the risk of insurer investment and underwriting activities. If so, a systematic relationship is likely between premium-to-surplus ratios and insurer asset mix, product mix, and other characteristics. If a systematic relationship exists between premium-to-surplus ratios

and insurer operating and financial characteristics, insurers that deviate from the estimated relationship may have higher or lower default probabilities than the average firm with similar characteristics. The regression results showed a significant positive relationship between premium volume and premium-to-surplus ratios and a strong negative relationship between premium-to-surplus ratios and the proportion of admitted assets represented by common stocks. However, the small number of insolvent insurers analysed obviously prevents firm conclusions concerning the potential value of the method as a solvency surveillance tool.

Powers (1995) employed a diffusion process to model insurer net worth and replaced the traditional emphasis on the probability of ruin with the use of a more general concept, the expected discounted cost of insolvency (EDCI). The EDCI is applied to the problem of constructing the regulator's objective function, where it represents a component of the expected present value of all future flows of funds for the equity owners and policy owners of the insurer. This objective function is then used to solve for the optimal rate of return and the optimal loss-to-net worth ratio.

Kim, Anderson, Amburgey, and Hickman (1995) used event history analysis to examine insurer insolvencies. The rate of insolvency of property-liability insurers specified by using an exponential model. For property-liability insurers, statistically significant factors with consistent signs in various versions of the exponential model included organizational age, premium growth, investment yields, underwriting results, expense ratios, loss reserve exposure, and realized and unrealized capital gains.

4.4 Recent developments in credit risk research

Arguably, the recent application of non-linear methods such as neural networks to credit risk analysis shows promise of improving on the older vintage credit-scoring models. Rather than assuming there is only a linear and direct effect from the X_i variables on Z (the credit score), or, in the language of neural networks, from the input layer to the output layer, neural networks allow for additional explanatory power via complex correlations or interactions among the X_i variables (many of which are non-linear).

Yet, neural networks pose many problems to financial economists. How many additional hidden correlations should be included? It is entirely possible that a large neural network, including large N non-linear transformations of sums of the X_i variables, can reduce Type I and II errors of a historic loan database close to zero. However, as is well known, this creates the problem of "over-fitting": a model that explains well in-sample may perform quite poorly in predicting out-of-sample. Finally, the issue of economic meaning is probably what troubles financial economists the most. For example, there is a problem raised when trying to explain the economic meaning of an exponentially transformed sum of the leverage ratio and the sales-to-total-assets ratio.

The option pricing work of Black, Scholes, and Merton neatly has also been extensively used in the credit risk modelling area. Junior and senior capital structure claims can be understood as options. Thus, one can determine the value of a firm's equity by reference to the underlying market value of the firm.

KMV Corporation (1995) has created an approach for estimating the default probability of a firm that is based conceptually on Merton's (1974) approach. In three steps, it determines the expected default frequency (EDF) measure for a company. In the first step, the market value and volatility of the firm are estimated from the market value of its stock, the volatility of its stock, and the book value of its liabilities. In the second step, the firm's default point is calculated from the firm's liabilities. Also, an expected firm value is determined from the current firm value. Using these two values plus the firm's volatility, a measure is constructed that represents the number of standard deviations from the expected firm value to the default point (the distance to default). Finally, a mapping is determined between the distance to default rate, based on the historical default experience of companies with different distance-to-default values. In the case of private companies, for which stock price and default data are generally unavailable, KMV uses essentially the same approach by estimating the value and volatility of the private firm directly from its observed characteristics and accounting data. These estimates, however, are based on public company data.

Because an EDF score reflects information signals transmitted from equity markets, it might be argued that the model is likely to work best in highly efficient equity markets conditions and might not work well in many emerging markets. This argument ignores the fact that many thinly traded stocks are those of relatively closely held companies. Thus, major trades by "insiders", such as sales of large blocks of shares (and thus, major movements in a firm's stock price), may carry powerful informational signals about the future prospects of the firm.

In sum, the option pricing approach to bankruptcy prediction has a number of strengths. First, it can be applied to any public company. Second, by being based on stock market data rather than "historic" book value accounting data, it is forward-looking. Third, it has strong theoretical underpinnings; because it is a "structural model" based on the modern theory of corporate finance and options, where equity is viewed as a call option on the assets of a firm.

Against these strengths are some weaknesses. First, it is difficult to construct theoretical EDFs without the assumption of normality of asset returns. Second, private firms' EDFs can be calculated only by using some comparability analysis based on accounting data and other observable characteristics of the firm. Finally, it is "static" in that the Merton model assumes that once management puts a debt structure in place, it leaves it unchanged – even if the value of a firm's assets has doubled. As a result, the Merton model cannot capture the behaviour of those firms that seek to maintain a constant or target leverage ratio across time.

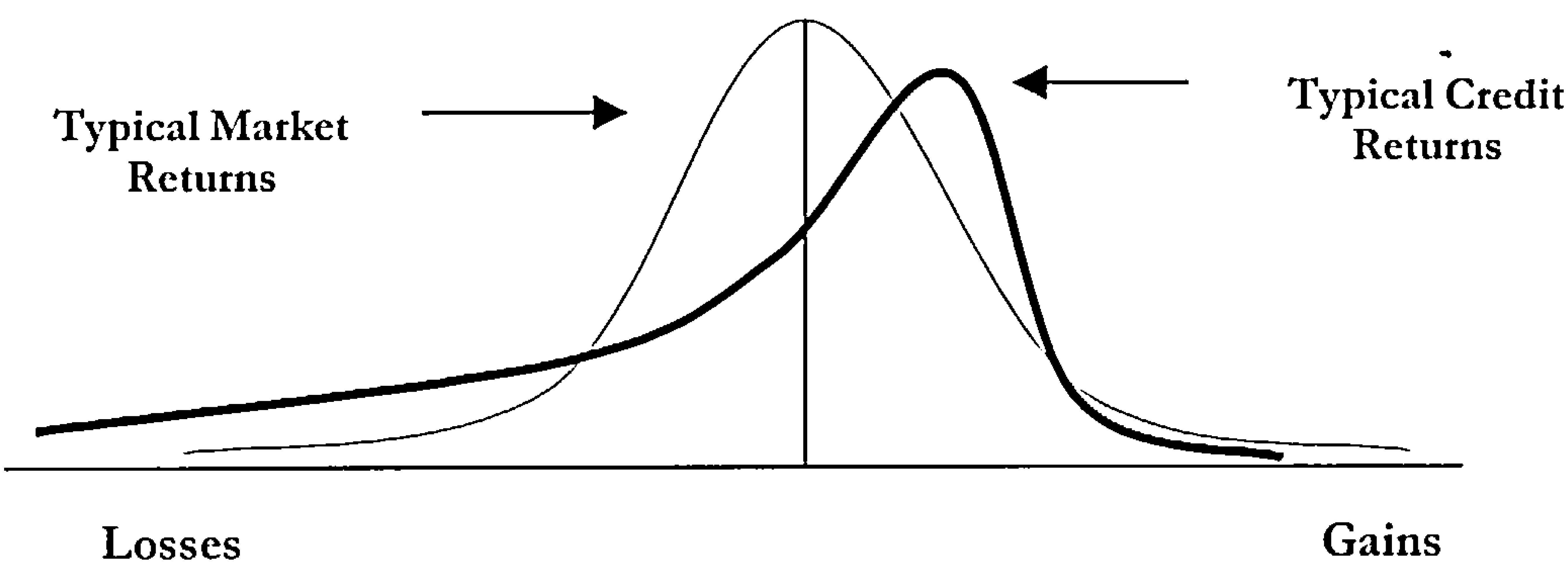
Another approach in identifying and quantifying the credit risk that is faced by a firm is J.P. Morgan's CreditMetrics (1997). CreditMetrics was introduced in 1997 as a value at risk (VAR) framework to apply to the valuation and risk of non-tradable assets such as loans and privately placed bonds. While RiskMetrics, which was developed by the same company, seeks to answer the question: "If tomorrow is a bad day, how much will I lose on tradable assets such as stocks, bonds, and equities?", CreditMetrics asks: "If next year is a bad year, how much will I lose on my loans and loan portfolio?". This methodology can be extended to any type of financial claims as receivables, loan commitments, financial letters of credit for which we can derive easily the forward value of the risk horizon, for all credit ratings.

However, because loans are not publicly traded, we observe neither the loan's market value nor the volatility of the loan value over the horizon of interest. By using available data on a borrower's credit rating, the probability that the rating will change over the next year (the rating transition matrix), recovery rates on defaulted loans, and credit spreads and yields in the bond (or loan) market, it is possible to calculate a hypothetical loan's market value and the volatility for any non-traded loan or bond, and, thus, a VAR figure for individual loans and the loan portfolio.

CreditMetrics methodology applies to portfolios and the primary reason to have a quantitative portfolio approach to credit risk management is that it addresses concentration risk more systematically. Concentration risk refers to additional portfolio risk that results increased exposure to one obligor or groups of correlated obligors. The changes in the portfolio's value are related to the eventual migrations in credit quality of the obligor, both up and downgrades as well as default.

In comparison to RiskMetrics, CreditMetrics poses two challenging difficulties. First, the portfolio distribution is far from being normal, and second measuring the portfolio effect due to credit diversification is much more complex than for market risk. While it is legitimate to assume normality of the portfolio changes due to market risk, it is no longer the case for credit returns which are by nature highly skewed and fat-tailed as shown in the Figure 4.1

Figure 4.1 Credit vs. Market returns



Source: CreditMetrics (1997)

The long downside tail of the distribution of credit returns is caused by defaults. Credit returns are characterised by a fairly large likelihood of earning a (relatively) small profit net interest earnings (NIE), coupled with a (relatively) small chance of losing a fairly large amount of investment. Across a large portfolio, there is likely to be a blend of these two forces creating the smooth but skewed distribution shape above.

Another problem is the difficulty of modelling correlations. For equities, the correlations can be directly estimated by observing high-frequency liquid market prices. For credit quality, the lack of data makes it difficult to estimate any type of credit correlation directly from history. CreditMetrics bases its evaluation on the joint probability of asset returns, which itself results from strong simplifying assumptions on the capital structure of the obligor, and on the generating process for equity returns.

A number of issues arise when we use the rating transitions assumed in the transition matrices in order to calculate the probabilities of moving to different rating categories (or to default) over the one (or more than one) year horizon.

The calculation of the transition numbers, which involves averaging one year transitions over the past data period, i.e. 20 years, is an important assumption about the way defaults and transitions occur. Specifically, CreditMetrics assumes that the transition probabilities follow a stable Markov process, which means that the probability that a bond or a loan will move to any particular state during this period independent of (not correlated with) any outcome in the past period. However, there is evidence that rating transitions are auto-correlated over time. For example, a bond or loan that was downgraded in the previous period has a higher probability (compared to a loan that was not downgraded) of being downgraded in the current period (Nickell, Perraudin, and Varotto, 1998). This suggests that a non-Markov process may better describe the rating transitions over time.

Another issue involves the transition matrix stability. The use of a single transition matrix assumes that transitions do not differ across borrower types (e.g., industrial firms versus banks, or the United States versus Japan) or across time (e.g., peaks versus troughs in the business cycle). Indeed, there is considerable evidence to suggest that important industry factors, country factors, and business cycle factors have an impact on rating transitions (Nickell et al., 1999). For example, when we examine a loan to a Japanese industrial company, we may need to use a transition matrix built around data for that country and industry.

A final issue relates to the portfolio of bonds used in calculating the transition matrix. Altman and Kishore (1997) found a noticeable impact of bond "ageing" on the probabilities calculated in the transition matrix. A material difference is noted, depending on whether the bond sample used to calculate transition is based on new bonds or on all bonds outstanding in a rating class at a particular moment in time.

Finally, quite recently ideas coming from insurance found their way into the new tools for credit risk measurement and management. Credit Suisse Financial Products (CFSP) has developed a model, called Credit Risk Plus (1997), similar to the one a property insurer selling household fire insurance might use when assessing the risk of policy losses in setting premiums. The idea is very simple, based on a portfolio of loans or bonds and their historic default experience, develop a table that can be used in a predictive sense for one-year, or marginal, mortality rates and for multiyear, or cumulative, mortality rates. Combining such calculations with loss given defaults can produce estimates of expected losses.

The model developed by CFSP stands in direct contrast to CreditMetrics in its objectives and its theoretical foundations. CreditMetrics seeks to estimate the full VAR of a loan or loan portfolio by viewing rating upgrades and downgrades and the associated effects of spread changes in the discount rate as part of the VAR exposure of a loan. Credit Risk Plus views spread risks as part of market risk rather than credit risk. As a result, in any period, only two states of the world are considered – default and non-default – and the focus is on measuring expected and unexpected losses rather than expected value and unexpected changes in value (or VAR) as under CreditMetrics. Thus, CreditMetrics is a mark-to-market model; Credit Risk Plus is a default mode model.

The second major difference is that, in CreditMetrics, the default probability in any year is discrete (as are the upgrade/downgrade probabilities). In Credit Risk Plus, default is modelled as a continuous variable with a probability distribution. An analogy from house fire insurance is relevant. When a whole portfolio of homes is insured, there is a small probability that each house will burn down, and (in general) the probability that each house will burn down can be viewed as an independent event. Similarly, many types of loans, such as mortgages and small business loans can be thought of in the same way, with respect to their default risk. Thus, under Credit Risk Plus, each individual loan is regarded as having a small probability of default, and each loan's probability of default is independent of the default on other loans. This assumption makes the distribution of default probabilities of a loan portfolio resemble a Poisson distribution.

4.5 Insurance ratings and the cost of capital

Insurance company ratings provided by private rating agencies are vitally important to investors, regulators, consumers, insurers, and insurance agents/brokers. Insurers use ratings in their advertising to assure buyers of the firm's strength. Insurance buyers use them in choosing their insurance companies and/or deciding how much they are willing to pay for insurance from particular firms. Brokers and agents often will not recommend coverage with non-rated insurers or insurers with ratings below some threshold of financial strength (Moody's, 1998), and many corporate insurance buyers require that all their insurers be highly rated. Strong financial ratings give insurers better access to capital markets. Ratings also provide a valuable tool for regulators in assessing the financial strength of insurers (Schwartz, 1994).

It should be noted that despite similarities, insurer ratings are quite different from corporate bond ratings. First, financial strength ratings are entirely optional in that there are no regulatory requirements to obtain a rating and agencies will not issue a full rating unless requested to do so by the insurer (with the exception of Standard & Poor's which does issue a type of purely quantitative rating even on insurers that do not apply for a rating). In addition, a bond rating applies to a particular debt issue, whereas an insurer rating applies to the entity itself and assesses the overall claims-paying ability of the insurer, since policyholder obligations must be met before payments are made to any other creditors or shareholders. Insurer ratings are particularly complex in that, unlike bond issues, which have fixed payments that are to be made at fixed times, claims payments involve financial obligations that are uncertain in both timing and amount.

The purposes of insurer ratings are also very different. Corporate bond ratings are used almost exclusively by investors and regulators, while the primary users of insurer financial strength ratings are insurance companies and insurance consumers and the agents/brokers who market insurance to consumers. This point is illustrated by the fact that the majority of insurers rated by A.M. Best are either mutual companies or privately-held stock firms. Additionally, the fact that many insurers with rated debt still choose to obtain one or more insurer financial strength ratings, and that these ratings often differ from their debt ratings, again illustrates the fact that bond ratings and insurer financial strength ratings do not serve the same purpose or measure precisely the same risk.

One interesting practical distinction between bond ratings and insurer ratings is that the bond rating agencies that rate a smaller proportion of bonds tend to issue higher ratings than the agencies that rate almost all bonds (Cantor and Packer, 1997), whereas the agencies (Moody's, S&P) that rate a smaller proportion of insurers tend to issue lower ratings than the agency (A.M. Best) that rates most insures. Another distinction is that there appears to be a greater divergence of opinion among rating agencies regarding insurer financial strength ratings than bond ratings (Ederington, 1986). Almost 90 percent of eligible property-liability insurers in 1995 applied for a rating from A.M. Best, while only about 18 percent applied for a rating from S&P and only 10 percent applied for a rating from Moody's. Given the overwhelming proportion of insurers that receive Best's Ratings, it appears that insurers consider obtaining at least one rating essential.

The literature on the determinants of insurer financial strength ratings is very limited. Pottier (1997) examines the determinants of Best's life insurer ratings and finds casual evidence suggestive of selection bias, but does not perform any formal tests for it or control for it econometrically. In addition, insurer ratings have been widely used as measures of insolvency risk and financial quality (Adiel, 1996; Anthony and Petroni, 1997; Cummins and Danzon, 1997; Pottier, 1998). These studies have only used Best's ratings.

Insurer ratings also have a direct impact on the cost of capital, since the primary source of debt capital to insurers is policy liabilities, and lower rated firms will likely have to sell their policies at lower prices compared to higher rated firms (Doherty and Tinic, 1981; Berger, Cummins, and Tennyson, 1992).

Since the cost of capital is a long-term concept, the intention would be to produce a figure that compensates equity investors over a long period of time. Although a company's cost of capital will change over time, the cost of capital for a company or an industry should be relatively stable from period to period unless there has been some dramatic structural reasons for the change.

An indirect way to determine the cost of capital would be to link the credit spreads of corporate bonds with the associated credit ratings for these bonds. As the credit spread compensates the holder of the debt instrument for expected losses, there should be a link between the credit spread and the credit rating class, given the fact that there exist ample evidence that rating categories indeed entail an indication of relative credit risk. Researchers have indeed shown that there exists a close relationship between credit rating classes and subsequent default experience. This is mirrored in empirical studies where it is always found that the credit spread widens at an increasing rate as credit rating worsens. This can for instance be seen in Duffee (1998) for US investment grade corporate bonds. However, the standard deviation of individuals bonds' credit spreads within a given rating category increases as the credit rating worsens. This indicates that not all bonds within the same rating class are assumed to bear the same credit risk. Apparently, the higher cross-sectional standard deviation in the lower rating classes indicates that rating agencies allow for more heterogeneity in these classes.

Theoretically, there is a relation between the term to maturity of the bond and its credit spread, which is also known as the term structure of credit spreads or credit spread curve (Merton, 1974; Jarrow, Lando, and Turnbull, 1997; Longstaff and Schwartz, 1995; Tychon, 1998). This relationship is not necessarily upward sloping. It can also be downward sloping or humped-shaped. The intuition behind the latter is that highly rated companies can hardly become better rated, but can get down-rated. The longer the term to maturity, the higher the probability that the credit rating of such a company increases. This explains that the credit spread can be lower for longer maturity bonds. Empirical evidence is available for this phenomenon (Sarig and Warga, 1989; Fons, 1994).

However, Helwege and Turner (1997) indicate that the downward sloping credit spread curve might be a consequence of using average credit spreads of bonds in a given credit rating class. A bias would be introduced when using sets of bonds with the same credit rating.

Both the seniority of a bond or loan and the collateral attached as security to it, have an impact on the credit spread because, arguably, both kinds of provisions will increase the recovery rate in case of default. Indeed, Izvorski (1997) finds that for defaulted US bonds debt seniority is one of the most important determinants of the recovery ratio, thus implying a lower yield for senior issues.

In general, the credit spreads on coupon bonds are not equal to credit spreads on zero-coupon bonds because of either a non-flat term structure or a non-flat credit spread structure. To the extent that the credit spread curve is upward sloping, higher coupon bonds will have lower credit spreads than lower coupon bonds with the same maturity (Litterman and Iben, 1991). Likewise, if the bond's duration shortens, e.g. because of an interest rate increase, the credit spread will decrease if the credit spread curve is upward sloping.

Other factors like the callability or other option features may be important to take into account. In efficient markets the value of these options is embedded in the bond's price and thus in the credit spread. The results in Duffee (1998) clearly indicate that the callability feature can dramatically change spread behaviour. Duffee finds that spreads are negatively related to changes in risk-free rates.

Also differences in liquidity may be important. To the extent that the riskless reference bond is more liquid than the risky bond, the spread between the two will also include a liquidity premium. When liquidity is measured as the issue's size, many authors find a negative relation between spread and size: the larger the size, the larger the issuer's liquidity, the lower the required yield and therefore the spread. Boardman and McEnally (1981) find a negative relation between size and yield for Baa or better-rated US corporate bonds. Also for highly levered transaction loans do Angbazo, Mei and Saunders (1998) find a negative relationship between size and their spread.

Finally, there are strong theoretical arguments to assume that there is a relation between credit spreads and the risk-free interest rate level. Duffee (1998) studies the relation between interest rate variables and credit spreads on non-callable corporate bonds. To this end he regresses monthly changes of credit spreads on changes in the three-month Treasury bill yield and changes in the slope of the term structure, as measured by the difference between the thirty-year constant-maturity Treasury yield and the three-month bill rate. Duffee finds significantly negative slope coefficients for both variables. For long maturities, the slopes of both variables are similar, thus canceling the influence of the three-month bill rate, leaving a negative association with long-term interest rate changes.

Arak and Corcoran (1996) also find a negative relation between yield spreads on privately placed issues and risk-frees rate when all variables are measured in levels. Fridson and Jonsson (1995), however, report that they did not find any relation between the level of Treasury rates and the spread on high-yield bonds, which are also below-investment grade.

Nearly all known credit risk management tools use data calibrated to the US situtation. Not only default rates, credit mitigation probabilities and recovery rates pertain to US, also empirical modeling of credit spreads is to a large extent confined to the US market.

Nevertheless, as markets become more global and corporate bond markets develop in other parts of the world, non-US data may be useful to ascertain the robustness of known results.

We obtained bond index data constructed by Lehman Brothers via their Fixed Income Database. Its indices are based on secondary market prices of bonds issued in the Eurobond market or in EMU-zone domestic markets and denominated in euro-or one of the currencies that joined the EMU. Besides direct government bond indices also investment grade corporate bond indices were used. The latter are based on publicly traded bonds, issued by companies domiciled in the European Union excludes convertible securities.

We want to focus on the behavior of credit spread through time for different ratings so we will work with the respective sub-indices produced by Lehman Brothers. All these indices are based upon the composite rating of Moody's and Standard & Poor's, if the issue is rated by both. If ratings do not coincide, an average rating is used which is 'rounded' downwards. The composition of each index is determined on the last business day of the previous month. During the month, each bond will stay in the index, regardless whether or not the bond is downgraded or upgraded. Also when bonds are called during the month, they are not removed from the index until the end of the month. A similar rule holds for changes in the amount outstanding during a month: face values are kept constant and are adjusted at the start of a new month.

Monthly data are available for the period 7/88 – 12/97. We computed the spread as the difference between the yield to maturity as reported by Lehman Brothers and a comparable government bond yield. The government bond yield is computed relative to the contemporaneous Benchmark Treasury yield curve as provided by Datastream.

In the next table summary statistics for changes in credit spreads are presented. To some extent these are more important in a risk management context, as it is important to understand how credit spreads behave through time.

Table 4.2 Summary statistics of yield spreads

		Average spread (%)	Standard deviation (%)	First order auto correlation	Second order auto correlation	Third order auto correlation	Ljung-Box Statistic at 20 lags
1-3 yrs	AAA	0.167	0.056	0.835	0.815	0.812	2876.3
	AA	0.279	0.029	0.853	0.784	0.740	2139.7
	A	0.381	0.080	0.642	0.517	0.481	851.5
	BBB	0.726	0.197	0.947	0.899	0.846	2835.1
3-5 yrs	AAA	0.184	0.051	0.937	0.915	0.897	4005.8
	AA	0.343	0.070	0.944	0.934	0.920	4804.5
	A	0.481	0.086	0.977	0.970	0.965	5255.1
	BBB	0.726	0.142	0.934	0.918	0.900	3844.1
5-7 yrs	AAA	0.230	0.048	0.934	0.919	0.898	3750.6
	AA	0.325	0.061	0.924	0.904	0.897	4321.5
	A	0.504	0.083	0.923	0.887	0.847	3442.5
	BBB	0.701	0.120	0.974	0.957	0.943	4375.3
7-10 yrs	AAA	0.311	0.081	0.975	0.964	0.951	4824.0
	AA	0.443	0.109	0.983	0.975	0.968	5273.6
	A	0.635	0.145	0.988	0.985	0.981	5506.7
	BBB	1.105	0.502	0.984	0.970	0.957	5609.9
10 + yrs	AAA	0.276	0.057	0.902	0.879	0.849	3056.4
	AA	0.420	0.095	0.971	0.952	0.943	4980.3
	A	0.733	0.187	0.991	0.987	0.984	5623.1

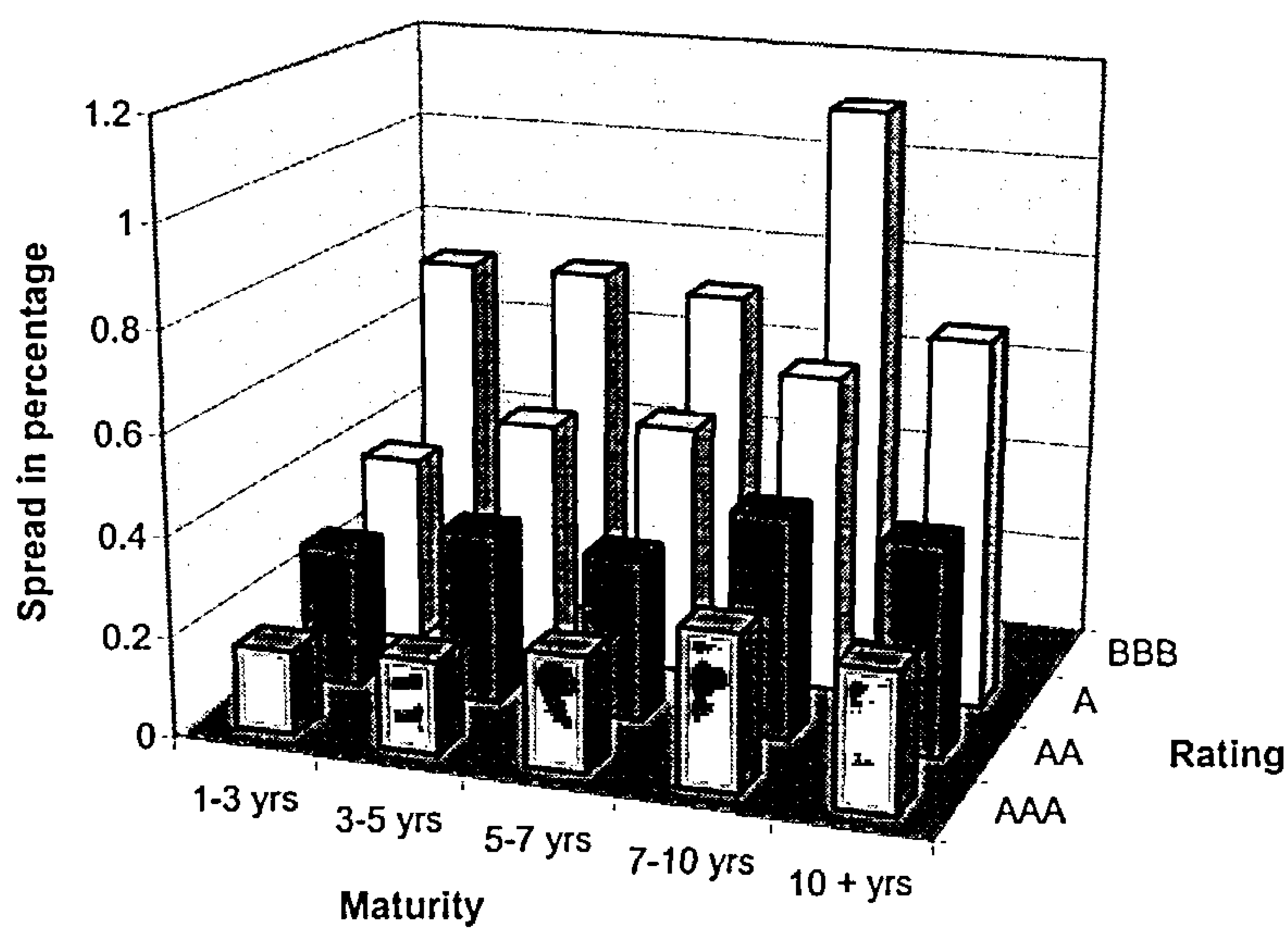
Note: Yield spreads are spreads relative to government bond yields with similar duration. Ratings are composite Moody's and Standard & Poor's ratings. Maturity buckets include the lower boundary and exclude the upper boundary. Statistics are computed from daily data over the period July 1988 through Dec 1997.

From Table 4.2, average spreads increase monotonically the lower the credit rating. This is also summarized in Figure 4.2. The relation is clearly not linear: the difference between the BBB-rated indices and AA-rated indices is generally much higher than between other adjacent rating classes. Remarkably, the spreads are considerably lower than the spreads reported by Duffee (1998). For AAA-rated bonds, our spreads vary between 17 basis points (bp) and 31 bp, whereas Duffee finds at least 67 bp. The picture is similar for the other rating categories. Of course, the sample used here covers a much sorter period than Duffee's. The average spreads do seem consistent with the findings of Pendrosa and Roll (1998) for US investment grade spreads covering the period 1995-1997. Unfortunately, they do not provide an estimate of the average spread on the US market, but the graphs they present, show spread levels similar to our data.

As far as the relation between spread and maturity is concerned, we generally find an upward sloping credit curve, as was also the case for the Duffee series (1998). Nevertheless, in some cases the relation is not monotone, e.g. '10 and more years' bucket for AAA and AA-rated bonds which can be assigned to a liquidity effect. In addition, it may be the case that relatively less credit risky issuers issue longer dated bonds.

Furthermore, as in Longstaff and Schwartz (1995b), the standard deviation of spreads is also increasing when credit rating deteriorates. No clear relation with maturity can be observed. Finally, the first lag autocorrelation coefficient of all spread series is relatively large, and often near 0.95 or higher. We do not present stationarity tests because of the reported low power of unit root tests on observation periods as short as this one. If we accept on economic grounds that the credit spreads series are stationary, it is clear from the high autocorrelation coefficients that they revert only slowly to their long-run average.

Figure 4.2 Average Credit Spreads by Rating Class and Maturity Bucket



In Table 4.3 summary statistics for changes in credit spreads are presented. It can be noticed that average changes in spreads are very small and insignificantly different from zero. This could have been expected as they measure the trend of credit spreads over the time period investigated.

The first three order autocorrelation coefficients of the spread changes are also shown. Without any exception, all first order coefficients are negative and usually significantly different from zero. This is in contrast to Duffee (1998) who uses monthly data and mostly finds positive autocorrelations. The negative autocorrelation may be a reflection of poor liquidity in the European corporate bond market. By frequently bouncing between the bid and the ask quote, negative autocorrelation may be introduced. High order autocorrelation does hardly seem present in the series.

Table 4.3 Summary statistics of yield spreads changes

		Average spread change (%)	Standard deviation (%)	Skewness	First order auto correlation	Second - order auto correlation	Third order auto correlation
1-3 yrs	AAA	0.028	3.240	-0.829*	-0.437*	-0.057	0.075
	AA	-0.020	1.587	0.629*	-0.266*	-0.086	-0.109
	A	-0.060	6.758	-0.068	-0.331*	-0.105	0.031
	BBB	0.255	6.391	2.966*	-0.068	0.040	-0.049
3-5 yrs	AAA	0.022	1.815	0.101	-0.319*	-0.048	-0.100
	AA	0.027	2.360	-0.116	-0.412*	0.026	-0.091
	A	0.033	1.861	-0.054	-0.359*	-0.026	-0.039
	BBB	0.093	5.131	-0.788*	-0.375*	0.009	0.095
5-7 yrs	AAA	0.011	1.762	-0.363*	-0.386*	0.044	-0.048
	AA	0.030	2.402	0.101	-0.371*	-0.081	-0.012
	A	0.032	3.239	-0.873*	-0.262*	0.015	-0.077
	BBB	0.086	2.766	-0.921*	-0.193*	-0.040	-0.049
7-10 yrs	AAA	0.033	1.820	-0.173	-0.283*	0.037	-0.026
	AA	0.047	2.003	-0.449*	-0.262*	-0.026	-0.029
	A	0.081	2.217	0.009	-0.350	0.022	-0.035
	BBB	-0.246	5.667	-1.446*	-0.200*	0.060	-0.054
10 + yrs	AAA	0.026	2.534	-1.050*	-0.383*	0.035	-0.042
	AA	0.062	2.290	-0.611*	-0.171*	-0.182*	-0.023
	A	0.123	2.555	0.329*	-0.293*	-0.044	0.054

Note: Yield spreads are spreads relative to government bond yields with similar duration. Ratings are composite Moody's and Standard & Poor's ratings. Maturity buckets include the lower boundary and exclude the upper boundary. Statistics are computed from daily data over the period 1 April 1998 through 17 May 1999. Asterisks denote autocorrelation, skewness or kurtosis coefficients more than two standard deviations away from zero.

From the skewness and kurtosis coefficients it is clear that spread changes are not normally distributed. This is both due to skewness and kurtosis. The skewness coefficients are statistically different from zero, and the kurtosis coefficients are significantly in excess of three, the value a normal distribution will show. Most skewness coefficients are negative. All kurtosis coefficients are significantly higher than three, which implies that the distributions have higher peaks and thicker tails than the normal distribution does.

4.6 Methodology and econometric considerations

We assume that the change in the spread $Spread_{mt}$ at time t of a bond in maturity group m can be attributed to the rating component $RATE_t$ and a unique component ε_t:

$$Spread_{mt} = \alpha + bRATE_t + \varepsilon_t$$

where $RATE$ = AAA, AA, A, or BBB, m = 1-3 yrs, 3-5 yrs, 5-7 yrs, 7-10 yrs, 10+ yrs.

Table 4.4 Relation between credit spreads and credit ratings

	Maturity groups				
	1-3 yrs	**3-5 yrs**	**5-7 yrs**	**7-10 yrs**	**10+ yrs**
Intercept	22.44	15.67	25.56	23.65	28.42
	(4.55)	(5.45)	(6.71)	(4.55)	(7.45)
Rate	-15.46	-16.87	-17.81	-18.34	-26.66
	(-6.71)	(-5.19)	(-4.20)	(-7.45)	(-7.49)
Adjusted R^2	0.43	0.45	0.59	0.62	0.57

From the above analysis, there is a clear relation between the credit spreads of corporate bonds and credit ratings. Since we have established this argument, the analysis can move forward in determining the appropriate financial ratios that convey important information about insurance companies to investors.

Different ratios can be easily computed for a public or a private company from the data in the company's balance sheet and income statement. It is often useful to have a model that combines all the relevant financial information about a company into a single number representing its credit quality. For example, such a model can be used to rate private companies in which a potential equity investment is considered. Using the same quantitative model for evaluating different companies provides us with an objective comparison of these companies.

A set of financial variables is used in order to identify the financial ratios that affect the rating of an insurance company. These ratios are hypothesised to influence the decision by the firm to obtain a rating or multiple ratings. According to the theory of financial intermediation, the principal role of credit rating agencies is the reduction of ex ante uncertainty or informational asymmetry about a firm's economic value and probability of financial distress (Ramakrishnan and Thakor, 1984; Millon and Thakor, 1985). Thus, the more likely investors, consumers, and regulators are to have different opinions about the true insolvency risk of an insurer, the greater the demand for and value of a financial strength rating.

Numerous variables are included in the model to proxy for the level of uncertainty about the firm's risk. Cantor and Packer (1997) argue that relatively high levels of underwriting profitability and leverage may be associated with greater uncertainty and thus a higher probability of obtaining an optional rating. Therefore, the model includes a leverage measure and a profitability measure. A measure of premiums is also included. While strong premium growth may be very positive for the financial health of the firm, growth sometimes is due to lower underwriting standards or under-pricing (Harrington and Danzon, 1994).

The loss reserve is probably one of the most common vehicles used to conceal the true condition of a financially troubled company. The loss reserve is an estimate of the amount that will ultimately be paid when the aggregate claims are settled; because it is an estimate, it is easy to understate it, thereby overstating the surplus. Moreover, in times of inflation, the losses may increase at a greater than anticipated rate, and the loss reserve may be understated in order to give the appearance of solvency until rates can be increased to rectify the situation. Therefore, a measure of the reserves is included in the model.

Cantor and Packer (1997) include the natural logarithm of long-term debt outstanding as a measure of the potential benefit of an additional rating. A variable is included in the model because the major benefit of obtaining a positive rating is a lower cost of debt, which should accrue in direct proportion to the amount of debt issued. Since in the present context ratings are for claim-paying ability on policies written rather than for long-term debt, the natural logarithm of net claims over net premiums is included following the same rationale. We expect that the higher this ratio is, the greater the likelihood of obtaining a negative rating.

The rest of the variables included represent factors that previous theory has indicated as important in determining insurer solvency risk. These include variables reflecting capitalisation (Kahane et al., 1986; Cummins and Derrig, 1989; Doherty, 1989), asset and liability risk (Kahane et al., 1986; Cummins and Derrig, 1989), liquidity (Kahane et al., 1986), size (Cummins et al., 1993; Cummins and Sommer, 1996), diversification

(Sommer, 1996), and reinsurance usage (Borch, 1974; Berger et al., 1992). Many of the variables are similar to measures used in articles studying the determinants of bond ratings, while others are explicitly stated as determinants of insurer ratings by the rating agencies themselves. Specifically, the variables include measures of liquidity, investment risk, use of reinsurance, leverage, and profitability.

We gathered financial information on 87 general insurance companies for years 1992-98. For each of these companies, we obtained the S&P credit ratings given in January of years 1992-1999, associating each rating number with the financial data for the prior year.

Table 4.5 Transformation of S&P's codes

S&P rating	Code
AAA	2
AA+	4
AA	5
AA-	6
A+	7
A	8
A-	9
BBB+	10
BBB	11
BBB-	12
BB+	13
BB	14
BB-	15
B+	16
B	17
B-	18
CCC+	19
CCC	21
CCC-	21

Standard & Poor's has developed credit ratings that may apply to an issuer's general creditworthiness or to a specific financial obligation. Standard & Poor's insurance ratings provide financial strength ratings that are prospective evaluations of an insurer's financial security to its policyholders. Long-term credit ratings range from 'AAA', reflecting the strongest credit quality, to 'D', reflecting the lowest. Long-term ratings from 'AA' to 'CCC' may be modified by the addition of a plus (+) or minus (-) sign to show the relative standing within the major rating categories. The S&P ratings were converted to a numerical credit code using the conversions given in the previous table (Table 4.5)

Data for financial ratios and other firm-specific variables are used from the A.M. Best's *Insight* database. In order to be included in the sample, an insurer must have financial data available on the *Insight* database necessary to calculate the various explanatory variables.

Since the data set represents cross-sectional and time-series data, we use the regression equation

$$Y_{it} = X'_{it}\beta + \eta_i + v_{it} \ , \ I = 1,\ldots, N \ , \ t = 1,\ldots, T \qquad (4.4)$$

where the subscripts i and t refer, respectively, to company number and year, Y_{it} are the credit codes, X_{it} are the financial ratios, η_i is fixed or random individual effects, and v_{it} is the error term independently and identically distributed over i and t with zero mean and variance σ_v^2. We omit time dummies for simplicity although they may capture unobserved aggregate effects.

In the case where observations on Y_{it} and X_{it} for $I = 1,\ldots,N$ and $t = 1,\ldots,T$ are available, an aggregate time series regression would threat η_i as part of the constant and thus unidentified, whilst a cross-section regression will yield a biased estimator of β if η_i is correlated with X_{it} across I.

In the case of the presence of fixed effects, β and η_i can be estimated consistently and efficiently by the following estimators which can be obtained by ordinary least squares (OLS) after the data are transformed by subtracting group means from each observation (Hsiao, 1986):

$$\hat{\beta}_{wg} = \left[\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)'\right]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i)\right] \tag{4.5}$$

and

$$\eta_i = \bar{y}_i - \hat{\beta}_{wg}\bar{x}_i, i = 1,..., N \tag{4.6}$$

where $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}, \bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$.

Because η_i is treated as a fixed constant, the estimator of β is called least-squares dummy-variable. This estimator is known also as the "within-groups estimator" or "covariance estimator", because it can be obtained using an appropriate transformation in the form of the orthogonal projection

$$Q = I_{NT} - P \tag{4.7}$$

where $P = \left[I_N - \frac{1}{T} i_T i_{\bar{T}}\right]$, with I as the identity matrix and i a column vector of ones. If we re-write (4.4) in the form

$$Y = X\beta + \eta + v \tag{4.8}$$

where Y, X, η, and v denote the (TN×1), (TN×k), (TN×1), and (TN×1) matrices respectively, we can pre-multiply (4.8) by Q yielding

$$QY = QX\beta + Qv \tag{4.9}$$

since $Q\eta = 0$. Given that X_{it} is uncorrelated with v_{is} for each s,t, then in the transformed specification QX and Qv are uncorrelated and thus we can apply OLS giving rise to the estimator

$$\hat{\beta}_{wg} = (X'QX)^{-1}X'QY \tag{4.10}$$

which coincides with (4.5) above.

In the random effects case η_i is assumed to be a random variable such that

$$u_{it} = \eta_i + v_{it} \tag{4.11}$$

where we suppose that $E(\eta_i) = E(v_{it}) = 0$ and $E(X_{it}\eta_i) = E(X_{it}v_{is}) = 0$. Under these conditions the OLS regression of Y_{it} on X_{it} (in levels) gives a consistent estimate of β.

However, the specification (4.11) generates autocorrelation and heteroscedasticity in u_{it} even if it is absent in v_{it}. This suggests that we seek a more efficient estimator than OLS. The appropriate GLS estimator of β, under these assumptions turns out to be a weighted average of the above within-groups estimator and the so-called "between-groups" estimator

$$\hat{\beta}_{bg} = (X'PX)^{-1}X'PY \tag{4.12}$$

that is

$$\hat{\beta}_{GLS} = D\hat{\beta}_{bg} + (I_k - D)\hat{\beta}_{wg} \tag{4.13}$$

where $D = (V_{bg} + V_{wg})^{-1}V_{wg}$, V_{bg} and V_{wg} are the covariance matrices of $\hat{\beta}_{bg}$ and $\hat{\beta}_{wg}$ respectively, and where the estimator $\hat{\beta}_{bg}$ can be obtained by applying OLS to the transformed model (Maddala, 1971)

$$PY = PX\beta + P\eta + Pu \qquad (4.14)$$

This GLS estimator is often known in the literature as the Balestra and Nerlove (1966) estimator, $\hat{\beta}_{BN}$ (with $\hat{\beta}_{BN} = \hat{\beta}_{GLS}$).

Hausman and Taylor (1981) show that if $\eta_i \sim iid(0, \sigma^2_\eta)$ and $v_{it} \sim iid(0, \sigma^2_v)$, a GLS transformation of (4.3) is given by

$$Y_{it} - (1 - \vartheta)\overline{Y}_i = \beta[X_{it} - (1 - \vartheta)\overline{X}_i] + \vartheta\eta_i + [v_{it} - (1 - \vartheta)\overline{v}_i] \qquad (4.15)$$

where $\vartheta = \sqrt{\dfrac{\sigma^2_v}{\sigma^2_v + T\sigma^2_\eta}}$, $\overline{Y}_i$ and $\overline{X}_i$ are once again the time means of the variables.

Estimates of σ^2_v and σ^2_η can be obtained as follows (Hausman and Taylor, 1981)

$$\hat{\sigma}^2_v = \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\tilde{Y}_{it} - \hat{\beta}_{wg}\tilde{X}_{it}\right]^2 \qquad (4.16)$$

and

$$\hat{\sigma}^2_\eta = \frac{1}{N}\sum_{i=1}^{N}\left[\overline{Y}_i - \hat{\beta}_{wg}\overline{X}_i\right]^2 - \frac{1}{T}\hat{\sigma}^2_v \qquad (4.17)$$

where $\tilde{X}_{it} = X_{it} - \overline{X}_i$ and $\tilde{Y}_{it} = Y_{it} - \overline{Y}_i$. The feasible GLS estimator can then be obtained by applying OLS to (4.15), having (4.16) and (4.17), to construct an estimate of ϑ.

When η_i is treated as a fixed constant, the model is referred to as a fixed effects model, whilst when η_i is treated as a random variable, it is called a random effects model. More generally, we can unify these two formulations, and we may assume from the outset that the effects are always random. What is crucial, however, is to investigate if η_i is correlated or not with the observed variables X_{it}. If η_i is correlated with X_{it}, the fixed effects model is viewed as one where investigators make inferences conditional on the effects that are in the sample. Whilst if η_i is not correlated with X_{it}, the random effects model is viewed as one where investigators make unconditional or marginal inferences with respect to the population of all effects.

Therefore, we have to compare the estimates in levels and deviations. Significant differences between the two indicate that correlated individual effects are omitted from the regression in levels. It is worth noting that this is equivalent to testing whether the effects are correlated or not with X_{it}.

To perform this experiment, we can use the traditional Hausman test (1978) based on the comparison between the within-groups estimator and the Balestra and Nerlove estimator:

$$h = \left[\hat{\beta}_{GLS} - \hat{\beta}_{WG}\right]' \left[\text{var}\left(\hat{\beta}_{wg}\right) - \text{var}\left(\hat{\beta}_{GLS}\right)\right]^{-1} \left[\hat{\beta}_{GLS} - \hat{\beta}_{wg}\right] \tag{4.18}$$

Under the null hypothesis that the effects are not correlated with the regressors, the $\text{plim}\left(\hat{\beta}_{GLS} - \hat{\beta}_{WG}\right) = 0$ and h is distributed as $\chi^2_{(k)}$ with k degrees of freedom.

4.7 Empirical results

Table 4.6 shows the parameter estimates obtained when we estimate our model. The variables that appear in the model include:

$$X_1 = \frac{\text{Statutory capital}}{\text{Total assets}},$$

$$X_2 = \frac{\text{Net income}}{\text{Total assets}},$$

$$X_3 = \frac{\text{Net premiums written in long - tail business}}{\text{Total net premiums written}},$$

$$X_4 = \frac{\text{Net reserves}}{\text{Shareholder funds}},$$

$$X_5 = \frac{\text{Net claims}}{\text{Net premiums written}},$$

$$X_6 = \frac{\text{Net commissions}}{\text{Net premiums written}},$$

$$X_7 = \frac{\text{Reinsurance ceded}}{\text{Net premiums written}},$$

$$X_8 = \frac{\text{Stocks in common stock investments}}{\text{Total invested assets}}$$

Because of the large number of variables found to be significant indicators of agency rating in past studies, a list of twenty-five variables (ratios) was compiled for evaluation. From the original list of variables, eight variables were selected as doing the best overall job together in determining the overall rating of an insurance company.

The coefficient on the inverse measure of leverage (capital to assets) is negative and significant both in the fixed effects and the random effects estimations. The significantly negative results are consistent with the hypothesis that higher leverage is associated with greater uncertainty, which in turn is associated with a negative impact on the ratings. Moreover, the rating reflects the fact that insurance companies have been punished for assigning too much capital in their books that has minimized the available funds for investment.

Also consistent with the uncertainty hypothesis is the positive and significant sign of the coefficient on the profitability variable. In this case the ratings have compensated the insurers that showed a better level of profitability overall. The assessment of a company's profitability performance is an integral part of the overall rating analysis. Although return on equity (ROE) is a vital performance benchmark, it can be easily influenced by the company's capital structure. Arguably, the key driver of profitability is the profit margin of the company is able to produce on its operating revenues. Furthermore, return on revenue (ROR) includes both an underwriting and an investment component and, hence, captures both sources of an insurance company's earnings. Although return on revenue is useful as a broad measure of earnings adequacy, it has its drawbacks. It does not differentiate between various product lines that often have different risks, some of which require higher levels of ROR for a certain standard performance than others.

Table 4.6 Fixed effects vs. random effects model

Variable	Fixed effects	Random effects
X_1	-0.02875	-0.035291
	(-1.883270)	(-1.798458)
X_2	0.073896	0.110796
	(1.68225)	(1.69324)
X_3	-0.010776	-0.017390
	(-0.219611)	(-0.386984)
X_4	-0.033870	-0.021098
	(-1.25823)	(-0.860464)
X_5	-0.048672	-0.042756
	(-2.79806)	(-2.81829)
X_6	-0.137429	-0.114145
	(-3.78026)	(-3.31443)
X_7	0.294564	0.293232
	(5.16532)	(5.23716)
X_8	0.122471	0.102669
	(2.91294)	(2.67139)

t-statistics in brackets

The reported p-value of the Hausman test is 0.1415 which supports the null hypothesis that the two vectors of coefficients (fixed and random effects) are identical. In this situation the fixed effects and random effects models become indistinguishable for all practical purposes.

The coefficient on stocks in common stock investments is positive and significant in both cases. Asset quality and investment performance are taken into account when the rating of a company is determined since premiums and deposits invested today must provide a yield sufficient to cover tomorrow's claims. A review of the insurer's asset allocation strategy among investments such as bonds, mortgages, preferred stock, real estate, common stock, collateralised mortgage obligations (CMOs), derivative instruments, and other invested assets, can highlight the quality of the company's investments. The assets are evaluated for credit quality and diversification. Of concern are asset concentrations by type and maturity, low credit quality, industry, geographic location, and within single issuers. An insurer's asset allocation is also important to determine how appropriate it is to support policyholder liabilities. Different asset classes have customary risk profiles and accompanying returns; thus, by choosing which asset to emphasize, a company preordains a large part of the return on its portfolio. The essence of this particular ratio is closely related with issues that might not look correlated, but in fact are, such as common and preferred stock issued by the same entity and perhaps convertible debt also issued by the same entity or a closely related family member. In this case, for instance, the nominal issuer might not be the same company, but if they are all part of the same family and control, a clear concentration can be developed.

Moreover, the predicted sign on the reinsurance variable can be ambiguous, depending on whether the use of reinsurance is seen as reducing uncertainty by shifting risk, or increasing uncertainty by making the insurer's financial health dependent on the financial health of its reinsurers. Although prudent use of reinsurance is often advisable, it can be misused in many fashions. Reinsurers' creditworthiness is always a concern since the deterioration of their financial strength will affect the financial stability of insurers that depend on them for extra financing. The results support the first argument since the coefficient on the variable measuring the percent of business ceded to reinsurers is positive and significant.

The coefficient on the variable measuring the percentage of business in long-tail lines (variable X_3) is not significant. The time value of money is significant component of the overall rating. This realigns writers of long-tail business with writers of short-tail business to the timing of payments of losses and associated capital needs. Adding short-tail business in the company's books can increase the level of capital needed in the short term due to short-term claim fluctuations.

Surprisingly, the variable representing the percent of net reserves is also negatively related and not significant. The reserve risk is the risk that past business will be less profitable than expected as the result of additional variability in estimating frequency and severity trends, as well as changes in economic, legal, and social conditions that can add further variability to claim costs. Furthermore, it measures the variability a company would expect to encounter in its reserve levels, given its lines of business, by comparing the present value of the actual claim runoff that has emerged to the reserves originally established for those claims by industry. However, in line with the negative relationship of the capital requirements, the percent of net reserves is negatively related to reflect the good economic conditions of the recent years.

The ratio of net commissions to net premiums written determines the percentage of direct written premiums that is paid to salespersons. A high ratio would show a high commission rate, which may be the result of an affiliate agency, or at least the payment of commissions higher than those normally paid in the industry. The variable is negatively related to the overall rating of the insurance companies. This is consistent with the expectation of a lower rating due to excessive commissions to agencies or salespersons.

Finally, the variable representing the amount of net claims as a percentage of the net premiums written is significantly negative in line with the stated expectation. The trend of this ratio from calendar year to calendar year helps determine whether the estimates originally made by management were accurate, high, or low. If the ratio has increased or fluctuated, original estimates were too high or too low. The accuracy of previous estimates provides a foundation for confidence in current estimates.

CHAPTER 5

SECURITISATION AND CAPITAL IN THE INSURANCE INDUSTRY

5.1 Introduction

Securitisation in its widest sense implies every such process which converts a financial relation into a transaction. History of evolution of finance and corporate law is replete with instances where relations have been converted into transactions. Examples of securitisation of relationships are cash flows from illiquid assets, such as car loans, mortgages, receivables, and leases, which are transformed into tradeable securities. This practice was introduced by Bank of America in 1977 and has been popularised through the issuance of mortgage-backed securities, which had a total volume of over $1 trillion outstanding in 1992.

In the sense in which the term is used in present capital market activity, securitisation has acquired a typical meaning of its own, called asset securitisation. It is taken to mean a device of structured financing where an entity seeks to pool together its interest in identifiable cash flows over time, transfer the same to investors either with or without the support of further collaterals, and thereby achieve the purpose of financing. Although the result of securitisation is financing, the entity securitising its assets it is not borrowing money, but selling a stream of cash flows that was otherwise to accrue it.

Thus, the meaning of securitisation is a blend of two forces: structured finance and capital markets. Securitisation leads to structured finance as the resulting security is not a generic risk in entity that securitises its assets but in specific assets or cash flows of such entity. Moreover, the idea of securitisation is to create a capital market product. The economic logic for securitisation is extremely powerful and the trend towards securitisation knows no limits. Capital markets are today a place where everything is traded: from claims over assets, to risks, and rewards.

The growing significance of securitisation to banks and other financial institutions has been attributed to several factors that fall into two major categories. First, sale of mortgages through securitisation is a form of regulatory arbitrage (Pavel, 1986; Greenbaum and Thakor, 1987, Pavel and Phillis, 1987; Kopff and Lent, 1988). Securitisation allows banking institutions to sell mortgages in capital markets and thereby frees up banks' capital for more lending. As a result, banks are able to originate loans (to earn fees) without permanently funding them.

Second, securitisation enables banks to sell mortgages at competitive prices and to re-deploy the sale proceeds in assets, which allows more effective management of interest rate risk and better diversification of asset portfolios (Kopff and Lent, 1988; Harvey, 1991).

Although the amount of securitisation is tremendous in banking and some other parts of the financial sector, it is rare in the insurance industry. Indeed, Prudential's sale of policyholder loans (PHLs), Cananwill's sale of premium loans, and Hanover Re's earthquake and wind and USAA's hurricane catastrophe bonds.

Furthermore, the introduction of catastrophe options, known as "cat" options, by the Chicago Board of Trade (CBOT), provides a new asset class to the potential investors. The characteristics of this new asset class, expected positive excess returns and portfolio diversification benefits, can provide an incentive to investors that will be able to earn a return in excess of the risk-free rate.

5.2 The securitisation process

In a securitisation transaction, at least four parties are involved: borrowers, originators (or sellers of assets), buyers of the assets, and investors in securities backed by the assets. The buyer can be a special purpose vehicle (SPV), established solely to purchase assets and to issue securities against the assets to investors. In this way, the underlying assets are isolated from the originators' other assets and liabilities. The originators usually act (and are compensated) as servicers of the sold assets for collecting and distributing interest and principal payments (Schwarcz, 1993).

An asset-backed security (ABS) can be structured as a pass-through or a pay-through. A pass-trough structure features equity financing with a trust passing interest and principal payments from the original borrowers to ABS investors. A pay-through transaction, however, raises capital by issuing both bonds and stocks, allowing active management of cash flows from the underlying assets to provide bondholders with stable cash flows. Guaranteed investment contracts (GICs), issued by banks or insurers, are commonly used for stabilising cash flows by guaranteeing the rate of return on unscheduled principal payments (Pavel, 1986).

To improve their marketability, most ABS issues have credit enhancement to protect investors against normal losses on the underlying assets. Credit enhancement takes one of the following forms:

1. a letter of credit issued by a bank;
2. a surety bond issued by an insurer;

3. reserve accounts established to collect the excess cash flows (the gross cash flows from the underlying assets minus the coupon payments to investors and servicing fees); and,
4. guarantee by the originator.

Greenbaum and Thakor (1987) purport that the originator possesses sufficient information about the borrower's quality to design an appropriate schedule of guarantee. This may be one reason why many ABS issues are guaranteed directly or indirectly by originators. Since investors prefer bonds with a fixed payment schedule (holding other factors, such as expected return, constant), many pay-through bonds (e.g., collaterized mortgage obligations) are structured into several "tranches" to provide more predictable cash flows and maturities.

In structuring a securitised transaction, the originator needs to decide whether to structure it as a sale of assets or as a loan from investors. For regulatory purposes, the assets are considered sold if the originator transfers to the buyer substantial "incidents" of ownership of the assets, which include the risks and benefits of owning the assets. If the assets are sold without resource, the risks and benefits of owning the assets are entirely transferred to the buyer. If some risks and benefits are retained by the originator, depending on the magnitudes of the retention, the transaction may be considered a loan from investors (Adelman and Lorence, 1989).

Sales recognition is important for two reasons. First, it allows realisation of gain or loss for tax purposes. This is particularly critical for transactions aimed at realising book loss of assets, such as Prudential's sale of policyholder loans. Second, it allows removal of the assets from the originator's balance sheet, and that removal frees up the originator's equity capital to support more lending.

5.3 Literature review

Although several professional books have dealt with the topic of securitisation, the academic literature on securitisation is not well developed. A general review of the reasons for securitisation is found in Calstrom and Samolyn (1993). Their analysis is based on a market-based rationale for loan sales with no resource. They emphasised the importance of internal bank funds as a determinant of local investment when bankers have a comparative advantage in screening and monitoring these projects. They showed that costly information and the attendant importance of bank capital in limiting on-balance sheet lending cause loan sales to arise. Loan sales are effectively a way to employ non-local bank capital to support local investments. The model characterises how outright loan sales can occur even when acquiring banks cannot perfectly screen the ex ante quality of the loans they are purchasing; purchasers assess that banks are selling loans because they do not have the capital to hold them. Thus, an important prediction emerges: banks that are capital-constrained in the face of high loan demand are more likely to engage in loan sales.

Although their model was referred to banks, the general framework on the financial market imperfections rather than on banks per se; the nature of the information produced by financial intermediaries can affect the form of external finance. Their results highlight that when the amount of capital affects an intermediary's marginal investment decision, asset sales may be an efficient way of funding local loans in times of high loan demand.

The one theoretical model published to date is an informational asymmetric model by Greenbaum and Thakor (1987) that predicts that banks will securitise their best assets, retaining their worst. Note that 'worst' is in terms of the expected value of the payoff of a loan. Greenbaum and Thakor's model suggests that an increase in regulatory taxes on banks has caused the rise of securitisation. This, however, does not explain the rise of securitisation within the non-banking sector, the critical role of rating agencies, and the rise of securitisation in jurisdictions where regulatory taxes are being reduced. They viewed the bank as an institution with a cost advantage in screening borrowers. Hence, banks are able to perform the loan origination function more efficiently than others. They showed that the bank's decision to fund a loan is affected by credit market informational asymmetries, the information processing technology, and by governmental intervention.

Clearly, with symmetric information regarding borrowers' payoff distributions and without governmental intervention, banks are indifferent between deposit funding and securitisation. However, with asymmetric information about borrowers' payoff distributions, and still without governmental intervention, banks will prefer securitisation for their best assets and deposit funding for their worst. Governmental deposit insurance and regulation will affect the bank's choice of funding mode under asymmetric information. Sufficiently low bank capital requirements in combination with sufficiently generous regulatory subsidies linked to footings will lead to the choice of the traditional deposit funding mode, regardless of the quality of borrowers. Thus, the incentive to securitise can be enhanced by third-party insurers and mutual funds. However, once again a sufficiently large footings-related regulatory subsidy can result in a preference for deposit funding. The choice of the funding mode will also be affected by information processing costs. For securitisation to be preferred these costs must be low enough.

Greenbaum and Thakor find that in an unregulated environment with asymmetric information, banks will securitise as well as fund some of their loans. Moreover, without contemporary information systems that support the servicing of large and complex asset pools and the trading of partitioned (stripped) claims against these pools, securitisation would be impossible. Thus, the information cost argument is facilitating. Technological advances have undeniably reduced the cost of producing liquidity. But this argument alone would lead to the liquefaction of intermediary assets without their necessary sale. Thus, with sufficiently low costs of liquefying assets, banks and thrifts could be expected to do so without disposing of them.

Lockwood et al. (1996) provide some evidence that supports Greenbaum and Thakor's hypothesis and give indirect evidence of the regulatory burden borne by banks. In a study of 294 securitisations by financial institutions, they show that securitisation increases shareholder wealth in well-capitalised banks and finance companies but reduces shareholder wealth in weak banks. They examine several propositions related to asset securitisation. Firstly, they test the proposition that securitisation leads to wealth effects for shareholders of issuing firms. Second, the proposition that wealth effects from securitisation differs by industry is examined by classifying securities into four groups using SIC codes (banks and thrift institutions, finance companies, automobile companies, and other industrial firms). Third, they test the proposition that wealth effects of the ABS differ on the basis of financial slack status of issuing firms. The hypothesis is based on the fact that the ABS announcement of firms with little financial slack will be viewed less favourably than the ABS announcements of firms with superior financial slack. They also test the proposition that wealth effects differ on the basis of type of asset being securitised (auto loans, credit card receivables, trade and lease receivables). Finally, Lockwood et al. test the proposition that the ABS issue leads to a change in market and interest rate risk of the issuing firms.

A sample of 294 public offerings was obtained consisting of 121 ABS issues by banks, 48 by finance companies, 65 by automobile companies, and 60 by other (non-auto) industrial companies. They proxy financial slack for the quarter preceding the ABS announcement as capital surplus plus retained earnings and they derive a relative firm slack variable as follows: first, they divide slack for the firm by the firm's market value (price per share times number of outstanding shares of common stock) for the quarter preceding the ABS announcement to derive a size-adjusted slack measure for the firm. Then, they repeat the procedure for the industry to which the firm belongs to derive a size-adjusted slack measure for the industry. Industry slack is the sum of the slack variables for all firms in the industry and industry size is the sum of the market values of all firms in the industry. Finally, the define the relative slack measure for each firm by dividing the firm's slack measure by the industry's slack measure.

In order to examine the wealth effects of the ABS announcement, standard event study methodology was employed. Daily returns for the 294 events were taken and the parameters of the market model were estimated using these daily returns over days –111 to –11 relative to the announcement day ($t = 0$),

$$R_{jt} = \alpha_j + \beta_j R_{mt} + e_{jt}, \ t = -111,\ldots,-11 \quad (5.1)$$

where R_{jt} is the return on security j for day t, R_{mt} is the return on the equally-weighted market index for day t, and e_{jt} is the random disturbance for security j, day t. The following 21 days (-10 through +10) were designated as the event period.

To examine the null hypothesis that the mean excess return across the n events equals zero, the following test statistics for excess return for event day t, ER_{pt}, and for cumulative excess returns over event days T1 through T2, $CER_p(T1, T2)$, were used,

$$ER_{pt}/S(ER_{pt}) \quad (5.2)$$

$$CER_p(T1, T2)/S(CER_p(T1, T2)) \quad (5.3)$$

where

$$ER_{pt} = \sum_{i=1}^{n} \frac{R_{it} - (\hat{\alpha}_i + \hat{\beta}_i R_{mt})}{n},$$

$$S(ER_{pt}) = \left(\sum_{t=-111}^{-11} \frac{ER_{pt}^2}{99} \right)^{1/2},$$

$$CER_p(T1, T2) = \sum_{t=T1}^{T2} ER_{pt},$$

$$S(CER_p(T1, T2)) = (T2 - T1 + 1)^{1/2} S(ER_{pt})$$

To determine the effect of securitisation on the wealth of stockholders of the issuing firms, excess returns were computed and examined for the entire sample and on sub-samples grouped by industry.

Differential effects based on financial slack and type of ABS were assessed using a cross-sectional OLS regression of the $CER_i(-1,0)$,

$$CER_i(-1,0) = \gamma_0 + \gamma_1 RS_i + \gamma_2 D_{2i} + \gamma_3 D_{3i} + \varepsilon_i, \; i = 1,2,\ldots, n \qquad (5.4)$$

where RS_i is the relative financial slack of the firm as of the quarter preceding the ABS announcement; D_{2i} equals one if the ANS is backed by auto loans, zero otherwise, and D_{3i} equals one if the ABS is backed by trade or lease receivables, zero otherwise; and ε_i is a random disturbance.

To examine the effect of securitisation on the market and interest rate risk of issuing firms, the following regression was performed on each security,

$$R_{it} = \alpha_i + b_{1i}R_{mt} + b_{2i}I_t + b_{3i}R_{mt}D_t + b_{4i}I_tD_t + u_t, \qquad (5.5)$$
$$t = -111,\ldots,-11,+11,\ldots,+120$$

where R_{it} is the daily stock return, R_{mt} is the return on an equally-weighted index, and I_t is the daily interest rate variable orthogonalised with the market index; that is, I_t is the residual for day t in the regression performed on days –111 through –11 and +11 through +120 of daily treasury bill rates on the market return. The variable D_t is a dummy variable equal to zero for days –111 through –11 and equal to one for days +11 through +120. The change during the post-event period (days +11 through +120) in systematic risk is measured by b_3, while the change in interest rate risk is measured by b_4.

Lockwood et al. reported four major findings. First, the effects of asset securitisation were found to be industry specific. These initial tests indicated that finance companies realised wealth gain and banks realised wealth loss at the time of ABS announcement.

Automobile and other industrial firms realised no change in wealth at the time of their ABS announcements.

Second, the wealth change at the time of the ABS announcement was found to be positively related to financial slack for banks. These findings indicated that banks that issue ABS at a time of capital weakness were viewed negatively by the market. To examine the slack proposition further, they compared the CERs of strong banks (high financial slack) with the CERs of weak banks (low financial slack). Strong banks experienced significant wealth gain, whereas weak banks experienced significant wealth loss.

Furthermore, they found that wealth effects of ABS were unaffected by the type of asset being securitised. There were no significant differential wealth effects across credit card, auto loans, and trade and lease receivables securitisations after controlling for the financial slack status of the issuing firms. Finally, the evidence suggested that market and interest rate risk dropped after the ABS announcement for automobile and finance companies. The evidence also indicated that interest rate risk dropped after the ABS announcement for strong banks. In contrast, there were a significant number of increases in both market and interest rate risk after the ABS announcement for weak banks. From their findings, it is clear that the market perceives benefits accruing to finance companies and strong banks that securitise assets. For these firms, securitisation may reduce the need for financing, offer fee income, and increase earnings even more in cases where securitisation is used to retire existing debt. ABS also may lead to a decrease in risk through a decreased reliance on debt financing and through earnings smoothing by an appropriate timing of receivable sales.

Moreover, the market perceives costs accruing to weak banks that securitise assets. Securitisation by weak banks may lead to over-collateralisation, high marginal cost to originate new loans, and reduction in loan portfolio size. The potential negatives also may led to a deterioration in the quality and stability of reported earnings, increasing risk for weak banks that securitised assets.

The role of rating agencies is generally neglected when academics consider the evolution of financial institutions' interaction with financial markets, yet it is critical. Jones et al. (1995) looks at Duff and Phelps' philosophy in assessing mortgage backed securities while Goldstein (1996) outlines how differing rating agencies' perception of probability of default versus severity of default affect ratings, but they do not mention securitisations.

Okabe (1998) describes securitisation as being an alternative financing source with an element of catastrophic risk off-lay for that part of risk not covered by the credit enhancements. The evolution of securitisation presents new challenges to investors in the debt and equity of the securitisers because the continuing exposure of securitisers to part of the securitised assets' risk is not evident from the post securitisation financial statements. Citing Moody's efforts to create alternative analytical techniques for securitisers (Foley and Foley, 1997), Okabe focuses particularly on the difficulties on interpreting the one-off gain from sale that securitisers must book under the recently promulgated Financial Standards Accounting Board Ruling 125 (FASB 1996). Lenders that use securitisation have created a variable in their financial statements – namely, the "residual" income that is captured in the gain-on-sale calculation – that affects the value of the asset backed debt as well as the corporate sponsor's debt and equity. In his analysis, Okabe showed how changes in the key variables used in calculating the gain-on-sale residual can affect the result and, in so doing, affect investors' evaluation of both the asset backed securities and corporate financial strength.

There are four major components to the calculation of the present value of the residual in a securitisation:

1. the structure and costs of the underlying ABS deal;
2. expected prepayments;
3. expected credit defaults; and
4. the present value discount rate.

Of these variables, only the structure of the ABS deal is fixed and unchangeable; each of the other variables is subject to interpretation and judgement. Okabe stressed the fact that the assumptions used to calculate the gain should represent management's best estimates as to the performance of the loan pool. Investors that apply methods that measure the level of risk in this receivable are better prepared to price this risk into the securities, whether they be ABS or corporate obligations.

Although credit-granting standards create a significant level of conformity in evaluating borrower behaviour, the probability of default of one specific loan is independent of the probability of default of another. The ability of an ABS trust to absorb this default risk is a function of the performance of the servicer and the funds expected to be available from the loan pool, which is reported by sponsoring lenders as the present value of the residual interest in the trust. As we mentioned above, there are only three variables that can affect the default rate and cash available – default rates, prepayments, and the valuation discount rate.

Default rates and volatility in credit loss performance can be measured in an objective, disciplined way through static pool analysis. These data may be available from the sponsoring lender or in the public domain through securitisation servicing reports. For amortising asset pools, the greatest determinant of available cash flow is prepayments, which affect a significantly larger portion of the loan pool than defaults. A change in prepayments causes a greater movement in the amount of excess cash flow than any other factor. Discount rates, while an important factor in determining the carrying value of the finance income receivable, affect the rate of return on the retained asset rather than the actual cash flow or asset performance.

A comprehensive analysis of the gain-on-sale calculation can provide both corporate and ABS investors with significant insight into the level of risk in the securities they own. It also offers a tool for determining whether returns justify the risks involved. However, residual analysis is not a panacea. Investors should continue to draw information and insight from every available source, including equity, fixed income, and debt rating analysts, security underwriters, credit enhancement providers, and, most important, the sponsoring lenders. These insights, when combined with residual analysis, provide an integrated analytic approach that gives investors the best opportunity to manage risk and maximise returns.

Another stream of the literature that considers similar concepts discusses asset sales. Lang et al. (1995) look at substantial non-fixed income asset sales from non-financial corporations. They argue that firms selling assets are looking for the cheapest source of funds for meeting their objectives – essentially an agency conflict story. They find that the stock market reaction to asset sales is positive only when the proceeds of the sale are paid out to shareholders. Their starting point is that management values firm size and control, so that it is reluctant to sell assets for efficiency reasons alone. For such management, a more compelling motivation to sell assets is that asset sales provide funds when alternative sources of financing are too expensive, possibly because of agency costs of debt or because information asymmetries make equity sales unattractive.

With this view, which they call the financing hypothesis of asset sales, the completion of an asset sale is good news about the value of the asset because if the value of the value had turned out to be low, the sale would not have taken place. Further, one expects the market to discount proceeds of asset sales retained by the firm in the presence of agency costs of managerial discretion since shareholders do not capture all of the value of the asset sold.

Lang et al. test the financing hypothesis of asset sales against the efficient deployment hypothesis of asset sales, where asset sales promote efficiency by allocating assets to better uses, and sellers capture some of the resulting gains. The efficient deployment hypothesis assumes that management maximises shareholder wealth. In contrast, the financing hypothesis assumes that management pursues its own objectives and, more specifically, values control and firm size. Since it values firm size, management has little incentive to sell assets unless it needs to raise funds and cannot do so cheaply on capital markets. Management may have to raise funds to reduce financial distress costs, to pay dividends to shareholders to prevent a takeover, or to undertake investments that it values but shareholders do not.

In order to investigate the financing hypothesis, they had to identify the use of the proceeds from asset sales, since this hypothesis specifies that the stock-price effect of the announcement of asset sales is related to the use of the proceeds. Since the analysis was focused in the differences between firms that were expected to pay out the proceeds and those that were not, Lang et al. used 93 sales made by 77 firms to determine why the asset was sold and how management expected to use the proceeds. The sample had 40 asset sales by 35 firms with proceeds paid out to creditors and/or shareholders and 53 sales by 43 firms with proceeds retained by the firm (the sample of 40 sales was called the 'pay-out sample' and the sample of 53 sales the 'reinvest sample').

If the financing hypothesis applies to the sales in the sample, one would expect the proceeds paid out to be used to pay down debt rather than to distribute cash to shareholders. If a firm is excessively levered in management's eyes, management has a strong motivation to sell assets to reduce leverage and avoid possible costs of financial distress. In contrast, management that values size and control seems unlikely to want to pay out the proceeds to the shareholders in the absence of pressures from the market for corporate control. The financing hypothesis also predicts that the market discounts the proceeds of successful asset sales when the proceeds are reinvested. In order to test this hypothesis, the authors compared the announcement abnormal returns for the pay-out sample and the reinvest sample. Finally, they employed a regression analysis framework of the abnormal return on firm and sale characteristics to identify any significant relation between the stock-price reaction and the use of the proceeds.

Their findings are consistent with the financing hypothesis. On average, firms benefit from announcing successful sales because a successful sale means that the firm received enough money to make the sale worthwhile. Further, proceeds are discounted when retained by the selling firm because of agency costs of managerial discretion. In the particular sample, firms selling assets typically were poor performers and they were more likely to pay out the proceeds when they find it difficult to service their debt. The average stock-price reaction to asset sales is positive and it is significantly higher for firms that pay out the proceeds. There was not, however, a direct link between abnormal returns and proxies for agency costs of managerial discretion. Though their evidence demonstrates the relevance of the financing hypothesis, it is also clear from the analysis and from the empirical results that it is difficult to evaluate the information conveyed by asset sales because asset sales convey news about the value of the asset sold, the intended use of the proceeds and, possibly, the firm's financial strength. Larger samples of possibly less significant asset sales might offer a way to disentangle these various effects with more precision and provide useful information on the relative importance of the financing hypothesis and of the efficient deployment hypothesis.

5.4 Securitisation in the insurance industry

In general, as we mentioned before, securitisation benefits borrowers by enhancing capital supply at competitive costs, it benefits investors by broadening the spectrum of investment options, and it benefits originators by generating underwriting fees and trading commissions. The benefits and costs that are unique to the insurance industry can take various forms.

Pavel and Phillis (1987) empirically show that regulatory taxes have an important impact on a bank's decision to sell loans. Similar regulatory taxes (including surplus and reserve requirements and premiums paid to guaranty funds) are imposed on insurers, and they can be avoided, in theory, with sales of assets and liabilities through securitisation. Shante

(1989) and Agostino and Cosgrove (1990) suggest that sales of loadings of renewal premiums through securitisation can improve surplus if the transaction is structured without recourse.

In 1988, General American Life "sold" a part of its renewal premiums to Citicorp. In 1987, Monarch Capital, the parent of Monarch Life Insurance Company, "sold" certain deferred charges on single premium variable life products to a syndicate of nine commercial banks. Neither company was allowed to treat its transaction as a sale of assets and both were required to set up reserves. Consequently, no surplus relief resulted. Thus, regulatory conservatism may be one reason that similar transactions have not been seen more often. Other reasons could be the lack of expertise in such transactions and the costs of securitising uncertain cash flows such as renewal premiums.

If regulatory and technological changes become favorable to securitisation, insurers (particularly mutual insurers because of their inability to raise capital in financial markets) with low expense ratios will benefit from securitisation by specializing in originating and servicing policies (i.e., Cananwill was able to originate and service more premium loans to earn fees by "churning" over the sale proceeds from securitised premium loans). Property-liability insurers, who often suffer from cycles of hard and soft markets partially due to availability of capital, will also benefit from securitisation (see Winter, 1988; Cummins and Danzon, 1992).

Moreover, securitisation can create liquidity. Sales of the securitised assets generate assets (cash) for the originator and, thus, liquidity. Another way is by converting illiquid assets into high-grade securities through tranching and credit enhancing. The high credit rating enhances the marketability of the securitised assets and thus increases the value of the underlying assets.

Another major benefit that incurred through securitisation is tax savings. Tax savings can originate from two sources: the tax shelter from realising the book loss of policyholder loans and the savings in equity tax for mutual insurers.

However, there are potential costs involved in the securitisation process of different assets. For example, sales of assets with market values lower than book values result in a lower equity-to-surplus ratio. This can be costly if capital requirements or optimal capital structure concerns induce the insurer to raise additional capital. Thus, while the realisation of book loss creates a tax shelter, it also generates the potential need for new equity capital.

An adverse selection problem could result from securitisation. Good risks would be securitised and sold off, and poor risks would be left on the insurer's books and insured by guaranty funds for a flat premium. This adverse selection might shift poor risks from insurers who securitise to other insurers, policyholders, or taxpayers. The risk-based capital requirements, however, would curtail such adverse selection.

Moreover, securitisation is a complex process that incurs underwriting, issuing, and credit enhancement costs. Among these costs, credit enhancement costs are the most significant for securitising insurance products, because a certain degree of cash-flow engineering is necessary to securitise complex insurance products, assets, or liabilities. From the perspective of costs and expertise involved in a transaction, insurance products or assets with more stable cash flows are more likely to be securitised.

5.5 Origins of insurance securitisation and the problem of risk transfer

Insurance securitisation most likely originated in the reinsurance capacity crisis that followed Hurricane Andrew, when the amount of coverage nearly halved and premium rates nearly doubled. Claims following Hurricane Andrew reached $18 billion, the most costly insured event in history. Another capacity and pricing crisis followed the Northridge Earthquake in California in 1994, which produced $11 billion in insured losses. After that, capacity surged and rates softened with the advent of new reinsurers in Bermuda and the revitalisation of the Lloyd's insurance syndication market.

From 1989 to 1995, total insured property losses in the U.S. reached $75 billion, nearly 50% more than total insured property losses in the preceding 40 years. In that same period, average insured loss per catastrophe exceeded $300 million, versus just $56 million in the earlier period. This explosion in insurance exposure was the result of several factors: rising population density in heavily insured areas, inflation of property values, and increased penetration of property/casualty insurance.

Broad diversification has become more elusive for property reinsurers as, globally, there has been little growth in insured values, apart from a few peril regions. Industry risk books abound with exposure to California earthquake, Southeastern U.S. hurricane, European windstorm and flood, and Japanese earthquake. Volatility in reinsurance pricing and capacity has followed increased population density and the growth of insurance demand concentrated in those few peril areas; insurers have begun to turn to capital markets and derivatives technology to smooth out those swings and to customise the terms of coverage, including multi-year cover.

By 1997, insurer USAA (United States Automobile Association), with exposures concentrated in Florida, Texas, Virginia, and the Carolinas, used a securitisation to secure cover for windstorm losses in states on the Atlantic and Gulf coasts. USAA entered into a single-occurrence, excess-of-loss reinsurance treaty with a Cayman reinsurer established specifically to provide that single coverage: the entire source of capital backing Residential Re, the special-purpose remote reinsurer, in underwriting the treaty was provided by an issuance of $477 million one-year notes. The note proceeds provided collateralised cover for an 80% quota share of up to $500 million in losses in excess of USAA's retention of $1 billion in losses. Thus, a capital markets issuance served to collateralise both the notes and the reinsurance treaty, creating a new source of high credit quality reinsurance capacity (on June 1, 2000, Residential Re issued its fourth annual instalment of one-year notes, with similar terms).

Investors have also shown interest in non-catastrophe insurance securitisations. In 1998, the special-purpose vehicle Mutual Securitisation plc issued two tranches of limited resource notes (£140 million class A1 due 2012 and £120 million class A2 due 2022) linked to the emerging surpluses, or profits, of National Provident Institution (NPI), a UK mutual life insurer. These notes respond to the mortality and lapse rates of a specified block of life policies and the management of a portfolio of assets supporting the policies. The main source of emerging surplus is embedded management charges levied on an asset portfolio. Repayment of the bonds will be linked to the repayment of a loan made to NPI by Mutual Securitisation; to repay the loan, NPI will depend upon the surplus associated with the defined book of policies.

Generally, catastrophe-linked securities is an emerging asset class of structured insurance risk products. These securities offer returns that are linked to the occurrence of catastrophic events such as earthquakes and hurricanes. They can provide investors with diversification from corporate and asset-backed securities at comparable or wider spreads. Issued through special purpose vehicles (just like the securitisation of regular assets), these securities offer an opportunity to participate directly in catastrophe risk without having to assume the operational risks inherent in securities issued by property and casualty insurance companies that underwrite this risk. Investing in "pure" catastrophe risk can also improve the risk/return profile of a diversified portfolio of assets because this risk is generally non-correlated with the systematic risks present in other securities markets.

In addition, an outgrowth of the need for additional reinsurance capacity following Hurricane Andrew (1992) and the Northridge Earthquake (1994), which in combination produced $45 billion in industry-wide insured losses (in 1997 dollars), catastrophe-linked securities provide insurers with a new form of reinsurance protection. In exchange for a reinsurance premium (i.e., interest on the securities), investors assume financial exposure to the risk that a catastrophe will strike and will generate insured losses above a certain level. If such a catastrophe occurs, catastrophe-linked securities investors would receive a reduced yield and/or lose part or all of their principal, and the insurer would receive a reinsurance claim payment. By transferring catastrophe risks to the capital markets in this manner, insurance companies are supplementing their use of traditional reinsurance and internal loss management mechanisms to reduce volatility in their financial statements and preserve overall liquidity.

Moreover, very little of the reinsurance in place provides protection against industry-wide losses for catastrophic events greater than $5 billion. That is, for a $50 billion catastrophic event, the overwhelming majority of the last $45 billion of losses (after the first $5 billion) are not covered by reinsurance. In a narrow sense, this is not surprising, given that the relatively small capital and surplus of the reinsurance industry ($26.7 billion for US reinsurers, $6.5 billion for Bermudan reinsurers, $7 billion for German reinsurers, and $16.8 billion for others). Thus, at present levels of capital, the world wide reinsurance industry is not capable of funding large event risks in the US alone, not to mention the rest of the world.

It is striking that so little reinsurance is available for large event losses. A number of studies have focused on the determinants of the corporate demand for insurance (Mayers and Smith, 1982; Mayers and Smith, 1990). These analyses explicitly recognise that while the primary motive for individuals' insurance purchases, risk aversion can partially explain the demand for insurance by closely held corporations and partnerships, it provides a deficient explanation for insurance purchases by widely held corporations. Mayers and Smith (1982) argue that the corporate demand derives from the ability of insurance contracts to (1) allocate risk to those of the firm's claimholders who have a comparative advantage in risk bearing, (2) lower expected transactions costs of bankruptcy, (3) provide real-service efficiencies in claims administration, (4) monitor the compliance of contractual provisions, (5) bond the firm's real investment decisions, (6) lower the corporation's expected tax liability, and (7) reduce regulatory constraints on firms.

Hoyt and Khang (1999) also study the corporate insurance purchase decision. They find evidence that tax effects, as outlined in Mayers and Smith (1982), are important. Specifically, insurance provides a faster adjustment of the depreciable basis of replacement property than does self-insurance. It also allows the firm to protect other tax benefits (investment tax credits, loss carry-forwards, etc.). Insuring depreciated assets can be profitable after tax. If the premium is a deductible expense, after tax it could be less than expected losses. At the same time if the recovery of a loss is not taxable income, after tax the expected recovery then exceeds the premium, with no offsetting book loss as the lost asset was depreciated. Hoyt and Khang found that insurance purchases increased significantly with increases in the ratio of cumulative depreciation to the historical cost of fixed assets.

In a subsequent study, Mayers and Smith (1990) examined reinsurance purchases by 1,276 property/casualty insurance companies and provided evidence that ownership structure matters. Generally, the less diversified the owners' portfolios, the greater the reinsurance purchases. Thus, as a confirmation of financial theory expectations, Lloyd's reinsure most, while widely held stocks reinsure least. Moreover, subsidiary and group relations affect the demand for reinsurance. They also provide evidence that size, credit standing, and geographic concentration reduce the demand for reinsurance and weak evidence that line-of-business concentration reduces reinsurance demand, as well. The estimated negative effect of geographic concentration suggests that the real-services argument is quantitatively important. However, their results face potential limitations. The power o the tests employed is reduced by the lack of information about the tax status of individual firms. The data are aggregated into lines of insurance; while within lines, policies are undoubtedly heterogeneous in their riskiness. With more detailed information about the risks of the specific policies sold by particular firms, tests with greater power are possible.

Thus, reinsurance and substitute risk transfer mechanisms in general, can be alternative forms of financing by allowing an insurer to write as if it had more surplus. If the servicing costs of debt exceed the loading elements of the reinsurance, risk transfer can increase the overall profitability. For mutual companies without access to either debt or equity markets, reinsurance can become the only viable financing arrangement.

Many observers, practitioners, and academics have argued that bringing catastrophic exposures directly to the capital markets can help reduce reinsurance prices and increase risk transfer. Mechanisms include cat-linked bonds, swaps, exchange traded options and futures, cat-linked issues of equity, etc.

5.6 Catastrophe risk management: Capital markets trends

Catastrophe risk can be viewed as composed of layers of risk from events with decreasing probability of occurrence and increasing magnitude of losses. Sophisticated modelling efforts have shown that catastrophic events occur in random intervals of time and less severe catastrophes occur with more frequency. Risk management of catastrophe losses varies from one insurer to another.

As insurers have increased use of advanced catastrophe modelling to predict losses, they have tended to purchase coverage equal to their probable loss under a severe loss scenario or, at a minimum, for losses in excess of 10% of their capital. However, as we mentioned before, large insurers find that protecting their balance sheet against an infrequent but large catastrophe is currently priced too high due to lack of capacity in the reinsurance industry for covering this type of risk. Hence, insurers are seeking capital market solutions to bridge this gap in capacity and to create a more efficient risk transfer mechanism.

As insurers explore alternative solutions for gaining additional reinsurance coverage, they have participated in several creative capital market developments including government initiatives, exchange traded derivatives and catastrophe-linked securities.

In response to reduced personal lines insurance availability after Hurricane Andrew, the US and state governments have created various funds to provide additional capacity in the event of a catastrophe. These include the Florida Hurricane Catastrophe Fund, the California Earthquake Authority, the Hawaii Hurricane Relief Fund, and others. These funds are set up to access the capital markets immediately after an event to provide additional funding either directly to homeowners or to insurance companies. These special funds are expected to provide incremental capacity to the property-casualty industry and potentially bridge part of the gap in reinsurance supply.

The introduction of exchange traded catastrophe risk contracts marked the entry of insurance risk products into the capital markets. Derivative contracts are offered on the Chicago Board of Trade (CBOT) and the Catex (newly formed electronic exchanges in New Jersey and Bermuda). These contracts have payoffs that depend on indices that measure insured losses from catastrophes in specific geographical regions. In principle, insurers can use these contracts to reduce their exposure to underwriting losses due to catastrophes.

The first insurance derivative contracts, introduced by CBOT in December 1992, were futures contracts based on the underwriting results of twenty-two insurers for the entire country and for three regions of the United States. (east, mid-west, and west). Trading volume for these contracts was anemic during their two years of existence, which has been attributed to several factors, including: (1) insurers/reinsurers' preference for option spreads (a long call position combined with a short call position with a higher exercise price) as opposed to future contracts because option spreads have payoff similar to catastrophe reinsurance contracts; (2) the lack of historical data on the underlying indices; and (3) the basis risk associated with broad geographical indices. The CBOT discontinued the futures contracts and now only trades option contracts.

These particular option contracts are based on Property Claims Services (PCS) indices, which track the aggregate amount of insured losses resulting from catastrophic events that occur in given regions and risk periods. After a catastrophic event occurs, PCS estimates the insured property damage by surveying a wide range of insurers regarding the dollar amount of claims they expect to receive (PCS defines a catastrophe to be an event that causes more than $25 million of insured losses to personal property, vehicles, boats, and business interruption). PCS also uses its own information about the value of the property in the affected countries and, in some cases, conducts its own on-the-ground survey of the damage.

Once PCS has made its assessment, it releases an official loss estimate for each of the states affected. The state losses are added to the appropriate regional and/or state indices (see Table 5.1) so that each index represents the total losses incurred in that region during the risk period.

Table 5.1 PCS Catastrophe indices

Region	Risk period	Contract months	States covered
Florida	Quarterly	Mar, Jun, Sep, Dec	FL
Texas	Quarterly	Mar, Jun, Sep, Dec	TX
California	Annual	Dec	CA
Eastern	Quarterly	Mar, Jun, Sep, Dec	Includes Northeastern and Southeastern regions
Northeastern	Quarterly	Mar, Jun, Sep, Dec	ME, NH, VT, MA, CT, RI, NY, NJ, PA, DE, MD, DC
Southeastern	Quarterly	Mar, Jun, Sep, Dec	VA, WV, NC, SC, GA, FL, AL, MS, LA
Midwestern	Quarterly	Mar, Jun, Sep, Dec	OK, AR, TN, KY, OH, MI, IN, IL, WI, MN, ND, SD, IS, NE, KS, MO
Western	Annual	Dec	HI, AK, WA, OR, CA, NV, AZ, NM, UT, CO, WY, MT, ID
National	Quarterly, Annual	Mar, Jun, Sep, Dec, Annual	All fifty states plus Washington DC

The risk period for each index is the time period over which losses are aggregated. For an index in which catastrophes are seasonal (as in the case of hurricanes and tornadoes), the risk period is quarterly and the options trade on a March, June, September, December cycle. For regions in which catastrophes are not seasonal (earthquakes), the risk period is annual and only a December contract is traded. Each index is zero at the beginning of its risk period and in increases by one point for each $100 million of insured property damage that occurs in the time period.

Following each risk period, moreover, there is a "loss-development" period of either six or twelve months. During this time, PCS will update the amount of damage that occurred during the risk period as more information becomes available. For example, if there were storms in Florida in the third quarter of 1997 that caused an aggregate of $10 billion of damage, the September 1997 Florida index value would be 100. But if during the fourth quarter, PCS determines that the storms actually caused $12.25 billion of damage, the September 1997 Florida index would be adjusted to 122.5, while the December 1997 Florida index would be unaffected by this change.

The development period is necessary due to the difficulty in making a timely and accurate assessment of the amount of damage that has occurred after a large catastrophic event. The development period thus insures that the indices are accurate reflections of the damages incurred, and that buyers receive the full benefits of their hedges.

Although the more narrowly defined geographical areas should help reduce the basis risk that is involved with the PCS catastrophe options, industry analysts continue to cite basis risk as one of the main shortcomings of these particular derivatives contracts (Major, 1997). Harrington and Niehaus (1999) provide evidence on the potential effectiveness of state-specific insurance derivatives in hedging underwriting risk by relating individual insurer groups' annual loss ratios for homeowners, commercial multiple peril, and fire insurance in twenty eastern and southern states. They find that state-specific PCS catastrophe derivatives can be effective hedges against variation in insurer by line and state loss ratios. They also suggest that industry by line and state loss ratios, on average, could provide more effective hedges than by line and state catastrophe loss ratios. Finally, they report that derivatives on state-specific catastrophe losses may allow homeowners insurers with different books of business in different states to construct materially more effective hedges than with a broader regional contract.

Apart from catastrophe options, catastrophe-linked securities have been issued by insurance companies that they want to transfer or sale insurance risk in the form of an investment security. These securities are issued for an expected maturity with the payment of coupon and retirement of principal dependent on occurrence of a catastrophic event with losses greater than a specified trigger during a defined risk or loss-occurrence period. As in other asset-backed transactions, the issuer sets up a special purpose vehicle (SPV) that is bankruptcy remote. The vehicle is generally set up offshore for regulatory and tax reasons and issues securities that carry the risk of catastrophe losses over a specified level. It then issues a back-to-back reinsurance contract to the insurer, thus enhancing the insurer's reinsurance coverage.

The security, like reinsurance, can be structured as a quota-share or an excess-of-loss issue. In the quota-share structure, the issuer shares with investors a fixed percentage of losses over the attachment point. In an excess-of-loss structure, investors absorb losses over the attachment point for the total amount of the issue (equivalent to the exposure amount in a reinsurance contract).

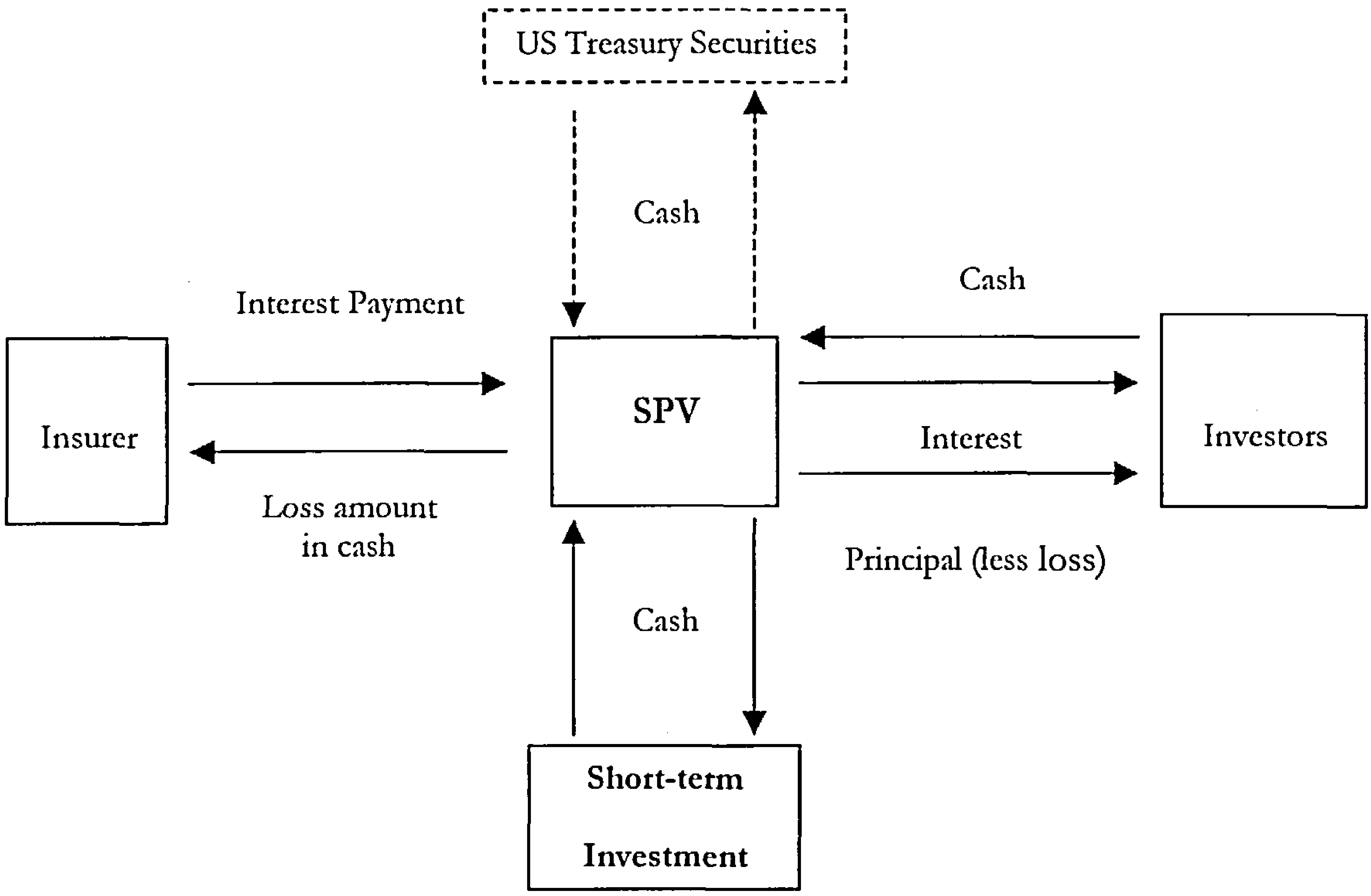
The underlying catastrophe can either be one type of event or a mix of events. Risks can be spread across geographic region, type of event, or underlying property type (residential, commercial, industrial, etc.). Only events that occur prior to the end of the specified loss occurrence period and result in losses in excess of the attachment point are considered loss events for the securities. Underlying losses in any specific transaction can be based either on the insurer's book of policies or on a basket of risks as measured, for example, by the PCS index.

If the contract is based on the actual loss experience of the cedant's own book of business, this is known as an indemnity-based contract, and closely resembles a traditional reinsurance program. An indexed structure, on the other hand, can make it easier for investors to analyse the risk, because they no longer need to understand the details of the cedant's business. However, the cedant can be exposed to basis risk, to the extent that its own exposure differs from that of the index used to determine the payoff of the contract. It is possible that in a given event, the cedant experiences large losses, while the losses in the index – and thus the payments to the cedant – are relatively small. Because of this basis risk, the cedant may want to pay less of a premium for such coverage. A final possibility for the contract structure is the use of physical parameters of the natural hazard, such as the magnitude and location of an earthquake, as the trigger (parametric structure).

The SPV invests cash raised from the issue in high quality, liquid, fixed income instruments (typically US Treasuries or AAA rated securities). This short-term portfolio is used to cover losses from events or to repay investors on maturity of the bond, and to provide a minimum rate of return (e.g., LIBOR, Treasury bill). The contract is structured like a cash-collateralised reinsurance contract, and unlike traditional reinsurance contracts, does not carry any credit risk of the reinsurer. The coupon of the catastrophe-linked securities includes a spread over the minimum rate earned by the short-term portfolio. The insurer pays the spread to the SPV, which passes through the total coupon payment to investors (see Figure 5.1).

The maturity of the security is based on the period during which a loss event can occur, called the risk period (or the loss occurrence period), and the time for computation of losses, called the development period. The development period may be up to one year, during which time the company receives final claims, surveys its policyholders' properties and determines total damage claims. Typically, loss estimates two to three months after the catastrophe give an indication of whether losses from the event have exceeded the trigger. However, the actual amount of losses is determined after the development period (i.e., after final claims are received).

Figure 5.1 Illustrative CLS structure



To attract a wider investor base, some structures provide protection of principal, with only coupon at risk. This is accomplished by establishing a structural feature which provides the investor with US Treasury STRIPS with a par value equal to the principal value of the catastrophe-linked security, upon occurrence of a qualifying catastrophe. Since principal for these securities is backed by US Treasuries, these securities will generally be rated higher than catastrophe-linked securities with principal at risk. Nevertheless, investors face the risk of earning little or no yield for the remaining period of the STRIPS. A related structure which has been considered involves swapping US

Treasury securities held by the SPV for surplus notes or equity of the insurer upon occurrence of a qualifying catastrophe.

The appeal of catastrophe-linked securities (i.e., cat bonds) has been documented in several different studies. Froot et al (1995) show that cat investments over-performed domestic bonds and that the returns on cat risks are less volatile than either stocks or bonds. They have examined pricing and claims data from actual reinsurance contracts (more than 2,000 contracts) brokered by Guy Carpenter Co. from 1970 to 1994. All of the contracts were excess-of-loss contracts (with a pre-specified maximum amount at risk). In order to calculate the return, they acted as though the investor put up an amount at the beginning of the contract year equal to two times the limit, thus accounting for a potential limit reinstatement. Insurers contribute the reinsurance premium and reinstatement premium (if any) into the same dollar pool. The authors assumed that all these funds were invested in US Treasury Bills, until and unless there is a drawdown due to the occurrence of an event. At the end of each year, the investor takes home all funds remaining in the account minus 1% of the limit, which the authors assumed that goes towards transaction fees. The investor's return is the excess above what an equal-sized investment in one-year Treasury bills would have returned.

Froot at al. reported that by investing in a portfolio of cat reinsurance contracts (weighted by limit), an investor would have earned 200 basis points above the Treasury Bill rate. In the best and worst year from 1970 to 1994, the excess return would have been 7.5% and –22.1%, respectively. Higher average returns were earned by the lower layers, which are the excess-of-loss contracts that are most frequently impacted by loss. Also, the returns on national exposures were generally higher than returns on regional exposures. However, they were more volatile too, which is expected given that national companies are, on average, more exposed to high risk areas (such as Florida, California, etc.).

Moreover, they also find that the correlation between cat risks and other asset classes was statistically indistinguishable from zero. The estimated correlation coefficients between cat exposures and other asset classes range from a low of –0.13 to a high of 0.21. by comparison, the correlation between US stocks and bonds is estimated to be much higher, at 0.40; between international stocks and international bonds at 0.45; and between international stocks and domestic stocks at 0.58. all of these latter correlations are statistically positive at standard levels of significance.

Another study by Canter et al. (1997) shows that a portfolio of ten prominent catastrophe reinsurance companies has a strong positive correlation (beta of 0.83) with the stock market movements. As a result, buying reinsurance company equity does not bring significant diversification benefits to the investors. In this respect, catastrophe-linked securities offer better diversification opportunities since they are expected to have near zero betas. They also correlated the yearly percent change in the S&P 500 index with the yearly percent change in the PCS national index using data from 1949 to 1994. They find that the correlation coefficient is insignificantly different from zero. Thus, options on PCS indices are expected to be zero-beta assets.

Litzenberger et al. (1996) demonstrate that returns on cat bonds are essentially uncorrelated with the market, making them excellent tools for portfolio diversification. They showed that adding small amounts of securitised reinsurance to diversified portfolios would enhance the risk/reward opportunities to investors. By considering the

Black and Literrman (1991) approach of the capital asset pricing model, they employed the Sharpe ratio of the portfolio in order to identify the attractiveness of this new asset class. The addition of a new security to an existing portfolio would be attractive if and only if it increases the portfolio's Sharpe ratio, i.e. the ratio of excess return to standard deviation of return. Over the sample period, March 1955 through December 1994, the correlations of the adjusted historical loss ratios (AHLR) with the returns on the S&P 500 index and a government bond index were 0.058 and 0.105 respectively. These numbers imply that the correlation with the securities embedded with the cat exposure option is slightly negative, because the return on these securities varies inversely with the AHLRs. Their results (Table 5.2) show that catastrophe reinsurance bonds are sufficiently attractive to warrant inclusion in a diversified bond or balanced fund.

Table 5.2 Excess returns for one-year cat notes

		Type of embedded CAT exposure	
	Weight of CAT note in enhanced portfolio	Excess of loss (bp)	Binary (bp)
S&P 500 index	1%	11	22
	2%	22	44
Bond	1%	2	4
	2%	4	8
Balanced (50% stock/50% bond)	1%	13	27
	2%	27	53
ROL		12.51%	19.25%
Offered return		7.94	10.85

Finally, Doherty (1997) identifies another two instruments in managing catastrophe risk. Apart from PCS options and catastrophe bonds, post-loss equity re-capitalisation and catastrophe equity puts are two instruments that can be used in order to mitigate the exposure faced by insurance companies. Contingent surplus notes (CSNs) are surplus notes than an insurer has purchased the right to issue to specific intermediaries or investors. The right may be contingent on certain events taking place, or it may be unconditional. An insurer that wants to use CSNs to access additional capital in the event of a catastrophe might arrange for a financial intermediary to set up an investment trust. Using the proceeds from the sale of trust notes or certificates to investors, the trust would invest in US government bonds or other liquid securities. The arrangement would give the insurer the right, under specified circumstances, to issue surplus notes to the financial intermediary in exchange for cash or liquid assets. The intermediary, in turn, would have the right to substitute the surplus notes for the securities held by the trust.

It is obvious that investors can earn higher returns by investing in CSN trusts than by investing directly in treasury securities. The trusts can pay higher returns as a result of the fees they collect on behalf of investors. To the extent that an insurer meets its obligations under its surplus notes, investors receive periodic payments of interest and principal, even after the insurer suffers substantial catastrophe losses.

On the other hand, insurers using CSNs may face high transaction costs. In addition to paying various fees, insurers incur the costs of giving intermediaries and investors the information those parties need to evaluate the risk they are assuming and their potential returns. Moreover, evaluating the probability that an insurer will repay surplus notes can be difficult, because the notes are subordinate to other claims on the insurer and because the insurer must get permission from the relevant regulatory authority to repay the notes.

Finally, catastrophe equity puts, or CatEPuts^{SM}, are a form of option that stock insurers can buy from investors. These options, developed by AON Corporation, give an insurer the right to sell a specified amount of its stock to investors at a predetermined price if catastrophe losses surpass a specified trigger. Thus, catastrophe equity puts can give insurers access to additional equity capital precisely when they need funds to cover catastrophe losses.

An insurance company that uses catastrophe equity puts faces a counterparty risk – the risk that the sellers of the catastrophe equity puts will not have enough cash available to purchase the agreed amount of the insurer's stock. Insurers can minimise that risk by buying catastrophe equity puts only from investors with superior credit ratings. Insurers can also include in the catastrophe equity puts language that requires the investors to collateralise the options if their credit ratings deteriorate.

Furthermore, an insurance company that uses catastrophe equity puts faces a risk that exercising its options will trigger a change in control of the company. The company can eliminate that risk by basing the catastrophe equity puts on non-voting shares, such as preferred stock.

On the other hand, investors selling catastrophe equity puts face the risk that they will end up owning shares of a firm that cannot survive. They can protect themselves against that possibility by including conditions that prevent insurers from exercising their catastrophe equity puts when they suffer losses so severe that they would still be impaired even after exercising their options and receiving the new capital. But, when catastrophe equity puts include such features, they provide less protection to insurers and their policyholders.

An investor that sells catastrophe equity puts faces the same stock market risk as investors that sell traditional options on stocks – the risk that unanticipated downward movement in the price of a stock will make the predetermined price specified in the put options less attractive than when the investor sold the puts. An insurer that buys catastrophe equity puts does not necessarily face a corresponding risk that unanticipated upward movement in the price of its stock would reduce the attractiveness of the predetermined price, because the insurer is free to decide not to exercise its catastrophe equity puts and instead to raise capital using other means.

5.7 Analysis of catastrophe bond structure and performance

When insurance securitisations were first considered, insurers were reluctant to disclose too much of their underwriting data. Cedants receive the most precise coverage from indemnified transactions, which respond directly to a specified group of policies, but many were reluctant to reveal their underwriting procedures or actual policy composition, beyond statutory filings.

An indemnified transaction – such as the four annual Residential Re issuances for USAA – reflects the underwriting and claims settlement process of the ceding company. In this case, the safe account was a trust, administered by the offshore SPV Residential Re. Residential Re had no business purpose other than to sell a one-year $400 million reinsurance contract to USAA and to issue $400 million in risk-transfer securities to fully collateralise that reinsurance (Froot and Seasholes, 1997). The proceeds of the issuance were held in a trust and invested in highly rated, short-term investments such as commercial paper. In the event of a catastrophe, the trustee would have sold the investments to cover 80% of USAA's losses in excess of $1 billion (until the $400 million is exhausted). In return for this reinsurance, Residential Re received a premium from

USAA of 600 basis points ($24 million). The premium, along with virtually all of the interest on the commercial paper, went directly to investors, regardless of whether USAA experienced a loss.

To fund the reinsurance, Residential Re issued securities of two types: principal variable and principal protected. If there was a loss, principal variable investors would have lost some or all of their initial investment. These notes paid interest at a rate of LIBOR plus 575 basis points (this is essentially interest plus the reinsurance premium of 600 basis points, less about 25 basis points for costs. The principal protected securities would have had their principal repayment delayed for ten years in the event of a loss, with a reduction in interest along the way. The principal protected securities paid LIBOR plus 273 basis points.

Following the risk of a covered peril event, the primary difficulty facing investors in indemnified notes is the existence of lengthy development periods, which are bond extensions that allow for the discovery of damage and the settlement of claims, a feature typical of insurance cover. Although the risk period ends on the scheduled maturity, at the option of the cedant, investors might have to wait two years or more to determine the disposition of their investment.

Some issuance bonds are linked not to the ceding insurer's business but to the behaviour of an industry-wide or geographic index, such as the data compiled by PCS in the U.S. Ceding insurers that issue indexed notes can be exposed to significant basis risk, to the extent that the index does not mimic cedant losses. Because it is generally easier to calculate an index than the final claims of the ceding insurer, indexed notes tend to have development periods under two years.

For example, Seismic Ltd. (Seismic), domiciled in the Cayman Islands, issued $145.5 million in 22-month notes. The note proceeds provide Bermuda-based Lehman Re with a source of indexed cover for cumulative insured losses in California if caused by earthquake and fire following earthquake, as determined by PCS. Seismic's payments to Lehman Re are tied to a reference schedule of cumulative insured quake losses. Under this schedule, principal in the collateral account will be paid out in increments of 10% of the original balance, per $1 billion in PCS estimated losses over $22.5 billion. Lehman Re may also extend the notes for a month at its sole discretion, but up to 18 months, in three-month segments, if PCS loss estimates for the fixed, 22-month period exceed certain hurdles, which grow over time.

This transaction is similar to a 1997 earthquake-linked two-year note issued by SR Earthquake Fund Ltd. There, Swiss Re sought PCS-indexed cover from California earthquakes. However, these notes were structured to respond to a large single event rather than to cumulative state-wide losses during the term of the notes.

Finally, notes can be structured parametrically, without reference to the cedant's business. Parametric notes make their payments based upon pricing calculations related primarily to the quantities associated with pertinent events, such as magnitude, intensity, and epicentre of an earthquake or wind speed, forward velocity, and country of landfall of a hurricane. This so-called synthetic indemnification could reduce basis risk to the cedant while nearly eliminating the development period for investors. Each indexed or parametric transaction must specify mathematically the relationship between the parametric formula or index and the claims against note principal by the ceding insurer.

In 1999, Concentric Ltd., domiciled in the Cayman Islands, issued $100 million five-year Japanese earthquake notes. Upon certain trigger-event earthquakes, the assets in the collateral account backing the notes will be used to make payments to Oriental Land, a Japanese real estate company, which is perhaps best-known as the owner and operator of Tokyo Disneyland. The amount of those payments will not be tied to indemnified losses of Oriental Land. Rather, they will be tied parametrically to earthquake depth, epicentre, and magnitude as measured by the governmental Japanese Meteorological Agency. Payments will be made under a fixed formula, so the further an earthquake's epicentre is from Tokyo Disneyland, the greater its magnitude must be to trigger a given level of payment.

Atlas Re, domiciled in Ireland, provided another recent parametric issuance: three-year multi-peril notes providing collateralised retrocessional capacity to French reinsurer SCOR S.A. The retrocessional agreement covers certain insured windstorm losses in seven European countries and certain insured losses due to earthquake and fire following quake in Japan and the contiguous U.S. Each year, SCOR will provide updated policy data to the peril modelling firm EQECAT. EQECAT will recalculate exposures, adjustment factors, and attachment points in three currencies. Payouts will be based on a subset of the subject business with high data resolution, then multiplied by an adjustment factor, to capture the likely effect on SCOR's whole subject business. EQECAT will determine the attachment points corresponding to a fixed annual probability and apply prevailing foreign exchange rates. EQECAT will also reset the per-event limit to maintain a constant loss exceedence probability. In this transaction, the notes will follow the fortunes of SCOR with respect to the subject business; in return, SCOR will retain a

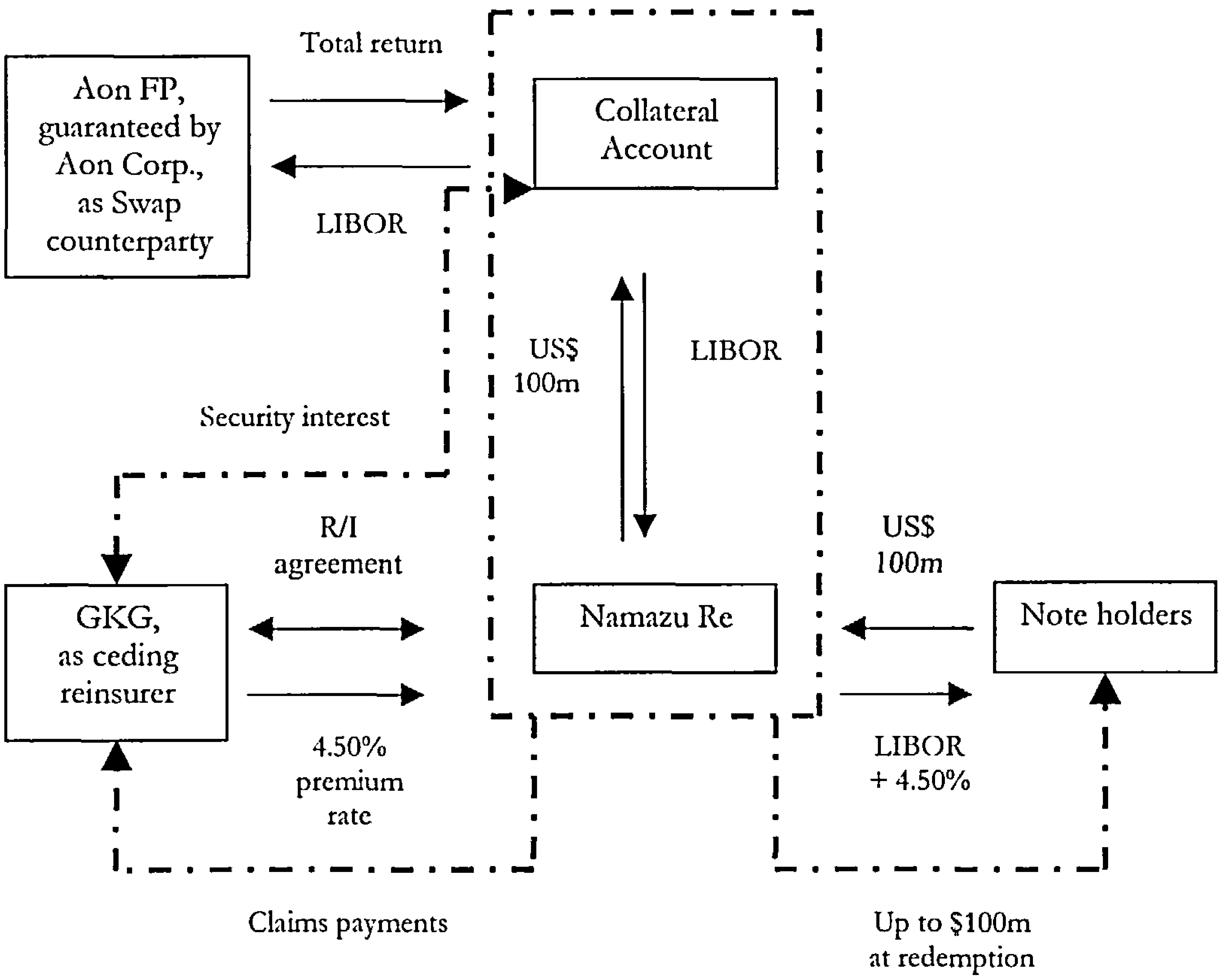
10% quota share in the coverage layer provided by the notes. As to extension risk in the notes, if adjusted covered losses reach 75% of the attachment point, SCOR may elect to delay maturity by 18 months.

Namazu Re is a special-purpose Cayman Islands Class B reinsurer that in 1999 issued $100 million five-year notes backing a reinsurance agreement with German reinsurer Gerling-Konzern Globale Ruckversicherungs-AG (Gerling). This agreement will provide payments to Gerling following certain earthquakes in Japan. Namazu Re was capitalised by issuing $5,000 in shares to a Cayman Islands charitable trust.

Under the terms of the reinsurance agreement, Gerling will cover the expenses of Namazu Re in connection with setting up the various vehicles and compensating the various service providers involved in the transaction (e.g., administrator, indenture trustee, and verification agent), in addition to the reinsurance premium, which provides the spread over LIBOR on the note coupons. Namazu Re has little capital and will rely on Gerling to pay its expenses, including the cost of any indemnifications provided by Namazu Re.

Namazu Re exists only to transfer the flows of a reinsurance treaty into bond flows; restrictions on its activities are important to isolate the collateral from other parties. Thus, Namazu Re's business will consist solely of the issuance of the notes and the execution and performance of the reinsurance agreement and related agreements and activities. Namazu Re will not engage in any other business, incur any other indebtedness, distribute its capital (other than its own liquidation after the notes have matured), or enter into any contract of insurance or reinsurance other than that with Gerling (see Figure 5.2).

Figure 5.2 Structure of the Namazu Re transaction



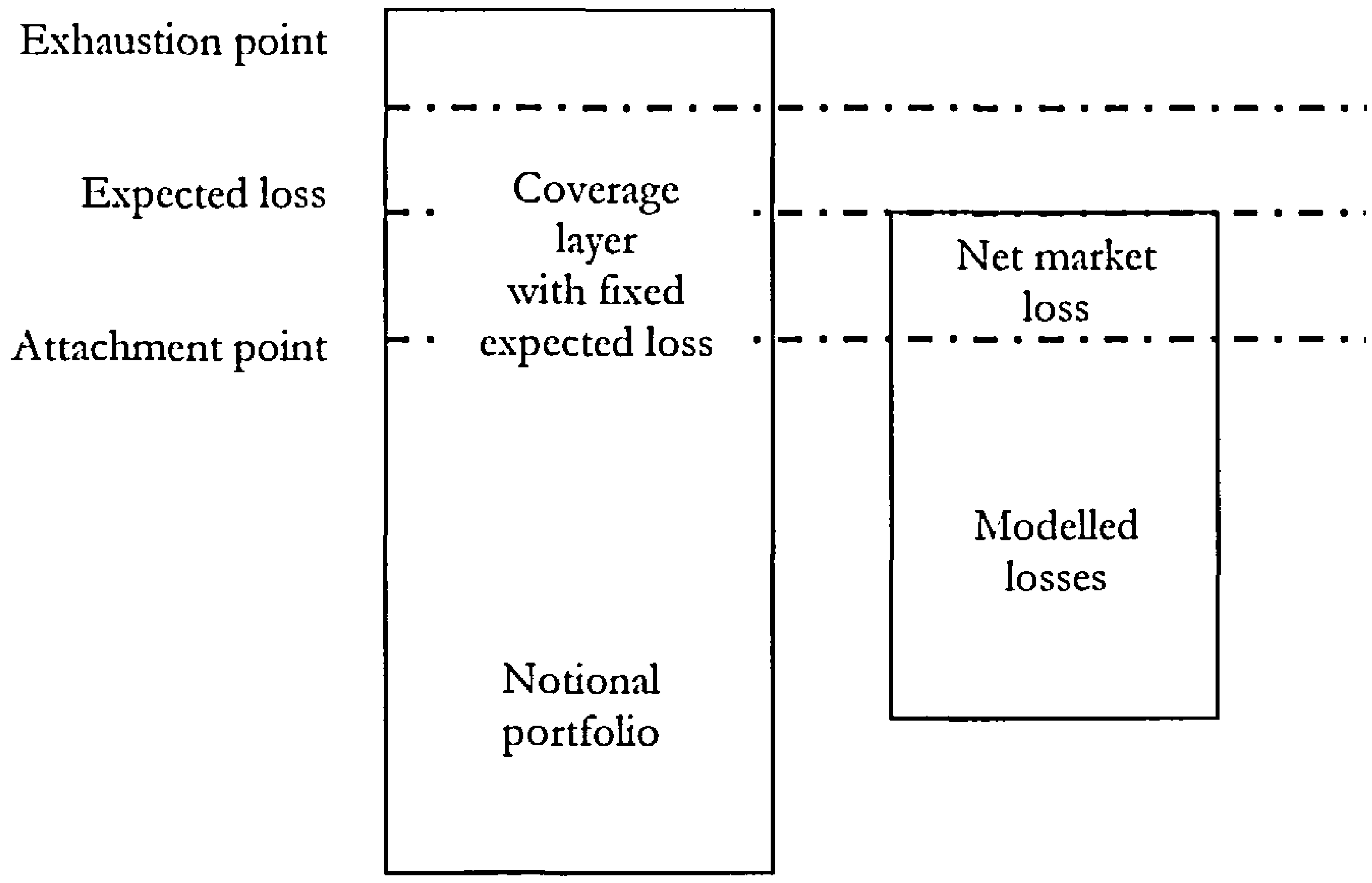
The note proceeds were deposited into a collateral account assigned by a deed of charge (under English law) to an indenture trustee and then invested in medium-term U.S. government and agency paper, high grade commercial paper, and AAA securities with a weighted average life of no more than five years.

Namazu Re also entered into a total-return swap with a AA counterparty, the primary purpose of which was to convert the return on the collateral account into a LIBOR coupon and to guarantee principal of the assets in the collateral account. The total return swap provides an enhancement of return on the collateral account. However, without the principal guarantee of AON Corp., the assets would have to be invested in short-term funds. In this transaction, investors are relying on the credit of AON Corp. for the LIBOR coupon and the return of principal before the assets in the collateral account mature.

The transaction was intended to mimic but not depend precisely on the indemnified losses of Gerling, so can be viewed as a synthetic indemnification. Semi-annually, Gerling will update the exposure data used by EQECAT, the peril modelling firm that assessed the transaction. EQECAT will then reset the layer covered by the reinsurance agreement, so that ex ante the transaction presents investors with a constant risk profile through time.

As a result, based not on a portfolio of policies but on a notional portfolio of locations and policy terms stored in EQECAT's Japanese earthquake model, EQECAT will calculate the loss level coinciding with a 1% per annum probability of attachment (the notional portfolio is denominated in Yen, though the bonds are in U.S. dollars; the translation is irrespective of currency exchange rates). EQECAT will reset the attachment point (the lower end of the layer) to maintain that 1% probability whenever there is a meaningfully large quake or whenever Gerling would like to update the notional portfolio. Initially, the attachment point was set to 9.30 billion Yen. EQECAT then sets the exhaustion point (the top of the layer) high enough that the layer running between the attachment point and the exhaustion point has an annual expected loss of 0.75%. The initial estimate was 12.583 billion Yen. Initially, a spike in damages up to that exhaustion point had a per annum probability of 32% (see Figure 5.3).

Figure 5.3 General Scheme of the Coverage Layer



The initial coverage layer is the 3.283 billion Yen, running from 9.300 billion Yen (attachment point) to 12.583 billion Yen (exhaustion point). The per-annum probability is 1% that damages on the notional portfolio reach 9.300 billion Yen and attach the note principal. The per annum probability is 0.32% that damages on the notional portfolio reach 12.583 billion Yen and exhaust the collateral account. The per annum expected loss is 0.75% of the notional portfolio, hence 0.75% of the note principal, since investors have a proportional interest (quota share) in the coverage layer.

5.8 Conclusions

Currently, most insurance securitisations address wind, earthquake, and mortality risks, with some issuance in auto lease residual values and interest in covering personal auto liability. This focus had been the result of specific reinsurance capacity constraints and was driven by the search for coverage.

In the future growing interest in the issuance of bonds tied to more generalised weather risks, such as excess heat or cold, excess rainfall or snowfall, and drought could be developed. These covers are not widely available at the moment, though many industries have large weather exposures, including agriculture, energy, property insurance, retailing, ski and other resorts, manufacturing, and real estate. According to the Chicago Mercantile Exchange, where energy-, weather-, and agriculture-related futures contracts are traded, about 20% of the U.S. GDP is vulnerable to weather or natural hazards.

Interest in life insurance securitisations will also remain strong. However, the use of assignments of life insurance policies in the secondary markets has been somewhat constrained by regulatory concerns. The NAIC has worked to address the possibility of unfair dealing in the assignment of elder life settlements. It is clear that legislation and oversight will pave the way for substantial growth in the securitisation of life insurance.

Moreover, different stakeholders can also join forces in promoting new financial instruments to supplement reinsurance for protecting insurers against catastrophic losses. The challenge is to convince investors that their chances of suffering large losses are relatively small compared to the expected return on their investment. This process is not any easy one, particularly if the investment community is unfamiliar with the types of risks against which they would be providing protection. The ambiguity associated with estimating future losses and the conflicts between experts on their assumptions for developing catastrophe models leave investors somewhat confused about what they are getting themselves into if they decide to commit funds to some of these new financial instruments.

Moreover, there is an opportunity to evaluate alternative strategies for managing the risks from natural disasters by taking advantage of a set of new developments in the areas of risk assessment, information technology, and catastrophe modelling. Turning first to risk assessment, by merging information derived from past records of earthquakes and hurricanes with an increased understanding of the characteristics of these hazards, scientists have been able to reduce the uncertainty about forecasting future events. With respect to damage estimation, engineers can better characterise the performance of different types of structures during hurricanes of different wind speeds and earthquakes of different magnitudes and intensities.

On the information technology side, the development of faster and more powerful computers enables investors and modellers to examine extremely complex phenomena in ways that were impossible even five years ago. Large databases can easily be stored and manipulated so that large-scale simulations of different disaster scenarios under alternative policies can now be undertaken.

New advances in catastrophe modelling provide an opportunity to combine scientific risk assessments with historical records to estimate the probabilities of disasters of different magnitudes and the resulting damage to the affected region. A catastrophe model is the set of databases and computer programs designed to analyse the effect of different scenarios on hazard-prone areas. The information can be presented in the form of expected annual losses based on simulations run over a long period of time (e.g., ten thousand years) or the effect of specific events (e.g., worst case scenarios). Several firms have developed catastrophe models and provide detailed analyses of their databases to the various parties concerned with these risks (e.g., insurance companies, reinsurers, government agencies, and disaster-prone communities).

Finally, catastrophe securities could be seen as a miniature reinsurer. The security is a package of catastrophe exposure and the capital needs to be set aside to cover this exposure without any credit risk. With the issuance of a catastrophe security, the market is presented with a very transparent type of deal. Assuming that all the information is given in the offering memorandum, investors can work out the probability of loss with the same accuracy as traditional reinsurers would be able to. Another reason to believe that this estimation of the probability is quite accurate is that the aggregation of information by an infinite number of market participants (a good proxy for capital markets) leads to the best achievable result based on all available information.

Having established the probability of loss, the market then prices in the cost of the capital, that in this case needs to be reserved for only one specific catastrophe event as described in the offering documents. The price for the risk is hence the market price. By definition this price must be very close to the cost of capital of a reinsurer that predominantly writes catastrophe business. The fact that the general level of catastrophe premium in the traditional reinsurance market is lower than the capital markets indicates value destruction (even when adjusted for the counter party credit risk).

The lack of focus on these key issues, which have had a profound impact on the development of financial management in modern business, has resulted in value destruction on a massive scale. Future research needs to address these facts and provide solutions that will price every little piece of balance sheet that an underwriter wants to put to work, to avoid further erosion of the equity value of the industry.

CONCLUSIONS AND FUTURE RESEARCH

The scope of this thesis is to investigate the different ways that capital is employed by insurance companies in their business and the various implications in a regulatory and risk measurement and management framework.

It has been identified that the current system used by insurance companies to report their regulatory level of capital is to a certain extent inadequate. In comparing the current European solvency margin system with the US risk-based system, one must conclude that the European model is inadequate. How should one proceed to draw the best from the risk-based models which are conceptually superior, without having the complexities, rigidities and costs of these systems? There is no simple answer to this question. One solution that suggests itself is a more collaborative approach between regulators and company management. Under such a system the regulatory authority would define risk models that would be acceptable, but would leave the detailed modelling to the companies themselves, with the oversight of an approved internal audit function within the company. Hence it would be possible to combine sufficient rigour without undue cost and inflexibility. An internal audit team combining actuarial, accounting, underwriting and economic expertise would seem appropriate for non-life insurance; perhaps an actuary alone would be sufficient for life insurance, since the task is less complex. The role of the regulatory authority would be to approve the models and to monitor the output of the models against benchmark standards.

Although the current European solvency margin system may appear crude compared to the US risk-based system, with weak theoretical foundations, its application is very straightforward. The dilemma between objective correctness and functionality is resolved in favor of the latter (Farny, 1997). The US risk-based capital model is significantly more tailor-made to an individual insurer risk profile, although its theoretical foundations are far from strong. It is also harder to apply in practice and imposes high internal compliance costs on insurers. There is a general skepticism as to whether the US-style risk-based model could be transferred to Europe. And this skepticism can be found among European supervisors in the Muller Report. Moreover, the US system is not only complex but is also very prescriptive even for insurance supervisors, since it has a detailed set of regulatory responses that are triggered if the capital base of an insurance company falls below the minimum risk-based capital level. One can argue that this degree of prescription builds undue rigidities into the regulatory process.

Regulatory capital is of course just one of a number of regulatory tools, alongside other supervision techniques and regulations covering conduct of business rules and client money. Ideally, the marginal costs and marginal benefits of all these tools applied in combination need to be judged, to try to decide on the most efficient mix.

Additionally, the imposition of capital rules in isolation is of little value if the regulator cannot be sure that the insurer has adequate systems to monitor and measure the risks that the capital standards are intended to limit, and that the firm's management are honest and competent. As well as choosing the right balance of regulatory tools, regulators need to take account of other mechanisms which are not part of the regulatory armoury, but which can also help them achieve their objectives. Market discipline is one important example: if the market is in a position to judge that a firm is weakly capitalised or poorly run, it may penalise the firm in various ways, such as discounting the firm's share price. And less tangibly, evidence of weak management can damage a company's reputation, which may make it harder for it to write new business. All these factors impose incentives, in different degrees on different companies, to operate in a sound and efficient manner, and to hold capital as a cushion against future losses.

These are some of the lessons that the Stage 2 of the EU solvency exercise. But it is clear that insurance supervisors and the insurance industry will also need to work together to construct a workable and cost-effective system for solvency of assessment, including capital adequacy, if the new computer-based risk modelling techniques are to be used. The fruitful interaction of regulators and the insurance industry has a good precedent in Europe: the development of the 3rd Insurance Directives. But what is also clear is that the current rethinking of capital adequacy and solvency within the banking sector by the Basel Committee on Banking Supervision will also have a key influence, as banking regulation itself moves away from its own static risk-based capital rules. A new solvency framework, including capital standards, is likely to emerge which has common features across the financial services sector, not least because of the need for level competitive playing-fields, as product and corporate convergence within the sector continues. However, the framework must be able to take into account the particular characteristics of the risks facing non-life insurers and those facing life insurers, if it is be effective.

Furthermore, another objective of this study is to investigate the presence and causes of the underwriting cycle in the United Kingdom and reveals several interesting findings.

The results of the autoregressive model largely support the existence of the underwriting cycle in the UK because underwriting cycles are found in the aggregate sector and all the lines that were tested.

The analysis of premium changes provides some support for the rational expectations/institutional intervention hypothesis although it is not able to gather enough evidence for the hypothesis that composite data collection, regulatory, policy renewal and accounting lags (Cummins and Outreville, 1987) have caused the underwriting cycle.

Moreover, the results generally differ from those found for the developed countries by Lamm-Tennant and Weiss (1997). This could be due partly to the fact that economic developments in these countries are different from the UK. It could also be due partly to the different level of regulatory control prevailing in these countries.

As UK is deemed to continue as the largest insurance and reinsurance financial centre, our findings pertaining to the underwriting cycle in UK would be useful to the existing insurers as well as those seeking to invest in the UK insurance market. One of the important findings from this study for the existing insurers and prospective entrants is that although the underwriting cycle does exist in the UK, the causes of it are different from those found in other countries. Therefore, they should take into account the differences when they enact measures to circumvent the detrimental effects of the underwriting cycle in the UK.

One of the shortcomings of this study is that, due to the lack of data and information, the analysis of premium changes is not able to provide support for the hypothesis that the underwriting cycle is caused by the institutional lags as advocated by Cummins and

Outreville (1987). Further, since only five lines of the insurance industry are included in this study, future research can extend the underwriting cycle test and analysis of premium changes to more lines when data are available. The cycle length of each line could be further analysed by using the cycle period analysis model proposed by Lamm-Tennant and Weiss (1997). Future research in this area could be particularly interesting if there are differences in the institutional structure and regulatory oversight of the industry in the UK.

In the third chapter, we investigate how credit ratings are connected to the credit risk faced by insurance companies. Insurer credit ratings also have a direct impact on the cost of capital, since the primary source of debt capital to insurers is policy liabilities, and lower rated firms will likely have to sell their policies at lower prices compared to higher rated firms.

Since the cost of capital is a long-term concept, the intention would be to produce a figure that compensates equity investors over a long period of time. Although a company's cost of capital will change over time, the cost of capital for a company or an industry should be relatively stable from period to period unless there has been some dramatic structural reasons for the change.

An indirect way to determine the cost of capital would be to link the credit spreads of corporate bonds with the associated credit ratings for these bonds. As the credit spread compensates the holder of the debt instrument for expected losses, there should be a link between the credit spread and the credit rating class, given the fact that there exist ample evidence that rating categories indeed entail an indication of relative credit risk.

We have indeed shown that there exists a close relationship between credit rating classes and credit spreads of corporate bonds.

Furthermore, we analysed the relationship between credit ratings and the financial ratios that highlight a financial healthy insurance company. The following variables have been employed in our estimations:

$$X_1 = \frac{\text{Statutory capital}}{\text{Total assets}},$$

$$X_2 = \frac{\text{Net income}}{\text{Total assets}},$$

$$X_3 = \frac{\text{Net premiums written in long - tail business}}{\text{Total net premiums written}},$$

$$X_4 = \frac{\text{Net reserves}}{\text{Shareholder funds}},$$

$$X_5 = \frac{\text{Net claims}}{\text{Net premiums written}},$$

$$X_6 = \frac{\text{Net commissions}}{\text{Net premiums written}},$$

$$X_7 = \frac{\text{Reinsurance ceded}}{\text{Net premiums written}},$$

$$X_8 = \frac{\text{Stocks in common stock investments}}{\text{Total invested assets}}$$

The coefficient on the inverse measure of leverage (capital to assets) is negative and significant in our estimations. The significantly negative results are consistent with the hypothesis that higher leverage is associated with greater uncertainty, which in turn is associated with a negative impact on the ratings. Also consistent with the uncertainty hypothesis is the positive and significant sign of the coefficient on the profitability variable.

The coefficient on stocks in common stock investments is positive and significant in both cases. Moreover, the predicted sign on the reinsurance variable can be ambiguous, depending on whether the use of reinsurance is seen as reducing uncertainty by shifting risk, or increasing uncertainty by making the insurer's financial health dependent on the financial health of its reinsurers. The results support the first argument since the coefficient on the variable measuring the percent of business ceded to reinsurers is positive and significant.

The coefficient on the variable measuring the percentage of business in long-tail lines (variable X_3) is not significant. Surprisingly, the variable representing the percent of net reserves is also negatively related and not significant.

The ratio of net commissions to net premiums written determines the percentage of direct written premiums that is paid to salespersons. A high ratio would show a high commission rate, which may be the result of an affiliate agency, or at least the payment of commissions higher than those normally paid in the industry. The variable is negatively related to the overall rating of the insurance companies. This is consistent with the expectation of a lower rating due to excessive commissions to agencies or salespersons. Finally, the variable representing the amount of net claims as a percentage of the net premiums written is significantly negative in line with the stated expectation.

It is crystal-clear that there is a close connection between the cost of capital assumed by the insurance companies and the credit rating assigned to them by the credit agencies. This affects not only the profitability of the insurance company but also the amount of capital that the company can employ in each day-to-day business functions.

Future research in the field must be focused on different ways that credit ratings can provide an into-depth analysis of the financial health of insurance companies and what strategic decisions have to be made in order to gain a favorable rating that will provide an easier way in accessing capital and improve financial performance.

Finally, in determining the way that insurance companies can refund themselves, we found that securitisation provides a straightforward way in accessing the capital markets. Securitisation benefits borrowers by enhancing capital supply at competitive costs, it benefits investors by broadening the spectrum of investment options, and it benefits originators by generating underwriting fees and trading commissions.

Currently, most insurance securitisations address wind, earthquake, and mortality risks, with some issuance in auto lease residual values and interest in covering personal auto liability. This focus had been the result of specific reinsurance capacity constraints and was driven by the search for coverage.

Securitisation can create liquidity by generating assets or by converting illiquid assets to high-grade securities that increases the value of the underlying assets. Moreover, securitisation can be used as a tax savings tool in a number of ways: the tax shelter from realising the book loss of policyholder loans and the savings in equity tax for mutual insurers.

In the future growing interest in the issuance of bonds tied to more generalised weather risks, such as excess heat or cold, excess rainfall or snowfall, and drought could be developed. These covers are not widely available at the moment, though many industries have large weather exposures, including agriculture, energy, property insurance, retailing, ski and other resorts, manufacturing, and real estate.

Interest in life insurance securitisations will also remain strong. However, the use of assignments of life insurance policies in the secondary markets has been somewhat constrained by regulatory concerns. The NAIC has worked to address the possibility of unfair dealing in the assignment of elder life settlements. It is clear that legislation and oversight will pave the way for substantial growth in the securitisation of life insurance.

Moreover, different stakeholders can also join forces in promoting new financial instruments to supplement reinsurance for protecting insurers against catastrophic losses. The challenge is to convince investors that their chances of suffering large losses are relatively small compared to the expected return on their investment. This process is not any easy one, particularly if the investment community is unfamiliar with the types of risks against which they would be providing protection. The ambiguity associated with estimating future losses and the conflicts between experts on their assumptions for developing catastrophe models leave investors somewhat confused about what they are getting themselves into if they decide to commit funds to some of these new financial instruments.

There is an opportunity to evaluate alternative strategies for managing the risks from natural disasters by taking advantage of a set of new developments in the areas of risk assessment, information technology, and catastrophe modelling. Turning first to risk assessment, by merging information derived from past records of earthquakes and hurricanes with an increased understanding of the characteristics of these hazards, scientists have been able to reduce the uncertainty about forecasting future events. With respect to damage estimation, engineers can better characterise the performance of different types of structures during hurricanes of different wind speeds and earthquakes of different magnitudes and intensities.

On the information technology side, the development of faster and more powerful computers enables investors and modellers to examine extremely complex phenomena in ways that were impossible even five years ago. Large databases can easily be stored and manipulated so that large-scale simulations of different disaster scenarios under alternative policies can now be undertaken.

Finally, new advances in catastrophe modelling provide an opportunity to combine scientific risk assessments with historical records to estimate the probabilities of disasters of different magnitudes and the resulting damage to the affected region. A catastrophe model is the set of databases and computer programs designed to analyse the effect of different scenarios on hazard-prone areas. The information can be presented in the form of expected annual losses based on simulations run over a long period of time (e.g., ten thousand years) or the effect of specific events (e.g., worst case scenarios). Several firms have developed catastrophe models and provide detailed analyses of their databases to the various parties concerned with these risks (e.g., insurance companies, reinsurers, government agencies, and disaster-prone communities).

Future research should explore the different ways that new structured deals can be employed in adding value to insurance companies balance sheets. Since securitisation represents a relatively new financial tool for insurance companies, researchers should explore alternative solutions that insurance companies can employ for gaining additional capital coverage, including exchange traded derivatives and catastrophe-linked securities.

The pricing and modelling of these instruments represents a research challenge that will be beneficial for both researchers and practitioners.

REFERENCES

1. Adelman, Charles M., and Roger D. Lorence, 1989, Securitising commercial mortgages, in: Phillip L. Zweig, ed., The Asset Securitisation Handbook, Homewood, Ill.: Dow Jones-Irwin, 298-334
2. Adiel, R., 1996, Reinsurance and the management of regulatory ratios and taxes in the property/casualty insurance industry, Journal of Accounting and Economics, 22: 207-240
3. Agostino, Dominick J., and William J. Cosgrove, 1990, The securitisation solution, Best's Review, Health/Life edition, 90: 44, 45, 88
4. A.M. Best Company, 1991, Best's Insolvency Study: Property/Casualty Insurers 1969-1990, Best's Review – Property/Casualty Insurance edition
5. A.M. Best Company, 1992, Best's Insolvency Study: Life/Health Insurers 1976-1991, Best's Review – Life/Health Insurance edition
6. A.M. Best Company, 1999, Best's Insurance reports – Property/casualty edition, Oldwick, N.J.: A.M. Best
7. Angbazo, L.A., J. Mei, and A. Saunders, 1998, Credit spreads in the market for highly leveraged transaction loans, Journal of Banking and Finance, 22, 1249-1282.

8. Anthony, J.H., and K.R. Petroni, 1997, Accounting estimation disclosures and firm valuation in the property/casualty insurance industry, Journal of Accounting, Auditing and Finance, 12: 257-281

9. Altman, E.I., 1968, Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy, Journal of Finance, 589-609.

10. Altman, E.I., 1993, Corporate Financial Distress and Bankruptcy, 2nd edition, NY: John Wiley

11. Altman, E.I., 1995, Predicting Financial Distress of Companies: Revisiting the Z-Score and Zeta Models, Working Paper, NYU Salomon Centre.

12. Altman, E.I., T.K.N. Baidya, and L.M.R. Dias, 1977, Assessing Potential Financial Problems for firms in Brazil, Working Paper, NYU Salomon Centre

13. Altman, E.I., R.G. Haldeman, and P. Narayanan, 1977, ZETA analysis: A new model to identify bankruptcy risk of corporations, Journal of Banking and Finance, 1: 29-54

14. Altman, E.I., and V.M. Kishore, 1997, Defaults and Returns onHigh-Yield Bonds: Analysis through 1997, Working Paper, NYU Salomon Centre

15. Ambrose, J.M. and J.A. Seward, 1988, Best's Rating, Financial Ratios and Prior Probabilities in Insolvency Prediction, Journal of Risk and Insurance, 229-244.

16. Arak, Marcelle and Patrick J. Corcoran, 1996, Yields on privately placed debt: examining the behavior, Journal of Portfolio Management, Spring, 88-96.

17. Balestra, P., and Nerlove, M., 1966, Pooling Cross Section and Time Series Data in the estimation of a dynamic model: The demand for natural gas, Econometrica, 34, 585-632.

18. Bank of International Settlements, 1999, A new Capital Adequacy Framework: a Consultative Paper . Basel Committee on Banking Supervision, Basel, June

19. Beaver, W.H., 1967, Financial Ratios as Predictors of Failure, Journal of Accounting Research, 4 (suppl.): 71-111.

20. Berger, Lawrence A., 1988, A model for the underwriting cycle in the property/liability insurance industry, Journal of Risk and Insurance, 55 (2): 298-306

21. Berger, L.A., J.D. Cummins, and S. Tennyson, 1992, Reinsurance and the liability insurance crisis, Journal of Risk and Uncertainty, 5: 253-272

22. Berger, A.N., Herring, R.J., Szego, G.P., 1995, The role of capital in financial institutions, Journal of Banking and Finance, 19: 393-430.

23. Black, Fisher, and Robert Litterman, 1991, Global asset allocation with equities, bonds, and currencies, Goldman, Sachs & Co.

24. Boardman, C.M., and R.W. McEnally, 1981, Factors affecting seasoned corporate bond prices, Journal of Financial and Quantitative Analysis, 16, 207-226.

25. Borch, K., 1974, Capital markets and the supervision of insurance companies, Journal of Risk and Insurance, 41: 397-405

26. Brockett, P.L., Cooper, W., Golden, L., and Utai Pitaktong, 1994, A Neural Network Method for Obtaining an Early Warning of Insurer Insolvency, Journal of Risk and Insurance

27. Browne, M.J., and Robert E. Hoyt, 1999, Economic and Market Predictors of Insolvencies in the Property-Liability Insurance Industry, Journal of Risk and Insurance, 62, No.2, 309-327

28. Browne, M.J., and Robert E. Hoyt, 1995, Economic and Market Predictors of Insolvencies in the Life-Health Insurance Industry, Journal of Risk and Insurance, 66, No.4, 643-659

29. Canter, Michael S., Cole, Joseph, B., and Richard L. Sandor, 1997, Insurance derivatives: A new asset class for the capital markets and a new hedging tool for the insurance industry, Journal of Applied Corporate Finance, 10 (3), 69-83

30. Cantor, R., and F. Packer, 1997, Difference of opinion and selection bias in the credit rating industry, Journal of Banking and Finance, 21: 1395-1417

31. Carlstrom, C., and K. Samolyn, 1993, Examining the microfoundations of market incentives for asset backed lending, Federal Reserve Bank of Cleveland, Economic Review, 29 (1), 27-38

32. Cooley, P.L., 1975, Bayesian and Cost Considerations for Optimal Classification with Discriminant Analysis, Journal of Risk and Insurance, 277-287.

33. Conference of the Insurance Supervisory Services of the Member States of the European Union, 1997, Report on Solvency of Insurance Undertakings. Brussels: the European Commission, April.

34. Credit Suisse Financial Products, 1997, Credit Risk Plus, Technical Document, London/New York

35. Cummins, David J. and J. Francois Outreville, 1987, An international analysis of underwriting cycles, Journal of Risk and Insurance, 54 (2): 246-262

36. Cummins, J.D., and R.A. Derrig, 1989, Financial models of insurance solvency, Norwell, MA.: Kluwer Academic Publishers

37. Cummins, David J., Scott E. Harrington, and Robert W. Klein, 1992, Cycles and crises in property/casualty insurance: Causes and implications for public policy, Kansas City, Mo.: NAIC

38. Cummins, J. David, and Patricia M. Danzon, 1992, Price shocks and capital flows in liability insurance, in: J. David Cummmins, Scott E. Harrington, and Robert W. Klein, eds., Cycles and crises in property/casualty insurance: Causes and implications for public policy, Kansas City, Mo.: NAIC

39. Cummins, J. David, Harrington, S.E., and Greg Niehaus, 1993, An economic analysis of RBC requirements in the Property-Liability insurance industry, Journal of Insurance Regulation, 11: 427-447.

40. Cummins, J. David, Harrington, S.E., and Robert W. Klein, 1995, Insolvency experience, Risk-Based Capital, and Prompt Corrective action in Property-Liability Insurance, Journal of Banking and Finance, 19: 511-527.

41. Cummins, J.D., and D.W. Sommer, 1996, Capital and risk in property/liability insurance markets, Journal of Banking and Finance, 20: 1069-1092

42. Cummins, J.D., and P.M. Danzon, 1997, Price, financial quality, and capital flows in insurance markets, Journal of Financial Intermediation, 6: 3-38

43. Cummins, J. David, Grace, M., and Richard D. Phillips, 1999, Regulatory solvency prediction in property-liability insurance: Risk-Based Capital, Audit Ratios, and Cash Flow Simulation, Journal of Risk and Insurance, 66: 417-458.

44. Deakin, E.B., 1972, A Discriminant Analysis of Predictors of Business Failure, Journal of Accounting Research, 10 (1): 169-179.

45. Denenberg, H., 1967, Is A-Plus Really a Passing Grade?, Journal of Risk and Insurance, 371-384.

46. Dickinson, G.M., 1997, Some Issues in Risk-Based Capital, The Geneva Papers on Risk and Insurance, No.82.

47. Doherty, N., and S. Tinic, 1981, Reinsurance under conditions of capital market equilibrium: A note, Journal of Finance, 36: 949-953

48. Doherty, Neil A., and Han Bin Kang, 1988, Interest rates and insurance price cycles, Journal of Banking and Finance, 12 (2): 199-214

49. Doherty, N., 1989, On the Capital structure of insurance firms, in: J.D. Cummins and R.A. Derrig, eds., Financial models of insurer solvency, Norwell, MA: Kluwer Academic Publishers

50. Doherty, Neil A., and James Garven, 1992, Insurance cycles: Interest rates and the capacity constraint model, Journal of Business, 68 (3): 383-404

51. Doherty, Neil A., 1997, Financial innovation in the management of catastrophe risk, Journal of Applied Corporate Finance, 10 (3), 84-95

52. Duffee, G.R., 1998, The relation between Treasury yields and corporate bond yield spreads, Journal of Finance, 53 (6), 2225-2242

53. Eck, James R., 1982, Detecting Financially Troubled Property-Liability Insurers, Journal of Business Research, 445-458.

54. Ederington, L.H., 1986, Why split ratings occur, Financial Management, 15:37-47

55. EU Insurance Committee, 2000. Reply to the IASC Issues Paper on Insurance Accounting, Brussels: European Commission, 20th July. DGMARKT/C2ULD(2000)404

56. European Commission: Internal Market DG, 1999, "Solvency Margin Review: a Working Document", December 21st, 1999. Markt/2067/99/Rev.2-EN

57. Evans, C.K., 1968, Basic Financial Differences of Substandard Automobile Insurers, Journal of Risk and Insurance, 489-513.

58. Farny, D., 1997, Security of Insurers: The American Risk-Based Capital Model versus the European Model of Solvability for Property and Casualty Insurers, The Geneva Papers on Risk and Insurance, No.82.

59. Fields, Joseph A., and Emilio C. Venezian, 1989, Interest rates and profit cycles: A disaggregated approach, Journal of Risk and Insurance, 56 (2): 312-319

60. Foley, T., and M. Foley, 1997, Alternative financial ratios of the effects of securitisation, Moody's Investor Services Global Credit Research, 3-15 September

61. Fons, Jerome S., 1994, Using default rates to model the term structure of credit risk, Financial Analysts Journal, 50 (5), 25-32.

62. Fridson, M.S., and J.G. Jonsson, 1995, Spread versus Treasuries and the riskiness of high-yield bonds, Journal of Fixed Income, 79-88.

63. Froot, Kenneth A., Murphy, Brian S., Stern, Aaron B., and Stephen E. Usher, 1995, The emerging asset class: Insurance risk, Guy Carpenter & Company Inc.'s Review of catastrophe exposures and the capital markets

64. Froot, Kenneth A., 1997, The limited financing of catastrophe risk: An overview, NBER Working Paper no. 6025, in K. Froot, ed., The Financing of Catastrophe Risk, 1999, University of Chicago Press, pp. 1-22.

65. Froot, Kenneth A. and Mark Seasholes, 1997, USAA: Catastrophe risk financing, Harvard Business School, Working paper no. N9-278-007

66. Fuller, Wayne A., 1976, Introduction to statistical time series, NY: John Wiley

67. Goldstein, A., 1996, What's in rating? Agency approaches to analysing MBS, Journal of Fixed Income, 5 (4), 61-73

68. Grace, M., Harrington, S.E., and Robert W. Klein, 1993, Risk-based capital standards and insurer insolvency risk: an empirical analysis, Paper presented at the 1993 Annual Meeting of the American Risk and Insurance Association.

69. Grace, Martin F., and Julie L. Hotchkiss, 1995, External impacts on the property/liability insurance cycle, Journal of Risk and Insurance, 62 (4): 738-754

70. Greenbaum, Stuart I., and Anjan V. Thakor, 1987, Banking funding modes: Securitisation versus deposits, Journal of Banking and Finance, 11: 379-401

71. Gron, Anne, 1992, Capacity constraints and cycles in property/casualty insurance markets, Journal of Law and Economics, 25: 110-127

72. Gron, Anne, 1994, Insurance markets with capacity constraints, Working paper, Chicago, Il: University of Chicago

73. Harmelink, Philip J., 1974, Prediction of Best's General Policyholders' Ratings, Journal of Risk and Insurance, 41: 621-632.

74. Harrington, Scott E. and Jack M. Nelson, 1986, A Regression-Based Methodology for Solvency Surveillance in the Property-Liability Insurance Industry, Journal of Risk and Insurance, 53: 584-605.

75. Harrington, S.E., and P.M. Danzon, 1994, Price cutting in liability insurance markets, Journal of Business, 67: 511-538

76. Harrington, Scott and Greg Niehaus, 1999, Basis risk with PCS catastrophe insurance derivative contracts, Journal of Risk and Insurance, 66 (1), 49-82

77. Hausman, J.A., 1978, Specification tests in Econometrics, Econometrica, 46, 1251-1271.

78. Hausman, J.A., and Taylor W.E., 1981, Panel data and unobservable individual effects, Econometrica, 49, 1377-1398.

79. Harvey, David M.W., 1991, Securitisation goes international, The Bankers Magazine, 174: 25-29

80. Helwege, J., and C.M. Turner, 1997, The slope of the credit yield curve for speculative-grade issuers, Research paper FRBNY (9725).

81. Hershbarger, Robert A. and Ronald K. Miller, 1986, The NAIC Information System and the Use of Economic Indicators in Predicting Insolvencies, Journal of Insurance Issues and Practices, 9: 21-43.

82. Hooker, N.D., Bulmer, J.R., Cooper, S.M., Green, P.A.G., and P.H. Hinton , 1995, Risk-Based Capital in General Insurance, Working paper, The Institute of Actuaries

83. Hoyt, R.E. and H. Khang, 1999, On the demand for corporate property insurance, Journal of Risk and Insurance

84. Hsiao, C., 1986, Analysis of panel data, Econometric Society Monograph, Cambridge University Press.

85. Izvorski, I., 1997, Recovery ratios and survival times for corporate bonds, IMF Working Paper, WP/97/84.

86. Jarrow, Robert A., David Lando, and Stuart M. Turnbull, 1997, A Markov model for the term structure of credit risk spreads, Review of Financial Studies, 10 (2), 481-523.

87. Jones, A., Hayssen, H., and J. Schneider, 1995, Rating of residential mortgage-backed securities, Journal of Fixed Income, 4 (4), 12-24

88. Kahane, Y., C.S. Tapiero, and L.Jacques, 1986, Concepts and trends in the study of insurer's solvency, in: J.D. Cummins and R.A. Derrig, eds., Financial models of insurer solvency, Norwell, MA.: Kluwer Academic Publishers

89. Kaplan, R.S., and G. Urwitz, 1979, Statistical models of bond ratings: A methodological inquiry, Journal of Business, 52: 231-261

90. Kenny, R., 1967, Fundamentals of Fire and Casualty Insurance Strength, Kenny Insurance Studies, Fourth Edition.

91. Kim, Y., Anderson, D., Amburgey, T., and James C. Hickman, 1995, The Use of Event History Analysis to Examine Insurer Insolvencies, Journal of Risk and Insurance, 94-110.

92. Klein, R.W., 1995, Insurance Regulation in Transition, Journal of Risk and Insurance, 62: 363-404.

93. KMV, 1995, Credit Monitor Overview, San Fransisco: KMV Corporation

94. Kopff, Gary J., and Jeremy Lent, 1988, Securitisation: Taking the business off the balance sheet, The Bankers Magazine, 171: 14-18

95. Lamm-Tennant, Joan and Mary A. Weiss, 1997, International insurance cycles: Rational expectations/Institutional intervention, Journal of Risk and Insurance, 64 (3): 415-439

96. Lang, L., Poulsen, A., and R. Stulz, 1995, Asset sales, firm performance, and the agency costs of managerial discretion, Journal of Financial Economics, 37, 3-37

97. Litterman, Robert, and Thomas Iben, 1991, Corporate bond valuation and the term structure of credit spreads, Journal of Portfolio Management, 17 (3), 52-64.

98. Litzenberger, Robert H., Beaglehole, David R., and Craig E. Reynolds, 1996, Assessing Catastrophe Reinsurance-linked securities as a new asset class, Journal of Portfolio Management, Special issue, 76-86

99. Lockwood, L., Rutherford, R., and M. Herrera, 1996, Wealth effects of asset securitisation, Journal of Banking and Finance, 20, 151-164

100. Longstaff, F.A., and E.S. Schwartz, 1995, A simple approach to valuing risky fixed and floating rate debt, Journal of Finance, 44 (4), 1085-1097.

101. Maddala, G.S., 1971, Econometrics, 2nd edition, McGraw Hill

102. Major, John A., 1997, Breaking open the new frontier: The Guy Carpenter Cat index, The Risk Financier, 1 (June): 7-8

103. Mayers, David, and Clifford W. Smith Jr., 1982, On the corporate demand for insurance, Journal of Business, 55 (2), 281-296

104. Mayers, David, and Clifford W. Smith Jr., 1990, On the corporate demand for insurance: Evidence from the reinsurance market, Journal of Business, 63 (1), 19-40

105. Merton, R.C., 1974, On the pricing of corporate debt: The risk structure of interest rates, Journal of Finance, pp. 449-470

106. Millon, M.M., and A.V. Thakor, 1985, Moral hazard and information sharing: A model of financial information gathering agencies, Journal of Finance, 40: 1403-1422

107. Morgan, J.P., 1997, CreditMetrics, NY: Technical Document

108. Moody's Investors Service, 1998, Corporate Bond Defaults and Default rates, Moody's Special Reports, NY: Moody's Investors Services

109. Moon, C.G., and J.G. Stotsky, 1993, Testing the differences between the determinants of Moody's and Standard & Poor's Ratings, Journal of Applied Econometrics, 8: 51-69

110. National Association of Insurance Commissioners, 1993, NAIC property/casualty risk-based capital formula exposure draft (Kansas City, MO).

111. National Association of Insurance Commissioners, 1995, The NAIC financial regulation standards and accreditation program (Kansas City, Mo: NAIC).

112. Nelson, R.E., 1971, Attrition Among Stock Companies in the Property-Liability Company Exits, Journal of Risk and Insurance, 357-366.

113. Nickell, P., W. Perraudin, and S. Varotto, 1998, Stability of rating transitions, Paper presented at the Bank of England Conference on Credit Risk Modeling and Regulatory Implications, London, September 21-22

114. Niehaus, Greg and Andy Terry, 1993, Evidence on the time series properties of insurance premiums and causes of the underwriting cycle: New support for the capital market imperfection hypothesis, Journal of Risk and Insurance, 60 (3): 466-479

115. Okabe, R., 1998, The impact of asset securitisation: A perspective for investors, Journal of Applied Corporate Finance, 11, 97-108

116. Pavel, Christine, 1986, Securitisation, Federal Reserve Bank of Chicago, Economic Perspectives, 10: 16-31

117. Pavel, Christine, and David Phillis, 1987, Why commercial banks sell loans: An empirical analysis, Federal Reserve Bank of Chicago, Economic Perspectives, 11: 3-14

118. Pedrosa, M., and R. Roll, 1998, Systematic risk in corporate bond credit spreads, Journal of Fixed Income, 8, 7-26.

119. Pinches, George E. and James S. Trieschmann, 1974, The Efficiency of Alternative Models for Solvency Surveillance in the Insurance Industry, Journal of Risk and Insurance, 41: 563-577.

120. Pinches, G.E. and J.S. Trieschmann, 1977, Discriminant Analysis, Classification Results and Financially Distressed P-L Insurers, Journal of Risk and Insurance, 289-298

121. Pottier, S.W., 1997, Life insurer risk characteristics and the rating process, Journal of Insurance Issues, 20: 111-130

122. Pottier, S.W., 1998, Life insurer financial distress, Best's ratings and financial ratios, Journal of Risk and Insurance, 65: 275-288

123. Ramakrishnan, R.T., and A.V. Thakor, 1984, Information reliability and a theory of financial intermediation, Review of Economic Studies, 51: 415-432

124. Richardson, J. and Michael Stephenson, 2000, Some aspects of regulatory capital, Occasional Paper Series, FSA.

125. Sanvincente, A.S., and F.L.C. Bader, 1996, Filing for financial reorganisation in Brazil: Event prediction with accounting and financial variables and the information content of the filing announcement, Working Paper, Sao Paulo University, Sao Paulo, Brazil

126. Sarig, O., and A. Warga, 1989, Some empirical estimates of the risk structure of interest rates, Journal of Finance, 44 (5), 1351-1360.

127. Schwarcz, S.L., 1993, Structured finance: a guide to principles of asset securitisation, 2nd edition, Practicing Law Institute, NY

128. Schwartz, R., 1994, Good, better, best, ReActions, 14:34-35

129. Shante, Michael E., 1989, Securitisation: New strategies for managing surplus, The Financial Reporter, April: 4-17

130. Sommer, D.W., 1996, The impact of firm risk on property/liability insurance prices, Journal of Risk and Insurance, 63: 501-514

131. Thornton, John H., and Joseph W. Meador, 1977, Comments on the validity of the NAIC Early Warning System for predicting failures among P-L insurance companies, CPCU Annals, 30: 191-211.

132. Trieschmann, James S. and George E. Pinches, 1973, A Multivariate Model for Predicting Financially Distressed Property-Liability Insurers, Journal of Risk and Insurance, 40: 327-338.

133. Venezian, Emilio, 1985, Ratemaking methods and profit cycles in property and liability insurance, Journal of Risk and Insurance, 52: 477-500

134. Webb, Bernard L., 1992, The property and liability insurance industry, Insurance Investment Management Handbook

135. Winter, Ralph A., 1988, The liability crisis and the dynamics of competitive insurance markets, Yale Journal of Regulation, 5: 455-500

136. Winter, Ralph, 1994, The dynamics of competitive insurance markets, Journal of Financial Intermediation, 3: 279-315